

Improving Productivity in Bioreactors through Control of Cell Heterogeneity

by

Michael Vitelli

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Chemical Engineering

Waterloo, Ontario, Canada, 2022

© Michael Vitelli 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner

Mario Jolicoeur
Professor, Dept. of Chemical Engineering, École
Polytechnique de Montréal

Supervisors

Hector Budman
Professor, Dept. of Chemical Engineering, University of
Waterloo

Mark Pritzker
Professor Emeritus, Dept. of Chemical Engineering,
University of Waterloo

Internal Members

Evelyn Yim
Professor, Dept. of Chemical Engineering, University of
Waterloo

Marc Aucoin
Professor & Associate Chair, Dept. of Chemical
Engineering, University of Waterloo

Internal-external Member

Maud Gorbet
Director, Dept. of Biomedical Engineering, University of
Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

I conducted and designed experiments, did all the chemical and biochemical analysis, the mathematical modeling and parameter estimation and wrote the drafts of all the papers and thesis. Professor Hector Budman contributed to the design of experiments, helped interpret the experimental data and reviewed all the manuscripts and thesis. Dr. Melih Tamer provided guidance at Sanofi Pasteur, gave insight from other projects completed on site that was related to our work and reviewed the manuscripts and thesis. Professor Mark Pritzker helped conceptualize and reviewed the manuscripts and thesis. Parts of section 2.3 in Chapter 2 are shortened from a journal submission (Vitelli *et al.*, 2021). Drafts of papers from the results of Chapters 4 and 5 are in the process of being prepared/submitted.

Abstract

Whooping cough, also referred to as pertussis, is a highly contagious bacterial respiratory tract disease. At Sanofi Pasteur, the fermentation step in the manufacturing of the vaccine for pertussis involves a sequence of reactors of increasing volume in which the final cell population from one reactor in the train is used to inoculate the following reactor. A main challenge with this operation is that the yield of the vaccine antigens can be highly variable. In particular, pertactin, which is generated in low levels and is highly variable relative to that of the other antigens of the vaccine, has a highly variable production rate thus posing a major bottleneck to the overall productivity. Based on the findings of previous studies by our group (Zavatti, 2019), oxidative stress appears to be related to the variability in productivity of antigens and pertactin in particular. To explain the observed variability of the process, we also hypothesize that the time profiles of dissolved oxygen, pH, temperature and aeration rates during fermentation may not accurately capture the presence of highly stressed cells within the cell population since they only reflect averaged measures of the cell population at any given time. Instead, only cytometric analysis of the heterogeneity of the cell population can provide a correct measure of the level of stress and its impact on productivity. Also, based on the hypothesis that population heterogeneity influences overall productivity, it was further hypothesized that the evolution of the fermentation process depends on the heterogeneity of the inoculum used to start the fermentation process. Then, it is argued that the process can be improved by using an inoculum that is tailored by means of cell sorting.

The current work focuses on the impact of oxidative stress and the growth profile of *Bordetella pertussis* and heterogeneity of intracellular concentrations for four main areas of application:

1. investigation of the possible origins of the oxidative stress in the manufacturing process.
2. development of a metabolic model describing the effect of oxidative stress on the growth of *B. pertussis*.
3. development of a coupled population balance-oxidative stress model to relate the heterogeneity of intracellular concentrations to *B. pertussis* growth profiles.
4. development of a protocol to sort *B. pertussis* on the basis of surface antigen concentration with the purpose of re-culturing.

Zavatti (2017) found that a high reactive oxidative species (ROS) at the beginning of the fermentation is associated with a low pertactin yield. One of the purposes of developing a mechanistic model is to determine whether the low growth rate is caused by high ROS levels or the two are merely correlated. This information could help to identify operating conditions that lead to high ROS. The model of Himeoka and Kaneko (2017) was developed to describe the explain the general behaviour during the lag, exponential growth, stationary and death phases without resorting to detailed mechanisms. In this study, we have adapted and extended this model to understand and describe the relation between cell growth, oxidative stress and NADPH under different oxidative conditions during the pertussis vaccine production. In view of the differences between *B. pertussis* to other bacteria, a main goal of the study is to assess via flask studies and model predictions whether the ROS level is the key determinant of growth under different ROS-inducing conditions rather than other factors such as substrate (glutamate) inhibition. Evaluation of the root-mean-square-error (RMSE) and Akaike information index (AIC) showed that the fit of

this new oxidative stress model to the experimental data was considerably better than that of a Contois-based model. The AIC is a particularly useful measure of the trade-off between model dimensionality and predictability in this case since the oxidative stress model involves a greater number of parameters than the Contois model. Also, the model and experimental data verified that high ROS levels at the beginning of flask culture is correlated to low growth profiles but is probably not the cause of the reduced biomass concentration.

Population balance models (PBMs) were formulated to describe the evolution in time of the cell population in terms of growth and oxidative stress. Flow cytometry data was used to gain insight into the distribution of important quantities (e.g., cell size, intracellular concentrations of metabolites) over the entire cell population. A coupled population balance-oxidative stress model was developed to predict distributions in cell size and intracellular glutamate, ROS, NADPH and NADP⁺ concentrations in shake flask cultures of *B. pertussis*. The major advantage of using a PBM is that it accounts for the distributions and can predict the heterogeneity of the cell population with respect to experimental conditions that are averaged out in bulk models. When comparing the coupled population balance – oxidative stress model to the bulk oxidative stress model, it is apparent that the PBM provides much better predictions of the intracellular ROS concentration. We hypothesize that due to the nonlinear relations between cell growth and oxidative stress, intracellular and cell surface quantities can be better modeled with population balance models.

Flow cytometry sorting is an extension of flow cytometry that enables cells to be sorted based on any property measured via flow cytometry. The actual sorting operation occurs downstream from the detectors that measure the light scattering and fluorescence energy. One of the hypotheses of the current work was that productivity can be enhanced by sorting a population of highly producing cells followed by cultivation of the sorted population. To avoid a lengthy a

validation of the manufacturing process by mutating *B. pertussis* cells to find a high pertactin-producing strain, a sort based on epigenetics was analyzed. However, the sorted cell populations were not able to maintain the properties for which they were selected, although the protocol was able to select *B. pertussis* cells which grew at a faster rate than the control seed provided by Sanofi.

Acknowledgements

First and foremost, I would like to show my sincerest gratitude to my supervisor Professor Hector Budman. He continuously supported me, provided guidance and used his invaluable time to help me throughout my PhD. His dedication to his research is inspiring and his work ethic towards supervising his students is exceptional. Thank you for taking me on as a student; you are an amazing role model.

I gratefully acknowledge Dr. Melih Tamer for supporting me throughout my PhD program. His role at Sanofi as my supervisor was instrumental and he always provided me resources and help for my research. I would like to thank Professor Mark Pritzker for his help conceptualizing our work and his amazing editorial skills for writing the thesis and manuscripts.

Thank you to my colleagues especially Piyush Agarwal, Xin Shen and Mariana Carvalho, for sharing their experience and who have been key to support me during my PhD.

Without the support of my family members, this work would not have been possible. My endless gratitude towards my parents, Joe and Cindy for bestowing their unconditional love and affection. Also, thank you to my siblings Steven and Alyssa, for their support and friendship. I would also like to thank my partner Etta Gunsolus for her support and encouragement.

I would also like to thank Professor Evelyn Yim, Professor Marc Aucoin, Professor Maud Gorbet and Professor Mario Jolicoeur who graciously accepted to be my thesis readers. Their insight, comments, and corrections are greatly appreciated.

Lastly, I would like to thank Natural Science and Engineering Research Council (NSERC) and Sanofi Pasteur for funding my research.

Table of Contents

List of Figures	xiii
List of Tables	xxii
Chapter 1 Introduction	1
Chapter 2 Literature Review	6
2.1 Vaccines and Whooping Cough.....	6
2.2 Oxidative Stress	8
2.3 Fluorescence-Activated Cell Sorting (FACS)	9
2.3.1 Isolation of naturally occurring cell sub-populations	13
2.3.2 Isolation of artificially induced cell subpopulations	20
2.3.3 Cell surface display systems	24
2.3.4 Flow Cytometry Droplet Sorting	27
2.3.5 Microfluidic FACS (μ FACS)	28
2.4 Population Balance Models	30
Chapter 3 Materials and Methods	35
3.1 Media formulation	35
3.2 Shake flask fermentations	36
3.3 Sample measurements.....	37
3.3.1 Optical density (OD ₆₀₀).....	37
3.3.2 Glutamate concentration	37
3.4 Fluorescence spectroscopy.....	37
3.4.1 Equipment	37
3.4.2 NADPH measurement	38
3.5 Flow cytometry	40
3.5.1 Equipment	40
3.5.2 <i>B. pertussis</i> gating.....	40
3.5.3 Viability stain.....	41
3.5.4 Oxidative stress	43
3.5.5 Surface pertactin antigen concentration	43
Chapter 4 Modeling the Effect of Oxidative Stress on <i>B. pertussis</i> Fermentations	46
4.1. Introduction.....	46

4.2 Model Outline	50
4.3 Materials and Methods.....	57
4.3.1 <i>B. pertussis</i> batch fermentations	57
4.3.2 Optical density (OD).....	58
4.3.3 Glutamate concentration	58
4.3.4 Fluorescence spectroscopy.....	59
4.3.5 Flow cytometry	59
4.3.6 Model fitting	61
4.4 Results and Discussion	63
4.4.1 Trends of the control <i>B. pertussis</i> flask culture	63
4.4.2 Glutamate and substrate inhibition of <i>B. pertussis</i>	68
4.4.3 Effect of initial OD on oxidative stress in <i>B. pertussis</i>	69
4.4.4 Effect of starvation on the growth rate of <i>B. pertussis</i>	71
4.4.5 Effect of seed batches	74
4.4.6 Calibration and validation of the oxidative stress model.....	75
4.4.7 Comparison of oxidative stress model to Contois growth model	97
4.8 Oxidative stress model development	99
4.9 Modeling biomass production in train of reactors	103
4.5 Conclusions.....	108
Chapter 5 Modeling Heterogeneity Using a Coupled Population Balance-Oxidative Stress Model	109
5.1 Introduction.....	109
5.2 Model Outline	112
5.2.1 Population balance general form	112
5.2.2 Model Formulation	114
5.2.3 Coupling of the population balance and oxidative stress models	118
5.3 Materials and Methods.....	123
5.3.1 <i>B. pertussis</i> batch fermentations	123
5.3.2 OD, glutamate and ROS measurements.....	124
5.3.3 Model fitting	124
5.3.4 Probability density function.....	126
5.4 Results and Discussion	128
5.4.1 Trends of the flow cytometry distributions.....	128

5.4.2 Characteristic curves	131
5.4.3 Re-binning and projection grid	135
5.4.4 Calibration and validation of population balance model	139
5.4.5 Comparison of bulk and population balance models	157
5.4.6 PBM and train of reactors	163
5.6 Conclusions	168
Chapter 6 Flow Cytometry Sorting of <i>B. pertussis</i>	169
6.1 Introduction	169
6.2 Materials and Methods	172
6.2.1 Streptomycin resistant <i>B. pertussis</i> seed	172
6.2.2 Flow cytometry sorting protocol	173
6.2.3 Re-Culture Protocol	173
6.2.4 Sample Measurements	174
6.3 Results	175
6.3.1 Oxidative stress of antibiotic resistant strain of <i>B. pertussis</i>	175
6.3.2 Flow cytometry sorting	176
6.3.3 Epigenetic effects of re-cultured cells	178
6.4 Discussion	181
6.5 Conclusions	183
Chapter 7 Conclusions and Future Work	185
7.1 Conclusions	185
7.1.1 Oxidative stress in flask experiments	185
7.1.2 Oxidative stress model	187
7.1.3 Coupled population balance – oxidative stress model	188
7.1.4 Flow cytometry sorting of epigenetic populations	190
7.2 Next Steps	190
References	193
Appendices	217
Appendix A Method of Characteristics	217
Appendix B Population Balance Model and Experimental Contours	219

List of Figures

2-1	Schematic of flow cytometry sorting operation.....	12
2-2	Application map for isolating a cell type of interest from heterogeneous samples. Pre-treatment of the sample may include dissociation and debris removal. Cells are then suspended and stained, most commonly by surface markers. The sorted cells can be used in a variety of applications including disease modeling, drug testing and characterization (i.e., lineage and gene expression)	14
2-3	Application map for isolating a cell type of interest from artificially induced cell populations. Pre-treatment of the sample may include library creation using mutagenesis. Most commonly the markers are inserted intracellularly via a plasmid. The sorted cells can be used in a variety of applications including screening for high producing cell lines or desired gene editing and for studying metabolic networks.....	21
2-4	Cell surface display.....	24
2-5	Application map for sorting via surface display. Cells are incubated with the marker library and commonly sorted based on binding affinity. The desired markers can be used in a variety of applications including screening vaccine or antibody candidates and marker selection for further testing.....	25
2-6	Water/ oil/ water droplet creation.....	27
3-1	Raw fluorescence of (a) control <i>B. pertussis</i> fermentation supernatant sample and (b) blank sample of phosphate buffered saline (PBS, 10 nM, pH 7.2).....	38
3-2	Filtered fluorescence spectra for control <i>B. pertussis</i> fermentation supernatant samples at (a) 24 hours and (b) 48 hours.....	39
3-3	Gating protocol for <i>B. pertussis</i> flow cytometry measurements. The gates are used as follows: (a) noise exclusion gate, (b) size inclusion gate and (c) doublet exclusion gate.....	41

3-4	Viability staining of a <i>Bordetella pertussis</i> sample.....	43
3-5	Flow cytometry of (a) unstained <i>Bordetella pertussis</i> , (b) negative stain of <i>Bordetella pertussis</i> with a control antibody conjugated to FITC and (c) <i>B. pertussis</i> stained with pertactin surface binding antibody conjugated to FITC.	45
4-1	Schematic representation of the mechanism described in the model. Adapted from Himeoka and Kaneko (2017)	51
4-2	Glutathione reductase pathway for reducing oxidative stress in the form of hydrogen peroxide.....	54
4-3	Evolution of (a) optical density (OD), (b) extracellular glutamate, (c) extracellular NADPH and (d) intracellular ROS concentrations in a <i>B. pertussis</i> flask culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L.....	64
4-4	Change in extracellular NADPH concentration per unit biomass of a control <i>B. pertussis</i> flask culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L.....	65
4-5	Viability PI-based staining of a control <i>Bordetella pertussis</i> flask culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L after a) 24 hours, b) 36 hours, c) 42 hours, and d) 48 hours.....	67
4-6	Variation of OD during fermentation at different initial glutamate concentrations over the (a) full 50 hours and (b) first 18 hours of incubation.....	69
4-7	Biomass growth curves obtained in control and 12-hour starved secondary flasks with media initially containing (a) 10.1 g/L and (b) 6.6 g/L glutamate.....	72
4-8	Viability staining of a control and starved <i>Bordetella pertussis</i> flask cultures with an initial glutamate concentration of 10.1 g/L: a) control - 0 hours, b) starved - 0 hours, c) starved - 6 hours, d) starved - 24 hours, e) starved – 36 hours, and f) starved - 48 hours.....	73
4-9	Comparison of <i>B. pertussis</i> seeds batches made in 2015 and 2018. a) Biomass (OD ₆₀₀) comparison, b) intracellular ROS comparison, c) viability of the 2015 batch seed and d) viability of the 2018 batch seed.....	75

4-10	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.28 and initial glutamate concentration of 10.1 g/L.....	80
4-11	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.36 and initial glutamate concentration of 9.8 g/L.....	81
4-12	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.13 and initial glutamate concentration of 10.7 g/L.....	82
4-13	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.....	83
4-14	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.27 and initial glutamate concentration of 12.0 g/L.....	84
4-15	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.12 and initial glutamate concentration of 12.4 g/L.....	85
4-16	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.38 and initial glutamate concentration of 6.6 g/L.....	86
4-17	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.16 and initial glutamate concentration of 6.6 g/L.....	87
4-18	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> seed (2015) culture with an initial OD of 0.10 and initial glutamate concentration of 10.0 g/L.....	88
4-19	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> seed (2015) culture with an initial OD of 0.08 and initial glutamate concentration of 10.3 g/L.....	89

4-20	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a <i>B. pertussis</i> seed (2018) culture with an initial OD of 0.11 and initial glutamate concentration of 10.6 g/L.....	90
4-21	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a starved (12 hours) <i>B. pertussis</i> culture with an initial OD of 0.32 and initial glutamate concentration of 9.3 g/L.....	91
4-22	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a starved (12 hours) <i>B. pertussis</i> culture with an initial OD of 0.28 and initial glutamate concentration of 9.6 g/L.....	92
4-23	Model validation of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.....	93
4-24	Model validation of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH and (d) intracellular ROS of a <i>B. pertussis</i> culture with an initial OD of 0.25 and initial glutamate concentration of 10.7 g/L.....	94
4-25	Model validation of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH and (d) intracellular ROS a starved (12 hours) <i>B. pertussis</i> culture with an initial OD of 0.31 and initial glutamate concentration of 6.8 g/L.....	95
4-26	Comparison of the biomass-substrate model and starved <i>B. pertussis</i> experimental results for OD ₆₀₀	98
4-27	Sanofi production process for whooping cough vaccine. A seed flask is used to inoculate a train of reactors of sizes 20 L, 200 L and 2000 L.....	104
4-28	Model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for two <i>B. pertussis</i> seed batches. Biomass curve of a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor. Glutamate concentration curve of e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor.....	106
4-29	Model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for <i>B. pertussis</i> cultures with media of 10 g/L and 10.5 g/L glutamate. Biomass curve of a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor. Glutamate	

	concentration curve of e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor.....	107
5-1	Gridding the cell volume and ROS concentration flow cytometry data of <i>B. pertussis</i> flask culture experiments for initial conditions in the population balance model. The shaded region contains 940 cells with 7-8 intensity cell volume and 0.06-0.07 intensity ROS concentration.....	123
5-2	Effect of distribution (a) mean and (b) standard deviation on the dependence of probability density function $f(X)$ on cell volume.....	128
5-3	Comparison of the dependence of the partitioning rate on cell volume according to a normal distribution (red and yellow) and fitted interpolated function (blue) for the probability density function.....	129
5-4	<i>B. pertussis</i> flask culture measurements: (a) optical density (OD), (b) extracellular glutamate concentration, (c) average cell volume, (d) average intracellular ROS concentration, (e) cell volume distributions and (f) extracellular ROS concentration distributions of a culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L.....	131
5-5	Characteristic curves: (a) cell volume, (b) intracellular glutamate concentration, (c) intracellular NADPH concentration and (d) intracellular ROS concentration for different initial cell sizes at the same initial ROS concentration intensity of 0.055.....	133
5-6	Characteristic curves: (a) cell volume, (b) intracellular glutamate concentration, (c) intracellular NADPH concentration and (d) intracellular ROS concentration for different initial ROS concentrations with cells with the same initial size (8.5).....	134
5-7	Characteristic curves for the same initial value of $B = 0.055$ and same initial values of A , S and C and for different initial values of cell size X : variation of (a) cell volume and (b) size ratio with time.....	135
5-8	Characteristic curves with the same initial concentration of ROS ($B_o = 0.105$): (a) cell volume and (b) number of cells to examine the effect of the partitioning rate on cells in the characteristic curve.....	136

5-9	Projection window and moving window characteristic curves for cell volume and intracellular ROS concentration over 6 hours.....	138
5-10	Re-binning characteristic curves after 6 hours onto the projection window.....	139
5-11	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration for a <i>B. pertussis</i> culture with an initial OD of 0.28 and initial glutamate concentration of 10.1 g/L.....	140
5-12	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration of a <i>B. pertussis</i> culture with an initial OD of 0.12 and initial glutamate concentration of 12.4 g/L.....	141
5-13	Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration of a <i>B. pertussis</i> culture with an initial OD of 0.13 and initial glutamate concentration of 10.1 g/L.....	142
5-14	Model validation of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration of a <i>B. pertussis</i> culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.....	143
5-15	Model and experimental flow cytometry distributions of cell volume for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.28 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	147
5-16	Model and experimental flow cytometry distributions of intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.28 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	148
5-17	Model and experimental flow cytometry distributions of cell volume for a <i>B. pertussis</i> shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	149

5-18	Model and experimental flow cytometry distributions of intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	150
5-19	Model and experimental flow cytometry distributions of cell volume for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	151
5-20	Model and experimental flow cytometry distributions of intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	152
5-21	Model validation and experimental flow cytometry distributions of cell volume for a <i>B. pertussis</i> shake flask culture with 8.6 g/L glutamate in the media and starting optical density of 0.14 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	153
5-22	Model validation and experimental flow cytometry distributions of intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 8.4 g/L glutamate in the media and starting optical density of 0.14 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.....	154
5-23	Model flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.25 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	156
5-24	Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.25 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	157

5-25	Comparison in model fitting between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a <i>B. pertussis</i> culture with an initial OD of 0.25 and initial glutamate concentration of 10.1 g/L.....	161
5-26	Comparison in model fitting between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a <i>B. pertussis</i> culture with an initial OD of 0.12 and initial glutamate concentration of 12.4 g/L.....	162
5-27	Comparison in model fitting between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a <i>B. pertussis</i> culture with an initial OD of 0.13 and initial glutamate concentration of 10.1 g/L.....	163
5-28	Comparison in model validation between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a <i>B. pertussis</i> culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.....	164
5-29	Population balance model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for <i>B. pertussis</i> cultures with media of 10 g/L and 10.5 g/L glutamate. Evolution of biomass coming from a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor. Evolution of glutamate concentration coming from e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor.....	167
5-30	Population balance model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for <i>B. pertussis</i> cultures with media of 10 g/L and 10.5 g/L glutamate. Distributions of cell volume and intracellular ROS concentration coming from a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor for a media with 10 g/L glutamate. Distributions of cell volume and intracellular ROS concentration coming from e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor for a media with 10.5 g/L glutamate.....	168
6-1	Comparison between control and streptomycin-resistant (strepR) strains of <i>B. pertussis</i> : (a) number distribution of initial ROS concentration and (b) evolution of growth of secondary cultures.....	177

6-2	Sorting gates of <i>B. pertussis</i> for (a) side scattering (R9 – low SSC, R10 – high SSC) and (b) PRN antibody concentration (R9 – low PRN, R10 – high PRN).....	178
6-3	Tryptic soy agar plated with low PRN surface concentration population.....	179
6-4	Plate properties of (a) side scattering distribution and (b) surface PRN concentration of sorted cell colonies.....	180
6-5	Growth of secondary cultures for control and sorted <i>B. pertussis</i>	182
B-1	Model flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	221
B-2	Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	222
B-3	Model flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	223
B-4	Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	224
B-5	Model validation of flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 8.6 g/L glutamate in the media and starting optical density of 0.14 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	225
B-6	Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a <i>B. pertussis</i> shake flask culture with 8.6 g/L glutamate in the media and starting optical density of 0.14 at (a) 6 , (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.....	226

List of Tables

2-1	Studies on sorting of heterogeneous cell populations including their gating strategies...	17
2-2	Studies that use surface display and flow cytometry sorting.....	26
2-3	Studies based on droplet flow cytometry sorting.....	28
3-1	Media components and normalized concentration.....	35
3-2	Growth factor components and normalized concentration.....	35
4-1	Definitions of concentrations used in the model.....	35
4-2	Parameters defined in model. The numerical values of the parameters are determined using the model and calibration set in section 4.4.....	57
4-3	Comparison of OD and ROS levels and biomass doubling time over the first 12 hours of fermentation at different initial glutamate concentrations.....	69
4-4	Comparison of OD and ROS levels over the first 12 hours of fermentation at different initial concentrations of glutamate and half the initial OD level.....	71
4-5	List of calibration conditions for <i>B. pertussis</i> oxidative stress model.....	76
4-6	List of validation conditions for <i>B. pertussis</i> oxidative stress model.....	76
4-7	Parameter estimation of oxidative stress model.....	77
4-8	Root mean-squared error of measured variables for model calibration and validation. The range of each variable is also shown to compare with the RMSE.....	96
4-9	Parameter estimation for Contois model.....	98
4-10	Comparison between Contois kinetic and oxidative stress models.....	99
4-11	Comparison of the fit, robustness and predictability of the Contois and Monod models and exponents on the flux for optimal oxidative stress.....	102

4-12	Initial conditions of intracellular NADPH and ROS concentrations of the seed flasks to model the seed batches from 2015 and 2018 in the train of reactors.....	105
5-1	General form population balance model definitions.....	114
5-2	Population balance model parameter definitions.....	121
5-3	Initial concentrations of intracellular components.....	123
5-4	Calibration and validation conditions for coupled population balance and oxidative stress model.....	124
5-5	Root-mean-squared-error (RMSE) of measured variables for model calibration and validation. The range of each compound is also shown to compare with the RMSE....	144
5-6	Adjusted root mean-squared error (RMSE) of averages and standard deviations of the cell volume and intracellular ROS distributions for model calibration and validation and the range of values and error.....	145
5-7	Kullback-Liebler divergence based on the 2D cell volume and intracellular ROS distributions at all time points.....	155
5-8	Parameter estimates of bulk and PBM oxidative stress models.....	158
5-9	Comparison of RMSE between bulk oxidative stress and coupled PB models.....	159
6-1	Flow cytometry sorting population averages and ranges for side scattering and PRN surface concentration.....	178
6-2	Average plate properties of sorted cell colonies 2 days after plating the sorted populations.....	181

Chapter 1 Introduction

Whooping cough, also referred to as pertussis, is a highly contagious bacterial respiratory tract disease. Although it can be prevented through vaccination at an early age, it is one of the world's leading causes of vaccine-preventable death (Dominguez *et al.*, 2017). The vaccine produced by Sanofi Pasteur, called 5-component pertussis, is an acellular vaccine which contains 5 antigens: types 2 and 3 fimbriae, filamentous hemagglutinin, pertactin and pertussis toxin.

The fermentation step in the manufacturing of the vaccine involves a train of reactors of increasing volume in which the final cell population from one reactor in the sequence is used to inoculate the following reactor. One of the problems with this operation is that the yield of the vaccine antigens can be highly variable. In particular, the production of pertactin which is low and highly variable relative to that of the other antigens of the vaccine poses a major bottleneck to the overall productivity. Based on the findings of previous studies by our group (Zavatti, 2019), oxidative stress appears to be related to the variability in productivity of antigens and pertactin in particular. Oxidative stress can be generally quantified by the total amount of reactive oxidative species (ROS) which include different forms of oxidative species such as super-oxides and hydrogen peroxide. Zavatti observed a correlation between the ROS levels and growth rate of *Bordetella pertussis* and argued that the productivity of antigen production is reduced in the presence of high ROS levels due to a slower growth rate. In the current work, we hypothesize that mathematical modelling of ROS formation in the presence of changing conditions in the process combined with measurements of ROS can provide a basis for optimizing operating conditions to mitigate the levels of ROS and their impact on the process. The results of Zavatti (Zavatti, 2019) did not indicate whether ROS is the direct cause for the changes in growth and productivity or

whether the ROS level and antigen growth rate are merely correlated. Furthermore, Zavatti could not clearly identify which conditions in the manufacturing process cause high ROS levels. We hypothesize in the current study that a mathematical model can serve to clarify these questions and help to identify operating conditions that lead to low ROS.

Mathematical models that involve ROS have been reported, particularly in the context of antibiotic resistance (Farha and Brown, 2013; Kohanski *et al.*, 2010; Yang *et al.*, 2019). A bulk metabolic model describing the effect of oxidative stress on the growth of *B. pertussis* is presented in Chapter 5. The model has been developed based on a previous theoretical model by Himeoka and Kaneko (2017) focused on the effect of starvation and the stationary growth phase of cells. To assess the validity of the mathematical model for this specific *B. pertussis* pathway, we have conducted an experimental flask study under different conditions of oxidative stress induced by high initial glutamate concentrations, low initial inoculum and secondary culturing following exposure to starvation conditions.

To explain the observed variability of the process, we also hypothesize that the evolution profiles of dissolved oxygen, pH, temperature and aeration rates during fermentation may not accurately capture the presence of highly stressed cells within the cell population since they only reflect averaged measures of the cell population at any given time. Instead, only cytometric analysis of the heterogeneity of the cell population can provide a correct measure of the level of stress and its impact on productivity. Thus, we hypothesize that quantifying and possibly manipulating the cell population heterogeneity is crucial in order to monitor and optimize process productivity. Flow cytometry is a sensitive and powerful tool that uses fluorescence and light scattering to simultaneously measure the physical and chemical properties of cells suspended in a fluid. Since cytometry is a high throughput process that characterizes individual cells at very high

rates, it has the capability to detect rare events in real time. Several parameters per cell can be measured simultaneously depending on the number of detectors available. Fluorescence-activated cell sorting (FACS) is a particular type of flow cytometry that enables cells to be sorted downstream from the detectors based on the measured scattering or fluorescence energy. It is possible to sort cells based on any parameter that can be measured by flow cytometry, including oxidative stress (Boone *et al.*, 2017; Llufrío *et al.*, 2018), cell viability (Huizer *et al.*, 2017; Sánchez-Luengo *et al.*, 2017) and cell surface markers (Yuan *et al.*, 2015; Yim *et al.*, 2014).

Since flow cytometry can capture the distribution of important quantities (e.g., cell size, intracellular oxidative species) over the entire cell population, we have extended the bulk oxidative stress model to a population balance-based one to describe the evolution of these distributions over time which are then correlated to the growth to elucidate the effect that population heterogeneity has on the process. Population balance models contain a set of equations which describe the behavior of a population of particles based on the analysis of behavior of single particles in the cell population. In a bioreactor, these models provide a deterministic description of the dynamic evolution of cell distributions (Fredrickson *et al.*, 1967). Furthermore, we hypothesize that the propagation of this population heterogeneity across the train of bioreactors composing the manufacturing process may be an additional factor in the growth and associated final productivity of the process. In Chapter 5, we present a coupled population balance – oxidative stress model to relate the distributions of cell volume and intracellular reactive oxidative species (ROS) to the growth of *B. pertussis*.

Following the recognition that the propagation of cell population heterogeneity is a crucial measure of the variability of the process, we suggest that the ability to control the inocula by cell sorting may help improve the process outcomes and optimize the productivity of pertactin. *B.*

pertussis has a regulon system that involves the Bvg⁺ phase characterized by expression of adhesins and toxins and the Bvg⁻ phase where virulence factors are not expressed. In a study done by Sanofi, it was determined that the surface antibody concentration of pertactin is much higher during the batch phase than the fed-batch phase of the fermentation, reflecting a change from a Bvg⁺ to Bvg⁻ phase (Martinez de Tejada *et al.*, 1998; Moon *et al.*, 2017). In Chapter 6, we sort cells based on surface pertactin concentration in the stationary phase to determine whether the epigenetic phenotype is maintained in the daughter cells of the sorted populations which could potentially lead to a new cell line of higher pertactin-producing cells. Also, based on reports that cell scattering is correlated to oxidative stress and that the latter affects or correlates with growth, we sort a cell sub-population based on scattering and re-culture it to assess the effect of this sorting procedure on cell growth.

The overall goal of the current work is to identify sources of variation in the yield of pertactin in the manufacturing process at Sanofi Pasteur. Following from the above discussion, this work pursues the following research objectives:

1. investigation of the following possible origins of the oxidative stress in the manufacturing process: i) differing concentrations of media components, ii) differing levels of initial biomass and iii) different seed batches.
2. development of a bulk metabolic model describing the effect of oxidative stress on the growth of *B. pertussis*. The model includes a mechanism for inhibition of *B. pertussis* growth by ROS (inhibitor) and the role of NADPH (byproduct) as a quencher of ROS.
3. development of a coupled population balance-oxidative stress model to relate the distributions of cell properties of interest (e.g., oxidative stress and cell size) to the growth of *B. pertussis* in shake flasks.

4. development of a protocol to sort *B. pertussis* with the purpose of re-culturing. This protocol will be used to sort particular populations (e.g., cells with high surface concentration pertactin) to determine whether the epigenetic phenotype is maintained in the daughter cells of the sorted populations.

Following the above objectives, the novel contributions of this work are as follows:

1. development and solution of a bulk oxidative stress model to describe the relation between cell growth, oxidative stress and NADPH under different oxidative conditions.
2. formulation and solution of a population balance model with the method of characteristics and a re-binning time interval aimed at fitting the distributions of cell size and intracellular ROS concentration obtained from flow cytometry.
3. development of a coupled population balance – oxidative stress model to determine whether the inclusion of heterogeneity in the ROS concentration improves the predictability of the model.
4. development of a protocol to sort sub-populations of *B. pertussis* based on scattering and cell surface antigen and testing the potential of this procedure for enhancing growth or productivity.

Chapter 2 Literature Review

Section 2.3 contains parts of a published review paper on flow cytometry sorting (Vitelli *et al.*, 2021).

2.1 Vaccines and Whooping Cough

Vaccines are biological substances that provide active immunity against particular diseases. They provide the immune system with an agent which resembles the disease-causing microorganism. By imitating infection, they pre-emptively tune the immune system to develop antibodies to protect against the disease. Thus, vaccines are administered to induce the immune system memory function. The types of currently available vaccines can be grouped into four categories: inactivated (polio), attenuated (tuberculosis), acellular (whooping cough) and mRNA (covid19). Inactivated vaccines use inactivated microorganisms which have been destroyed by chemicals or radiation. Attenuated vaccines contain live microorganisms that have been cultured under conditions that disable their virulence. Acellular vaccines contain protein subunits called antigens which adhere and interact with a specific target cell. Antigens are targeted and bound by specifically tailored antibodies produced by the immune system of the organism. Once antibodies for a particular antigen are produced by the immune system, they are translated whenever the antigen is reintroduced into the organism. In the case of mRNA vaccines, the mRNA instructs the body's cells to spike the antigen found on the surface of the target cell. The rest of the process is the same as acellular vaccines once the antigen is translated.

Whooping cough, also referred to as pertussis, is a highly contagious respiratory tract disease. The causative agent of this illness is the gram-negative bacterium *B. pertussis*, which was

first isolated by Bordet and Gengou in 1906. Symptoms of whooping cough initially resemble the common cold followed at a later stage of the disease by uncontrollable violent coughing fits that are discernable by a typical "whoop" sound made while inhaling air. Without treatment, the duration of the disease can range from several weeks to several months; severe complications can lead to pneumonia, seizures, brain damage or death. Since children less than 1 year in age are more likely to suffer the worst symptoms and consequences, whooping cough remains one of the world's leading causes of vaccine-preventable deaths (Dominguez *et al.*, 2017). An estimated 50 million cases and 300,000 deaths are reported every year (World Health Organization).

The acellular pertussis vaccine produced by Sanofi Pasteur is referred to as five-component pertussis which are the antigens fimbriae (types 2 and 3), filamentous hemagglutinin, pertactin and pertussis toxin. Fimbriae are submicroscopic proteinaceous appendages that protrude from the cell surface. They are comprised of two major subunits that bind to sulphonated sugars located in the respiratory tract (Babu *et al.*, 2001). Filamentous hemagglutinin is a main appendage of *B. pertussis* and facilitates adherence of the cells to the respiratory epithelial cells. Pertactin is an outer membrane protein from the autotransporter family, which is also involved in adhesion to the host cell. Autotransporter secretion is one of the mechanisms developed by gram-negative bacteria and occurs when the protein is able to transport autonomously through the outer cell membrane (Leo *et al.*, 2012). Pertussis toxin is one of the major protein toxins secreted by *B. pertussis* and glycoconjugate molecules on the surface of target cells (Witvliet *et al.*, 1989). These antigens have been selected because they are able to elicit a strong immune response in both human and animal models (Locht, 1999; Noofeli, 2007).

2.2 Oxidative Stress

Oxidative stress occurs when an imbalance occurs between the production of free radicals and the capacity of the system to counteract these compounds (Zheng *et al.*, 2009). The free radicals are reactive oxidative species (ROS) such as singlet oxygen, hydroxyls, superoxides, peroxides, hydroperoxides and the non-radical hydrogen peroxide which are continuously produced during aerobic metabolism (Eruslanov and Kusmartsev, 2010). Whether ROS functions as signaling molecules or causes oxidative damage depends on the delicate equilibrium between its production and scavenging (Sharma *et al.*, 2012). ROS also serves as second messengers (i.e., intracellular signaling molecules secreted by the cell in response to exposure to extracellular signaling molecules) in a variety of cellular processes including those that help to tolerate various environmental stresses; however, in excess it can be harmful to proteins, lipids and DNA (Eruslanov and Kusmartsev, 2010).

NADPH is important for aerobic organisms to survive since it is a source of reducing equivalents that serve to protect against ROS. The enzymes NAD⁺ kinase and NADP⁺ phosphatase are known to regulate the levels of NAD⁺ and NADP⁺ and therefore play pivotal roles in controlling the communication between metabolic networks that produce NADH and NADPH and are involved in the mechanism of diffusing oxidative stress (Singh *et al.*, 2007). Oxidative stress has also been found to promote NADPH production which in turn tends to diminish oxidative stress by facilitating reactions that promote a reducing environment (Singh *et al.*, 2007; Grose *et al.*, 2006).

Nutrient excess has been shown to cause oxidative stress in mammalian cells and bacteria (Wellen and Thompson, 2010; Wang and Levin, 2009). The tricarboxylic acid (TCA) cycle oxide

nutrients and the resulting electrons are transferred to produce NADH and FADH₂. The electrons are eventually donated to molecular oxygen via the electron transport chain (ETC) where incomplete reduction produces superoxides. Therefore, when the breakdown of components from the TCA cycle exceeds the capacity of the electron chain cycle (as is the case with excess nutrients), ROS production increases and can lead to oxidative stress (Wellen and Thompson, 2011).

2.3 Fluorescence-Activated Cell Sorting (FACS)

Flow cytometry is a sensitive and powerful tool that uses fluorescence and light scattering to simultaneously measure the physical and chemical properties of cells suspended in a fluid. Cytometry is a high throughput process that characterizes individual cells at very high rates and thus can detect rare events in real time. Several cellular parameters can be measured simultaneously depending on the number of detectors available. The determination of these parameters is based on the availability of a wide range of fluorescent dyes and the ability to conjugate antibodies to fluorophores. Flow cytometry can characterize gene expression (Wylot *et al.*, 2015; von Recum-Knepper *et al.*, 2015; Finegersh and Homanics, 2016), cell viability (Huizer *et al.*, 2017; Sánchez-Luengo *et al.*, 2017; Arias-Fuenzalida *et al.*, 2017), target binding affinity (Wilson *et al.*, 2017; He *et al.*, 2018), cell surface markers (Yuan *et al.*, 2015; Crouch and Doetsch, 2018; Brooks and Seagroves, 2018), status of the cell in the cell cycle (Ferullo *et al.*, 2009; ter Huurne *et al.*, 2017) and oxidative stress (Boone *et al.*, 2017; Llufrío *et al.*, 2012). The main disadvantages of flow cytometry are that many dyes exhibit similar emission spectra (Jahan-Tigh

et al., 2012) and RNA integrity and cell viability may be compromised by the toxicity of these dyes (Jager *et al.*, 2018; Ramesh *et al.*, 2015).

Cell samples are usually subjected to a pre-treatment which can include centrifugation, tissue dissociation and/or resuspension in a buffer. The suspended cell samples are then stained with fluorescent conjugated antibodies and/or dye(s) and introduced into the flow cytometer. Flow cytometers typically contains two fluids travelling in the same direction toward a flow cell: one containing the cell suspension and the other containing sheath fluid. Hydrodynamic focusing occurs as the slower stream containing the cell suspension is confined within the faster flowing sheath fluid stream to produce a laminar flow profile for the cells. Ideally, the particles in the cell suspension line up in single file to be characterized one at a time by focusing a laser beam on each passing cell and measuring the energy re-emitted as fluorescence by the excited fluorophores to detectors located downstream.

The detectors are most commonly photomultiplier tubes (PMT) which convert the emitted or scattered light into amplified electrical pulses that are processed by appropriate electronics to extract information such as pulse's height, width and area. Forward and side scattering are related to particle size and cell membrane complexity (granularity), respectively. The emitted light is made up of fluorescence usually from antibodies conjugated to fluorophores, naturally fluorescent molecules or externally added fluorescent dyes. The fluorophores may be located intracellularly or on the cell membrane. More than one fluorophore may be monitored simultaneously; the total number depends on the number of detection channels that are available. Since multi-colour detection can be complicated by fluorescence spillover from one channel to another, correct compensation for this effect is crucial for accurate measurement. A significant amount of literature on this subject is available (Szalóki and Goda, 2015; Nguyen *et al.*, 2013).

Flow cytometry sorting is an extension of flow cytometry that enables cells to be sorted downstream from the detectors that measure the fluorescence energy. Figure 2-1 shows a schematic of the flow cytometry sorting operation. In this operation, the stream is broken into droplets at the breakoff point just downstream from the detection region. The drop delay is the time it takes for a cell to go from the detection region to the breakoff point. This is carefully chosen so that the droplets containing the cells of interest will be properly identified for sorting. Shortly after passing by the detector, a charge is acquired by the cells that have the properties of interest for sorting as they traverse a charging ring or adapter. The charged droplets are then separated into different collection bins by passing them through an electrostatic field applied by an electrostatic deflection system consisting of charged plates. Droplets may contain no cells, a single cell or several cells. Droplets that do not contain cells will not be measured by the detectors and will pass into the waste located directly below the incoming stream. Droplets that have more than one cell are classified as doublet events. In order to describe rare cell events, it is important to properly discriminate doublet events from true single cells (Stadinski and Huseby, 2020; Cossarizza *et al.*, 2019). Also, droplets containing cells may be too close together to be properly separated. In order to maintain purity, if droplets containing cells are too close together (i.e. droplets containing intensity measurements) the charger will choose not to charge the droplets and they will be passed into the waste bin, leading to some loss of product.

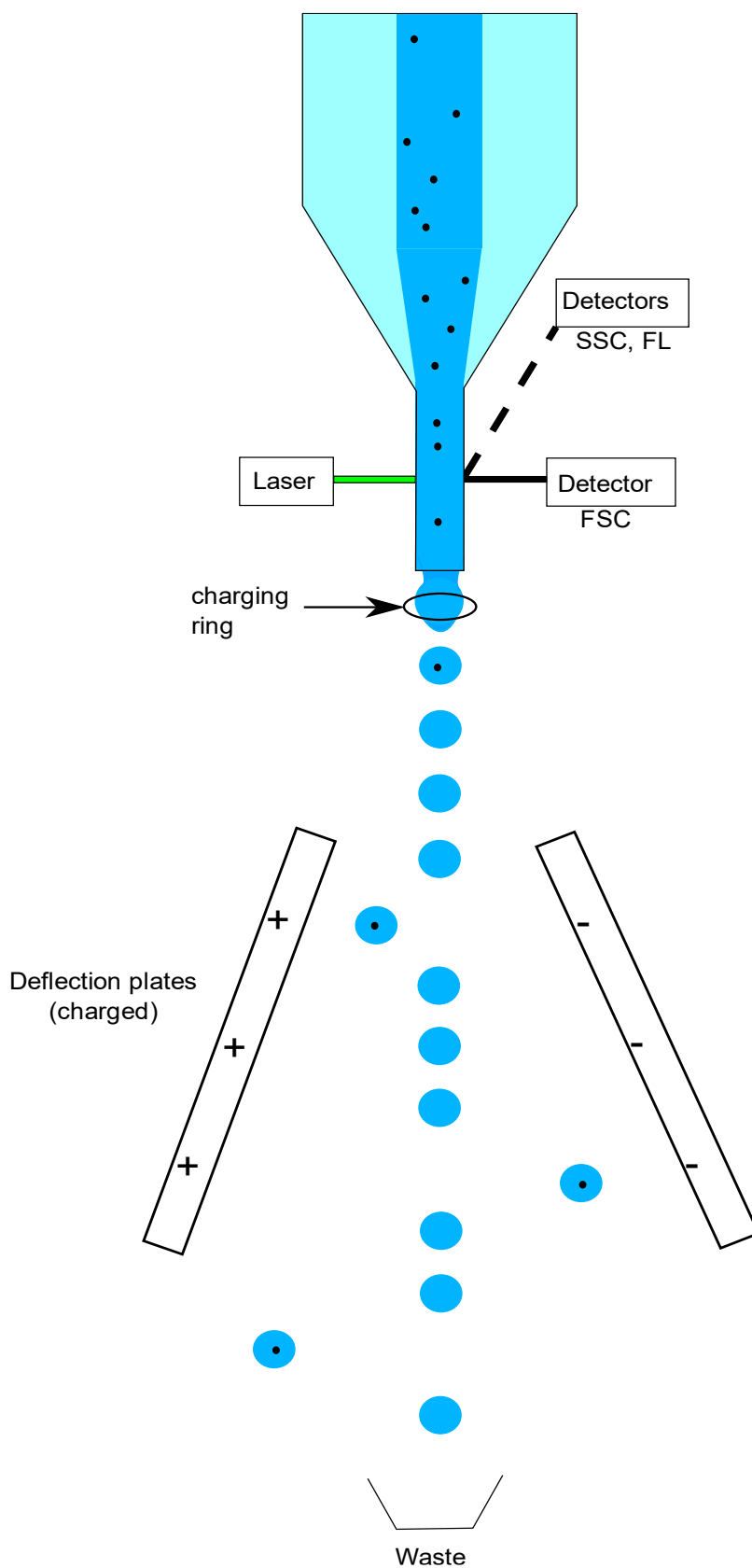


Figure 2-1: Schematic of flow cytometry sorting operation

Flow cytometry sorting has become a powerful tool to monitor, screen and separate cells based on any chemical or physical property that can be measured by flow cytometry. This has led to many applications of cell sorting such as isolation of particular cell types from mixtures, screening of mutant libraries or specific antibodies and determination of the likelihood for viability in a particular cell environment. One of the objectives of the research described in this thesis is to identify potential uses of cell sorting to improve the fermentation process of *B. pertussis*. Toward this goal, this section discusses established and cutting-edge methods and applications of flow cytometry sorting, with particular focus on their use in the pharmaceutical industry. The applications to be reviewed can be grouped as follows: i) cell type isolation, ii) high throughput screening, iii) cell surface display, iv) droplet FACS and v) microfluidic FACS. Each of the following sub-sections is devoted to one of these applications followed by a final sub-section discussing potential future applications.

2.3.1 Isolation of naturally occurring cell sub-populations

A major application of flow cytometry sorting is the isolation of sub-populations from heterogeneous samples. For example, tissues can be carefully dissociated into single cells for sorting of sub-populations. Common downstream practices with sorted cells include further culturing or transcriptomic evaluation. In this section, we outline recent studies that have employed FACS to target cells for further use and discuss separation strategies used to isolate sub-populations, as shown schematically in Figure 2-2.

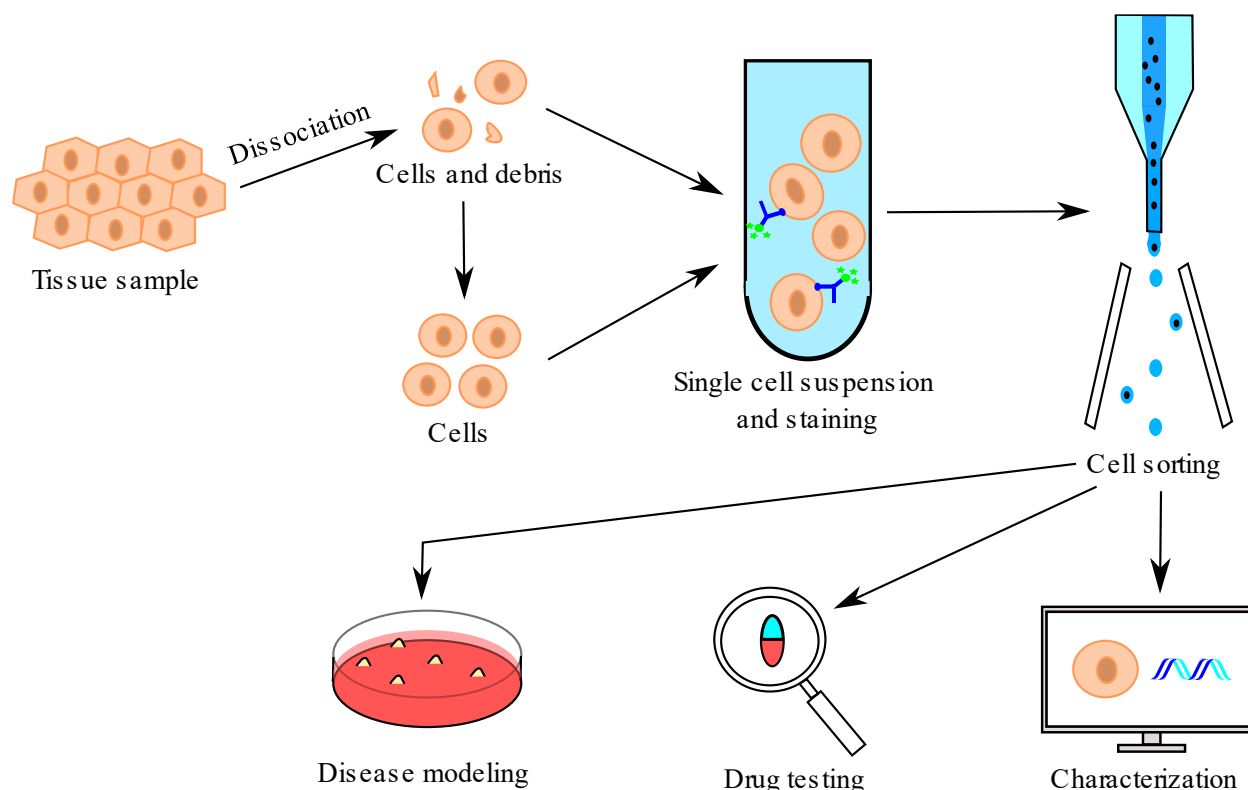


Figure 2-2: Application map for isolating a cell type of interest from heterogeneous samples. Pre-treatment of the sample may include dissociation and debris removal. Cells are then suspended and stained, most commonly by surface markers. The sorted cells can be used in a variety of applications including disease modeling, drug testing and characterization (i.e., lineage and gene expression).

Cell sorting can be used as a pre-treatment process to isolate a homogeneous cell type or sub-population for further culturing. This is particularly relevant to this thesis where one of the goals is to sort cells on the basis of a particular property and inoculate a fermenter with the sorted population in order to increase the productivity of biomass growth and potentially antigen production. In this application, the key challenge is that the exposure of cells to stressful operating conditions imposed by flow cytometry sorting may compromise cell viability and the ability to grow in subsequent culturing steps. This has turned out to be a main challenge for the cell sorting work conducted in this thesis, as shown in Chapter 6. While many cells are insensitive to the

sorting procedure (Mollet *et al.*, 2007; Andra *et al.*, 2020), the shear stress caused by the volumetric fluid flow and nozzle can cause significant cellular damage (Mollet *et al.*, 2008; Ma *et al.*, 2002). The effect of shear stress in bioreactors on cell viability has been widely studied. For example, Lange *et al.* (2001) conducted experiments in capillary tubes under different controlled shear conditions and found that microorganism viability is significantly reduced when subjected to a shear stress above a certain threshold. Specific methods can be followed to assess the fluid shear stress produced by the sorting protocol. For example, a sensor developed by Varma *et al.* (2015) can be used to measure fluid shear stress pathway activation and evaluate its impact on the experimental protocol including the equipment used for sorting and the assay of interest. Other experimental conditions to consider include the sample flow rate which can affect the precision of the measurements (Cossarizza *et al.*, 2019) and stain concentration which can cause cell damage if excessive (Fried *et al.*, 1982). Furthermore, surface markers that are used for sorting may induce apoptosis, e.g., anti-Sialec F in eosinophils (Zimmermann *et al.*, 2008). Careful selection of the buffer can improve sorting outcomes such as the integrity of RNA post-sorting (Kunnath-Velayudhan *et al.*, 2018; Nilsson *et al.*, 2014).

Optimization of the protocol to be used for cell sorting is a lengthy and technical procedure that involves many variables including sample preparation, fluorescent dye or fluorescent conjugated antibody selection and concentration, sample concentration, sample flow rate, buffer selection, gating procedure and compensation if more than one fluorophore is detected. For this reason, cell sorting protocols are reported regularly in great detail (Póvoa *et al.*, 2020; Buehler *et al.*, 2012; Abujarour *et al.*, 2013; Lv *et al.*, 2016). A protocol for *B. pertussis* will be developed in this work and presented in Chapter 6.

Stem cells which can differentiate along lineage pathways and self-renew *in vivo* are ideally suited for critical applications in the treatment of injuries and disease (Wang *et al.*, 2018). Thus, procedures to separate viable pure stem cell or progenitor cell populations by use of flow cytometry sorting are popular and continue to evolve quickly. Protocols are continually being improved in terms of digestion techniques, available markers and pre-treatments to remove materials that have adverse effects on sorting pure populations (Wylot *et al.*, 2015; Ishii *et al.*, 2017; Maesner *et al.*, 2016).

Instead of tissue samples, flow cytometry sorting can be used to sort antibody-producing white blood cells (B-cells) from blood samples. For example, Sadreddini *et al.* (2016) investigated the effect of cell sorting on the transformation efficiency of antibody-producing memory B-cells. The B-cells were transformed by the Epstein-Barr virus into immortal lymphoblastoid cells that proliferated indefinitely (Price and Dave, 2013). The researchers were able to effectively isolate human memory B-cells using surface markers for B-cells conjugated to fluorophores. They further evaluated the efficiency of the sorting procedure based on the responses of the sorted cells to tetanus neurotoxin. The combination of Epstein-Barr virus transformation and cell sorting was found to have very low efficiency for large scale antibody production and so other strategies such as antibody engineering and display technology were proposed as preferred methods for large scale production.

Many studies have been concerned with the effects of the flow cytometry sorting procedure on cellular metabolism, integrity and viability. Richardson *et al.* (2015) detected short-term transcriptional artifacts, while Finegersh *et al.* (2016) found that intracellular staining which requires fixation and permeabilization of cells can compromise the integrity of RNA. Fortunately, options are available to address some of these problems. Kunnath-Velayudhan *et al.* (2018)

reported that the use of a high salt buffer and then a buffer-containing RNase inhibitor allowed the RNA to remain intact for fixed and permeabilized murine T cells.

Table 2-1 summarizes the literature on cell sorting of homogeneous cell types or sub-populations from heterogeneous mixtures. The most common sorting strategies make use of surface marker antibodies since different cell types and even sub-populations of cell types have surface markers that are lineage-specific. Once the lineage markers are determined, sorting to yield high purity cells is possible following optimization of the pre-treatment procedure and sorting protocol. In Chapter 6, a sorting procedure based on the surface antigen concentration of pertactin will be used to assess whether the epigenetic phenotype is maintained in the daughter cells of the sorted populations.

Table 2-1: Studies on sorting of heterogeneous cell populations including their gating strategies.

Sorted Cell	Sample	Sorting Strategy	Purpose	Reference
Motor neuron	Induced pluripotent stem cell	Surface marker antibodies and intracellular fluorescent protein	ALS disease modeling and drug testing	(Toli <i>et al.</i> , 2015)
Rod shaped ventricular myocytes	Adult mouse heart tissue	Viability dye and surface marker antibodies	Analysis of proteins	(López <i>et al.</i> , 2017)
Endothelial cells	Mouse CNS tissue	Surface marker antibodies and viability dye	Protocol for sorting pure endothelial cells	(Wylot <i>et al.</i> , 2015)
Mesenchymal stem cells	Bone marrow and adipose tissue	Surface marker antibody linked to VEGF expression	Control of desired level of a transgene	(Helmrich <i>et al.</i> , 2018)

T cells	Murine splenocyte (spleen tissue)	Surface marker antibodies and viability dye	Buffer effects on RNA integrity	(Kunnath-Velayudhan <i>et al.</i> , 2018)
Satellite glial cells (SGCs)	Rat dorsal root ganglion (neuron)	Antibodies against intracellular epitopes of SGC specific proteins	Function and changes following disease or injury	(Jager <i>et al.</i> , 2018)
Skeletal muscle satellite cells	Mouse skeletal muscle	Surface marker antibodies	Comparison of muscle digestion techniques	(Ishii <i>et al.</i> , 2017)
T and B cell subpopulations	Horse blood	Surface marker antibodies	Lineage gene expression analysis	(Tomlinson <i>et al.</i> , 2018)
Neurospheroidal cells	Glioblastoma cell line	Forward/ side scatter and viability dye	Characterization of minor populations in cell lines	(Immanuel <i>et al.</i> , 2018)
Eosinophils	Mouse lungs	Viability dye and surface marker antibodies	Isolation strategy not involving apoptotic markers	(Geslewitz <i>et al.</i> , 2018)
Endothelial cells and pericytes	Mouse brain tissue	Surface marker antibodies	Comparison of functional differences	(Crouch and Doetsch, 2018)
Hematopoietic progenitor, circulating endothelial, and outgrowth endothelial cells	Peripheral blood and umbilical cord blood samples	Surface marker antibodies, viability dye, and front/side scatter	Gene expression of subpopulations and analysis of markers	(Huizer <i>et al.</i> , 2017)

Keratinocytes, and T cell subpopulations	Human skin cell samples	Surface marker antibodies	Sequencing of cell subpopulations	(Ahn <i>et al.</i> , 2017)
Satellite cells	Adult mouse skeletal muscle tissue	Surface marker antibodies and viability dye	Gene expression profile and regenerative property	(Maesner <i>et al.</i> , 2016)
Primitive endoderm and epiblast cells	Embryonic stem cells	Surface marker antibodies and viability dye	Protocol and gating strategy	(Hölzenspies <i>et al.</i> , 2015)
Non-pericytic adipose progenitor cells	Mouse subcutaneous inguinal white adipose tissue	Surface marker antibodies	Investigation of self-renewal properties	(Frazier <i>et al.</i> , 2016)
Medium spiny neurons	Nucleus accumbens	Intracellular fluorescent protein	Study of epigenetic mechanisms	(Finegersh and Homanics, 2016)
Circulating tumor cells	Gastric cancer patients	Surface marker antibodies	Comparison of biological behaviour between gastric cancer patients and cell line	(Yuan <i>et al.</i> , 2015)
Dendritic cell subpopulations	Atherosclerotic plaques	Surface marker antibodies	Evaluation of cell liberation from plaques	(Van Brussel <i>et al.</i> , 2015)

2.3.2 Isolation of artificially induced cell subpopulations

In this sub-section, we outline recent studies in which flow cytometry sorting has been used to screen and sort out valuable mutants induced by artificially induced sub-populations based on different methods such as random mutagenesis and nuclease genome editing. High throughput screening strategies have been established to improve the performance of industrial microorganisms. Advances in flow cytometry sorting and other microfluidic techniques have significantly improved the screening efficiency (Zeng *et al.*, 2020). The critical advantages of flow cytometry sorting are that the screening scale (i.e. number of cells analyzed) can be much higher than that attained by other sorting methods and it has the ability to carry out measurements on individual cells. Another high throughput method with similar advantages is automated colony picking. However, colony pickers are better suited for applications for secreted products as flow cytometry sorting measures productivity based on surface protein expression which is not always strongly correlated to secreted levels (Nakamura and Omasa, 2015). Also, flow cytometry sorting is a more technically difficult method than colony picking since optimization of the protocol plays a large role in the outcome of cell sorting, as previously discussed in section 2.3.1. However, colony picking is not suited to protein engineering applications (Fabritius *et al.*, 2018) and has a much slower screening efficiency than flow cytometry sorting (Zeng *et al.*, 2020). Applications where sub-populations are artificially induced by either random mutagenesis or nuclease genome editing are discussed separately below. Figure 2-3 provides a roadmap of the typical strategy used for sorting sub-populations of interest.

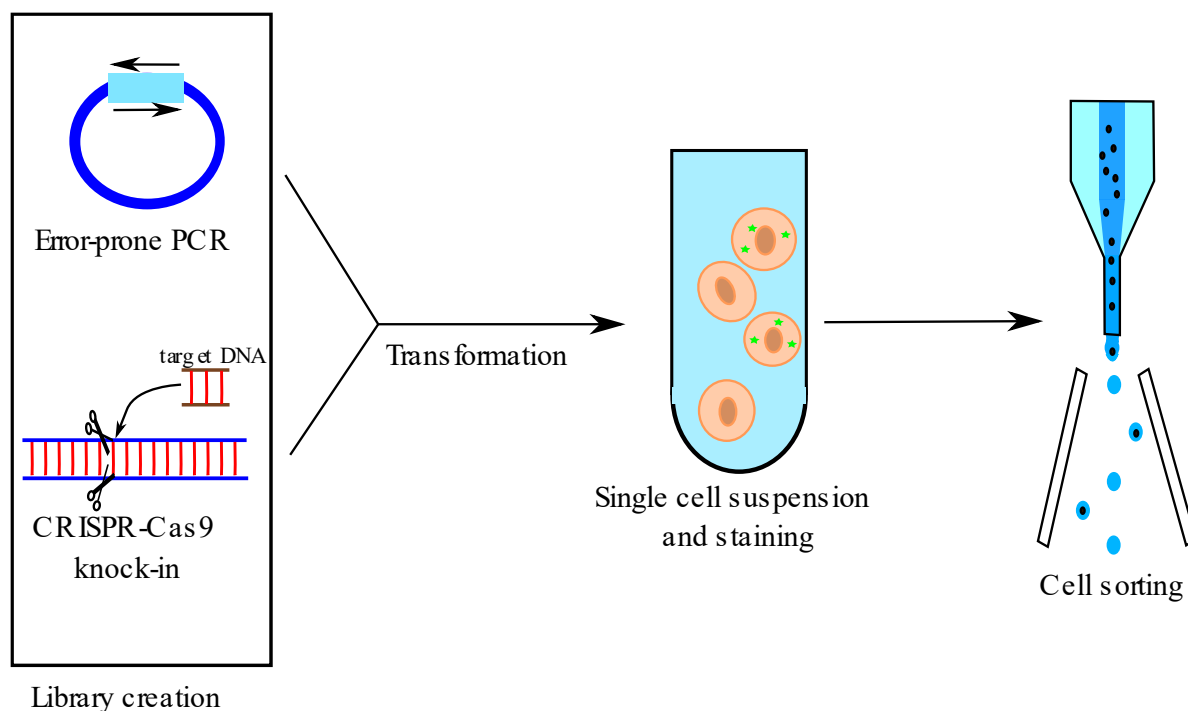


Figure 2-3: Application map for isolating a cell type of interest from artificially induced cell populations. Pre-treatment of the sample may include library creation using mutagenesis. Most commonly the markers are inserted intracellularly via a plasmid. The sorted cells can be used in a variety of applications including screening for high producing cell lines or desired gene editing and for studying metabolic networks.

2.3.2.1 Random mutagenesis

Random mutagenesis is a powerful tool to modify enzymes, proteins or metabolic pathways in order to improve their function (Labrou, 2018). By physical (e.g. ultraviolet (UV) irradiation) or chemical means (Ghatnekar, 1964; Kodym and Afza, 2003), it introduces into the genome random mutations that can vary in terms of the number of mutations and location where they occur. To screen for mutants resulting from physical mutagenesis sources, sorting for viable cells by flow cytometry sorting can greatly decrease the size of the library. Flow cytometry sorting can also be used to directly recover desired mutants by using a fluorescent dye specific to the property of interest (Liu *et al.*, 2016).

Chemical mutagenesis is much more common. The error-prone polymerase chain reaction (PCR) utilizes a polymerase that lacks proof-reading ability and so can cause mutations to occur for a gene of interest. The locations of mutations are determined by the upstream and downstream PCR primer-binding sites on the gene (McCullum *et al.*, 2010) and the number of mutations per template gene depends on the number of doublings (Wilson and Keefe, 2005). Site saturation mutagenesis is used to determine residue-function relationships and probe the biochemical mechanism of an enzyme/protein. In this case, mutations occur at either a single or a set of residues located at catalytic sites of an enzyme. The residues are randomized to produce all possible amino acids at a chosen position and generate protein variants that cannot be produced naturally (Siloto and Weselake, 2012).

The size of the library of mutants can be increased by combining mutagenesis methods. Prodrugs which are widely used in cancer medicine are inert when administered but become highly cytotoxic when activated (Denny, 2003). Copp *et al.* (2014) developed a high-throughput screening platform to design prodrug-activating enzymes using site saturation and error-prone PCR mutagenesis methods. They fused a fluorescent protein to a SOS promoter which becomes activated when DNA is damaged. The amount of this fluorescent protein served as a measure of the amount of DNA damage caused by the prodrug in *E. coli* through activation by nitroreductase. Copp *et al.* (2017) later used this flow cytometry sorting protocol to select enhanced cancer prodrug activation variants.

Other forms of random mutagenesis can be used to induce genes to participate in specific reaction pathways. Random transposon mutagenesis allows genes to be transferred into a host organism. In this case, these genes can cause mutations by interrupting or modifying the function of the extant gene on a chromosome (Simon *et al.*, 1983).

Sorting mutant cells with the purpose of re-culturing is an important aspect of disease modeling but flow cytometry sorting protocols can compromise viability. Instead of re-culturing sorted cells, Ramesh *et al.* (2015) developed a recovery method to directly isolate plasmid DNA from flow cytometry-sorted populations which expressed the desired phenotype for subsequent transformation. This protocol was used to screen two *E. coli* libraries and led to a 5-fold increase in the recovery of the plasmid compared to recovering the sorted cells.

2.3.2.2 Nuclease Genome Editing

Nuclease genome editing enzymes such as zinc finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) and clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) are powerful tools to add or remove genes from the genome of cells (Gai *et al.*, 2013). These protocols rely on nucleases that are engineered to cut specific genomic target sequences. Flow cytometry sorting is widely used to screen for cells that have been successfully edited with these techniques to improve the rate of gene editing in cells (Yumlu *et al.*, 2017) and isolate cells which have been edited successfully (He *et al.*, 2015). Cells that over-express or under-express genes of interest may have undesirable effects such as compromised viability. It is essential to ensure that the nucleases are expressed within a sufficient fraction of cells so that the desired amount of the genome is edited since it has been observed that transfection levels can vary more than 100-fold (from CRISPR-Cas9 system) within cell populations in programmable nuclease constructs (Lonowski *et al.*, 2017).

2.3.3 Cell surface display systems

Cell surface display systems utilize expressed recombinant proteins that are incorporated into the host cell membrane to display a protein of interest on the surface of the cell. The recombinant DNA encodes an anchoring motif (carrier protein) fused to the protein of interest (passenger protein) and uses the machinery of the host cell to translate and fold the proteins. A schematic showing the steps of the cell surface display procedure explained above is shown in Figure 2-4. The carrier protein, usually a membrane protein with a filamentous structure (Wan *et al.*, 2002), anchors to the cell membrane to allow the fused passenger protein to point outside the cell membrane for easy extracellular targeting.

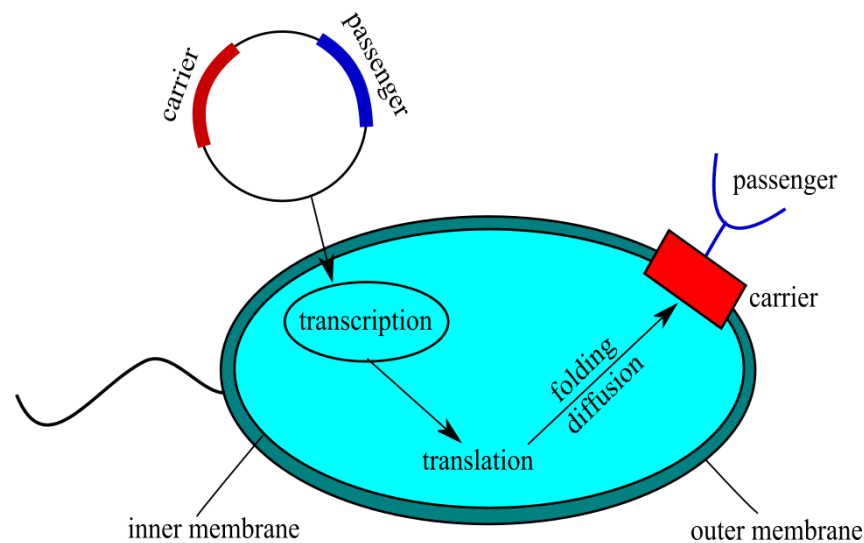


Figure 2-4: Cell surface display

Flow cytometry sorting is a very efficient high-throughput method to screen cells with desired antibody-protein affinities. A major advantage of using flow cytometry sorting in combination with the surface display approach is that it is easy to label the passenger protein by fluorescence since it is located on the cell surface. Figure 2-5 provides a roadmap of the typical strategy used in surface display applications.

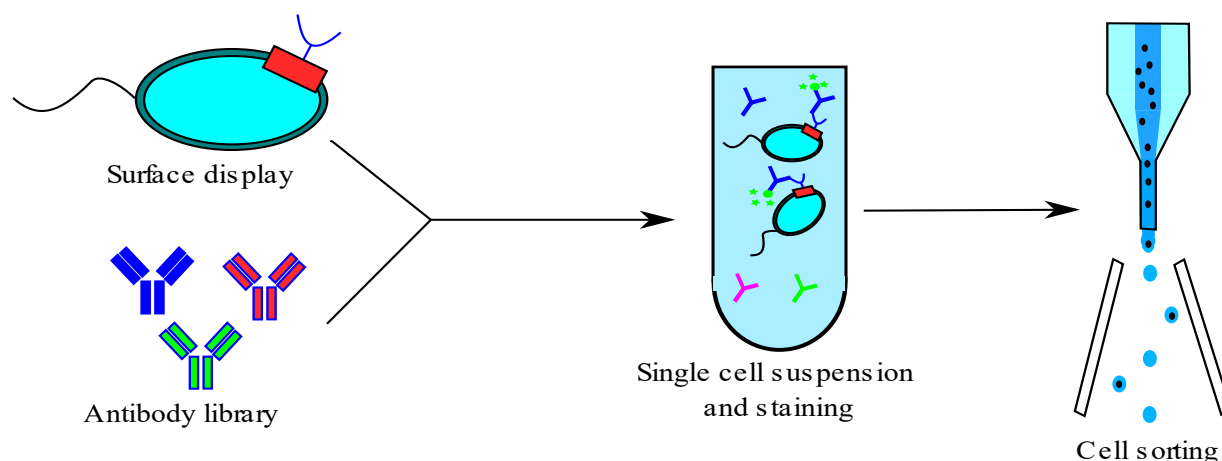


Figure 2-5: Application map for sorting via surface display. Cells are incubated with the marker library and commonly sorted based on binding affinity. The desired markers can be used in a variety of applications including screening vaccine or antibody candidates and marker selection for further testing.

The efficiency of surface display systems is strongly affected by the characteristics of the carrier protein, passenger protein and host cell and the fusion method (Sandkvist and Bagdasarian, 1996). The display of a phage surface was the first surface display system developed (Smith, 1985). The application of this procedure to bacteria rather than phages has the advantages of being easier to cultivate with selection markers without contamination, displaying random peptides using the flagella occurring in bacteria and facilitating screening by flow cytometry since they are larger in size (Lunder *et al.*, 2005; Hu *et al.*, 2018). However, despite the advantages of cell display in bacteria, phage surface display has been more commonly utilized because of its ability to select more effective sequences for binding peptides and generate larger libraries (Lunder *et al.*, 2005; Hu *et al.*, 2018). The display of a yeast surface display was developed as an alternative to that of prokaryotic cells since it solved the problem of expression biases against some eukaryotic proteins and led to more efficiently folded proteins (Boder and Wittrup, 1997). Cell-display technology has many applications, which include vaccine development (Chen *et al.*, 2018; Xu *et al.*, 2016), antibody library screening (Sun *et al.*, 2016; Yim *et al.*, 2014) and enzyme library screening

(Deweid *et al.*, 2018; Chen *et al.*, 2011). Table 2-2 outlines some of the literature using flow cytometry and surface display.

Table 2-2: Studies that use surface display and flow cytometry sorting.

Display Type	Sorting Strategy	Purpose	Reference
Yeast surface display	Fluorescent antibody	Vaccine candidate	(Chen <i>et al.</i> , 2018)
Bacterial (<i>E. coli</i>) surface display	Fluorescent antibody	Screen potential antibodies for H1N1	(Yim <i>et al.</i> , 2014)
Bacterial (<i>E. coli</i>) surface display	Fluorescent antibody	Screen antigen-specific antibodies	(Xu <i>et al.</i> , 2016)
Bacterial (<i>straphylococcal</i>) surface display	Fluorescent antibody	Screen for therapeutics against various cancers	(Åstrand <i>et al.</i> , 2016)
Yeast surface display	Fluorescent antibody	Screening hapten-specific antibodies	(Sun <i>et al.</i> , 2016)
Bacterial (<i>straphylococcal</i>) surface display	Fluorescent antibody	Affibody screening for amyloid β peptide	(Lindberg <i>et al.</i> , 2015)
Hybridoma surface display	Fluorescent antibody	Screening for hybridoma cells that can produce hapten-specific antibodies	(Dippong <i>et al.</i> , 2017)
Yeast surface display	Fluorescent antibody	Increase sortase A binding affinity and activity	(Chen <i>et al.</i> , 2011)
Yeast surface display	Fluorescent antibody	Optimize enzyme activity	(Deweid <i>et al.</i> , 2018)
Yeast surface display	Fluorescent antibody	Optimize designed receptor for p21-activated kinase I	(Butz <i>et al.</i> , 2014)

2.3.4 Flow Cytometry Droplet Sorting

The measurement of secreted products by flow cytometry sorting has been previously used to determine the concentration of intracellular species since the cells are separated from the supernatant following the sorting operation. When active transport processes are involved in the secretion of the product, its intracellular concentration will not be highly correlated to the total productivity as it is preferentially retained by the cell. A technique has been developed to encapsulate single cells in a water-oil-water (w/o/w) double emulsion where the inner water-phase contains the media and secreted products (Caen *et al.*, 2018; Baret *et al.*, 2009). This has been accomplished by a microfluidics device that injects a fast-moving aqueous phase containing the cells into a slower-traveling oil phase to encapsulate the latter, as shown in Figure 2-6.

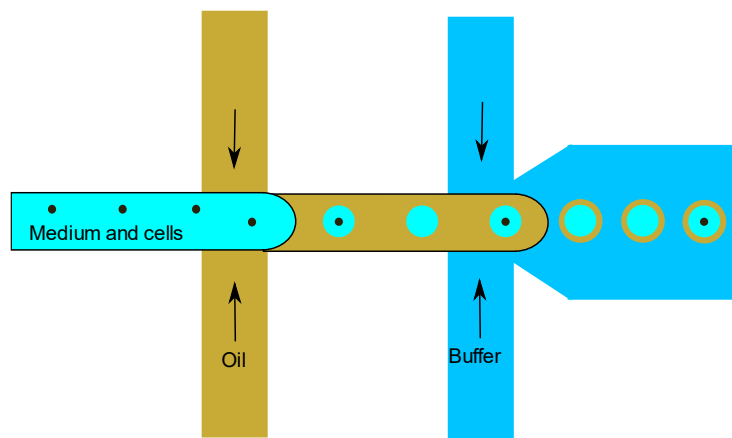


Figure 2-6: Water/ oil/ water droplet creation

The major advantage of double emulsion encapsulation is that it prevents the secreted metabolites or products from being transported away from the cells and traps them within the droplet for detection. Table 2-3 outlines the literature using droplet sorting.

Table 2-3: Studies based on droplet flow cytometry sorting.

Display Type	Sorting Strategy	Purpose	Reference
<i>Yarrowia lipolytica</i>	Riboflavin autofluorescence	Screen high producer of riboflavin	(Wagner <i>et al.</i> , 2018)
<i>Lactococcus lactis</i>	Riboflavin autofluorescence	Investigate cause of increased riboflavin production	(Chen <i>et al.</i> , 2017)
<i>Saccharomyces cerevisiae</i>	Enzymatic reaction	Enrichment of cellulase expressing cells	(Ostafe <i>et al.</i> , 2013)
<i>Saccharomyces cerevisiae</i>	Enzymatic reaction	Screen high confusing xylose strains	(Wang <i>et al.</i> , 2014)
<i>Escherichia coli</i>	Enzymatic reaction	Screen high producers of L-lactate	(Zhou <i>et al.</i> , 1997)
<i>Saccharomyces cerevisiae</i>	Fluorescent reporter protein	Detailed description of genotypes and phenotypes	(Terekhov <i>et al.</i> , 2017)
Dye only	Fluorescent dyes	Screen suitable markers for w/o/w droplets	(Hai and Magdassi, 2004)
<i>Escherichia coli</i>	Enzymatic cleavage	Screen for potential antibodies, enzymes, or proteins	(Ma <i>et al.</i> , 2016)

2.3.5 Microfluidic FACS (μ FACS)

Cell sorting has also been accomplished using miniaturized devices, some of which carry out FACS by taking advantage of a range of physical properties. These smaller devices are termed microfluidic FACS (μ FACS) and fall within the category of *lab-on-a-chip* technologies. The use of μ FACS has several advantages over that of conventional FACS in that it eliminates potentially biohazardous aerosols, sample loss and cross-contamination due to the closed-platform designs and disposable chips. In addition, μ FACS devices are portable, sample processing protocols are

simplified, reagent and equipment costs are lower and rapid analysis is possible, making them very attractive for both industrial and healthcare (point-of-care) applications.

Broadly, μ FACS techniques can be organized into four overarching categories based on the physical principles exploited for cell manipulation (Shields et al., 2015): i) electrokinetic (e.g.; Wang et al., 2010), ii) acoustophoretic (e.g., Li et al., 2019; Mutafooulos et al., 2019; Ren et al., 2018), iii) optical (e.g., Wu et al., 2012) and iv) mechanical (Cheng et al., 2017; Cho *et al.*, 2010). Similar to traditional FACS, the operation of these devices involves serial interrogation by laser light, real-time classification and command-driven sorting. Recently, comprehensive reviews specific to microfluidic cell sorting have been published by Shields et al. (2015) and Shen et al. (2019).

Many of the μ FACS technologies are still at the proof-of-concept or prototype stage and only a few have demonstrated the performance or validation required for integration into clinical practice. For example, Cheng *et al.* (2017) tested the validity of a μ FACS chip with piezoelectric actuators to enrich GFP-expressing HeLa cells from a population including non-fluorescent HeLa cells. Cho *et al.* (2010) also used a μ FACS chip with a piezoelectric actuator to enrich fluorescent human erythroleukemic cells from a mixture containing stained and unstained erythroleukemic cells. In another proof-of-concept experiment, Wu *et al.* (2012) used pulse laser-triggered μ FACS to sort nalm-6 human pre-B cells and B lymphoma Ramos cells with high viability and purity. Although promising, μ FACS requires further development before it can be used commercially. Some of the limitations include low throughput rates since the devices usually consist of a single channel and have short lifespans due to clogging.

μ FACS has the potential to directly affect health care, particularly for personalized medicine in cancer treatment. Hirai *et al.* (2015) presented a μ FACS for bladder cancer diagnosis. The μ FACS system was able to meet requirements for use in biological and clinical applications and has several advantages over existing methods since it can be operated by non-specialist researchers, provides early and non-invasive cancer diagnostics and can isolate cancer cells from low-grade and low-stage cancers. Ma *et al.* (2017) demonstrated a benchtop μ FACS system that made use of a highly focused traveling surface acoustic wave beam to sort cells upon fluorescence interrogation. The acoustic μ FACS system was used to isolate fluorescently labeled MCF-7 breast cancer cells from diluted whole blood samples with high purity and extremely high viability after sorting.

2.4 Population Balance Models

Most biotechnological processes are regulated through closed loop control of lumped or average quantities such as dissolved oxygen, average substrate concentration or pH. However, due to the heterogeneity of the cell population, e.g., variability in cell size, cell age, intracellular concentrations, control approaches that are based on only the average culture properties may not guarantee consistent process outcomes. Population balance models are one approach to describe the heterogeneity of the cell population and its effect on culture growth and productivity. By correlating the evolution of the cell population to productivity and inputs that can be manipulated, these models have the potential to improve and optimize processes.

Population balance models contain a set of equations which describe the behavior of a population of particles from the analysis of behavior of single particles. In a bioreactor, these

models provide a deterministic description of the dynamic evolution of cell distributions. Population balance models can be mass-based if the properties used to describe the state of the cell follow the conservation of mass or age-based if the age of cells is used to differentiate each cell within the population. The general form of a population balance model assuming that cell death is negligible and that cells grow in one stage is shown in Eqn (2.1) (Fredrickson *et al.*, 1967; Ramkrishna, 1979).

$$\begin{aligned} \frac{\delta N(\mathbf{x}, t)}{\delta t} + \nabla_{\mathbf{x}}(\mathbf{r}(\mathbf{x}, S)N(\mathbf{x}, t)) + \gamma(\mathbf{x}, t)N(\mathbf{x}, t) + DN(\mathbf{x}, t) \\ = 2 \int_x^{x_{max}} \gamma(\mathbf{x}, t)p(\mathbf{x}, y, z)N(\mathbf{x}, t)dx \end{aligned} \quad (2.1)$$

The number of cells from the entire population at each physiological state is described as the time-dependent variable $N(\mathbf{x}, t)$. The first term on the left-hand side corresponds to the accumulation of cells at a level of physiological properties (\mathbf{x}) where the latter can be one or more cell properties of interest such as levels of growth, oxidative stress, DNA content, etc. The second term accounts for the rate of gain/loss of cells with the physiological state due to chemical reactions and to changes in cell volume. Correspondingly, the chemical reactions rates or growth of cells are represented by $r(\mathbf{x}, S)$ which depends on the available substrate concentration and the physiological state of the individual cells. The third term accounts for the rate of loss of cells due to cell division leading to the birth of smaller cells and depends on the partitioning rate $\gamma(\mathbf{x}, t)$. The fourth term on the LHS is the dilution rate which represents the rate at which cells leave the bioreactor. The right-hand side term represents the rate of birth of cells with the physiological property level from the division of larger cells. This birth process is a function of a partition probability density function $p(\mathbf{x}, y, S)$.

One of our objectives in the current thesis is to compare the predictability of a population balance model with a similar bulk kinetic model that uses average cell properties. However, this section discusses other uses of population balance models focusing on the solution of the partial differential equation, physiological properties of cells used in the models and the impact of including heterogeneity in the model.

The solution of population balance models typically is numerically challenging since the model consists of a system of coupled partial integro-differential equations and ordinary differential equations. For example, Mantzaris *et al.* (1999) derived a general mass-based population balance model that can consider any number of cellular components of interest and can be applied with equal or unequal partitioning among the components and fixed or changing substrate concentration. The limitation of this model was the high computational time and memory required for its solution when using a time-explicit finite difference method consisting of a hybrid of leapfrog and Lax-Friedrichs schemes. However, due to the continuous improvements in computational power, the use of population balance models to describe cell state distributions has shown a recent upsurge. Solutions have been reported using a number of numerical techniques including finite differences methods (Mantzaris *et al.*, 2001a), finite element methods (Mantzaris *et al.*, 2001b) and hybrid methods such as a combination of the method of partial characteristics and method of moments (Durr and Waldherr., 2018).

The greatest value of population balance models is to capture the heterogeneity of cell populations and thereby effectively predict experimental outcomes for cases where the effects of heterogeneity are significant. For example, Morchain *et al.* (2017) presented a population balance model for bioreactor dynamics that combined liquid-cell mass transfer and population heterogeneity and was subsequently compared to experimental data. The advantage of the model

was that the specific growth rate varied across the cell population. The authors proposed a law for the probability that cells growing at a specific rate produce daughter cells with a different growth rate. The actual growth rate of individual cells was based on a comparison between substrate uptake and micro-mixing characteristic times. Simulated results for situations of a pulse addition of substrate in a fed-batch culture and the transition from batch to continuous mode were in quantitative agreement with experimental observations.

Population balance models have also been used to formulate hypotheses about a system. For example, Durr *et al.* (2016) investigated the influence of cell-to-cell variability on the maximum virus yield. The multi-dimensional population balance model was composed of a detailed description of the intracellular viral replication cycle and also accounted for the expected heterogeneity in the host cell population. The simulations suggested that modifying the cell lines (e.g. gene knockout) would increase maximum virus concentration and reduce the time needed to achieve the maximum concentration. However, the predictions of this study were not verified by comparison to experimental data.

Cell growth is a very important and heavily studied modeling parameter in many biological systems. With population balance models, it is possible to determine how cells uptake the available substrate to gain weight and engage in cell division. Quedeville *et al.* (2018) reported a 2-dimensional population balance model able to distinguish between growth based on cell number and growth based on cell mass when equilibrium is disrupted. In this model, cell division is driven by the cell size. This model was solved using a first order finite-volume method and a Monte-Carlo algorithm. The authors were able to demonstrate that growth in mass and growth and number are only equivalent at steady state in a continuous bioreactor and at the exponential growth phase during a batch process.

Flow cytometry can be used to measure heterogeneity within cell cultures. However, due to experimental limitations it may be difficult to measure many subsets from a process. Durr *et al.* (2015) proposed a procedure to incorporate limited measurement information from small scale culture experiments into a multidimensional population balance model. Their algorithm was based on an efficient moment approximation and a numerical solution of the corresponding population balance model with the method of characteristics. The authors were able to apply the algorithm to the production of poly(3-hydroxybutyrate) (PHB) in bacterial cultures using three reaction rates (cell growth, PHB synthesis and PHB metabolization). This model algorithm is only applicable in cases where cell death and cell division are negligible.

Although many population models have been based on the assumption of time-invariant concentrations of key metabolites, the population balance equations can be coupled with equations describing time-varying cell metabolism to improve predictability. For example, Pigou and Morchain (2015) coupled a hydrodynamic model, population balance model for growth rate and metabolic model which predicts the reaction rates depending on the state of individual cells. The cells were distinguished in terms of the capabilities by which daughter cells inherit their growth rate. A discretization method was used to solve the population balance equation which was applied to predict acetate production of batch and fed-batch cultures. The authors showed that coupling the metabolic and population balance models strongly improves the prediction of acetate production compared to the metabolic model alone.

Chapter 3 Materials and Methods

3.1 Media formulation

Sanofi Pasteur makes 2X concentrated media and 100X concentrated growth factor solutions used in the fermentations of *B. pertussis*. The normalized formulation of media and growth factor solutions used at Sanofi Pasteur are found in Tables 3-1 and 3-2. The solutions for the shake flask experiments contain 0.5 volume fraction of the media solution and 0.01 of the growth factor solution. The concentration of monosodium glutamate was varied between 6.6 – 12.4 g/L to assess its effect on oxidative stress and biomass.

Table 3-1: Media components and concentration for a 2X concentrated solution

Media Components	Concentration (normalized)
monosodium glutamate	20 g/L
sodium chloride	0.25
potassium phosphate monobasic	0.05
magnesium chloride	0.01
potassium chloride	0.02
tris base	0.15
casamino acid	1

Table 3-2: Growth factor components and concentration for a 100X concentrated solution

Growth Factor Components	Concentration (normalized)
L-cysteine HCl	0.04
nicotinic acid	0.004
ascorbic acid	0.4
reduced glutathione	0.15
heptakis beta cyclodextrin	0.01
FeSO ₄ (H ₂ O) ₇	1
CaCl ₂ (H ₂ O) ₂	0.02

3.2 Shake flask fermentations

B. pertussis fermentations were carried out in 250 mL flasks (polycarbonate, sterile, with a vented cap having 0.22 µm pore-size PTFE membrane from VWR, Canada) with a working volume of 60 mL by incubation at 36°C and 200 rpm for various times depending on the experiment. The compositions of media and growth factor components used are listed in section 3.1. The primary flasks were inoculated with 0.6 mL of a working batch seed (provided by Sanofi Pasteur), while the secondary flasks were inoculated with product from the primary flasks through dilution resulting in an OD₆₀₀ in the range between 0.1 – 0.4. Samples were collected at various times from the flasks to measure biomass via optical density, glutamate concentration via BioProfile, NADPH via fluorescence spectroscopy and ROS via flow cytometry. After measuring optical density, the cells and supernatant were separated by centrifugation at 10,000 G for 3 minutes.

Shake flask experiments were conducted as part of two different studies: i) determination of the relation between cell growth, oxidative stress and NADPH levels under different oxidative conditions using bulk and population balance models and ii) assessment of the impact of sorting populations via FACS of high/low cell complexity and surface PRN antibody concentration on growth and potential antibody yield.

3.3 Sample measurements

3.3.1 Optical density (OD₆₀₀)

Each sample collected from the shake flask culture was diluted 20-fold in a 0.9% saline solution and its OD level was determined via optical spectrophotometry by measurement of its absorbance at 600 nm.

3.3.2 Glutamate concentration

The supernatant obtained after centrifugation (section 3.2) was collected in order to determine the glutamate concentration (after 20-fold dilution) using the chemistry module of a BioProfile FLEX Analyzer (Nova Medical). Glutamate was measured using an amperometric electrode containing immobilized enzymes in its membrane (Derfus *et al.*, 2009). In the presence of oxygen, the enzyme membranes produce hydrogen peroxide which is then oxidized at a platinum anode held at constant potential. The resulting current is proportional to the sample concentration.

3.4 Fluorescence spectroscopy

3.4.1 Equipment

Analysis of the supernatant for extracellular NADPH concentration by fluorescence excitation-emission spectroscopy was carried out with a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies). The supernatant samples were diluted 20 times and

analyzed using polymethylmethacrylate cuvettes having a slit width of 5 nm and PMT (photomultiplier tubes) of 800 V for the peak measurements.

3.4.2 NADPH measurement

Determination of the NADPH concentration involved measurement in the excitation and emission ranges of 330-370 nm (10 nm intervals) and 300-600 nm (1 nm intervals), respectively, (Figure 3-1a). The blank used contained phosphate buffered saline (PBS, 10 mM, pH 7.2) and its spectra was subtracted from the sample spectra (Figure 3-1b). The large intensity peaks occurring in every emission involve first-order Rayleigh scattering which occurs when the emission wavelength of oscillating molecules is close to the excitation wavelength of the incident light frequency. To eliminate this effect, we replaced the fluorescence intensity with values of 0 over a range of 20 nm centered at the excitation frequency (Andersen 2005). The results of this method to eliminate the effects of Rayleigh scattering is shown when comparing Figure 3-1b and Figure 3-2a.

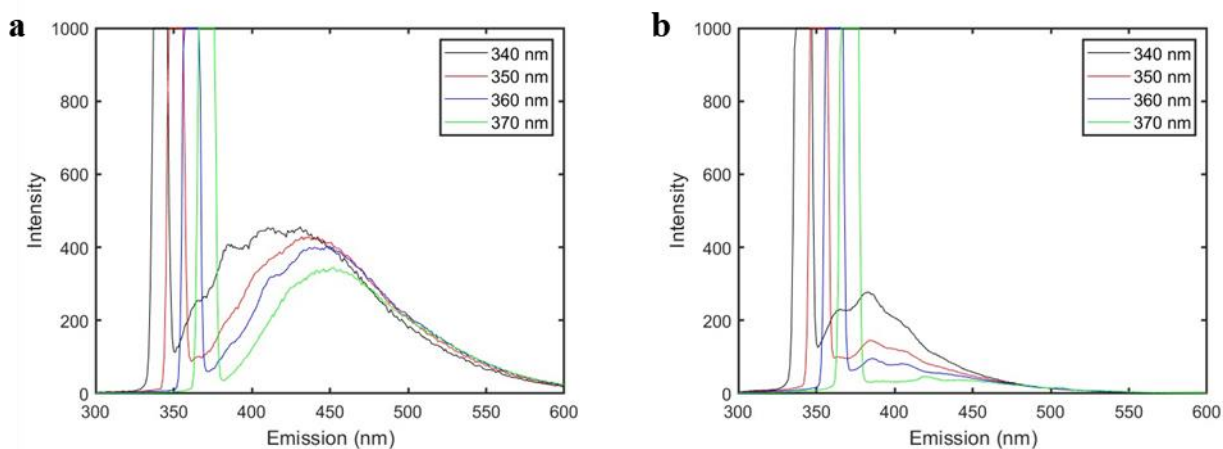


Figure 3-1: Raw fluorescence of (a) control *B. pertussis* fermentation supernatant sample and (b) blank sample of phosphate buffered saline (PBS, 10 mM, pH 7.2).

The filtered fluorescence spectra for two *B. pertussis* fermentation supernatant samples is shown in Figure 3-2. NADPH absorbs at two excitation wavelengths: 260 nm and 340 nm; both of these excitations produce an emission at 460 nm. The excitation at 260 nm occurs because of the adenine group and the excitation at 340 nm arises due to the absorption by the pyridine ring (Rover *et al.*, 1998). However, since NADH also fluoresces at an excitation wavelength of 260 nm (Lakowicz, 2006), we used the peak at an excitation wavelength of 340 nm for detection of NADPH in this work. Another peak within the measured range is located at the excitation/emission pair of 370/480 nm. Since many compounds such as flavins involved in redox reactions fluoresce in this range (Croce and Bottiroli, 2014), we did not use this peak in any of the models in this work.

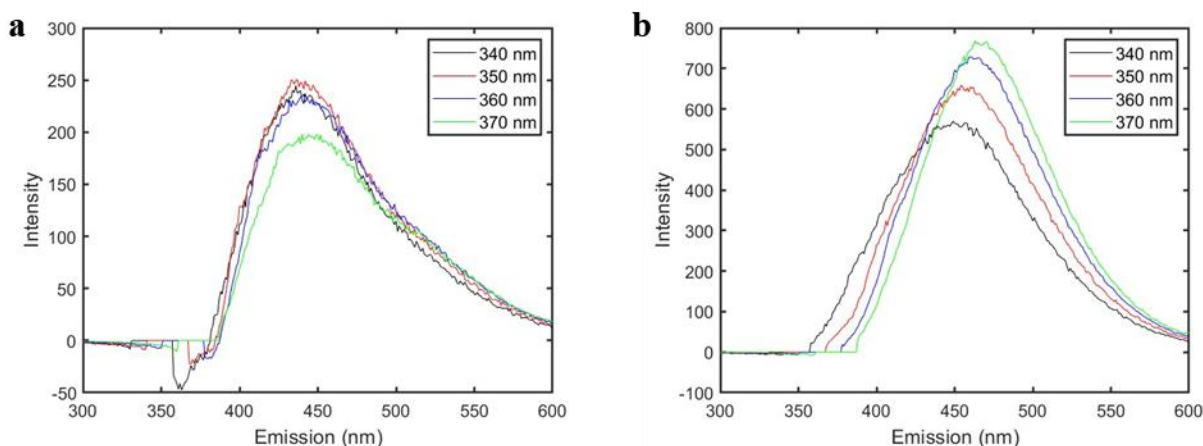


Figure 3-2: Filtered fluorescence spectra for control *B. pertussis* fermentation supernatant samples at (a) 24 hours and (b) 48 hours.

3.5 Flow cytometry

3.5.1 Equipment

Flow cytometry was conducted using an S3e Cell Sorter (Bio Rad). This instrument has 2 excitation lasers at 488 nm (primary) and 640 nm (secondary) with emission filters to detect forward scattering (FSC) and side scattering (SSC) and four channels to detect fluorescence emission (526/48 (FL1), 593/40 (FL2), 670/30 (FL3) and 700 LP (FL4)). All flow cytometry measurements use a total of 20,000 events per sample at a rate of 1000 events per second.

3.5.2 *B. pertussis* gating

Prior to obtaining flow cytometry data, a gating strategy must be implemented to eliminate noise, collect *B. pertussis* cells and exclude doublets. Figure 3-3 depicts the flow cytometric gating protocol used during measurements of single *B. pertussis* cells. In the first gate (Figure 3-3a), the side scattering area is plotted along the x-axis and the fluorescent area in the FL1 detector is plotted along the y-axis. The goal of this gate is to eliminate events which contribute to the scattering intensity but not to fluorescence. The second gate (Figure 3-3b) plots the area of forward scattering along the x-axis and the area of side scattering along the y-axis. This gate locates the region of *B. pertussis* size and complexity and also roughly checks the sample for contamination. Since other bacteria present may have a similar size and complexity, it is critical to ensure they are not collected. This is done by plating on Bordet-Gengou (BG) agar on which *B. pertussis* can grow and Tryptic (Trypticase) Soy Agar (TSA) plates on which *B. pertussis* cannot. The last gate (Figure 3-3c) is used to exclude droplets that contain more than one cell. The width of the side scattering is plotted along the x-axis and the height of the side scattering is plotted along the y-axis. The sample solution is dilute ($OD_{600} = 0.05$) and therefore a narrow gate can be used at the location

where most cells are located. Any cells with a larger side scattering width are considered to contain more than one cell since the height of the side scattering peak will not change with multiple cells in a droplet but the side scattering width will be about double the size for a droplet containing 2 cells. The cells found in Figure 3-3b must pass through the gate found in Figure 3-3a and the cells found in Figure 3-3c must pass through both gates in Figures 3-3a and b. Therefore, the 20,000 events collected for measurements of viability, oxidative stress and surface antigen concentration have already been subjected to the gating protocol in Figure 3-3.

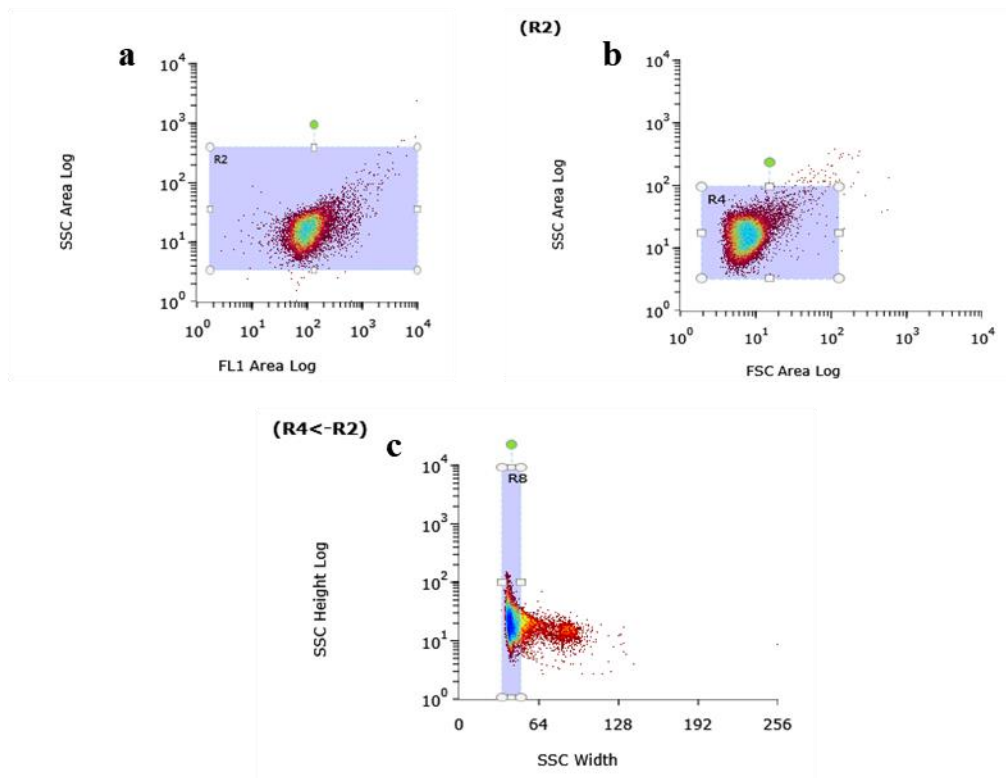


Figure 3-3: Gating protocol for *B. pertussis* flow cytometry measurements. The gates are used as follows: (a) noise exclusion gate, (b) size inclusion gate and (c) doublet exclusion gate.

3.5.3 Viability stain

Viability was measured using SYTO9 (Invitrogen, Thermo Fisher Scientific, USA) and propidium iodide (PI, Invitrogen Thermo Fisher Scientific, USA). SYTO9 is a green fluorescent

nucleic acid dye and can stain live and dead cells, while PI is a red nucleic acid dye that is only able to enter the cells if the membrane has been damaged. Re-suspended cells (1 mL) were stained with 0.12 μ M of SYTO9 and 3 μ M of PI and then incubated in the dark for 10 min prior to measurement.

The fluorescent emissions from SYTO9 and PI were measured using the green fluorescence FL1 and red fluorescence FL3 detectors, respectively. SYTO9 and PI exhibit excitation/emission pairs at 483/503 nm and 530/625 nm, respectively. A sample viability stain is shown below in Figure 3-4. The fluorescence signal from SYTO9 is shown in FL1 and plotted along the x-axis, while the fluorescence signal from PI is shown in FL3 and plotted along the y-axis. The plot of FL3 area versus FL1 area is divided into four quadrants: Q1 represents dead cells, Q2 represents live membrane-damaged cells, Q3 represents the live healthy cells and Q4 has no meaning. The gates are determined by comparing the fluorescence of a dead cell culture (cells treated with 70% ethanol and placed in the freezer for 10 minutes) and a live cell culture. Cells that fall into Q4 are technically live cells. This occurs because a small difference in the fluorescence intensity exists from sample to sample. However, because the wavelength ranges selected as gates are fixed, some live cells will fall outside the Q3 quadrant. The goal is to locate the gates at wavelengths that minimize the amount of cells that fall in this quadrant.

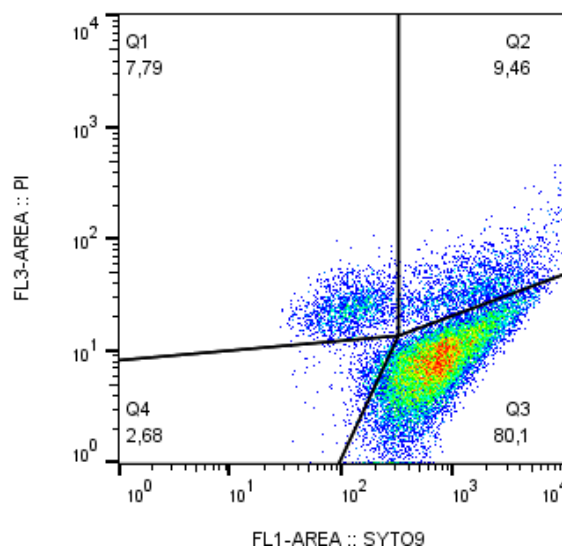


Figure 3-4: Viability staining of a *B. pertussis* sample.

3.5.4 Oxidative stress

Oxidative stress was measured using the fluorescein derivative carboxy-2',7'-dichlorodihydrofluorescein diacetate (carboxy-H₂DCFDA) from Sigma-Aldrich (Canada). A solution containing 1 mM of this reagent was prepared by dissolution in DMSO and added to the samples to yield a final concentration of 50 μ M. The reagent is non-fluorescent and cell-permeable but is deacetylated by esterases intracellularly and converted to the fluorescent DCF when it reacts with hydrogen peroxide and other oxygen derivatives (Halliwell and Whiteman, 2004). To measure the ROS distribution, the FL1 diode was used for DCF emission (green fluorescence) which fluoresces at an excitation/ emission pair of 503/523 nm.

3.5.5 Surface pertactin antigen concentration

To measure the surface concentration of pertactin, an antibody provided by Sanofi (PRN 3-16) was conjugated to FITC for fluorescence detection. A FITC conjugation kit (Abcam, Canada) was used to conjugate the antibody to the fluorophore. The antibody was diluted to a

concentration of 1 µg/ml in DMSO and 1 µL modifier reagent was added to 10 µL antibody solution and mixed gently. The resulting solution was added into FITC-lyophilized material, mixed and incubated in the dark for 15 minutes. After incubation, 1 µL Quencher reagent was added to every 10 µL of antibody used and mixed gently. The conjugated antibody was stored in the dark in a -20°C freezer.

A 1 mL sample was taken from the shake flask and diluted to an OD₆₀₀ of 0.05. The sample was then centrifuged at 4,000 G for 4 minutes and the supernatant discarded. A 25 µL of antibody solution (1µg/mL) was then added to each tube and incubated for 1 hour in the dark at 4°C. Then 1 mL of 1x PBS was added to each tube and centrifuged at 4,000 G for 4 minutes to wash the sample. The supernatant was discarded and the pellet re-suspended in 400 µL PBS.

Along with the PRN 3-16 antibody, a negative control was also conjugated to FITC. This antibody does not attach to *B. pertussis* and is used to check for nonspecific binding of the conjugation. Figure 3-5 shows the unstained *B. pertussis* cells, negative control stain and PRN 3-16 antibody staining. As shown in Figure 3-5b, no cells after the negative control stain are located in Q3 representing the PRN-positive region, reflecting that the negative control does not attach to *B. pertussis*.

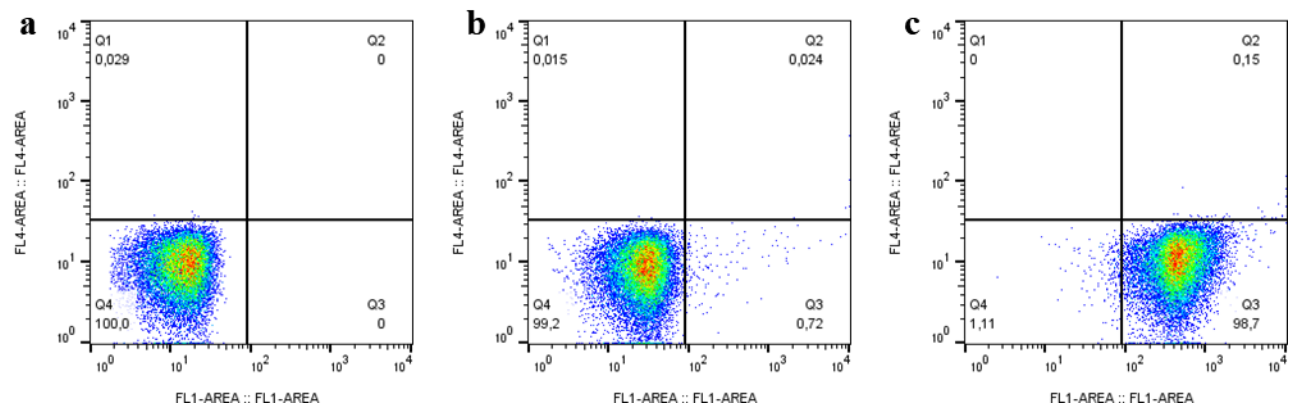


Figure 3-5: Flow cytometry of (a) unstained *B. pertussis*, (b) negative stain of *B. pertussis* with a control antibody conjugated to FITC and (c) *B. pertussis* stained with pertactin surface binding antibody conjugated to FITC.

Chapter 4 Modeling the Effect of Oxidative Stress on *B. pertussis* Fermentations

4.1. Introduction

Whooping cough is a highly contagious respiratory tract disease. The causative agent of this illness is the gram-negative bacterium *B. pertussis*, which was first isolated by Bordet and Gengou in 1906 (Bordet and Gengou, 1906). Early symptoms of whooping cough resemble the common cold, but then evolve into uncontrollable violent coughing fits, discernable by the "whoop" sound made as air is inhaled. Prevention of this disease relies on early vaccination.

The acellular pertussis vaccine produced by Sanofi Pasteur contains five antigens: fimbriae (types 2 and 3), filamentous hemagglutinin, pertactin and pertussis toxin. The manufacturing process of this vaccine is divided into two main stages: i) upstream comprising two parallel trains of three consecutive reactors of increasing volume where the bacteria are grown and ii) downstream purification of the five target antigens in several steps. In a previous study, the causes of batch-to batch variation in the production of the antigens were identified in the upstream fermentation process (Zavatti, 2014). High levels of nicotinamide adenine dinucleotide phosphate (NADPH) were detected in the supernatant of low antigen-producing fermentations. Since NADPH is known to be involved in detoxification reactions required to combat oxidative stress, it was hypothesized that the occurrence of high oxidative stress may negatively affect the productivity (Singh *et al.*, 2007). To further corroborate this point, Zavatti (2019) conducted a series of 2 L and 20 L bioreactor experiments in which the oxidative stress was induced by either

the addition of hydrogen peroxide or a low inoculum dose. These experiments further confirmed that high oxidative stress is correlated to high levels of NADPH, low cell growth and low productivity (Zavatti *et al.*, 2019).

Oxidative stress results when an imbalance occurs between the formation of free radicals from oxygen and the capacity of the system to safely counteract these compounds (Zheng *et al.*, 2009). Reactive oxidative species (ROS) cause oxidative stress and are found as singlet oxygen, hydroxyls, superoxides, peroxides, hydroperoxides and the non-radical hydrogen peroxide which are continuously produced during aerobic metabolism (Eruslanov and Kusmartsev, 2010; Fasnacht and Polacek, 2021). Whether ROS function as signaling molecules or cause oxidative damage depends on the delicate equilibrium between its production and scavenging (Sharma *et al.*, 2012). ROS also serves as second messengers (i.e., intracellular signaling molecules secreted by the cell in response to exposure to extracellular signaling molecules) in a variety of cellular processes including those that promote the tolerance of various environmental stresses; however, when present in excess, they can be harmful to proteins, lipids and DNA (Eruslanov and Kusmartsev, 2010; Storz 1999).

Living organisms utilize counteracting mechanisms involving enzymes (catalase and superoxide dismutase), small proteins (thioredoxin and glutaredoxin), and molecules (glutathione) (Kashmiri and Mankar, 2014) to protect themselves against oxidative stress. The presence of NADPH is important for aerobic organisms to survive oxidative stress since it is a reducing agent that serves to protect against ROS. The enzymes NAD⁺ kinase and NADP⁺ phosphatase are capable of regulating the levels of NAD⁺ and NADP⁺, play pivotal roles in controlling the communication between metabolic pathways that produce NADH and NADPH and are involved in the mechanism of diffusing oxidative stress (Singh *et al.*, 2007). For example, glutathione

reductase can detoxify ROS and use NADPH to replenish its reducing power (Lushchak, 2001). Oxidative stress has also been found to promote NADPH production which in turn tends to diminish oxidative stress by promoting reactions that produce a reductive environment (Singh *et al.*, 2007; Grose *et al.*, 2006).

Mathematical models that involve ROS have been reported, particularly in the context of antibiotic resistance (Farha and Brown, 2013; Kohanski *et al.*, 2010; Yang *et al.*, 2019). Additionally, models have been used to study the effect of oxidative stress on culture outcomes. For example, a model was developed to predict the amount of ROS production in *Escherichia coli*. This model showed that ROS generation can be increased in *Escherichia coli* and lead to higher bacteria susceptibility to oxidative attack via antibiotics (Brynildsen *et al.*, 2013). In another study focused a mechanism that enables the rapid increase in NADPH by exposing *Escherichia coli* to hydrogen peroxide (Christodoulou *et al.*, 2018). This study revealed inhibition of the pentose phosphate pathways which occurs due to the inactivation of glyceraldehyde 3-phosphate dehydrogenase by ROS and allosteric inhibition of the first pentose phosphate pathway enzyme by NADPH. Oxidative stress was also induced to increase industrial triacylglycerol (TAG) production in *Rhodococcus opacus* PD630 (Sundararaghavan *et al.*, 2020).

Several studies on bacteria have reported that the inhibition of glycolysis allows cells to divert flux into the pentose phosphate pathway to promote NADPH synthesis and enhance the protection against oxidative stress (Christodoulou *et al.*, 2018; Sundararaghavan *et al.*, 2020; Mullarky, 2015). On the other hand, *B. pertussis* has been shown to behave differently than other bacteria and consequently is expected to respond differently to oxidative stress. For example, gluconeogenesis is driven by glutamate as the main nutrient in *B. pertussis* rather than by glycolysis. In fact, it has been shown that *B. pertussis* can synthesize most amino acids and be

grown in media containing mostly glutamate and proline which is also converted to glutamate (Stainer and Scholte, 1970). Also, NADPH is produced from the conversions of isocitrate and glutamate to alpha-ketoglutarate and malate to pyruvate, but it is not generated in the pentose phosphate pathway as in *Escherichia coli* (Kanehisa and Goto, 2000).

This chapter presents a metabolic model describing the effect of oxidative stress on the growth of *B. pertussis*. The model is based on a previous theoretical one developed by Himeoka and Kaneko (2017) focused on the effect of starvation and the stationary growth phase of cells. It should be emphasized that the theoretical model of Himeoka and Kaneko was proposed as a generic description of growth inhibition due to the interplay between promoters of growth (ribosomes) and waste products (misfolded or mistranslated proteins). In the current study, we adapt the model to describe a particular mechanism of inhibition of *B. pertussis* growth by ROS (inhibitor) and the role of NADPH (byproduct) as a quencher of ROS. To assess the validity of this model for this specific *B. pertussis* pathway, we have conducted an experimental flask study in which ROS, NADPH, glutamate and biomass concentrations are measured by flow cytometry, spectro-fluorescence, Bioprofile and spectrophotometry, respectively. The experiments are conducted under different conditions of oxidative stress induced by high initial glutamate concentrations, low initial inoculum and secondary culturing following exposure to starvation conditions. Variability in inoculum size or initial glutamate concentration occurs commonly in the manufacturing process due to measurement errors, fluctuations in media composition including casamino acids and differences in seed batches. Also, the behavior of the culture following exposure to starvation conditions is of particular industrial interest since the vaccine is manufactured in a fermentation train of bioreactors of increasing volume. Consequently, the exposure to starvation conditions in one bioreactor may have a significant impact on the following

bioreactors in the train. The main goal of the proposed model is to understand and describe the relations between cell growth, oxidative stress and NADPH under different oxidative conditions. In view of the differences between *B. pertussis* to other bacteria, a main goal of the study is to assess via flask studies and model predictions whether the ROS or another factor such as substrate (glutamate) inhibition level is the key determinant of growth under different ROS-inducing conditions.

Other potential uses of this model include: i) model-based optimization of cell growth, ii) identification of possible sources of ROS by combination with models describing other metabolic pathways that are known to produce oxidative stress and iii) model-based estimation of ROS levels that is difficult to determine from online measurements of biomass and NADPH.

The remainder of the chapter is organized as follows. Section 4.2 presents the mathematical model and its application to the oxidative stress pathway in *B. pertussis*. Section 4.3 describes both the experimental and theoretical methods used in the work and the materials used in the experiments. Section 4.4 presents results including comparison of experimental data and model predictions and possible implications of the model on the manufacturing process. Section 4.5 presents conclusions.

4.2 Model Outline

The general model of Himeoka and Kaneko (2017) was developed to describe the behaviour during the lag, exponential growth, stationary and death phases without resorting to detailed mechanisms. It includes the formation of a complex between the following two types of components: component A which autocatalytically promotes cell growth and component B which

inhibits the autocatalytic process by forming a complex (component C) with component A. This process is shown schematically in Figure 4-1. In the original version of the model, Himeoka and Kaneko hypothesized that component A may correspond to ribosomes. Ribosomes are a well-studied autocatalytic species that promote growth through the synthesis of proteins (Maitra and Dill, 2015; Scott *et al.*, 2010). Component B may represent waste products (i.e. misfolded or mistranslated proteins) or other molecules that are produced with the aid of component A but do not facilitate growth. For example, the production of proteins in *E. coli* related to the stationary phase such as HPF and YfiA is induced under stress conditions such as starvation (Maki *et al.*, 2000; Ueta *et al.*, 2008).

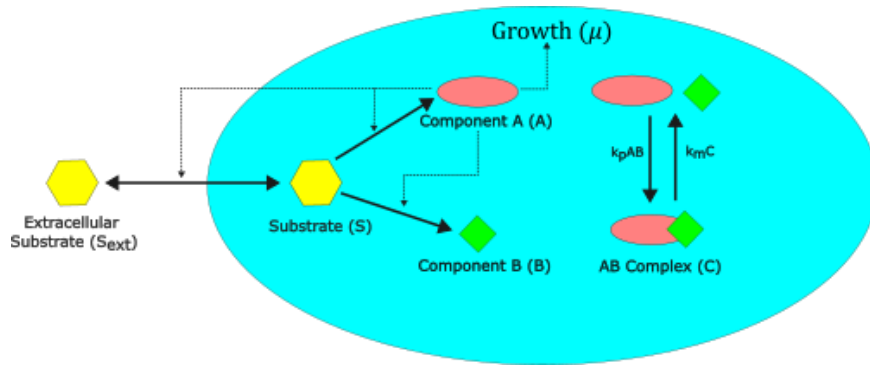


Figure 4-1: Schematic representation of the mechanism described in the model. Adapted from Himeoka and Kaneko (2017).

Once cells enter the stationary phase, a certain time span known as the lag time is generally required for growth to recover after resources are supplied. The originally reported model was used to quantitatively relate the lag time to both the starvation time and the maximal growth rate and satisfy the already-known growth laws in the exponential phase. It also described the dependence of lag time on the intensity of the starvation process. For example, if the external concentration of the substrate is reduced quickly, the lag time is short compared to a situation where the external concentration is reduced more slowly. The fact that some of the phenomena

explained by the model of Himeoka and Kaneko were also observed in our experimental studies for *B. pertussis* (e.g. longer lags following long starvation periods) motivated the use of a modified version of this model in the current work.

Bacteria are consistently challenged to adapt to changes in nutrient availability and stress conditions. The manner in which bacteria adapt to stasis caused by lack of nutrients or other stressful conditions has led to the idea that defenses against increased levels of intracellular oxidative stress are essential traits for non-growing cells (Dukan and Nystrom, 1998). In non-growing cells, the ROS concentration should accumulate since they are not diluted by an increase in cell volume (McDougald *et al.*, 2002). Cells with an elevated concentration of ROS should be much more sensitive to additional oxidative stress imposed by external factors. It has been reported that a major overlap exists between the cell use of global regulators to deal with both starvation and oxidative stress (McDougald *et al.*, 2002). One pathway of adaptation to starvation in bacteria involves the intracellular signal guanosine tetraphosphate which accumulates in response to nutritional deficiency and controls the macromolecular synthesis for entry into starvation and non-growth (Chatterji and Ojha, 2001). One pathway of adaptation to starvation in bacteria involves the intracellular signal guanosine tetraphosphate which accumulates in response to nutritional deficiency and controls the macromolecular synthesis for entry into starvation and non-growth (Chatterji and Ojha, 2001). Guanosine tetraphosphate metabolism is mediated by two global regulatory proteins: ribosome-associated enzymes that respond to a lack of amino acids and bifunctional enzyme-inhibiting enzymes responsible for the breakdown of guanosine tetraphosphate (Cashel *et al.*, 1996; Cashel, 2000). Increased amounts of guanine tetraphosphate activates the synthesis of RNA polymerase sigma (Gentry *et al.*, 1993) which controls genes essential for the starvation stress defense in many bacterial species (Hengge-Aronis, 1996). For

example, RNA polymerase sigma regulates the gene *bolA*. This gene is involved in the morphological adaptation of starved *E. coli* cells that affects the membrane structure and therefore plays a role in osmotic shock and oxidative stress (Santos *et al.*, 1999).

NADPH is extremely important to reduce oxidative stress through reactions that produce a reductive environment (Grose *et al.*, 2006). To avoid reactive oxygen intermediates produced during the respiratory burst, bacteria produce enzymes, such as catalase and superoxide dismutase, which detoxify peroxides by transforming superoxide radicals into hydrogen peroxide and oxygen (Farr and Kogoma, 1999). NADPH also contributes to the proper functioning of enzymes such as superoxide dismutase, glutathione peroxidase and catalase. *B. pertussis* produces a catalase and a Fe-superoxide dismutase (Khelef *et al.*, 1996). NADPH is known to be tightly bound to catalase and to offset the ability of the hydrogen peroxide which is the substrate of catalase to convert the enzyme to an inactive state. In the process, the bound NADPH becomes NADP^+ and is replaced by another molecule of NADPH (Kirkman *et al.*, 1999). Another method of reducing oxidative stress in *B. pertussis* is through reduced glutathione. Reduced glutathione scavenges superoxide and hydroxyl radicals non-enzymatically and serves as an electron donor to several enzymes involved in ROS detoxification (Pannala *et al.*, 2013). NADPH is then used to replenish the glutathione to its reduced state, as shown in Figure 4-2.

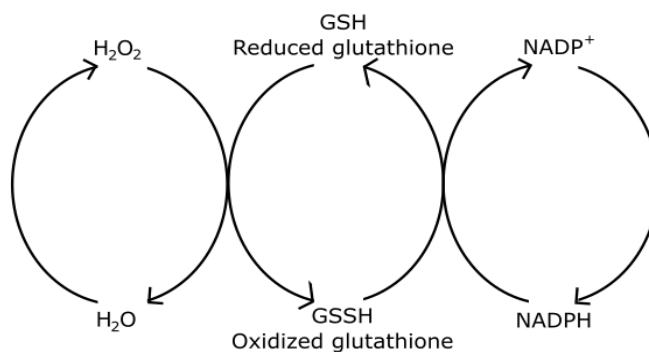


Figure 4-2: Glutathione reductase pathway for reducing oxidative stress in the form of hydrogen peroxide

NADPH is not only vital for the anti-oxidative defense mechanisms of most organisms, but it is also the driving force of most biosynthetic enzymatic reactions, including those responsible for the biosynthesis of all major cell components such as DNA and lipids (Arnér and Holmgren, 2000; Koh *et al.*, 2004; Singh *et al.*, 2008;). NADPH is produced during the metabolism of glutamate entering the TCA cycle and in the conversions of isocitrate to alpha-ketoglutarate and malate to pyruvate. While NADPH is used only for biosynthetic reactions and cannot be used to generate ATP, *B. pertussis* contains transhydrogenase (Parkhill *et al.*, 2003) which converts NADPH into NADH. NADH can be converted into ATP via the electron transport chain. Also, ATP production via oxidative phosphorylation is not effective in the absence of NADPH (Singh *et al.*, 2007).

Considering the relationship between starvation and oxidative stress as well as the role NADPH plays in the synthesis of major cellular components and quenching of ROS, we propose a modified version of the model of Himeoka and Kaneko to describe the interplay between the main nutrient (glutamate), ROS and NADPH. Since component A promotes growth and is used to quench component B in their general model, we hypothesize that NADPH (quencher) can be

defined as component A and ROS (inhibitor) as component B. As discussed above, NADPH is instrumental in quenching the cell of ROS and promoting cell growth (Singh *et al.*, 2007). Since component C is the product of the reaction of NADPH with ROS and NADPH is converted to NADP^+ by quenching ROS, we denote NADP^+ in the model as component C. Lastly, glutamate is defined as S since it is the major substrate in the media.

We have modified and adapted the model reported by Himeoka and Kaneko (2017) to our study by including the following additional elements: a substrate inhibition term, exponents in the expressions for the fluxes of A and B defining their dependence on the intracellular substrate concentration, a degradation term for ROS and a mass transfer coefficient for the transport of NADPH into the supernatant. The substrate inhibition term involving glutamate and exponents on the fluxes were determined by calibration of the model to experimental data that we collected (see sections 4.2 and 4.6, respectively). The degradation of ROS is included since these species react very rapidly with other compounds if not quenched by an antioxidant. The mass transfer term has been added to account for the secretion of NADPH into the supernatant. A reported pathway involving the secretion of catalase bound to NADPH into the environment (Wan *et al.*, 2017) supports the inclusion of the mass transfer term.

The model given by Eqns (4.1) – (4.10) below is based on transient mass balances of species and rate expressions for the reactions described schematically in Figure 4-1. Table 4-1 summarizes the definitions of the symbols used for the concentrations of various species, while Table 4-2 defines the model parameters. It should be noted that glutamate (substrate S) contributes to both the production of NADPH (A) and ROS (B) through the terms $F_A A$ (equal to the growth μ) and $F_B A$, respectively, in this model. Both these terms introduce autocatalytic effects on A (NADPH) by which a higher level of NADPH promotes further growth (μ) and substrate

consumption. Both the terms F_A and F_B are functions of the intracellular substrate concentration

S. The term $k_p AB$ accounts for the co-metabolism of NADPH and ROS.

$$\frac{dS_{ext}}{dt} = -\frac{\mu X}{Y_{x/s}} \quad (4.1)$$

$$\frac{dX}{dt} = \mu X - DX \quad (4.2)$$

$$\frac{dS}{dt} = -F_A A - F_B A + \alpha A (S_{ext} - S) - \mu S \quad (4.3)$$

$$\frac{dA}{dt} = F_A A - k_p AB + k_m C - \mu A - K_m A \quad (4.4)$$

$$\frac{dA_{ext}}{dt} = (K_m + D) A X \quad (4.5)$$

$$\frac{dB}{dt} = F_B A - k_p AB + k_m C - \mu B - d_B B \quad (4.6)$$

$$\frac{dC}{dt} = k_p AB - k_m C - \mu C \quad (4.7)$$

$$F_A = \frac{\left(\frac{v}{1 + S/K_{si}} \right) S}{KX + S} \left(\frac{S}{K_t + S} \right)^{1.5} \quad (4.8)$$

$$F_B = \frac{\left(\frac{v}{1 + S/K_{si}} \right) S}{KX + S} \left(\frac{K_t}{K_t + S} \right)^{1.5} \quad (4.9)$$

$$\mu = \alpha F_A A \quad (4.10)$$

Table 4-1: Definitions of concentrations used in the model.

Compound	Model Variable
X	Biomass (<i>B. pertussis</i>) intensity
S	Intracellular glutamate concentration (mmol/L)
S _{ext}	Extracellular glutamate concentration (mmol/L)
A	Intracellular NADPH concentration (mmol/L)
A _{ext}	Extracellular NADPH concentration (mmol/L)
B	ROS concentration (mmol/L)
C	NADP ⁺ concentration (mmol/L)

Table 4-2: Parameters defined in model. The numerical values of the parameters are determined using the model and calibration set in Section 4.4.

Symbol	Meaning	Units
μ	rate constant for biomass growth	$\frac{1}{hr}$
$Y_{x/s}$	yield coefficient	$\frac{intensity\ biomass}{g\ glutamate}$
D	rate constant for biomass death	$\frac{1}{intensity\ biomass\ hr}$
F_A	flux of substrate conversion into A	$\frac{1}{hr}$
F_B	flux of substrate conversion into B	$\frac{1}{hr}$
k_p	rate constant for $A+B \rightarrow C$	$\frac{1}{mM\ hr}$
k_m	rate constant for $C \rightarrow A + B$	$\frac{1}{hr}$
d	rate of degradation of B (ROS)	$\frac{1}{hr}$
ν	rate constant for the generation of NADPH	$\frac{1}{mM}$
K_{si}	rate constant for inhibition	$\frac{1}{hr}$
K	saturation constant for F_A and F_B	mM
K_t	phenomenological constant changing A/B balance	mM
K_m	transport function for A across the cell membrane into the supernatant	$\frac{1}{intensity\ biomass\ hr}$
α	unit conversion constant (=1)	$\frac{1}{mM}$

4.3 Materials and Methods

4.3.1 *B. pertussis* batch fermentations

B. pertussis fermentations were carried out in 250 mL flasks (polycarbonate, sterile, with a vented cap having 0.22 μm pore-size PTFE membrane from VWR, Canada) with a working volume of 60 mL by incubation at 36°C and 200 rpm for 48-60 hrs. Each experiment was carried

out with two flasks operating in parallel under identical conditions. The formulation of media and growth factor components used are found in section 3.1 in Chapter 3. The manipulated variables that have been found to affect the oxidative stress in the culture include the initial concentration of cells, initial concentration of glutamate, duration of starvation time prior to inoculation and changes in *B. pertussis* working seeds produced approximately every 3 months from a seed bank (Zavatti 2014; Wellen and Thompson, 2010). The two seeds used in this work were generated in different years – 2015 and 2018. Two separate 1 mL samples were collected at a time for measurement during the course of the fermentations. Each of the 1 mL samples was measured for its OD, glutamate, NADPH and ROS concentrations. The data shown are averages over the four measurements obtained on each of the 1ml-samples collected from the two identical flasks at each time. The cells and supernatant were separated by centrifugation at 10,000 G for 3 minutes. Glutamate and NADPH concentrations were measured by BioProfile and fluorescence spectroscopy, respectively. The cells were re-suspended in 1 mL phosphate buffered saline (PBS, 10 mM, pH 7.2) and diluted to an OD of 0.05 for flow cytometry measurements of the ROS level. Cell growth was monitored by optical density.

4.3.2 Optical density (OD)

The sample was diluted by a factor of 20 in 0.9% saline solution and OD obtained using optical spectrophotometry by measurement of the absorbance at 600 nm.

4.3.3 Glutamate concentration

The supernatant from the centrifugation step above was collected to determine the glutamate concentration (after 20-fold dilution) using the chemistry module of a BioProfile FLEX Analyzer (Nova Medical). Glutamate was measured using an amperometric electrode that contains

immobilized enzymes in its membranes (Derfus *et al.*, 2009). In the presence of oxygen, the enzyme membranes reduce glutamate to produce hydrogen peroxide which is then oxidized at a platinum anode held at constant potential. The resulting flow of electrical current is proportional to the sample concentration.

4.3.4 Fluorescence spectroscopy

Analysis of the supernatant for extracellular NADPH concentration by fluorescence excitation-emission spectroscopy was carried out with a Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies). Each of the supernatant samples was diluted 20 times and analyzed in polymethylmethacrylate cuvettes using a slit width of 5 nm and PMT of 800 V for the peak measurements. Each sample was analyzed over the excitation and emission ranges of 330-370 nm (at 10 nm intervals) and 300-600 nm (at 1 nm intervals). The blank used was phosphate-buffered saline (PBS, 10 nM, pH 7.2). It should be noted that the region of fluorescence analyzed is generally correlated to the amount of extracellular NADPH present (Rover *et al.*, 1998).

4.3.5 Flow cytometry

Oxidative stress was measured using the fluorescein derivative carboxy-2',7'-dichlorodihydrofluorescein diacetate (carboxy-H₂DCFDA) from Sigma-Aldrich (Canada). A solution containing 1 mM of this reagent was prepared by dissolution in DMSO and then added to the samples to yield a final concentration of 50 μ M. The reagent is non-fluorescent and cell-permeable but is deacetylated by esterases intracellularly and converted to the fluorescent DCF when it reacts with hydrogen peroxide and other oxygen derivatives (Halliwell and Whiteman, 2004).

Viability was measured using SYTO9 (Invitrogen, Thermo Fisher Scientific, USA) and propidium iodide (PI, Invitrogen Thermo Fisher Scientific, USA). SYTO9 is a green, fluorescent nucleic acid dye and can stain live and dead cells while PI is a red nucleic acid dye that is only able to enter the cells if the membrane has been damaged. Re-suspended cells (1 mL) were stained with 0.12 μ M of SYTO9 and 3 μ M of PI and then incubated in the dark for 10 min prior to measurement.

Flow cytometry was conducted using an S3e Cell Sorter (Bio Rad). This instrument has 2 excitation lasers at 488 nm (primary) and 640 nm (secondary) with emission filters to detect forward scattering (FSC) and side scattering (SSC) and four channels to detect fluorescence emission (526/48 (FL1), 593/40 (FL2), 670/30 (FL3) and 700 LP (FL4)). To measure the ROS distribution, the FL1 diode was used for DCF emission (green fluorescence) which fluoresces at an excitation/emission pair of 503/523 nm. This measurement involved a total of 20,000 events per sample at a rate of 1000 events per second. The average value obtained from 20,000 events was used in the model. Fluorescence emission for SYTO9 was measured using the green fluorescence FL1 detector and for PI using the red fluorescence FL3 detector. SYTO9 and PI have excitation/emission pairs at 483/503 nm and 530/625 nm, respectively. The same total events (20,000) and rate (1000 events per second) was used for each sample. Viability was included in the model by only considering cells that had not experienced membrane damage. Dead cells were considered to be PI-positive. The percentage of PI-positive cells in the 20,000 events was determined and the OD was adjusted to only include viable cells. The details regarding the gating procedure used to measure oxidative stress and viability can be found in section 3.5 in Chapter 3.

4.3.6 Model fitting

The model given in Eqns (4.1) – (4.10) was calibrated using the `fmincon` subroutine in MATLAB. This function is designed to find the minimum of a nonlinear multivariate function. For the problem in this study, the sum of the root mean-square error (RMSE) of the measured variables (biomass, extracellular glutamate concentration, intracellular ROS concentration and extracellular NADPH concentration) was minimized. To ensure that all four variables have comparable effects on the outcome of the model, the RMSE of each variable was standardized by dividing it by its corresponding mean, as shown in the following expression:

$$RMSE = \sum \frac{\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}}{\mu_i} \quad (4.11)$$

where y_i is the measured value, \hat{y}_i is the predicted variable, n is the number of calibration samples and μ_i is the mean of each variable.

The model was calibrated with 115 samples obtained under 13 different flask conditions and validated with 26 samples collected from an additional 3 conditions. The OD, glutamate, NADPH, and ROS concentrations were measured for each of these samples. The parameters obtained by fitting the model to the experimental data include the various rate constants (k_p , k_m , D , d , v , $Y_{x/s}$), mass transfer coefficient of A (K_m), substrate inhibition constant for glutamate (K_{si}), saturation constants (K , K_t), initial concentration of component A (intracellular NADPH) and the ROS and NADPH concentration factors to convert their measurement intensities into concentration units. Since the intracellular NADPH concentration is not measured, it is not possible to specify its initial level and so is included among the parameters to be fitted in the model. These terms are defined in Table 4.2 appearing in section 4.2.

Selection criteria are important tools to obtain a model that simultaneously has the appropriate structure and dimensionality. These criteria assess whether a fitted model maintains an optimal balance between its predictive accuracy and number of fitting parameters. The Akaike information criteria (AIC) introduced by Akaike (1973) estimates the prediction error from the extension of the maximum likelihood principle, i.e.,

$$AIC = -\log(L) + 2k \quad (4.12)$$

where L is the maximum likelihood function and k is the number of predicted model parameters in the model. The lower the value of the AIC, the better the model fitting is in terms of structure and predictability. Assuming normal errors, the maximum likelihood function can be replaced by the residual sum-of-squares (RSS) (Cavanaugh and Neath, 2019), i.e.,

$$AIC = -n \log\left(\frac{RSS}{n}\right) + 2k \quad (4.13)$$

where n is the number of samples to which the model is fit. However, this AIC may become inaccurate when the sample size is small ($n/p < 40$ with p representing the number of fitted parameters) (Yang, 2019). In this case, the corrected version of AIC shown below in Eqn (4.14) can be used.

$$AIC = -n \log\left(\frac{RSS}{n}\right) + 2k + \frac{2k(k+1)}{n-k-1} \quad (4.14)$$

4.4 Results and Discussion

4.4.1 Trends of the control *B. pertussis* flask culture

Control conditions were used to mimic those in the first reactor in the train at the Sanofi production facility. These conditions were run using the Sanofi media formulation containing 9.8 g/L of glutamate. This involved inoculating a secondary flask with cells collected during the exponential growth phase from the primary flask. Figure 4-3 shows the (a) OD, (b) glutamate, (c) extracellular NADPH and (d) intracellular ROS concentrations measured over the 48 hours of reaction time.

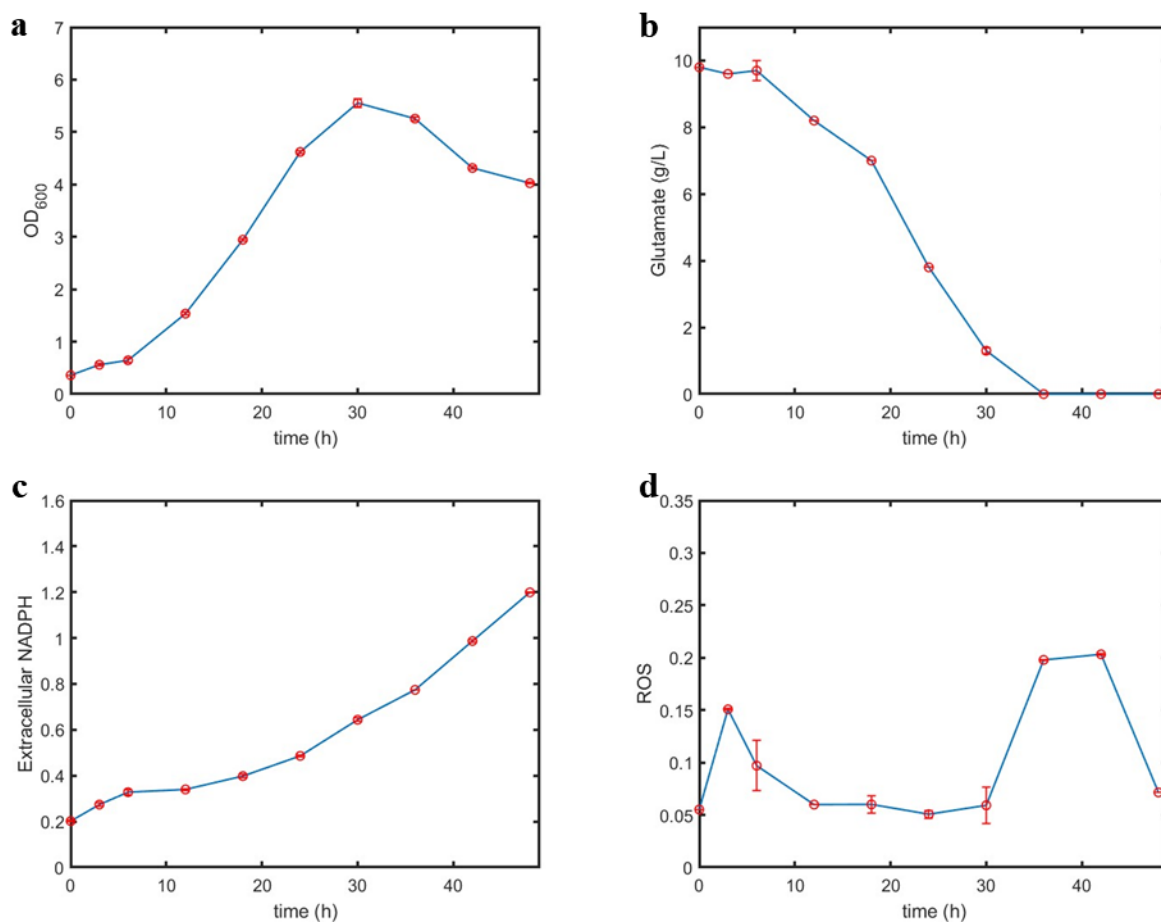


Figure 4-3: Evolution of (a) optical density (OD), (b) extracellular glutamate, (c) extracellular NADPH and (d) intracellular ROS concentrations in a *B. pertussis* flask culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L.

A short lag phase in this culture is observed because the inoculating cells are taken from the exponential growth phase of the primary culture. It is noteworthy that the ROS concentration exhibits a spike rise during the lag phase. In a study by Rolfe *et al.* (2012), it was suggested that oxidative damage in the lag phase is caused by a combination of increased intracellular iron concentration (found in media) and newly available oxygen. Also, the rate of hydrogen peroxide generation was found to be 5 to 10 times greater when a bacterial culture shifts from the lag to exponential growth phase (González-Flecha and Demple, 2001). In general, when the amount of

NADH and FADH generated by the TCA cycle exceed the capacity of the electron transport chain, this leads to high ROS concentration levels (Zhao et al, 2019).

During the exponential growth phase, the ROS concentration drops to its lowest value where it remains constant. Figure 4-4 shows the change in extracellular NADPH concentration per unit biomass with time. As shown in this figure, the amount of NADPH excreted per cell remains constant after the lag phase.

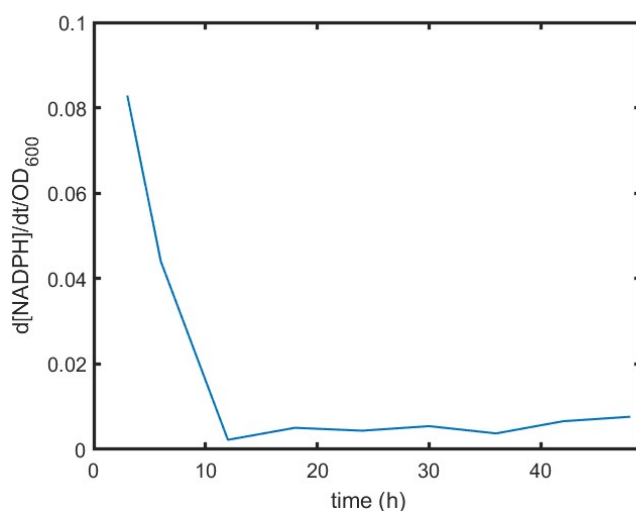


Figure 4-4: Change in extracellular NADPH concentration per unit biomass of a control *B. pertussis* flask culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L.

During the stationary phase (after ~30 hours), the concentration of ROS builds up again (Figure 4-3d). This results in oxidative damage, particularly to DNA, lipids and proteins (Eruslanov and Kusmartsev, 2010; Storz 1999). Stationary cells are particularly vulnerable to oxidizing conditions when they lack energy and material to repair or replace the damaged molecules (Schurig-Briccio *et al.*, 2009) or to oxidative stress when ROS cannot be quenched. For example, it has been shown that growth under starvation conditions generates oxidative stress. The increased level of ROS comes from the imbalance between ROS production and antioxidant

mechanisms (Aubron *et al.*, 2012). Figure 4-5 shows the viability of *B. pertussis* cells of the control culture during the exponential growth, stationary and death phases as measured with flow cytometry. An explanation on how to read the flow cytometry figures is discussed in section 3.5.3 in Chapter 3. The different quadrants in Figure 4-5 correspond to different ranges of cell viability values according to their PI values. As shown in Figure 4-5a, almost all the cells are viable during the exponential growth phase (when the growth rate is equivalent to the maximum growth rate) and so they are found in quadrant 3. However, when the cells are in the stationary phase (Figures 4-5b and c), the numbers of both dead cells (quadrant 1) and membrane-damaged live cells (quadrant 2) increase. When the cells reach the death phase (Figure 4-5d), the amount of PI-positive cells rises; 50.9% of the cells have membrane damage and 21.4% of cells are dead after 48 hours into the incubation. At the 36-hour and 42-hour time points, the intracellular concentration of ROS rises sharply (Figure 4-3d), possibly due to lipid damage in the membrane, as has been observed in previous experiments conducted under starvation conditions (Arts *et al.*, 2015). The decrease of RO during the death phase may be due to its reaction with other cellular components or due to a severely damaged membrane which allows it to leak into the extracellular environment.

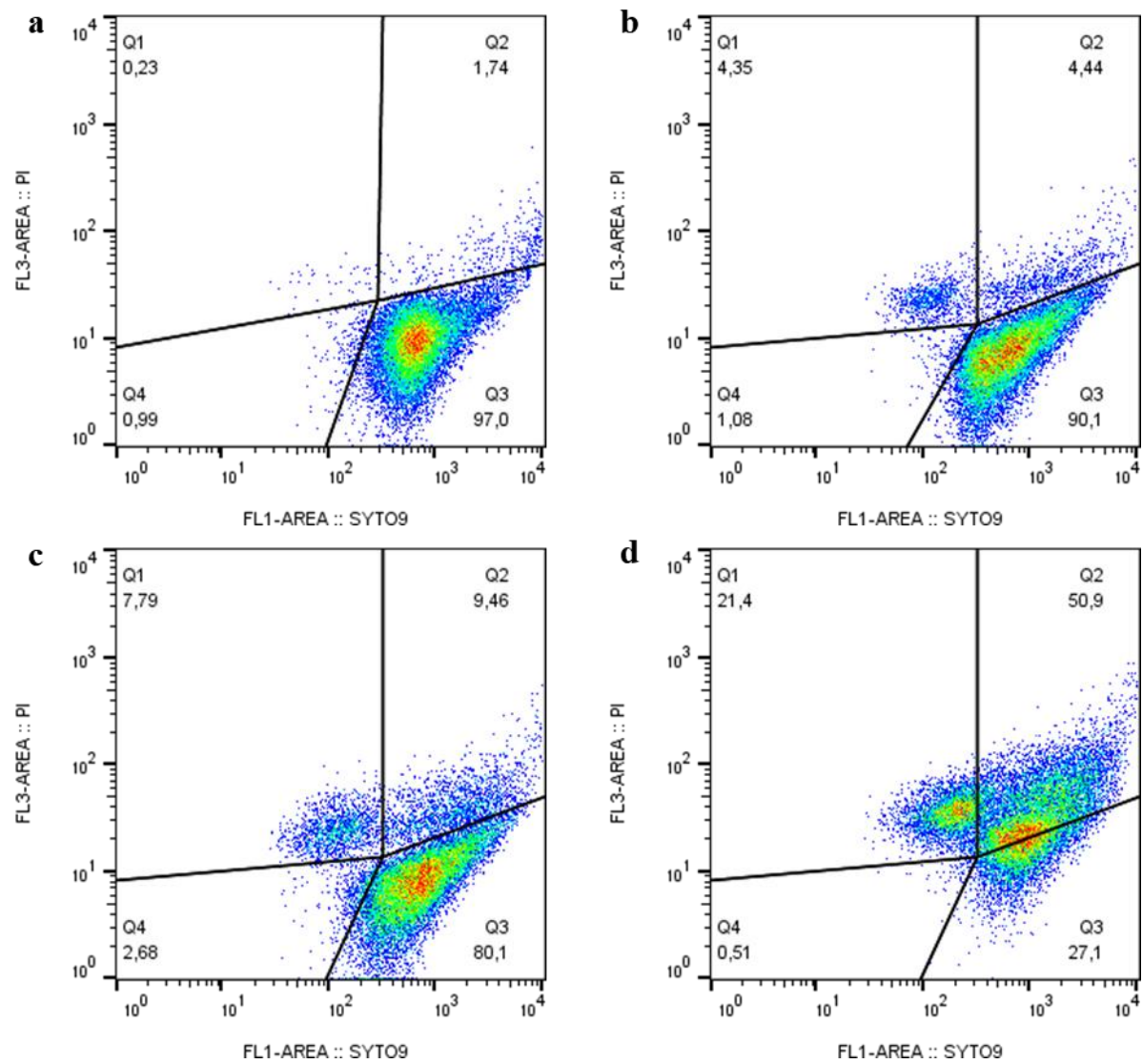


Figure 4-5: Viability PI-based staining of a control *B. pertussis* flask culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L after a) 24 hours, b) 36 hours, c) 42 hours, and d) 48 hours.

4.4.2 Glutamate and substrate inhibition of *B. pertussis*

As shown in Figure 4-6 and Table 4-3, the growth rate of biomass increases as the initial concentration of glutamate in the media decreases. For example, the doubling time of the bacteria is shortened by ~0.86 hours when the initial concentration of glutamate is reduced from 11.6 g/L to 6.6 g/L as shown in Table 4-3. Also, the intracellular ROS concentration during the lag and early exponential phases rises as the initial concentration of glutamate increases (Table 4-3). Since a larger lag appears between the start of the fermentation and the time that glutamate is first consumed at the higher glutamate concentration, we hypothesize that glutamate inhibition is occurring. Accordingly, we have added a substrate inhibition term into the mathematical model to account for this experimental observation. Oxidative stress caused by nutrient excess has been reported in both mammalian and bacterial cells (Wellen and Thompson, 2010; Wang and Levin, 2009). The tricarboxylic acid (TCA) cycle nutrients and the resulting electrons are transferred to produce NADH and FADH₂. The electrons are eventually donated to molecular oxygen via the electron transport chain (ETC), but incomplete reduction can lead to the production of superoxides. Therefore, when the breakdown of components from the TCA cycle exceeds the consumption rate capacity of the electron chain cycle as is the case when nutrients are in excess, ROS production increases and can lead to oxidative stress (Wellen and Thompson, 2011). It has also been shown that high levels of NADH and FADH may inhibit the fluxes in the TCA cycle (Vemuri *et al.*, 2006). Since glutamate is consumed primarily in the TCA cycle, this process will be inhibited if the resulting NADH and FADH reach high enough levels.

On the other hand, the data in Figure 4-6 indicate that the amount of biomass produced at higher initial glutamate concentrations eventually catches up and surpasses the level attained when the glutamate concentration is lower. Such a result is not unexpected since larger overall glutamate

consumption should yield a higher level of biomass production. The practical implication of this observation is that a fed-batch operation would be best to increase the growth rate of *B. pertussis* by reducing the impact of initial substrate inhibition by glutamate which is the main carbon source in this process.

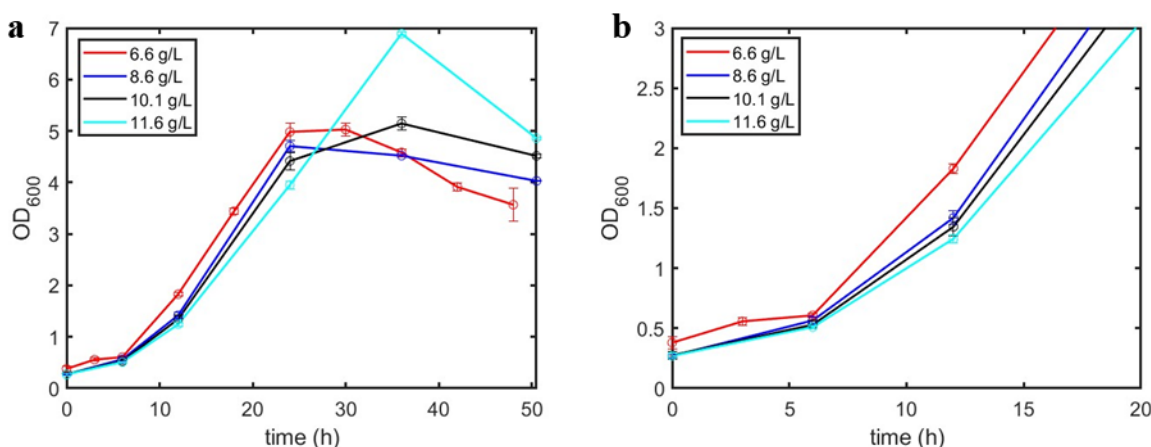


Figure 4-6: Variation of OD during fermentation at different initial glutamate concentrations over the (a) full 50 hours and (b) first 18 hours of incubation.

Table 4-3: Comparison of OD and ROS levels and biomass doubling time over the first 12 hours of fermentation at different initial glutamate concentrations

Initial glutamate concentration (g/L)	OD ₆₀₀ (intensity)			Average ROS (intensity)			Doubling time (hrs)	Growth Rate (biomass/hr)
	0 h	6 h	12 h	0 h	6 h	12 h		
6.6	0.38	0.61	1.63	101	102	85	5.37	0.1291
8.6	0.27	0.57	1.42	138	123	87	6.01	0.1152
10.1	0.28	0.53	1.35	155	136	87	6.10	0.1137
11.6	0.27	0.51	1.25	174	141	96	6.23	0.1112

4.4.3 Effect of initial OD on oxidative stress in *B. pertussis*

Experiments carried out with lower initial density cultures of *B. pertussis* (i.e., small size inoculum) exhibited higher oxidative stress during the lag phase and early portion of the

exponential growth phase (Table 4-4) although had a higher biomass growth rate. The difference in oxidative stress with respect to the control conditions described in the previous section was unexpected since the intracellular concentration that can generate stress should be independent of inoculum size for the same extracellular nutrient concentration. Also, no clear correlation is observed between the oxidative stress and biomass growth for conditions with different initial biomass concentrations. The phenomenon of quorum sensing provides a possible explanation for the increased stress in low inoculum cultures and the small impact of oxidative stress on growth. Quorum sensing is the regulation of gene expression in response to fluctuations in cell population density and involves quorum-sensing bacteria that produce and release chemical signal molecules called auto-inducers (Miller and Bassler, 2001). Even minute amounts of an auto-inducer can alter gene expression. Most quorum sensing-controlled processes are unproductive when undertaken by an individual bacterium acting alone but become beneficial when carried out simultaneously by a large number of cells (Waters and Bassler, 2005). *B. pertussis* is reportedly able to use quorum sensing (Serra *et al.*, 2008; Van Beek *et al.*, 2018) although no evidence exists that they use quorum sensing in response to oxidative stress. However, other bacteria have been reported to use quorum sensing as a protective mechanism against oxidative stress. For example, *Burkholderia pseudomallei* uses quorum sensing to regulate the response to oxidative stress (Lumjiaktase *et al.*, 2006). A non-specific DNA-binding protein DpsA was found to play a key role in protecting *B. pseudomallei* from oxidative stress mediated by organic hydroperoxides. Such protective action could explain the observation of higher stress in low inoculum cultures when the secretion of a smaller amount of protective protein would be expected than when the inoculum level is higher. It has also been reported that oxidative stress drives the selection of quorum sensing mutants in *Staphylococcus aureus* populations (George *et al.*, 2019). Inhibition with respect to biomass

concentration has been added in the model in Eqns (4.7) and (4.8) to account for the effect of inoculum size on oxidative stress levels.

Table 4-4: Comparison of OD and ROS levels over the first 12 hours of fermentation at different initial concentrations of glutamate and half the initial OD level.

Initial glutamate concentration (g/L)	OD ₆₀₀ (intensity)			Average ROS (intensity)			Growth Rate (biomass/hr)
	0 h	6 h	12 h	0 h	6 h	12 h	
8.4	0.14	0.27	0.70	159	121	83	0.1392
8.6	0.27	0.57	1.42	138	123	87	0.1152
10.4	0.13	0.25	0.64	171	141	85	0.1412
10.1	0.28	0.53	1.35	155	136	87	0.1137

4.4.4 Effect of starvation on the growth rate of *B. pertussis*

In order to simulate starvation conditions that may occur in a train of fermenters, the growth of a culture was continued in a primary flask for approximately 10 hours after glutamate had been completely depleted. The culture collected at that time was then transferred to a secondary flask for sampling and testing. Figure 4-7 presents comparisons between the growth curves in the starved secondary flasks with those obtained in a control flask operated with the same media at initial glutamate concentrations of 10.1 and 6.6 g/L. As shown in Figure 4-7, the lag phase (taken as the time duration for the initial biomass population to double) in the case of the starved flask is approximately 36 hours longer at both glutamate levels. However, given that the OD in the starved case remains very low and does not vary over the lag phase, we infer that the growth rate is negligible and so the death rate must also be very small. Interestingly, the initial concentration of glutamate in the media does not affect the lag phase of the starved cases.

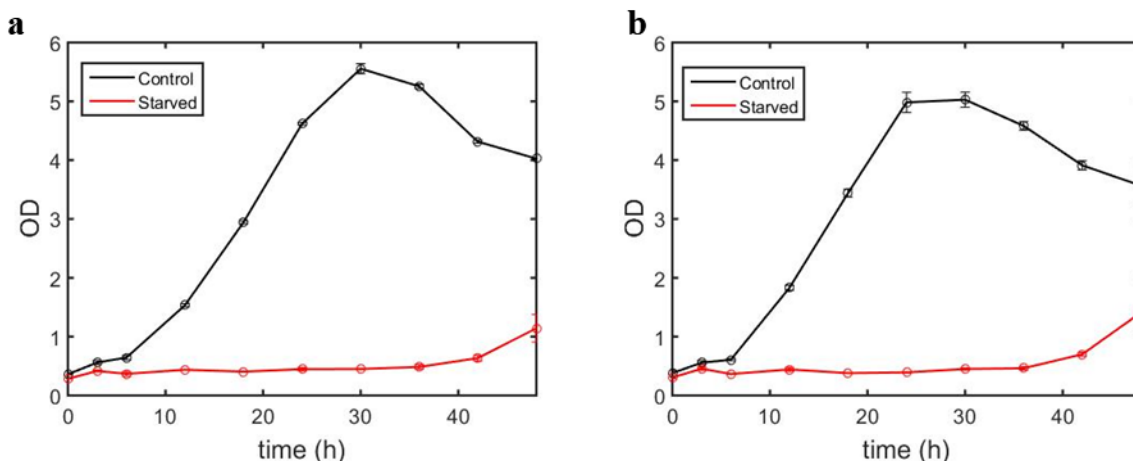


Figure 4-7: Biomass growth curves obtained in control and 12-hour starved secondary flasks with media initially containing (a) 10.1 g/L and (b) 6.6 g/L glutamate.

Figure 4-8 depicts the PI-based viability of the culture obtained under control and starved conditions at a glutamate concentration of 10.1 g/L. Compared to the control, the culture obtained under starved condition exhibits a larger percentage of dead cells (53.4% compared to 1.19%) and membrane-damaged live cells (20% compared to 3.65%). Over the next 6 hours in the starved condition (Figure 4-8c), the membrane-damaged cells are either repaired or die, as shown by the decrease in the population of cells in Q2 and the simultaneous increase in the percentage of cells in Q1 and Q4. As the 24-hour and 36-hour points are reached, a larger percentage of cells becomes viable (Figures 4-8d-e), as reflected by the increase of the population in Q3. When the biomass concentration begins to increase at the end of the lag phase, 90.1% of the cells have become viable (Figure 4-8f).

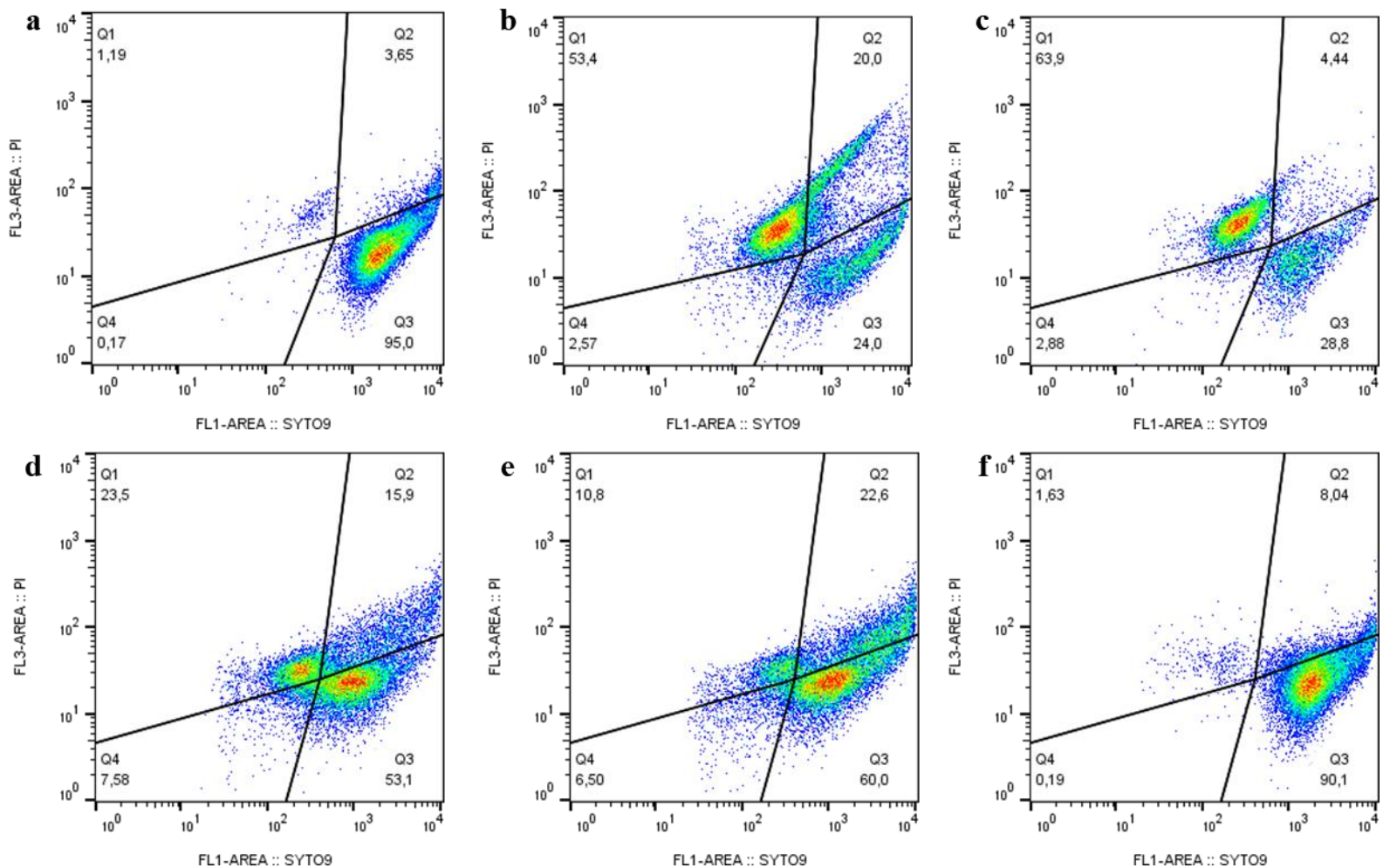


Figure 4-8: Viability staining of a control and starved *B. pertussis* flask cultures with an initial glutamate concentration of 10.1 g/L: a) control - 0 hours, b) starved - 0 hours, c) starved - 6 hours, d) starved - 24 hours, e) starved – 36 hours, and f) starved - 48 hours.

4.4.5 Effect of seed batches

Two seed batches have been compared to assess the batch-to-batch variability as well as the effect of the length of time that the seed batches are stored in a -80°C environment prior to incubation. The procedure followed to generate these seed batches is as follows. A library seed was re-suspended in a 6 mL flask, plated for 4 days and inoculated in another flask with a working volume of 1 L for 24 hours. Glycerol was then added to the resulting culture (10% by volume) and 2 mL aliquots were placed in a -80°C freezer for storage. The OD₆₀₀ and number of colony-forming units were measured several times during the protocol for quality control. Figures 4-9a and b depict the growth curves, intracellular concentration of ROS of the seeds using the Sanofi media composition. Seed 2 (produced in 2018) has a shorter lag phase than seed 1 (2015) which can be explained by the difference in viability between the seeds. Seed 2 contains 24.9% more live cells than seed 1 (comparison of Figures 4-9c and 4-9d). This could be due to some unknown difference in the procedures followed to obtain the seed batches or the fact that seed 1 was stored at -80°C for three years longer than seed 2. Since exposure to cryogenic temperatures slows down deterioration but does not prevent it entirely, the viability of cells likely decreases the longer they are stored prior to incubation (Walters *et al.*, 2004). The rate of deterioration depends on several factors including the species involved, protocol, aging and additives (Walters *et al.*, 2014; Hubalek, 2003; Simone, 1992).

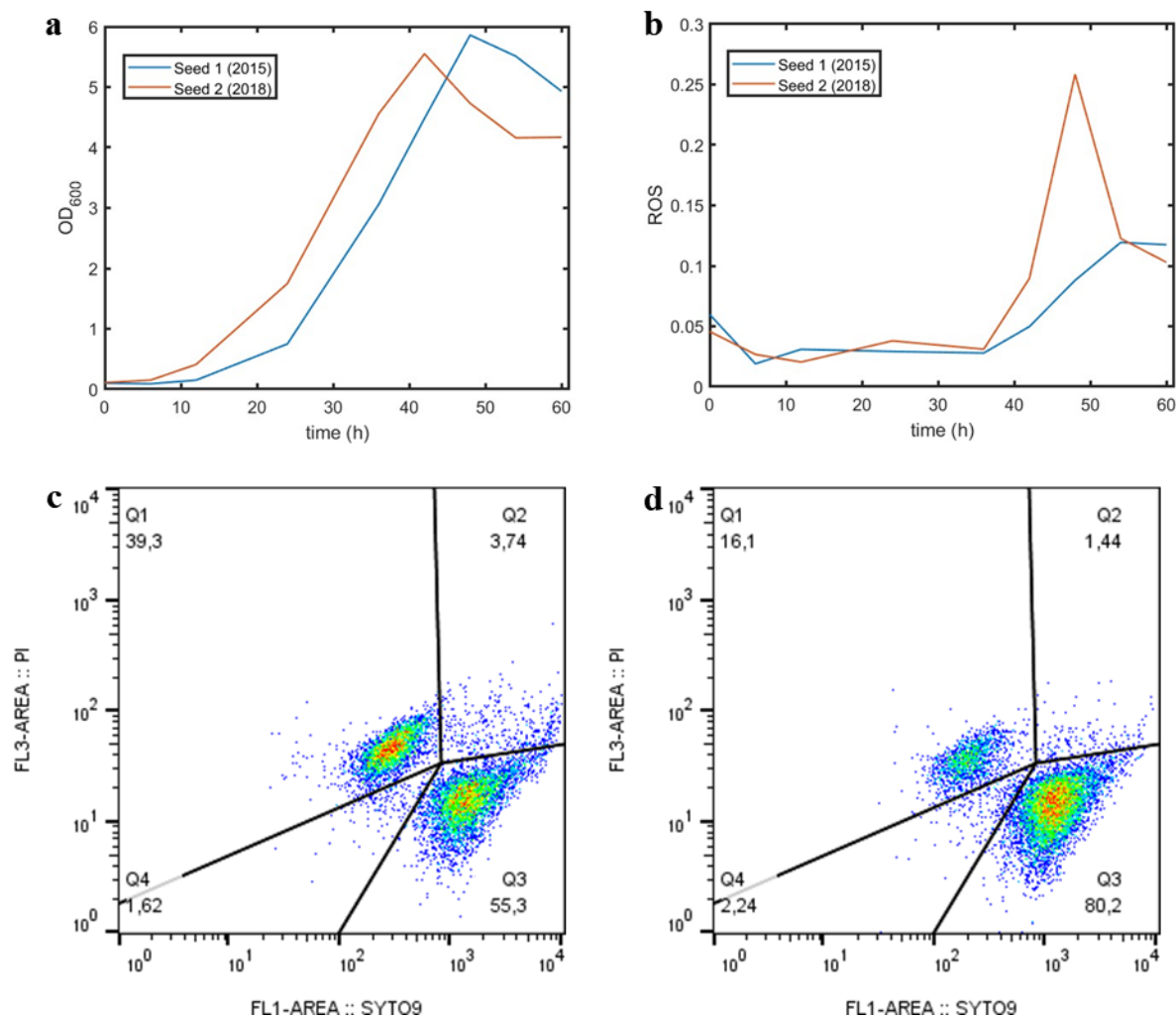


Figure 4-9: Comparison of *B. pertussis* seeds batches made in 2015 and 2018. a) Biomass (OD₆₀₀) comparison, b) intracellular ROS comparison, c) viability of the 2015 batch seed and d) viability of the 2018 batch seed.

4.4.6 Calibration and validation of the oxidative stress model

The oxidative stress model has been calibrated by fitting to experimental data obtained under 13 flask conditions and validated with data from an additional 3 conditions. The calibration and validation conditions are listed in Tables 4-5 and 4-6, respectively.

Table 4-5: List of calibration conditions for *B. pertussis* oxidative stress model

Trial	OD	Glutamate concentration (g/L)	Length of Primary (hr)	Seed Batch
1	0.28	10.1	48	1
2	0.36	9.8	48	1
3	0.13	10.7	48	1
4	0.14	8.6	48	1
5	0.27	12.0	48	1
6	0.12	12.4	48	1
7	0.38	6.6	48	1
8	0.16	6.6	48	1
9	0.10	10.0	N/A	1
10	0.08	10.3	N/A	1
11	0.11	10.6	N/A	2
12	0.32	9.3	60	1
13	0.28	9.6	60	1

Table 4-6: List of validation conditions for *B. pertussis* oxidative stress model

Trial	OD	Glutamate concentration (g/L)	Length of Primary (hr)	Seed Batch
1	0.27	8.4	48	1
2	0.25	10.7	48	1
3	0.31	6.8	60	1

The system of equations (Eqns (4.1) – (4.10)) comprising the model is found in section 4.2, while the definition of each parameter obtained by fitting the model is given in Table 4-2. The numerical values of the parameters so estimated are listed in Table 4-7. The fitting procedure is described in section 4.3.6. This mechanism has not been previously applied to the growth of *B. pertussis* and therefore the parameter values cannot be compared to those obtained in other studies. However, the initial concentration of NADPH (A_o) and intracellular ROS concentration (Table 4-7) are similar to values found in the literature for other bacteria. For example, a starved culture of *Corynebacterium glutamicum* exhibited a NADPH concentration in the range of 0.127 – 0.185 mM and increased to 0.313 mM after the addition of glucose (Goldbeck *et al.*, 2018), while the

concentration range in 3 strains of *E. coli* was found to be 0.25 – 0.45 mM (Shen *et al.*, 2021). The difference in starved NADPH concentration may be due to a difference in the duration in starved-nutrient conditions. In the experiments presented in this section, the period of starvation is ~10 hours while the period is not reported by Goldbeck *et al.* (2018). If the starvation period in this earlier study was shorter, one would expect the initial NADPH concentration to be higher than that obtained here (Table 4-7). Intracellular ROS in *E. coli*, predominantly found as H₂O₂, was found to have an intracellular concentration <0.1 μ M and substantially below 1 μ M when toxic conditions exist (Seaver and Imlay, 2001). According to our model, the generation of extracellular NADPH comes from cell lysis rather than excretion. This conclusion is reached by examination of the two terms contributing to the extracellular accumulation of NADPH that appear on the right-hand side of Eqn (4.5). A comparison of the fitted numerical values of the rate constant D for lysis and the coefficient K_M for mass transfer of NADPH into the supernatant shows that the excretion rate is several orders of magnitudes smaller than the death rate. Therefore, the secretion pathway was removed from the model to reduce the number of parameters.

Table 4-7: Parameter estimation of oxidative stress model

Parameter	Value
v	$16.0 \frac{mM}{hr}$
K	$1.72 mM$
K_t	$1.73 mM$
k_p	$0.536 \frac{1}{mM hr}$
k_m	$0.001 \frac{1}{hr}$
$Y_{x/s}$	$0.643 \frac{intensity\ biomass}{g\ glutamate}$
D	$0.0099 \frac{1}{intensity\ biomass\ hr}$

K_{si}	0.126 mM
d_B	$0.109 \frac{1}{hr}$
K_M	$1.42E-8$
	$\frac{1}{intensity\ biomass\ hr}$
A_o (secondary)	[0.5-0.787] mM
A_o (starved)	[0.004-0.01] mM
A_o (seed 1)	0.112 mM
A_o (seed 2)	0.350 mM

A comparison of the experimental and model calibration curves for the evolution of OD, extracellular glutamate concentration, extracellular NADPH and intracellular ROS concentrations over time is given in Figures 4-10 – 4-22. A comparison of the experimental and model validation curves is provided in Figures 4-23 – 4-25. The following paragraphs provide a discussion of the common behaviour in the experimental and model results, success of the model fitting and implications of the observed trends.

The concentration of ROS increases sharply as the growth of *B. pertussis* leaves the lag phase and enters the exponential growth phase (Figure 14-10c). The culture experiences oxidative stress as it is transferred from a stationary-phase environment to a fresh oxygenated medium which affects the growth rate and duration of the lag phase (Cuny *et al.*, 2007). In Figure 4-10c, another increase in ROS is observed later during the stationary phase. This pattern is consistent in all of the experiments involving a 48-hour primary flask and seed flasks. Starvation of an essential nutrient triggers the production of catalases and other oxidative stress proteins. Therefore, oxidative damage to biomolecules increases during the stationary phase as the defense mechanism against oxidative stress eventually fails (Dukan and Nyström, 1999). In Figure 14-21a, the cells in the starved condition experience a very long lag phase but they eventually recover. High

concentrations of ROS are observed at the beginning of the fermentation when the cells begin to recover (Figure 4-21c). It is hypothesized that cells starved for a long period in the primary flask must alter their DNA superstructure, restructure their cellular morphology, reorganize their global metabolism and repair oxidatively damaged biomolecules in order to begin to divide again (Bertrand, 2019). A comparison of the experimental and model validation curves is provided in Figures 4-23 – 4-25. The validation experimental results shown found in Figures 4-23 – 4-25 exhibit a good fit between the experimental and model data, similar to that observed with the calibration sets.

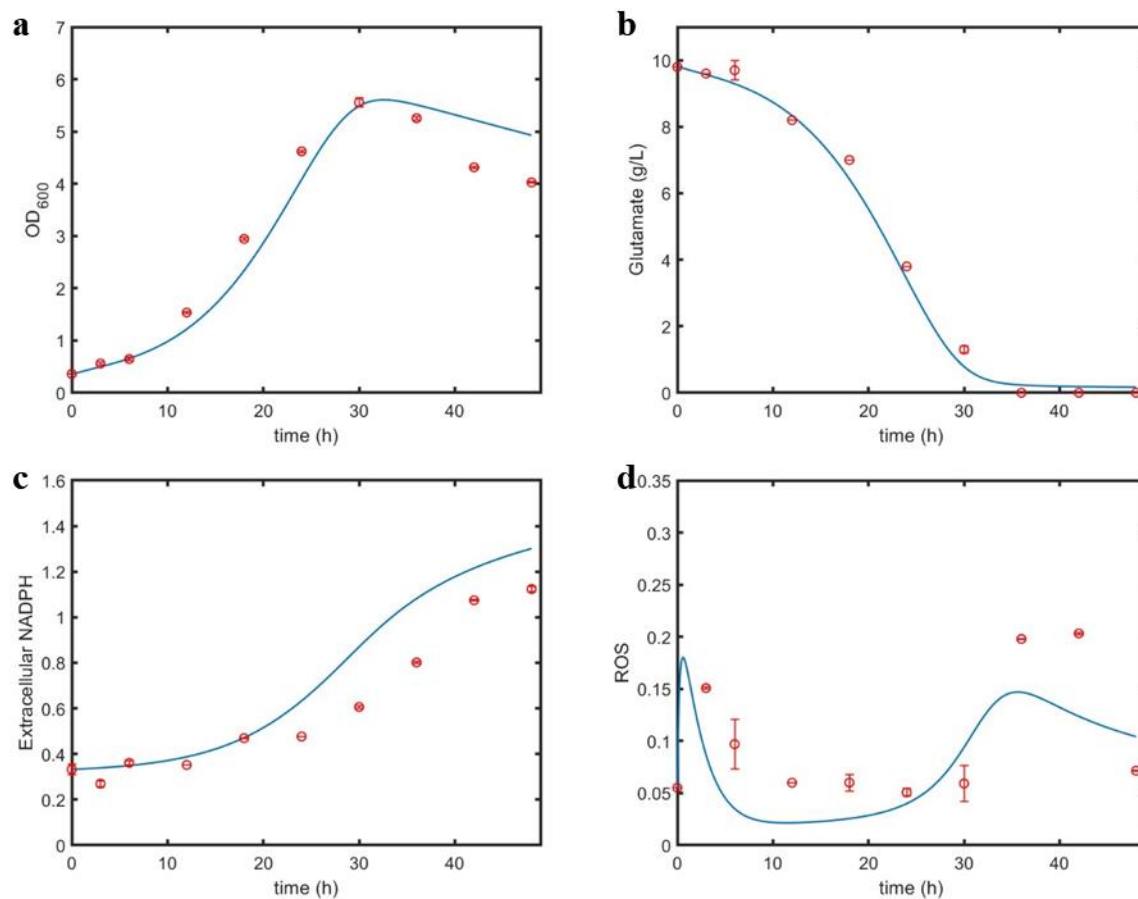


Figure 4-10: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.28 and initial glutamate concentration of 10.1 g/L.

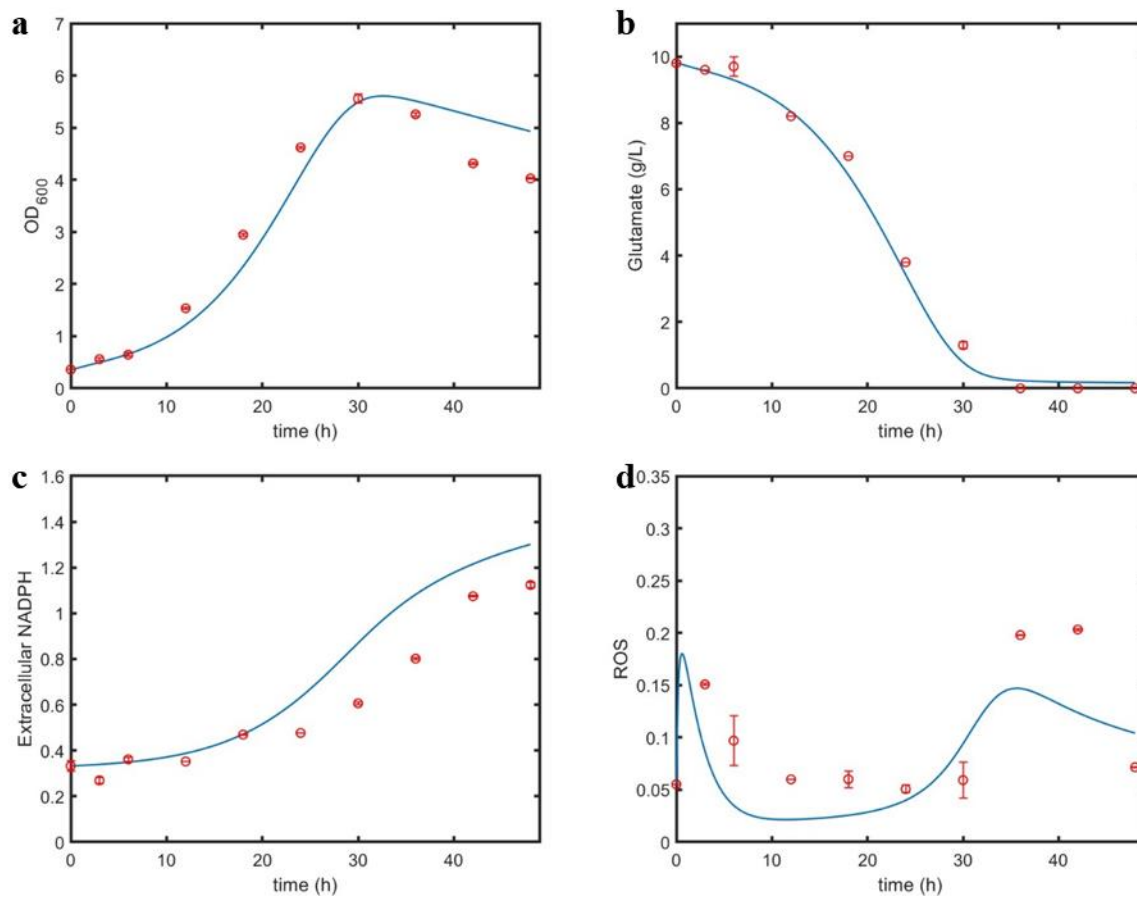


Figure 4-11: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.36 and initial glutamate concentration of 9.8 g/L.

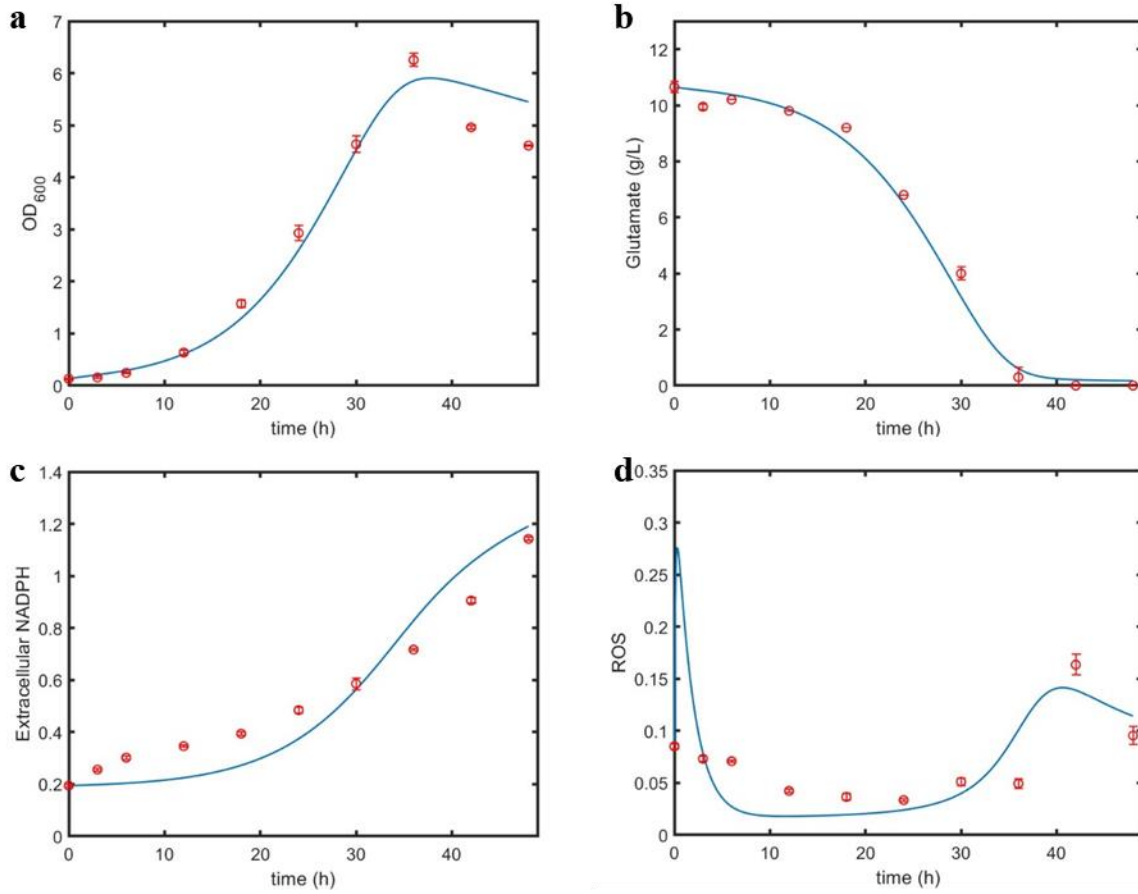


Figure 4-12: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.13 and initial glutamate concentration of 10.7 g/L.

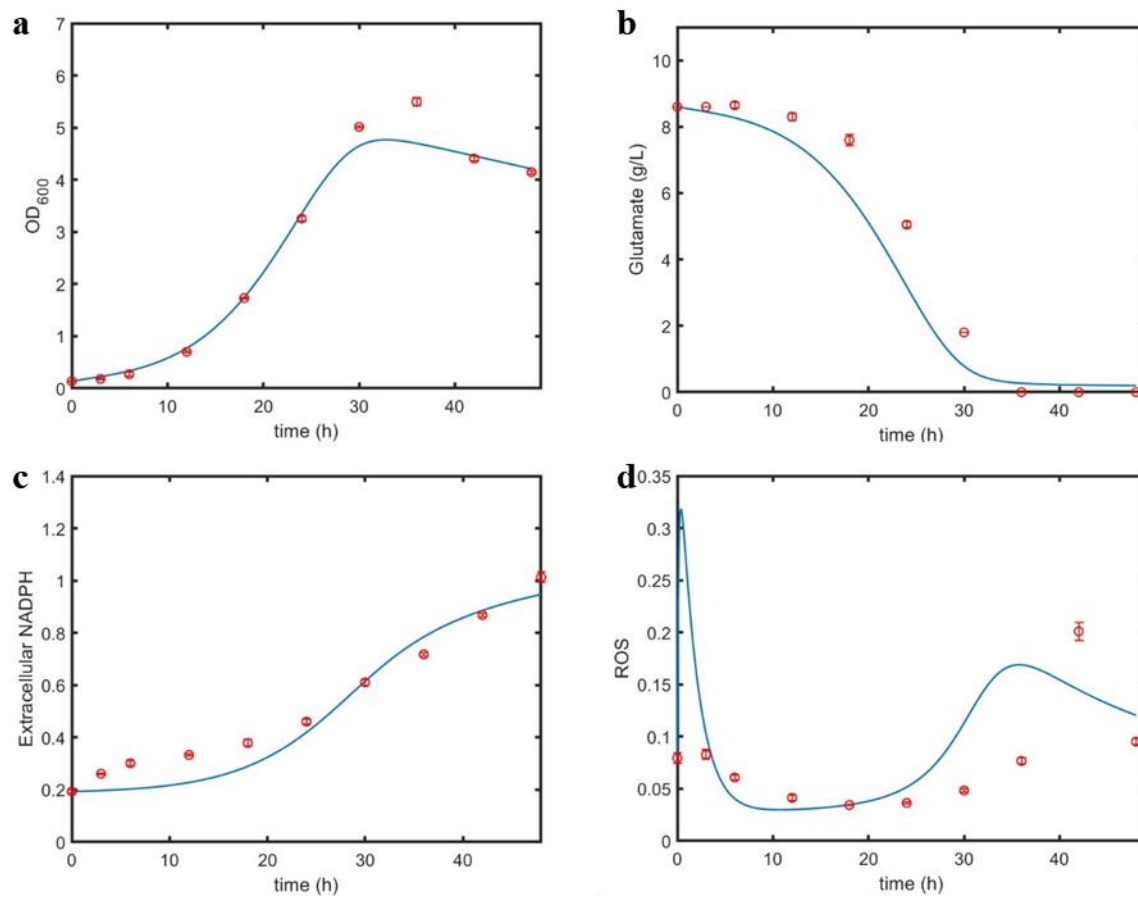


Figure 4-13: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.

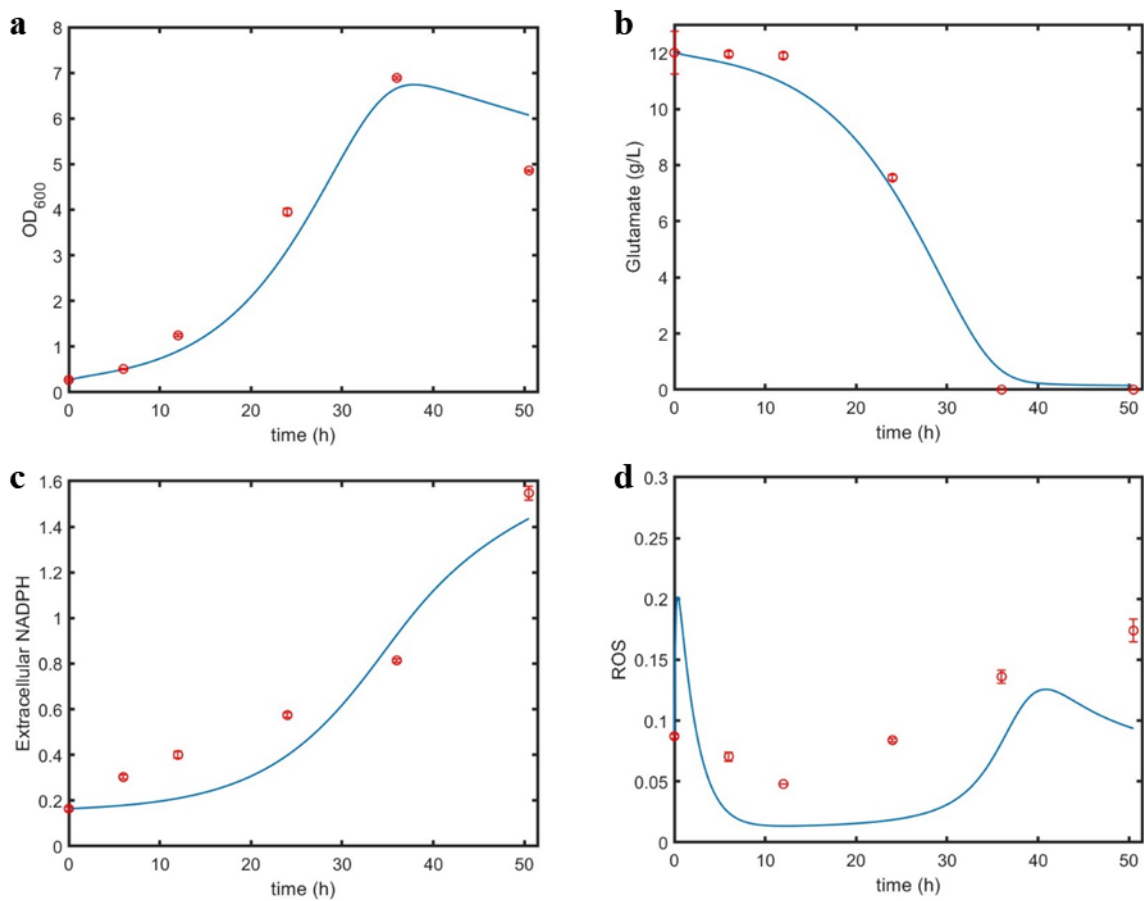


Figure 4-14: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.27 and initial glutamate concentration of 12.0 g/L.

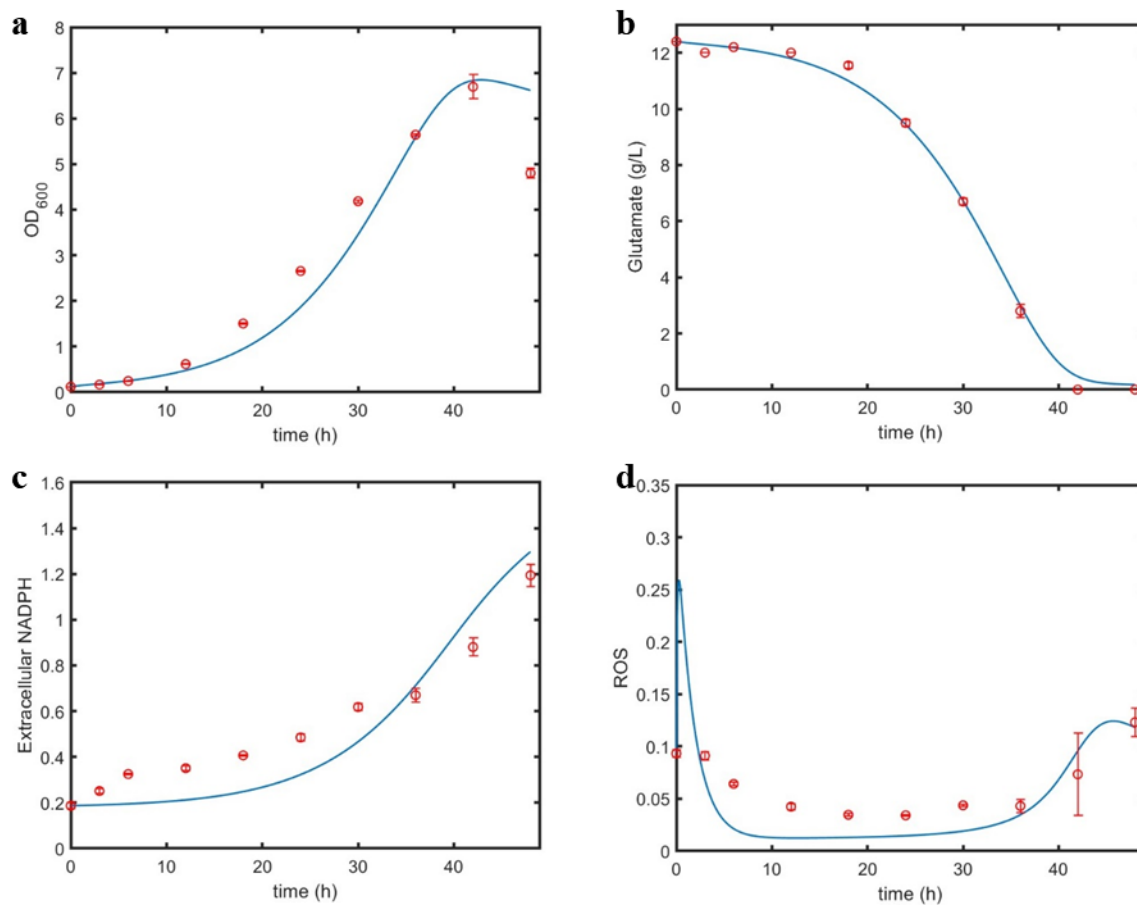


Figure 4-15: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.12 and initial glutamate concentration of 12.4 g/L.

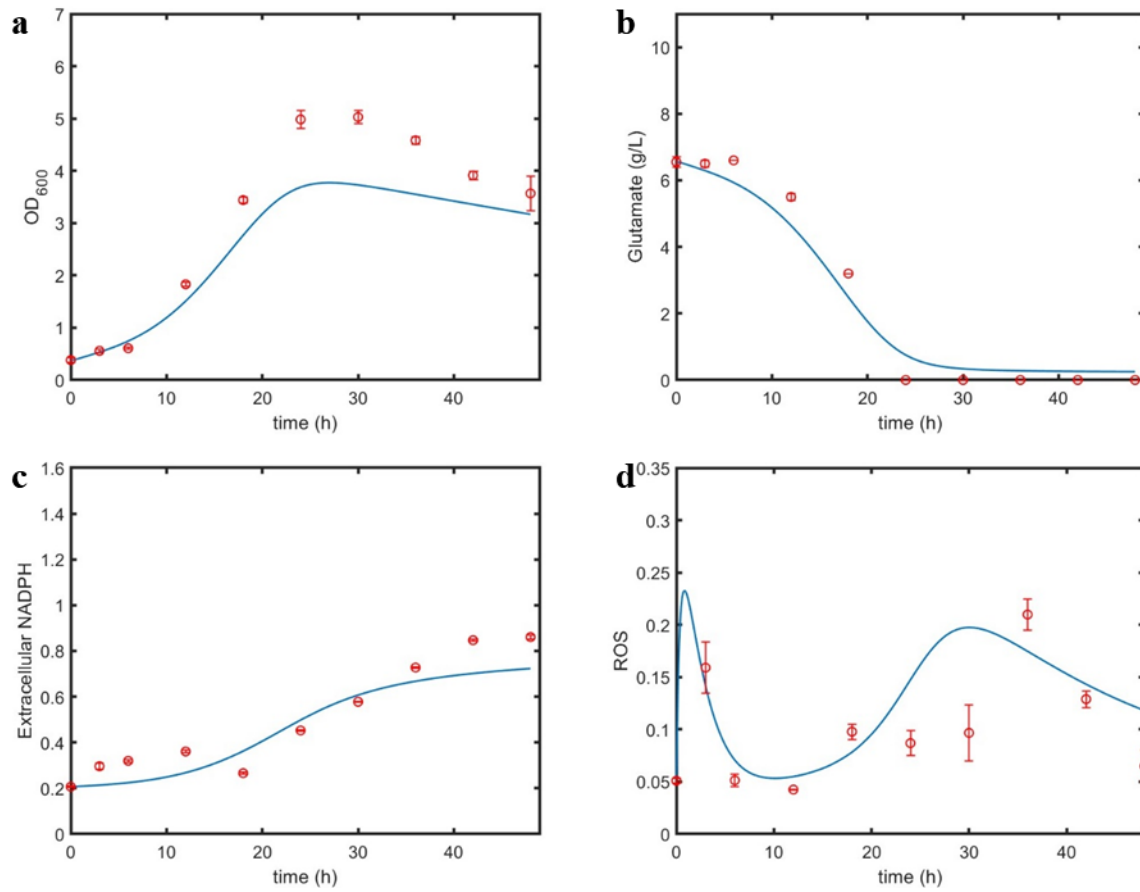


Figure 4-16: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.38 and initial glutamate concentration of 6.6 g/L.

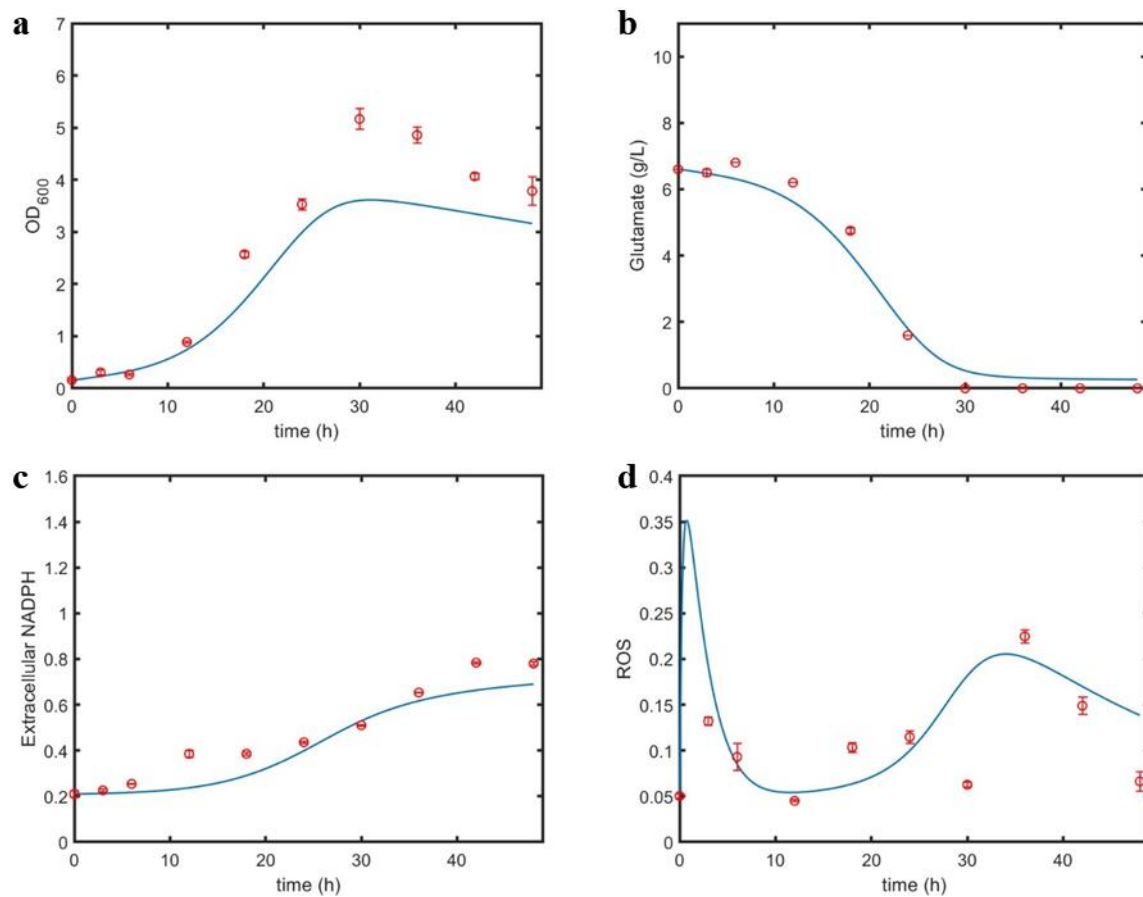


Figure 4-17: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.16 and initial glutamate concentration of 6.6 g/L.

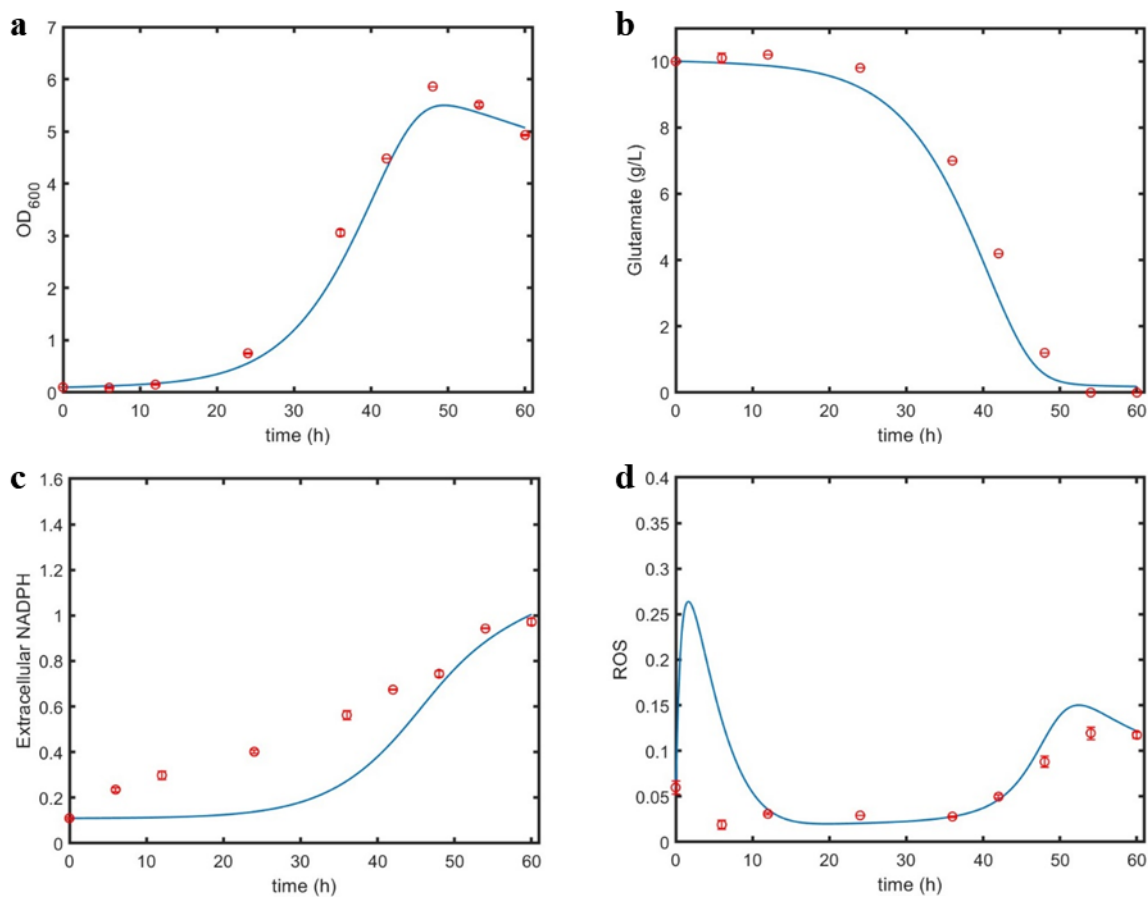


Figure 4-18: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* seed (2015) culture with an initial OD of 0.10 and initial glutamate concentration of 10.0 g/L.

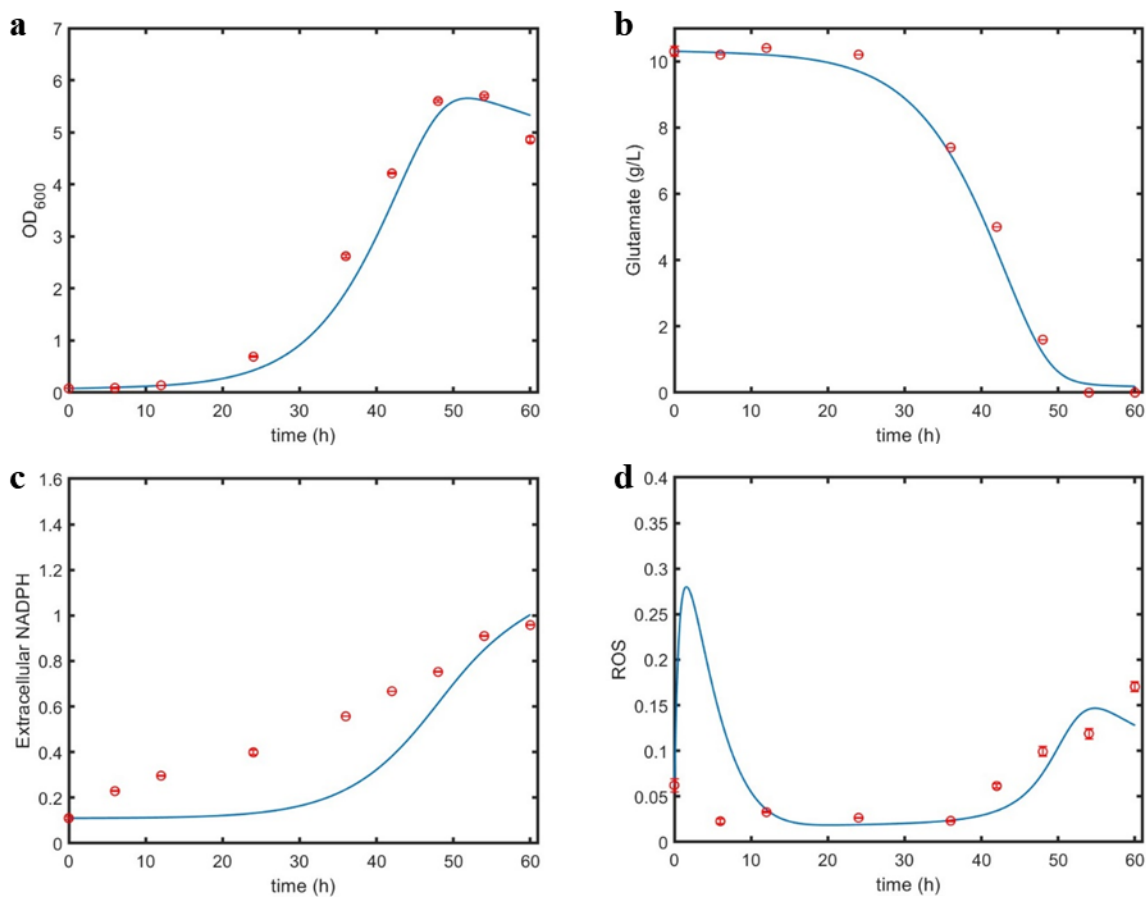


Figure 4-19: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* seed (2015) culture with an initial OD of 0.08 and initial glutamate concentration of 10.3 g/L.

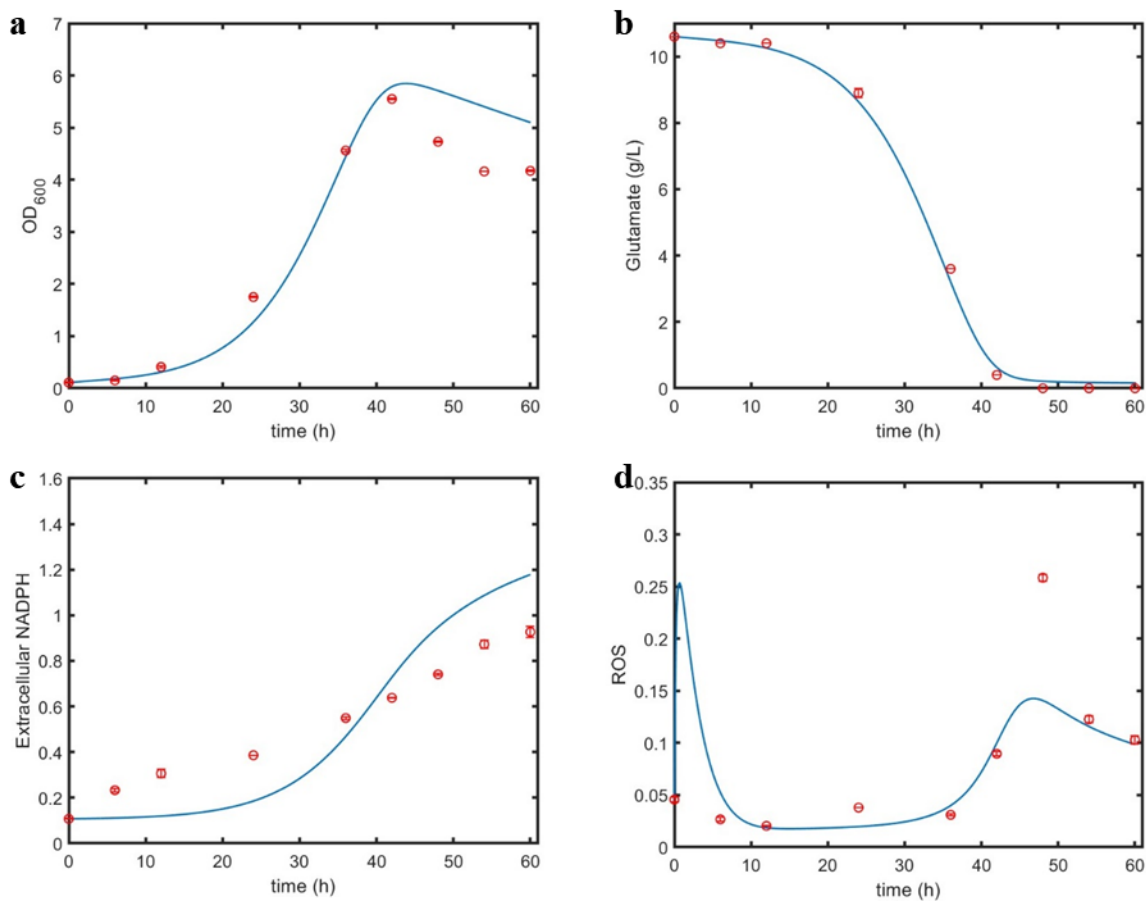


Figure 4-20: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a *B. pertussis* seed (2018) culture with an initial OD of 0.11 and initial glutamate concentration of 10.6 g/L.

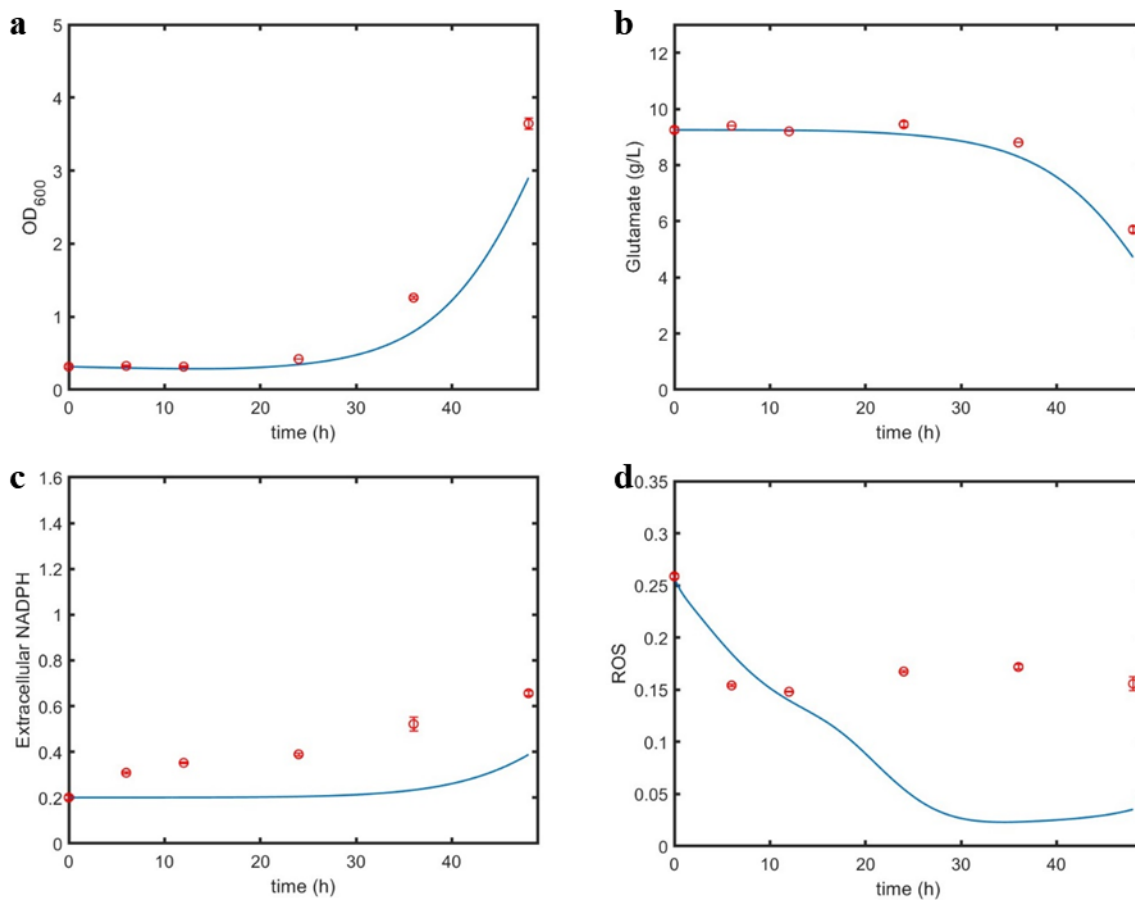


Figure 4-21: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a starved (12 hours) *B. pertussis* culture with an initial OD of 0.32 and initial glutamate concentration of 9.3 g/L.

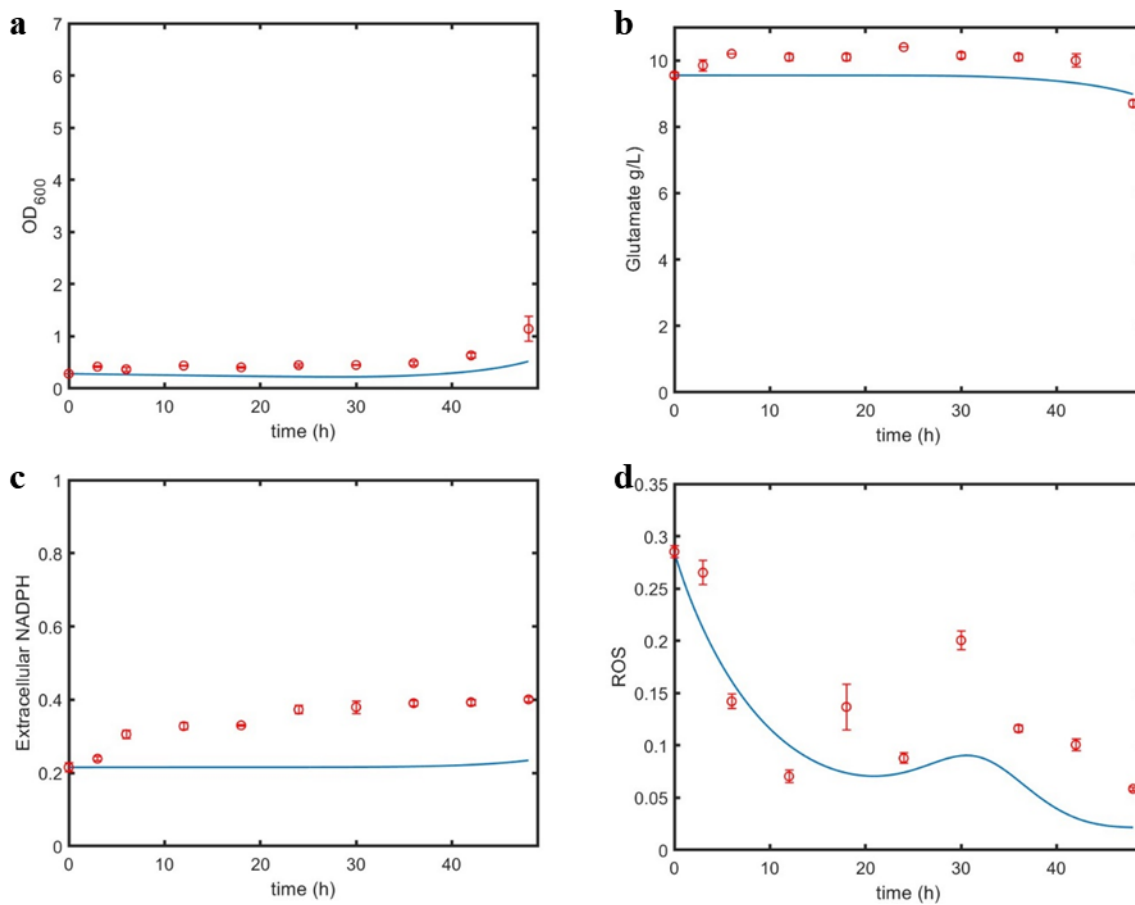


Figure 4-22: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH, and (d) intracellular ROS of a starved (12 hours) *B. pertussis* culture with an initial OD of 0.28 and initial glutamate concentration of 9.6 g/L.

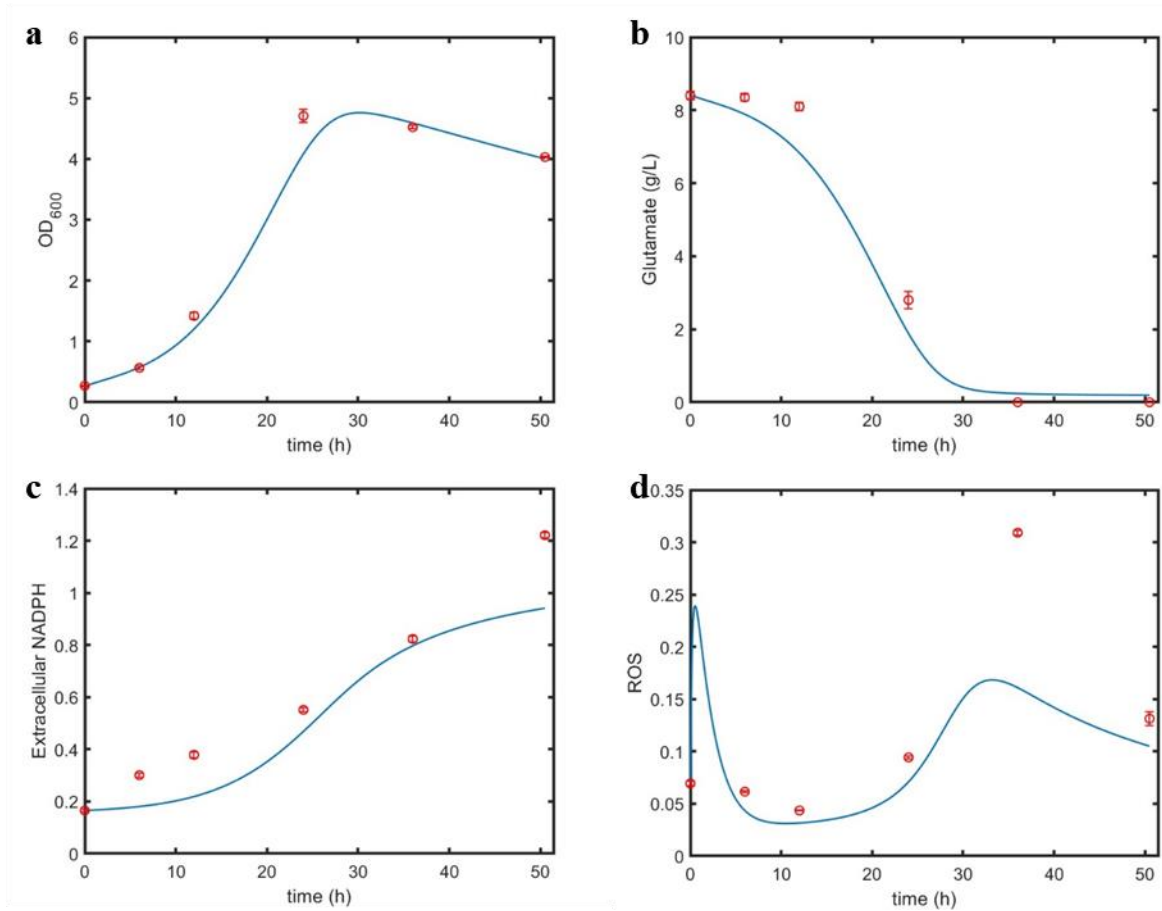


Figure 4-23: Model validation of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.

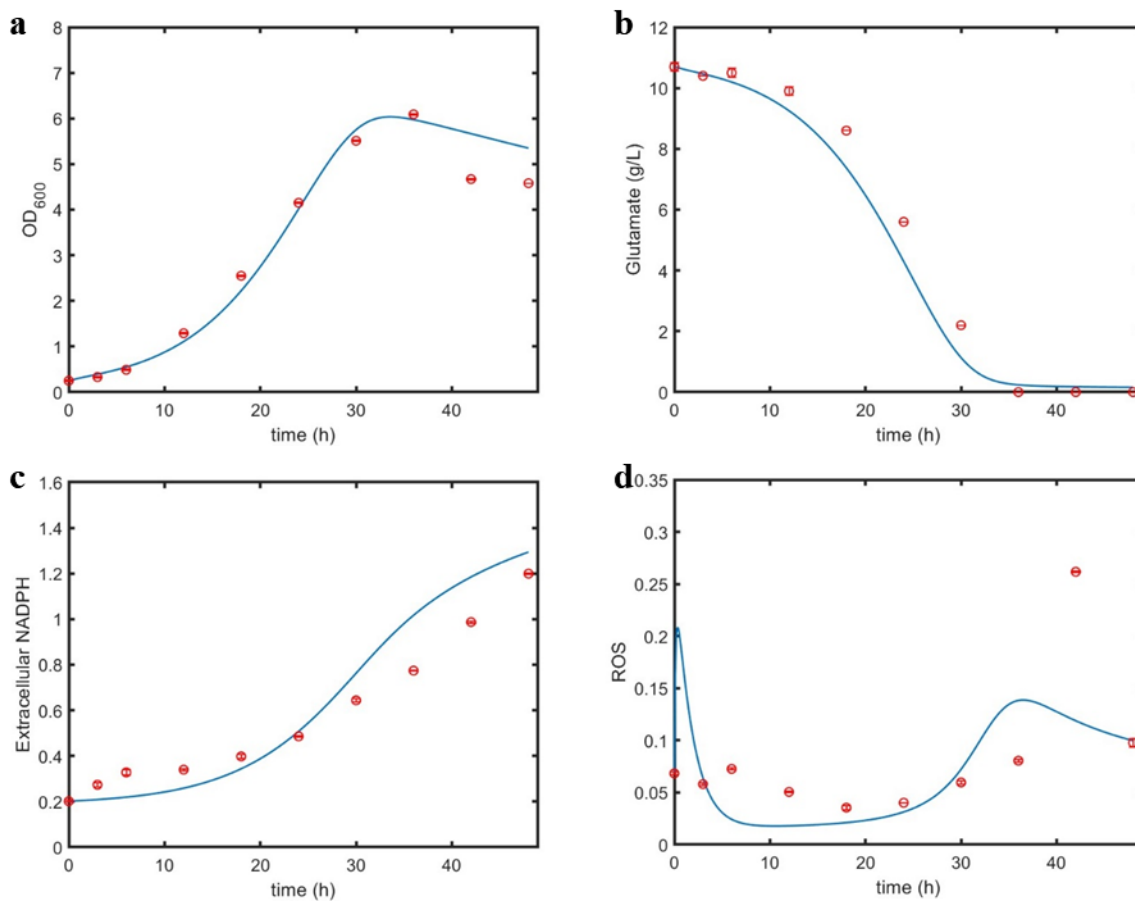


Figure 4-24: Model validation of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH and (d) intracellular ROS of a *B. pertussis* culture with an initial OD of 0.25 and initial glutamate concentration of 10.7 g/L.

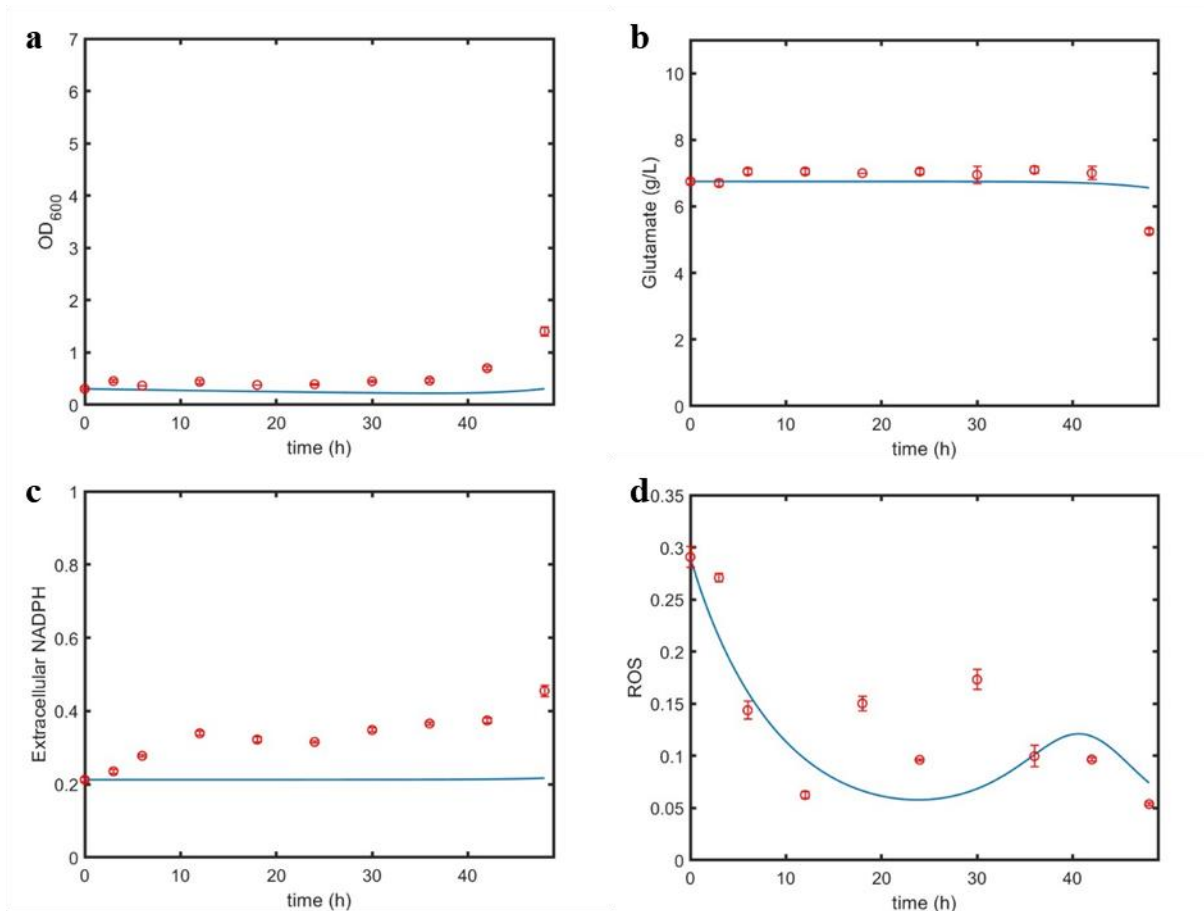


Figure 4-25: Model validation of (a) optical density (OD), (b) extracellular glutamate concentration, (c) extracellular NADPH and (d) intracellular ROS a starved (12 hours) *B. pertussis* culture with an initial OD of 0.31 and initial glutamate concentration of 6.8 g/L.

The total root-mean squared errors (RMSE) for the fits of the calibrated and validated values are presented in Table 4-8 below. The RMSE for all conditions are summed and calculated separately for each measured variable (i.e., OD₆₀₀, extracellular glutamate concentration, intracellular ROS concentration and extracellular NADPH concentration. The RMSE for the calibration and validation are very similar indicating that the noise is not over-fitted in the calibration step. The model provides good fits for the OD, glutamate and NADPH concentrations for all experimental conditions, but poorer fits for ROS concentration. The RMSE for OD, glutamate and NADPH concentrations is less than 10 % of the experimental data ranges for these

variables while the RMSE for ROS concentration is equivalent to 11.8% of the full range of experimental data values that this quantity spans over the course of the incubation. On the other hand, the model does not fit the behaviour during the starvation experiments as well, as shown in Figures 4-21c and 4-25. Following exposure to starvation conditions in the primary flask, the concentration of ROS is very high at the beginning of the secondary flask step presumably leading to many membrane-damaged cells. The fit of the model could be improved by adding a term for ROS generation in the death phase or when cells experience membrane-damage. However, this was not done to avoid over-parameterization of the model and to maintain a focus on the other phases of the model. Also, the model is not able to predict the amplitude of the ROS increase during the stationary phase after glutamate is depleted since it does not include any mechanism to form ROS other than the $S \rightarrow B$ step (Figure 4-1) and little glutamate is available to be converted into ROS during starvation.

Table 4-8: Root mean-squared error of measured variables for model calibration and validation.

The range of each variable is also shown to compare with the RMSE.

	OD₆₀₀ (intensity)	Extracellular glutamate concentration (g/L)	Intracellular ROS concentration (intensity)	Extracellular NADPH concentration (intensity)
RMSEC	0.524	0.495	0.047	0.136
RMSEV	0.387	0.622	0.054	0.130
Range	[0-7]	[0-12.5]	[0-0.4]	[0-1.4]

Flow cytometry enables single cell measurements to determine the distribution of the population with respect to the measured variables. While the current work only considers the mean of the ROS distribution, it is possible to use the cytometry data to model the evolution of the distribution, but this is beyond the scope of the model presented in this chapter. A population

balance model presented later in the thesis is combined with the model presented in this chapter to describe such a ROS distribution across the cell population.

4.4.7 Comparison of oxidative stress model to Contois growth model

Contois growth kinetics is one the most widely used mathematical models for predicting the growth of microorganisms by considering the growth rate to be a function of the substrate concentration and population density of the culture (Contois, 1952). One approach to assess the usefulness of the oxidative stress model proposed in this study is to compare it to a conventional substrate-biomass model with Contois kinetics that does not contain the mechanisms associated with oxidative stress. The equations associated with this biomass-substrate model are given in Eqns (4.15) – (4.17). The same technique described in section 4.3.6 is used to calibrate and validate the Contois model (i.e. `fmincon` subroutine to minimize the RMSE of OD and glutamate concentration). Terms for substrate inhibition of glutamate and cell death have been included for consistency between the models. The Contois model parameters and values are shown in Table 9. As shown in Figure 4-26, this simplified model is not able to predict growth following starvation since it does not have a mechanism in place to generate the extended lag phase of starved cells. The key reason that the proposed model is able to predict the longer lag is that NADPH is mostly depleted at the end of the primary flask following starvation. Then, a larger time is needed in the secondary flask to build up the necessary amount of NADPH to start growth according to the autocatalytic effect included in Eqn (4.10). Data obtained from the 60-hour primary flask (see Tables 4-5 – 4-6) were not used to calibrate and validate the Contois kinetic growth model because the error associated with the starvation experiments is so high the fit of the other experiments is affected to minimize the total error.

$$\frac{dX}{dt} = (\mu_M - D)X \quad (4.15)$$

$$\frac{dS}{dt} = -\frac{\mu_M X}{Y_{x/s}} \quad (4.16)$$

$$\mu_M = \frac{\left(\frac{v}{1 + S/K_{si}} \right) S}{K_M X + S} \quad (4.17)$$

Table 4-9: Parameter estimation for Contois model

Parameter	Value
v	$0.1743 \frac{mM}{hr}$
K_M	$0.7554 mM$
K_{si}	$54.87 mM$
$Y_{x/s}$	$0.8083 \frac{\text{intensity biomass}}{g \text{ glutamate}}$
D	$0.0236 \frac{1}{\text{intensity biomass hr}}$

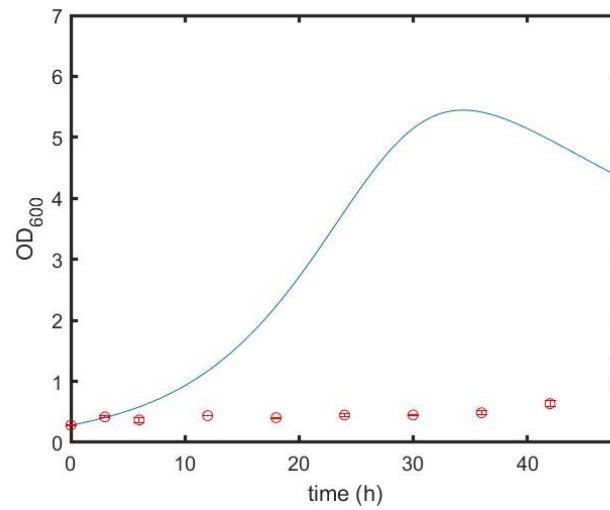


Figure 4-26: Comparison of the biomass-substrate model and starved *B. pertussis* experimental results for OD₆₀₀.

The oxidative stress model and simple biomass-substrate model are compared on the basis of the standard RMSE for the extracellular glutamate concentration and biomass (OD_{600}) as well as AIC, as shown in Table 4-10. The oxidative stress model improves the predictability of both the biomass and the extracellular glutamate concentration. The improved fit of the oxidative stress model is particularly significant regarding the extracellular glutamate concentrations as reflected by the significant reduction of the RMSE compared to that obtained using the Contois model. The AIC criterion (Eqn (4.14)) is used to assess the trade-off between model dimensionality and predictability. As shown in Table 4-10, the AIC is much lower for the oxidative stress model, indicating that the additional parameters and mechanism included in this model are needed to improve the predictability of the model compared to the simple biomass-substrate model. This further supports the hypothesis that the mechanism of oxidative stress is needed to predict the culture behavior when exposed to starvation conditions.

Table 4-10: Comparison between Contois kinetic and oxidative stress models

Model	OD ₆₀₀ (intensity)		Extracellular Glutamate Concentration (g/L)		AIC
	RMSE calibration	RMSE validation	RMSE calibration	RMSE validation	
Contois	0.667	0.490	0.984	0.969	-121.4
Oxidative Stress	0.542	0.387	0.495	0.622	-496.9

4.8 Oxidative stress model development

The development of the oxidative stress model went through many iterations in order to obtain the model structure with the best predictability and robustness. For example, we hypothesized that different type of kinetics (Contois vs Monod) can be used to describe the rate of reaction of the substrate. Also, it was of interest to test whether the model should include a direct effect of oxidative stress on cell growth and death rates. The selection of the initial guesses for the

model parameters also has a large impact on the final estimates listed in Table 4-7 as some lead to saddle points (local minima model solutions) for summed RMSE found in Eqn (4.11) or physically impossible solutions (i.e., solutions with negative concentrations or complex numbers). Therefore, we carried out numerous iterations in which the model equations with different initial guesses were solved to obtain the best fit to the experimental data based on the summed RMSE for model parameter estimates converging to an absolute minimum.

The Monod model relates the specific growth rate to substrate concentration (Monod, 1949). In comparison, the Contois model adds an additional term to the relationship as the specific growth rates are functions of substrate concentration and population density (Contois, 1952). Both model variations are widely used to describe bacterial growth in bioreactors. The fluxes to convert substrate in NADPH and ROS in Eqns (4.8) and (4.9) are functions of the glutamate concentration. The last term in Eqns (4.8) and (4.9) enables the flux to be affected as more glutamate is converted into NADPH (at high intercellular glutamate concentrations) or ROS (at low intracellular glutamate concentrations). The level of conversion can be controlled by raising the substrate concentration to an exponent in this term. We changed this exponent from the original value of 1 used by Himeoka and Kaneko (2017) in their model because a comparison of the model predictions and experimental data suggests that the flux from $S \rightarrow B$ has to yield a bigger spike in the concentration of B when S is low in order to fit the experimental data better. This exponent was varied arbitrarily from 1 to 3. Table 4-11 shows the predictability and robustness of the oxidative stress mechanism using Monod and Contois equations as well as the different exponents in the expressions of the fluxes from glutamate to NADPH and ROS (Eqns (4.8) and (4.9)). Contois kinetics fit better than Monod kinetics regardless of the exponent used for glutamate as shown by the lower magnitude of the AIC values. The exponents have less effect on the predictability of the

model as shown by similar standard errors across the measured variables and similar AIC. However, when the exponent becomes 3 or higher, the model performs less well as shown by the increase in AIC and increase of standard errors associated with the intracellular ROS concentration.

The inclusion of additional terms was explored in the oxidative stress model to determine if the intracellular ROS concentration has a greater inhibitory effect on growth or an impact on the death rate of cells. The adjustments to the death rate and growth rate are shown in Eqns (4.18) and (4.19), respectively. The effect on the death rate of cells can be evaluated by comparing the results presented earlier in this communication with those obtained by setting the parameter k_d associated with the dependency of death rate on intracellular ROS concentration to zero. The death rate was not affected by ROS concentration because the fit of the model yielded a value of zero for k_d when minimizing the RMSE of the measured variables the solution. The effect on cell growth can be determined by setting the parameter K_B (associated with the inhibition of growth due to intracellular ROS concentration) to increasing values. Based on the AIC values obtained from these simulations, we conclude that intracellular ROS only has an indirect effect on the growth rate by participating in a competitive side reaction that consumes some of the NADPH that otherwise would be available for cell growth.

$$D = D_o + k_d B \quad (4.18)$$

$$\mu = \frac{F_A A}{1 + \frac{B}{K_B}} \quad (4.19)$$

Table 4-11: Comparison of the fit, robustness and predictability of the Contois and Monod models and exponents on the flux for optimal oxidative stress.

Model	Exponent	AIC	OD ₆₀₀ (intensity)		Intracellular glutamate concentration (g/L)		Extracellular NADPH concentration (intensity)		Intracellular ROS concentration (intensity)	
			RMSEC	RMSEV	RMSEC	RMSEV	RMSEC	RMSEV	RMSEC	RMSEV
Monod	1	-448.3	0.590	0.461	0.551	0.731	0.149	0.121	0.050	0.062
Contois	1	-493.0	0.506	0.374	0.449	0.864	0.133	0.147	0.052	0.068
Contois	1.5	-496.9	0.524	0.387	0.495	0.622	0.136	0.130	0.047	0.054
Monod	2	-451.3	0.586	0.434	0.527	0.769	0.141	0.126	0.056	0.062
Contois	2	-488.3	0.518	0.407	0.476	0.728	0.139	0.135	0.052	0.052
Contois	3	-473.5	0.522	0.412	0.488	0.770	0.141	0.132	0.060	0.053
Contois with Eqn (4.19)	1.5	-386.2	0.638	0.567	0.594	0.893	0.158	0.142	0.086	0.072

4.9 Modeling biomass production in train of reactors

Batch-to batch variation in the production of antigens in the upstream fermentation process is a well-known problem. The trends observed in the case of fermentations that exhibit a low pertactin yield include an increase in the concentration of ROS at the beginning of the culture (Zavatti, 2019) and low OD₆₀₀ (Agarwal, 2021). Variability in inoculum cell density or initial glutamate concentration are common occurrences in the manufacturing process due to measurement errors, media composition including casamino acids and differences in seed batches. The oxidative stress model can be used to predict the effects of different seeds and initial glutamate concentration on the biomass during the train of reactors used in the Sanofi operation (or any other process for that matter) of the whooping cough vaccine to gain insight on potential sources for antigen variation. The fermentation process of *B. pertussis* includes two parallel trains of three consecutive reactors of increasing volume which are inoculated by seed flasks, as shown in Figure 4-27. Differences in flask and bioreactor culture kinetics may arise due to the imperfect control of agitation, aeration and pH in the bioreactors. In addition, this will lead to some inaccuracy in the values of the parameters estimated by fitting the model to experimental data. An ideal way to examine some of these effects is to apply the oxidative stress model presented in this chapter. This provides a nice example of an important use of this model.

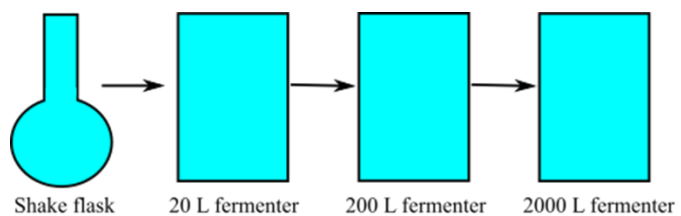


Figure 4-27: Sanofi production process for whooping cough vaccine. A seed flask is used to inoculate a train of reactors of sizes 20 L, 200 L and 2000 L.

The *B. pertussis* seeds (seed 1 generated in 2015 and seed 2 generated in 2018) were compared in Figure 4-28 for biomass and glutamate concentration in the train of reactors process using the set of parameters found in Table 4-7. The difference between these two seeds is captured by a difference in the initial concentration of intracellular NADPH and ROS. The initial ROS concentration was measured via flow cytometry and the NADPH concentration was determined for both cases by calibrating the oxidative stress model. These initial values are shown in Table 4-12. The length of the lag phase in the 2015 seed is longer than the 2018 seed due to the higher initial concentration of ROS and lower initial concentration of NADPH (Table 4-12). This resulted in a lower concentration of biomass and higher concentration of glutamate at the end of the seed flask (Figures 4-28a and e). However, the model simulations show that the difference in growth between the two seeds diminishes as the seeds continue down the train of reactors. By the 2000 L reactor, almost no difference is observed in the biomass of the two seeds. It should be noted that a major assumption made in carrying out these simulations is that the two seeds exhibit the same growth rates which of course may not be correct. Further experiments should be completed to assess this assumption.

The media composition can also vary from batch to batch. While measurement errors can occur, another source of increased glutamate concentrations can come from casamino acids found

in the media. A comparison of the effect of two media compositions with initial glutamate concentrations of 10 g/L and 10.5 g/L on the biomass and glutamate concentration in the train of reactors process is shown in Figure 4-29. Although this difference in initial concentration is small, it is typical of the batch-to-batch variability that would occur in an industrial setting. The condition with 10.5 g/L glutamate in the media yields a lower biomass in all the reactors of the train; however, by the end of the 2000 L bioreactor, the biomass produced matches that achieved when the initial glutamate concentration is 10.0 g/L. It has been shown in a previous study that the area under the growth curve has a greater effect on pertactin production than the final biomass concentration (Agarwal, 2021). Therefore, the glutamate concentration of the media may be one of the major causes of variation found in the upstream process at Sanofi.

Table 4-12: Initial conditions of intracellular NADPH and ROS concentrations of the seed flasks to model the seed batches from 2015 and 2018 in the train of reactors

Seed Year	Initial Intracellular NADPH Concentration (mmol/L)	Initial Intracellular ROS Concentration (intensity)
2015	0.112	0.0623
2018	0.345	0.0456

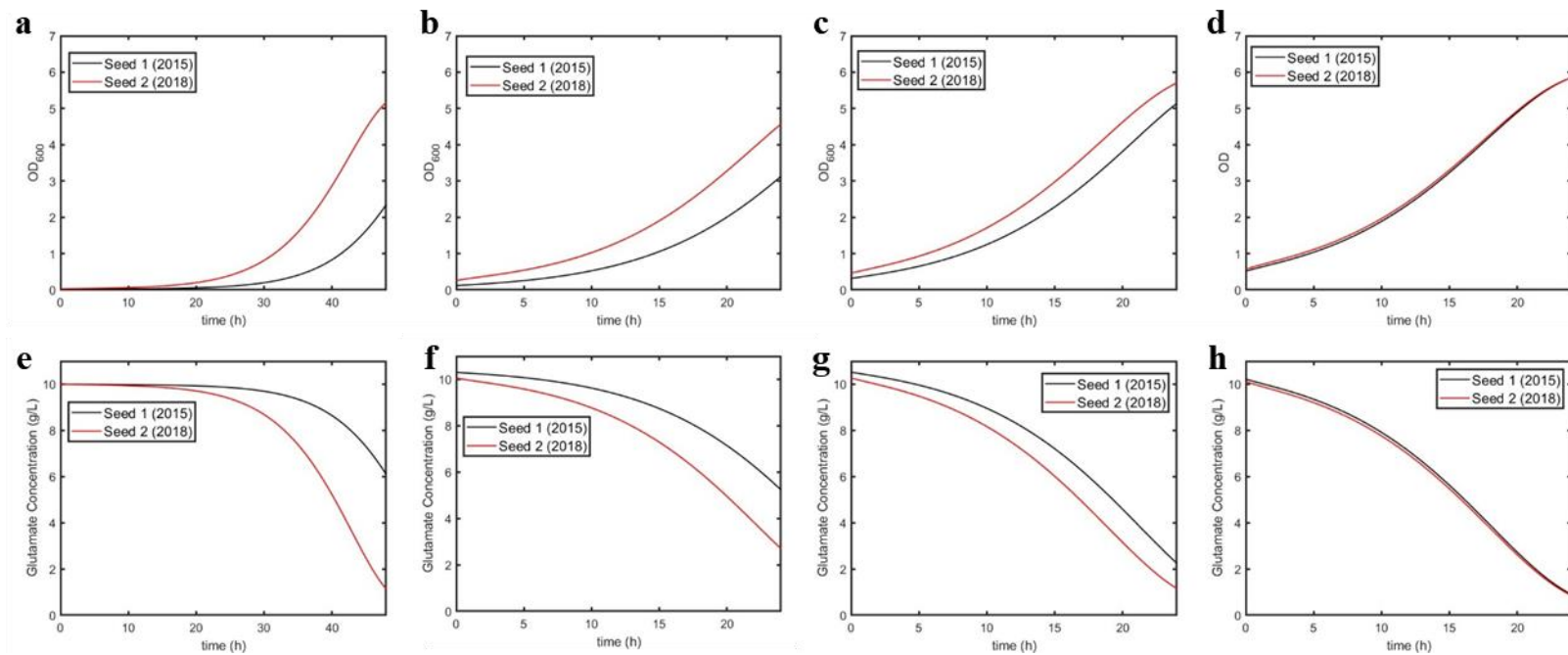


Figure 4-28: Model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for two *B. pertussis* seed batches. Biomass curve of a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor. Glutamate concentration curve of e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor.

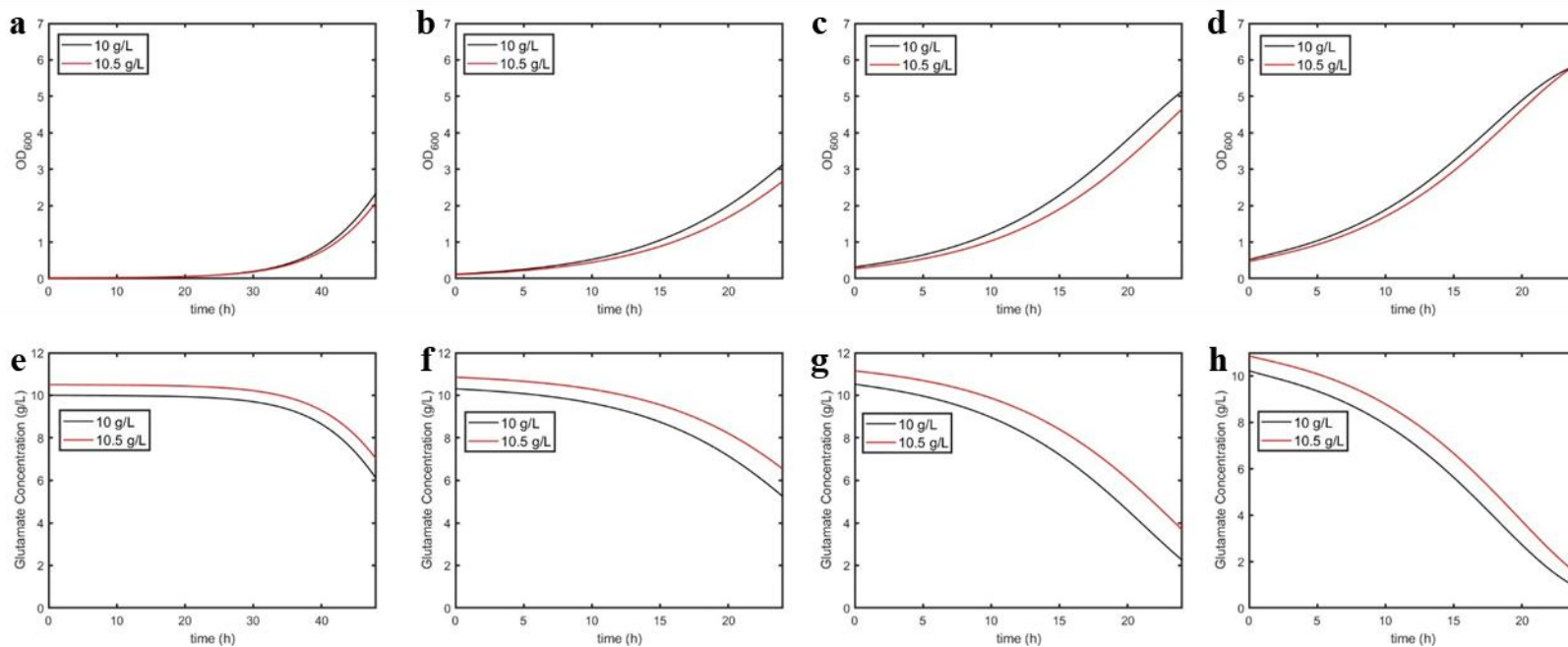


Figure 4-29: Model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for *B. pertussis* cultures with media of 10 g/L and 10.5 g/L glutamate. Biomass curve of a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor. Glutamate concentration curve of e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor.

4.5 Conclusions

A metabolic model to predict the growth of *B. pertussis* that accounts for oxidative stress was calibrated and validated against data collected during flask fermentation experiments. The conditions analyzed in these experiments cover a range of initial glutamate concentrations, initial OD and starvation conditions. The experimental data gave clear evidence for substrate inhibition due to glutamate and the resulting increase in oxidative stress. The model was able to describe different scenarios of oxidative stress conditions and the effects of glutamate levels on substrate inhibition and oxidative stress. The experiments also show that oxidative stress has a low impact on the growth rate of *B. pertussis*. Quorum sensing may explain this effect when the inoculum size is small. Following the comparison of different model structures and comparisons of model predictions to data, it was concluded that intracellular ROS only has an indirect effect on the cell growth rate by reacting with NADPH and thereby reducing the amount of NADPH otherwise available for growth. It is shown that a simple substrate-biomass model cannot describe the evolution of the substrate and biomass especially during the secondary fermentation that follows a starvation period in the primary fermentation. The ability of the model to describe the impact on growth following starvation periods makes it suitable to describe and possibly optimize the operation of a train of fermenters used in the vaccine manufacturing process. The predictive ability of the model for ROS and extracellular NADPH may be further enhanced by more accurate modelling of the ROS levels during starvation periods.

Chapter 5 Modeling Heterogeneity Using a Coupled Population Balance-Oxidative Stress Model

5.1 Introduction

Heterogeneous conditions within a cell population is a common occurrence in bioreactors and is typically caused by the processing of cell cultures in vessels of varying sizes, initial process and operating conditions and effects of mixing. Cell populations in bioreactors are commonly heterogeneous in properties such as size, yield and growth due to spatial gradients in culture conditions such as the dissolved gas concentration, pH, nutrient concentration and shear rate (Lara *et al.*, 2006). These spatial gradients typically arise because of deficient mixing or the location of aeration and nutrient inlets, especially in large-scale operations. When cultured in a heterogeneous environment, cells are exposed to varying conditions as they travel through the various zones of a bioreactor. Such changes can affect cell metabolism, yield and kinetics of the reactions of interest. Heterogeneity of the cell population may also arise in well mixed operations due to heterogeneity in the initial conditions, e.g. intracellular concentrations, cell size, stage within the cell cycle of growth, etc. Due to this heterogeneity, each cell in the population is at a different stage in its life cycle at any particular time and so contains different quantities of proteins, DNA, RNA and other components.

Most biotechnological processes are generally regulated through closed loop control of lumped or average quantities such as dissolved oxygen, average substrate concentration or pH. However, due to the heterogeneity of the cell population, control approaches that are based on only the average culture properties may not necessarily guarantee consistent process outcomes. Several studies (e.g. Le *et al.*, 2012, Enfors *et al.*, 2001) have shown that bioprocesses operated under precise dissolved oxygen and pH control may still yield significantly different final biomass and product titers due to metabolic shifts resulting from minute variations in either initial or environmental conditions. Due to this sensitivity, it is imperative to understand the nature and causes of the cell heterogeneity since relatively minor differences in the distribution of the cell population may lead to large variability in the resulting levels of growth and productivity.

For example, the whooping cough (*B. pertussis*) vaccine manufacturing operation at Sanofi involves a train of reactors of increasing volume in which the final cell population from one step is used to inoculate the next reactor in the sequence. We hypothesize in this work that the variability in cell population may propagate and be even amplified moving down the train of bioreactors, especially in the presence of metabolic shifts. Population balance models (PBMs) can be formulated to describe the evolution in time of these distributions which can then be correlated to the final productivity. PBMs contain a set of equations which describes the heterogeneous behavior of a population of particles from analysis of the behavior of single particles. In a bioreactor, such analysis provides a deterministic description of the dynamic evolution of cell distributions. Since flow cytometry measures the properties of individual cells, the distributions obtained in flow cytometric measurements can then be used to calibrate a population balance model. The flow cytometry data can be used to gain insight into the distributions of important quantities (e.g., cell size, intracellular concentrations of metabolites) over the entire cell

population. The resulting PBM can then be used to predict the heterogeneity of the cell population with respect to experimental conditions that were not used to calibrate the model.

Recently, interest in applying population balance models to describe cell state distributions has grown. Quedeveille *et al.* (2018) used a population balance model to differentiate between growth in mass from growth in the number of cells. The cell length and the rate of anabolism were chosen as the physiological properties in the population balance model. This model was able to demonstrate that growth in mass and cell number are only equivalent at steady state in a continuous bioreactor and during the exponential growth phase in a batch culture. Durr *et al.* (2016) analyzed the influence of cell-to-cell variability on important process variables for influenza vaccine production. Jerono *et al.* (2021) used a population balance model to determine parameters related to kinetics and cell division in a batch fermentation of yeast. The model revealed a good agreement between the model predictions and measurements of biomass, glucose and ethanol.

The objective of the current study is to present a population balance model that makes use of flow cytometry data to predict distributions in cell size and intracellular glutamate, ROS (reactive oxidative species), NADPH and NADP⁺ in shake flask cultures of *B. pertussis*. Our focus on oxidative stress related properties is motivated by our hypothesis that ROS is a main cause of variability in productivity of antigens observed in the whooping cough vaccine manufacturing process. The population balance model is then coupled to the oxidative stress model presented in Chapter 4 based on a reaction mechanism that relates glutamate concentration, extracellular NADPH levels and growth to oxidative stress levels. Flow cytometric data for cell size and intracellular ROS concentration collected from the cell cultures are used to calibrate the model. Since ROS is known to have a significant effect on protein synthesis, this model should be useful in predicting the production of antigens when the culture is exposed to different oxidative stress

conditions. Although a quantitative correlation between ROS and specific antigen productivity for each cell is not available at this point, the model can be used to improve our understanding of the fermentation conditions that favor the propagation of cells under conditions of lower ROS concentration which are expected to be superior producers of antigens than cells subject to higher ROS levels. This could be ultimately used to minimize fluctuations in the production of the desired antigen in the upstream process and provide means of increasing the final yield of the entire process.

The remainder of this chapter is organized as follows. Section 2 presents the development and coupling of the population balance model and oxidative stress model. Section 3 describes the materials and methods used during the experiments and algorithm used to fit the model to the experimental data. Section 4 presents the flow cytometry distributions obtained from the experiments, analysis of experiments and model predictions, comparison of the population balance and bulk model predictions and possible implications of the model on the manufacturing process. Finally, the conclusions from this phase of the project are given in section 5.

5.2 Model Outline

5.2.1 Population balance general form

The general form of a population balance equation is a partial differential equation as shown below in Eqn (5.1) (Fredrickson *et al.*, 1967; Ramkrishna, 1979). The definitions of the parameters appearing Eqn (5.1) are given in Table 5-1. The population balance model calculates the distributions of cell properties of interest. Examples of such properties can be the location of

the cell in the bioreactor (considering mixing and the concentration within the space), physiological properties such as the intracellular concentration of components or the size of the cell. In this work, cell size and the intracellular concentrations of glutamate, NADPH, ROS, and NADP⁺ are the physiological properties to be determined by the PBM.

$$\begin{aligned} \frac{\delta N(\mathbf{x}, t)}{\delta t} + \nabla_{\mathbf{x}}(\mathbf{r}(\mathbf{x}, s)N(\mathbf{x}, t)) + \gamma(\mathbf{x}, t)N(\mathbf{x}, t) + \theta N(\mathbf{x}, t) \\ = 2 \int_x^{x_{max}} \gamma(\mathbf{x}, t)p(\mathbf{x}, y, z)N(\mathbf{x}, t)d\mathbf{x} \end{aligned} \quad (5.1)$$

Four terms appear on the LHS of Eqn (5.1). The 5 physiological properties of interest in the current problem (i.e., intracellular molecules of glutamate, NADPH, ROS, and NADP⁺ and cell size) make up the vector \mathbf{x} . The first term in Eqn (5.1) accounts for the accumulation of the number of cells N at a specific level x of the physiological state. The second term accounts for the gain (or loss) of the number of cells N with a physiological state \mathbf{x} due to reactions occurring as the cells evolve through the cell cycle. The third term accounts for the loss of cells with physiological state x due to cell division leading to the birth of smaller cells. The fourth term is the dilution rate due to cells with physiological state \mathbf{x} leaving the bioreactor. The term on the RHS of Eqn (5.1) represents the rate of birth of cells with the physiological state \mathbf{x} arising from the division of each cell into 2 daughter cells.

The population balance model typically consists of a first order partial integro-differential equation as shown in Eqn (5.1) coupled to a set of ordinary differential equations that describe the temporal evolution of the vector. The solution to this system of the PDE and ODE equations can be obtained using with a variety of numerical methods. Mantzaris *et al.* (2001) compared several finite difference algorithms for the solution of the general formulation of the PDE. In this chapter,

we proposed a novel approach for solving the model based on the method of characteristics which offers some computational advantages as described below. The population balance model was calibrated and validated using data collected during flask inoculation experiments on *B. pertussis*.

Table 5-1: General form population balance model definitions

Variable	Definition
$\mathbf{x}(t)$	Vector of physiological property of cells
$X(t)$	Size of the cell (intensity)
t	time (hr)
$S(t)$	substrate concentration (mmol/L)
$N(\mathbf{x}, t)$	number of cells with state \mathbf{x} at time t (cells)
$\mathbf{r}(\mathbf{x}, S)$	Vector of reaction rates of individual cell components
$\gamma(\mathbf{x}, t)$	partitioning rate (hr^{-1})
θ	dilution rate (hr^{-1})
$p(\mathbf{x}, y, z)$	probability density function for cell division

5.2.2 Model Formulation

For the development of the model and its calibration with the data from fermentation flask experiments, we make the following assumptions:

1. Each parent cell splits the physiological property \mathbf{x} evenly between two daughter cells.

Since the physiological properties involving intracellular species (glutamate, NADPH, ROS and NADP^+) are split evenly, the concentration of these components in the daughter cells will be the identical to those in the mother cell. The cell size of the mother cell will be divided evenly between the 2 daughter cells. Therefore, the term for the birth of cells in the physiological state level is simplified as shown in Eqn (5.2), i.e.,

$$2 \int_x^{x_{max}} \gamma(\mathbf{x}, t) p(\mathbf{x}, 2\mathbf{x}, S) N(\mathbf{x}, t) d\mathbf{x} = 4\gamma(2\mathbf{x}, t) N(2\mathbf{x}, t) \quad (5.2)$$

Equal partitioning simplifies the birth of cells with physiological property \mathbf{x} because instead of a probability that any cell with a higher level of physiological property (from levels of \mathbf{x} to \mathbf{x}_{\max}) dividing into a cell with physiological property level of \mathbf{x} , the birth rate of cells is now limited to cells that have twice the physiological property ($2\mathbf{x}$). This eliminates the integral, probability density function and determines that the cells with $N(2\mathbf{x}, t)$ divide into cells with property \mathbf{x} . The probability density function ($p(\mathbf{x}, 2\mathbf{x}, S)$) can be a variety of function depending on the assumptions made (Mantzaris *et al.*, 2001a).

2. Since flask experiments are conducted in batch operating mode, no dilution term appears, i.e.,

$$\theta = 0 \quad (5.3)$$

3. Cell death is negligible. This assumption which has been made to simplify the computations, is also very reasonable since the experiments discussed in Chapter 4 show that during the exponential growth phase the percentage of the total number of cells that are dead is <3 %.
4. The various species involved in biomass production over the course of the incubation react according to the mechanism described in the oxidative stress model presented in Chapter 4.
4. The process for coupling the population balance model and oxidative stress model is shown in section 5.2.3.
5. The cell growth rate is proportional to its size, i.e.,

$$\mathbf{r}(\mathbf{x}, S) = \mu X(t) \quad (5.4)$$

where the proportionality factor μ is the growth rate .

6. The partitioning rate contains a probability density function $f(x)$ that is a function on the size of the cell ($X(t)$) (Mantzaris *et al.*, 1999):

$$\gamma(\mathbf{x}, t) = \frac{f(X)}{1 - \int_0^X f(X) dX} \mu_X(t) \quad (5.5)$$

where $f(X)$ is the division probability density function only on cell size. The probability density function is taken to be a left-hand side truncated Gaussian distribution with a mean and standard deviation. It is renormalized in the denominator in order to be a probability density function (Eakman et al., 1966). The division rate depends on the substrate concentration through the growth rate function (Tsuchiya *et al.*, 1966).

Simplification of Eqn (5.1) based on these assumptions yields the population balance equation in the form of the PDE given in Eqn (5.6), subject to the initial conditions $N(\mathbf{x}, t = 0) = N_o$.

$$\frac{\delta N(\mathbf{x}, t)}{\delta t} + \nabla_{\mathbf{x}}(\mathbf{r}(\mathbf{x}, s)N(\mathbf{x}, t)) + \gamma(\mathbf{x}, t)N(\mathbf{x}, t) - 4\gamma(2\mathbf{x}, t)N(2\mathbf{x}, t) = 0 \quad (5.6)$$

To complete the model, equation (5.6) is coupled to a set of ordinary differential equations that describe the reaction rates of the various species. Generally, systems involving a PDE coupled to a set of ODEs is computationally expensive when treated by finite differences methods. Consequently, we propose to solve the system of equations by the method of characteristics in this study. This method is a technique for solving first order partial differential equations by transforming the PDE in Eqn (5.6) into a family of ordinary differential equations. The goal is to transform coordinates from (\mathbf{x}, t) to a new system of coordinates in which the partial differential equation becomes a system of ordinary differential equations that yield the solution along trajectories in the x - t plane that are referred to as characteristic curves. The transformation of our model using the method of characteristics is shown below in Eqn (5.7) – (5.9). To simplify the presentation of the method here, we assume the vector \mathbf{x} consists of a single component $x(t)$, i.e., a one-dimensional problem. In the next section, the presentation will be extended to account for a

multi-dimensional problem. In Eqns (5.7) – (5.9), the physiological property $x(t)$ is transformed into a family of characteristic curves $C_x(t)$, the number of cells from $N(x, t)$ into $v(t)$ and the number of cells with double the physiological property $2x(t)$ from $N(2x, t)$ into $v_2(t)$ as follows in Eqns (5.7) – (5.9). The full derivation using the method of characteristics is shown in Appendix A. $C_x(t)$ is the solution of $x(t)$ for a particular initial condition x_o . Since the system involves a distribution of initial states x_o , the model determines a family of solutions $C_x(t)$ corresponding to the different values of x_o . Therefore, the method transforms the original PDE into the solution of many ODEs defined by Eqn (5.7) where each ODE is solved with a different value of x_o . Since the method of characteristics applied in the current study is used to solve the problem for each initial condition in a moving system of coordinates with velocity $r(x, s)$, then it can be shown that $\nabla_x(r(x, s)N(x, t)) = 0$ (Lim, 2005). This latter term represents the convection of the quantity $N(x, t)$ with a velocity $r(x, s)$. Since the relative velocity with respect to a system of coordinates that moves with velocity $r(x, s)$ is zero, the corresponding convection term is zero in the moving system of coordinates. The elimination of this term is a key advantage of the method of characteristics over other PDE solution methods since this convection term often causes numerical stability problems.

$$\frac{dC_x(t)}{dt} = r(C_x(t), t); C_x(0) = x_o \quad (5.7)$$

$$v(t) = N(C_x(t), t); v_2(t) = N(2C_x(t), t) \quad (5.8)$$

$$\frac{dv(t)}{dt} + (\gamma(C_x(t), t))v(t) - 4\gamma(2C_x(t), t)v_2(t) = 0 \quad v(0) = N(x_o, t) \quad (5.9)$$

The set of equations to solve the population balance model is shown in Eqns (5.10) – (5.13). Along the characteristic curves the number of cells at a particular state dependent only on time and on the initial condition, as shown in Eqns (5.11) – (5.12). The total biomass is defined in Eqn

(5.10) as the weighted sum of the number of cells of a particular volume. The consumption of the substrate in the bulk is a function of the total biomass, as shown in Eqn (5.11). The substrate consumption rate also includes a substrate inhibition term found in the study of *B. pertussis* flask cultures with differing glutamate concentration in the media, as shown in section 4.4.2 in Chapter 4. Eqn (5.12) describes the change in the cell size $X(t)$ according to the growth rate μ . Finally, Eqn (5.13) describes one of the ODEs obtained from the method of characteristics according to Eqn (5.9) for a particular characteristic curve with initial condition $C_x(0) = x_o$. The growth rate μ in Eqns (5.12) and (5.13) depends on all the variables of the problem and so couples the population balance model with the oxidative stress model to be presented in section 5.2.3. Table 5-2 lists the parameter definitions for the population balance model.

$$N_b(t) = \sum X(t)v(t) \quad (5.10)$$

$$\frac{dS_{ext}}{dt} = -\frac{1}{Y} \left(\frac{\left(\frac{v_b}{1 + S_{ext}(t)/K_{si}} \right) S_{ext}(t)}{kN_b(t) + S_{ext}(t)} \right) N_b(t) \quad (5.11)$$

$$\frac{dX}{dt} = \mu X(t) \quad (5.12)$$

$$\begin{aligned} \frac{dv}{dt} = & -\mu \left(\frac{f(X(t))}{1 - \int_0^X f(X(t))dX} X(t) \right) v(t) \\ & + 4\mu \left(\frac{f(2X(t))}{1 - \int_0^{2X} f(X(t))dX} 2X(t) \right) v_2(t) \end{aligned} \quad (5.13)$$

5.2.3 Coupling of the population balance and oxidative stress models

The method of characteristics presented in the previous section requires solving the characteristic curves for each of the variables in the state vector \mathbf{x} with ODEs of the type of Eqn

(5.7). Since \mathbf{x} in this problem contains five state variables, i.e., intracellular glutamate, NADPH, NADP^+ and ROS concentrations and cell size, a family of characteristic curves must be generated for each of these variables. The oxidative stress model presented in Chapter 4 is used to generate these curves. To summarize the phenomena described by this model, it describes the reactions by which glutamate is consumed by the cell and used for the production of NADPH (A) or ROS (B). The relative fluxes of glutamate for these two pathways depends on the concentration of glutamate inside the cell. If the glutamate concentration is high, the flux to produce NADPH dominates; if the glutamate concentration is low, the flux for ROS dominates. As discussed in Chapter 4, NADPH promotes cell growth, while ROS inhibits growth by reacting with NADPH to form NADP^+ and lowers the amount of NADPH available for growth. Eqns (5-14) – (5-24) present the coupled population balance and oxidative stress model. Table 5-2 lists the definitions of the parameters used in these equations.

$$N_b(t) = \frac{\sum Xv}{\sum X_o v_o} OD_{600,initial} \quad (5.14)$$

$$\frac{dS_{ext}}{dt} = -\frac{1}{Y} \left(\frac{\left(\frac{v_b}{1 + S_{ext}/K_{si}} \right) S_{ext}}{kN_b + S_{ext}} \right) N_b \quad (5.15)$$

$$\frac{dX}{dt} = \mu X \quad (5.16)$$

$$\frac{dS}{dt} = -F_A A - F_B A + A(S_{ext} - S) - \mu S \quad (5.17)$$

$$\frac{dA}{dt} = F_A A - k_p AB + k_m C - \mu A \quad (5.18)$$

$$\frac{dB}{dt} = F_B A - k_p AB + k_m C - \mu B - d_B B \quad (5.19)$$

$$\frac{dC}{dt} = k_p AB - k_m C - \mu C \quad (5.20)$$

$$\frac{dv}{dt} = -\mu \left(\frac{f(X(t))}{1 - \int_0^X f(X(t)) dX} X \right) v + 4\mu \left(\frac{f(2X(t))}{1 - \int_0^{2X} f(X(t)) dX} 2X \right) v_2 \quad (5.21)$$

$$F_A = \frac{\left(\frac{v_{max}}{1 + S/K_{Si}} \right) S}{KN_b + S} \left(\frac{S}{K_t + S} \right)^{1.5} \quad (5.22)$$

$$F_B = \frac{\left(\frac{v_{max}}{1 + S/K_{Si}} \right) S}{KN_b + S} \left(\frac{K_t}{K_t + S} \right)^{1.5} \quad (5.23)$$

$$\mu = F_A A \quad (5.24)$$

Eqns (5.16) – (5.24) are solved for every set of initial conditions, i.e.,

$$\mathbf{x} = [X(0) \ S(0) \ A(0) \ B(0) \ C(0)]^T \quad v(0) = N(\mathbf{x}(0))$$

Table 5-2: Population balance model parameter definitions

Variable	Definition	Units
N_b	total biomass concentration	<i>intensity</i>
t	time	<i>hr</i>
S_{ext}	extracellular substrate concentration	<i>mmol/L</i>
X	cell size	<i>intensity</i>
S	intracellular glutamate concentration	<i>mmol/L</i>
A	intracellular NADPH concentration	<i>mmol/L</i>
B	intracellular ROS concentration	<i>mmol/L</i>
C	intracellular NADP ⁺ concentration	<i>mmol/L</i>
v	number of cells with a certain size X and intercellular concentrations of components	<i>cells</i>
v_2	number of cells with double the size of v with the same intercellular concentration of components	<i>cells</i>

μ	growth rate of individual cells	$\frac{1}{hr}$
$f(X(t))$	probability density function for the partition rate of cells of size X	-
Y	yield coefficient	$\frac{intensity\ biomass}{g\ glutamate}$
v_b	maximum specific growth rate for biomass	$\frac{1}{intensity\ biomass\ hr}$
k	half saturation constant	$\frac{1}{hr}$
K_{si}	rate constant for substrate inhibition	$\frac{1}{hr}$
F_A	flux of substrate conversion into A	$\frac{1}{mM\ hr}$
F_B	flux of substrate conversion into B	$\frac{1}{hr}$
k_p	rate constant of $A+B \rightarrow C$	$\frac{1}{hr}$
k_m	rate constant of $C \rightarrow A + B$	$\frac{mM}{hr}$
d	rate of degradation of B (ROS)	$\frac{1}{hr}$
v_{max}	rate constant for the generation of NADPH	mM
K	saturation constant of $F_A + F_B$	mM
K_t	A and B balance constant	$\frac{1}{intensity\ biomass\ hr}$

In the population balance model, Eqns (5.16) – (5.20) determine the distributions arising from the heterogeneities occurring in the population of cells. These equations generate the characteristic curves, each with their own set of initial conditions. Flow cytometry is used to measure the individual cell volumes and intracellular ROS concentrations. Figure 5-1 shows the contour of the number of cells with a particular combination of cell volume and ROS concentration. In order to simplify the number of different initial conditions, the cells were classified into discrete groups represented by the grid in Figure 5-1. All the cells contained within each square grid, referred to as a bin, were summed and the values corresponding to the middle point of the grid were defined as the initial cell volume and ROS concentration. For example, 940

cells lie within 7-8 intensity cell volume and 0.06-0.07 intensity ROS concentration (highlighted on figure). All 940 cells found in this bin are considered to have initial conditions of 7.5 for cell volume and 0.065 for ROS concentration. A total of 462 initial conditions combining 21 different initial cell volumes and 22 initial ROS concentrations were considered. Since the initial concentrations of intracellular glutamate, NADPH and NADP⁺ were not measured in these experiments, each was assumed to be the same in all 462 initial bins in the model. Thus, these initial concentrations were assumed to be the same in all of the cells regardless of size and ROS level. The initial concentrations of these components are summarized in Table 5-3 and the model solves only for the initial concentrations of NADPH and NADP⁺. The initial concentration of glutamate assumed was the value obtained in the primary flask by the bulk oxidative stress model shown in Chapter 4.

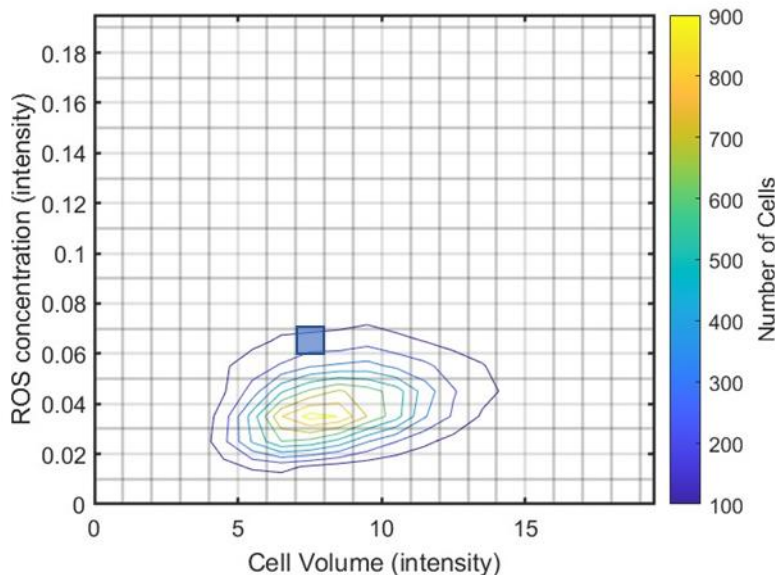


Figure 5-1: Gridding the cell volume and ROS concentration flow cytometry data of *B. pertussis* flask culture experiments for initial conditions in the population balance model. The shaded region contains 940 cells with 7-8 intensity cell volume and 0.06-0.07 intensity ROS concentration.

Table 5-3: Initial concentrations of intracellular components

Component	Initial Concentration (mmol/L)
S	0.104
A	0.2780
C	0.0420

5.3 Materials and Methods

5.3.1 *B. pertussis* batch fermentations

The data from *B. pertussis* flask experiments used to develop the oxidative stress model in Chapter 4 are also used in the development of the population balance model. The experimental procedure is briefly described here for clarity. *B. pertussis* fermentations were carried out in 250 mL flasks (polycarbonate, sterile, with a vented cap having 0.22 μm pore-size PTFE membrane from VWR, Canada) with a working volume of 60 mL by incubation at 36°C and 200 rpm for 36 hrs. Two 1-mL samples were collected every six hours to measure OD via spectrophotometry, glutamate concentration via BioProfile and ROS concentration via flow cytometry. Four experiments were conducted (3 for model calibration and 1 for model validation). The differences in the conditions for these cultures include the starting OD and glutamate concentration in the media. Table 5-4 lists the conditions used for the calibration and validation of the model.

Table 5-4: Calibration and validation conditions for coupled population balance and oxidative stress model

Trial	Set	OD ₆₀₀ (intensity)	Glutamate concentration (g/L)
1	Calibration	0.28	10.1
2	Calibration	0.13	10.7
3	Calibration	0.12	12.4
4	Validation	0.14	8.6

5.3.2 OD, glutamate and ROS measurements

A detailed description of the methods used for these measurements is provided in Chapter 3. Two sets of measurements (OD, glutamate and ROS concentrations) were made based on the two 1-ml samples collected at each interval. Thus, each data point shown represents an average of four measurements, i.e., duplicate samples from each of the two flasks. Following are summaries of the experimental methods used.

To measure OD, the sample was diluted by a factor of 20 using a 0.9% saline solution and OD was obtained using optical spectrophotometry by measurement of the absorbance at 600 nm. The supernatant was collected after centrifugation at 10,000 G for 3 minutes in order to determine the glutamate concentration (after 20-fold dilution) using the chemistry module of a BioProfile FLEX Analyzer (Nova Medical).

Oxidative stress was measured using the fluorescein derivative carboxy-2',7'-dichlorodihydrofluorescein diacetate (carboxy-H₂DCFDA) from Sigma-Aldrich (Canada). Samples with an OD of 0.05 were prepared and carboxy-H₂DCFDA was added to the samples to obtain a final concentration of 50 μ M. Flow cytometry was conducted using an S3e Cell Sorter (Bio Rad). To measure the ROS distribution, the FL1 diode was used for DCF emission (green fluorescence) which fluoresces at an excitation/ emission pair of 503/523 nm. A total of 20,000 events per sample were collected at a rate of 1000 events per second.

5.3.3 Model fitting

The coupled model (Eqns (5.14) – (5.24)) was calibrated using the *fmincon* function (subroutine) in MATLAB. This function is designed to find the minimum of a nonlinear

multivariate function of the square errors between predicted and measured quantities. In particular, the sum of the root-mean-square-error (RMSE) of the measured variables (biomass and extracellular glutamate concentration) as well as the RMSE of the mean and standard deviation of the flow cytometric distributions of cell volume and intracellular ROS concentration were minimized. To ensure that all the variables have comparable effects on the outcome of the model, the RMSE of each variable was standardized by dividing it by the corresponding mean, as shown in the following expression:

$$RMSE = \sum \frac{\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}}{\mu_i} \quad (5.25)$$

where y_i is the measured value, \hat{y}_i is the predicted variable, n is the number of calibration samples and μ_i is the mean of each variable.

The model was calibrated with 30 samples collected under 3 different flask conditions and validated with an additional 10 samples collected from a different condition. OD and glutamate concentration measurements were obtained for each sample (calibrated to a total 30 samples with 10 used in validation). The flow cytometry data were divided into 462 values and modelled via the method of characteristics. The average and standard deviations of the distributions for size and standard deviation were calculated based on the model results of 1776 differential equations corresponding to the number of different combinations of initial ROS concentration and cell size conditions. For each such combination, a different characteristic curve was calculated. The model parameters estimated by the fitting were the various rate constants (reaction rate constants, substrate inhibition rate constant, growth rates, partitioning rate and degradation rate constant for

ROS), initial concentration of component A (intracellular NADPH) and the ROS concentration factor to convert intensity into a concentration unit. Intracellular NADPH was not measured and therefore was considered to be a calibration parameter. These terms are defined in Table 5-2 in section 5.2.3.

The Kullback-Leibler divergence (Kullback and Leibler, 1951) has been often used as a measure of the disparity between two probability distributions $P(x)$ and $Q(x)$. Kullback-Leibler divergence (Eqn (5.26)) is calculated as the sum of probability of each event in $P(x)$ multiplied by the logarithm of the ratio of the probability of the event in $P(x)$ to the probability of the event in $Q(x)$ shown in Eq (5.26). In the case of the flow cytometry data for cell volume and ROS concentration, $P(x)$ represents the fraction of cells in a bin predicted by the model and $Q(x)$ corresponds to the fraction of cells in a bin obtained from the experiments. The sum involved in Eqn (5.26) is taken over all 464 bins. The closer the Kullback-Leibler divergence is to 0, the more similar are the measured and predicted population distributions. A value of 0 corresponds to perfect agreement between the distributions.

$$D_{KL} = \sum P(x) \ln \frac{P(x)}{Q(x)} \quad (5.26)$$

5.3.4 Probability density function

The partitioning rate in the population balance model contains a probability density function $f(X)$ as shown in Eqn (5.27). The probability density function is defined in terms of 2 quantities – the mean (m) and variance of the distribution. The dependence of the probability density function $f(X)$ on cell volume is shown below in Figure 5-2. A change in the mean of this function shifts the curve horizontally (Figure 5-2a), whereas a change in the standard deviation (σ) changes the width of the curve (Figure 5-2b). As the standard deviation increases, so does the

width of the curve. The partitioning rate also depends on the substrate concentration, as evident from Eqn (5.27) (Tsuchiya *et al.*, 1966; Mantzaris *et al.*, 1999).

$$\gamma = \frac{f(X)}{1 - \int_0^X f(X)dx} \left(\frac{\left(\frac{v_{max}}{1 + S/K_{si}} \right) S}{KX + S} \left(\frac{S}{K_t + S} \right)^{1.5} \right) AX \quad (5.27)$$

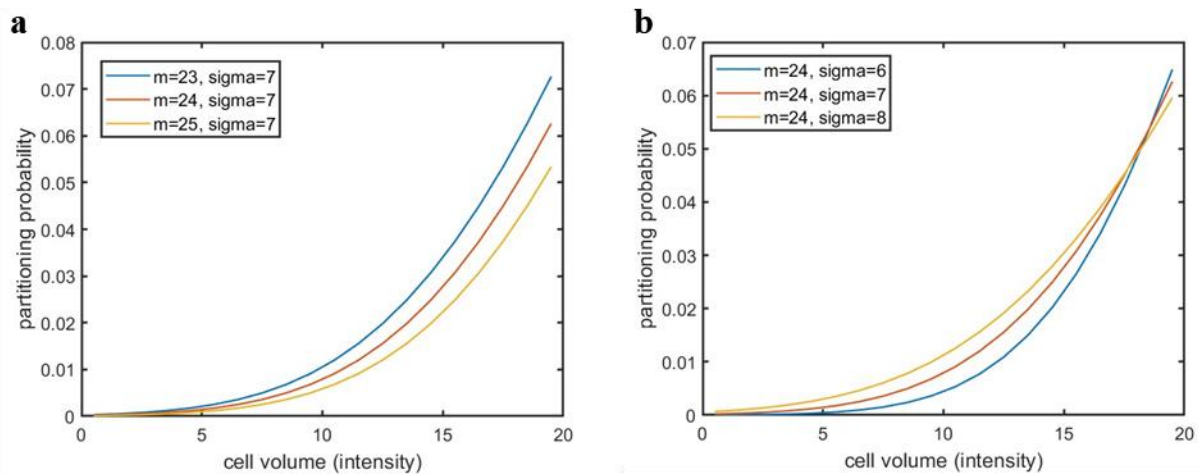


Figure 5-2: Effect of distribution (a) mean and (b) standard deviation on the dependence of probability density function $f(X)$ on cell volume.

While the probability density function was initially assumed to be normal with a prescribed mean and a standard deviation, the fitting of the distribution to the experimental data was found to be challenging due to the nonlinearity of the problem and the occurrence of local minima in the parameter fitting optimization procedure. After failing to satisfactorily fit the model to the experimental data for a plethora of conditions for the mean and standard deviation of the probability density function, we decided to use an interpolated function (with 8 points) as shown in Figure 5-3. The interpolated function allowed for a steeper increase in the probability of division with increasing cell size while still having a low probability of cell division before the sharp rise close to size 12. The interpolated curve for γ applies to the data exactly as does the probability

density function $f(X)$; a one-to-one correspondence exists between these two quantities for each cell volume X . The use of this interpolated function allowed for more control of the partitioning rate in the parameter estimation procedure leading to a better overall fit of the model. It should be noted that tuning the probability function affects the distribution of cell size but not the other states since the reaction rates depend on the total number of cells N_b , but not on individual cell sizes (see Eqns (5.22) and (5.23)).

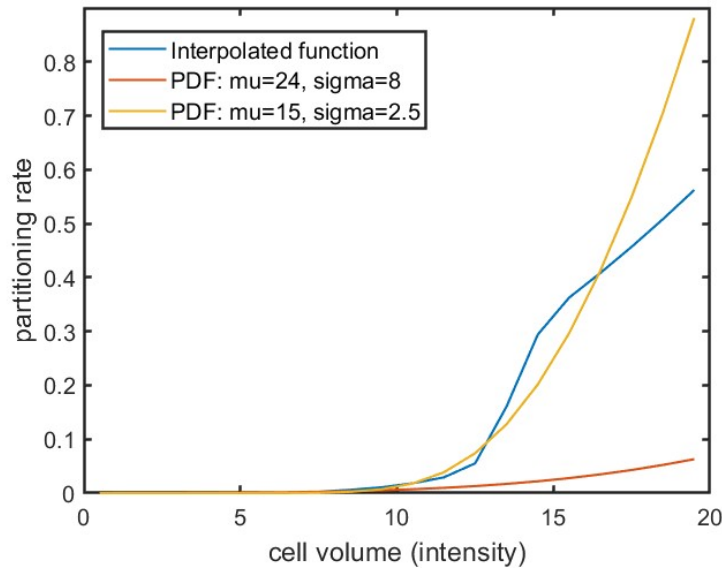


Figure 5-3: Comparison of the dependence of the partitioning rate on cell volume according to a normal distribution (red and yellow) and fitted interpolated function (blue) for the probability density function.

5.4 Results and Discussion

5.4.1 Trends of the flow cytometry distributions

The same control conditions as in Chapter 4 were used to mimic the first reactor in the Sanofi manufacturing train. This condition involved the use of the Sanofi media formulation

containing 9.8 g/L glutamate. The experiment was conducted in a secondary flask of *B. pertussis* which was inoculated by cells collected during the exponential growth phase from the primary flask. The measured values of biomass and extracellular glutamate concentration, average cell volume, intracellular ROS concentration, cell volume distribution and intracellular ROS concentration distribution obtained under these control conditions are presented in Figure 5-4.

The cells have a larger cell volume and a wider distribution during the lag and stationary phases (Figures 5-4c and e), whereas they have a smaller volume and narrower distribution during the exponential growth phase. Generally, cells increase in size during the lag phase and their metabolism is targeted to prepare the molecules necessary for replication (Rolfe *et al.*, 2012). Usually, the average cell size decreases during the stationary phase as cells stop growing while division can still occur (Kolter *et al.*, 1993). On the other hand, pH can also play an important role in cell division; alkaline pH can suppress division activation leading to an increase in cell size (Mueller *et al.*, 2020). Since the pH was not controlled during this batch flask, the pH of the culture increases over time. However, the effect of pH was not incorporated into the model in this study.

The increase in ROS concentration during the lag and stationary phases was discussed in detail in section 4.4.1 in Chapter 4. The distributions reflect the same pattern as the ROS concentration is lowest during the exponential growth phase and highest during the lag phase (Figures 5-4d and f). An interesting observation is the occurrence of a tail in the ROS concentration distribution of the 36-hour sample (Figure 5-4f). While many cells contain a low level of ROS and are preparing for survival mode at this point, some older cells that are unable to cope with nutrient starvation may still contain high amounts of ROS. For example, starvation of glutamate may cause depletion of NADPH which is an essential reactant in anti-oxidative reactions.

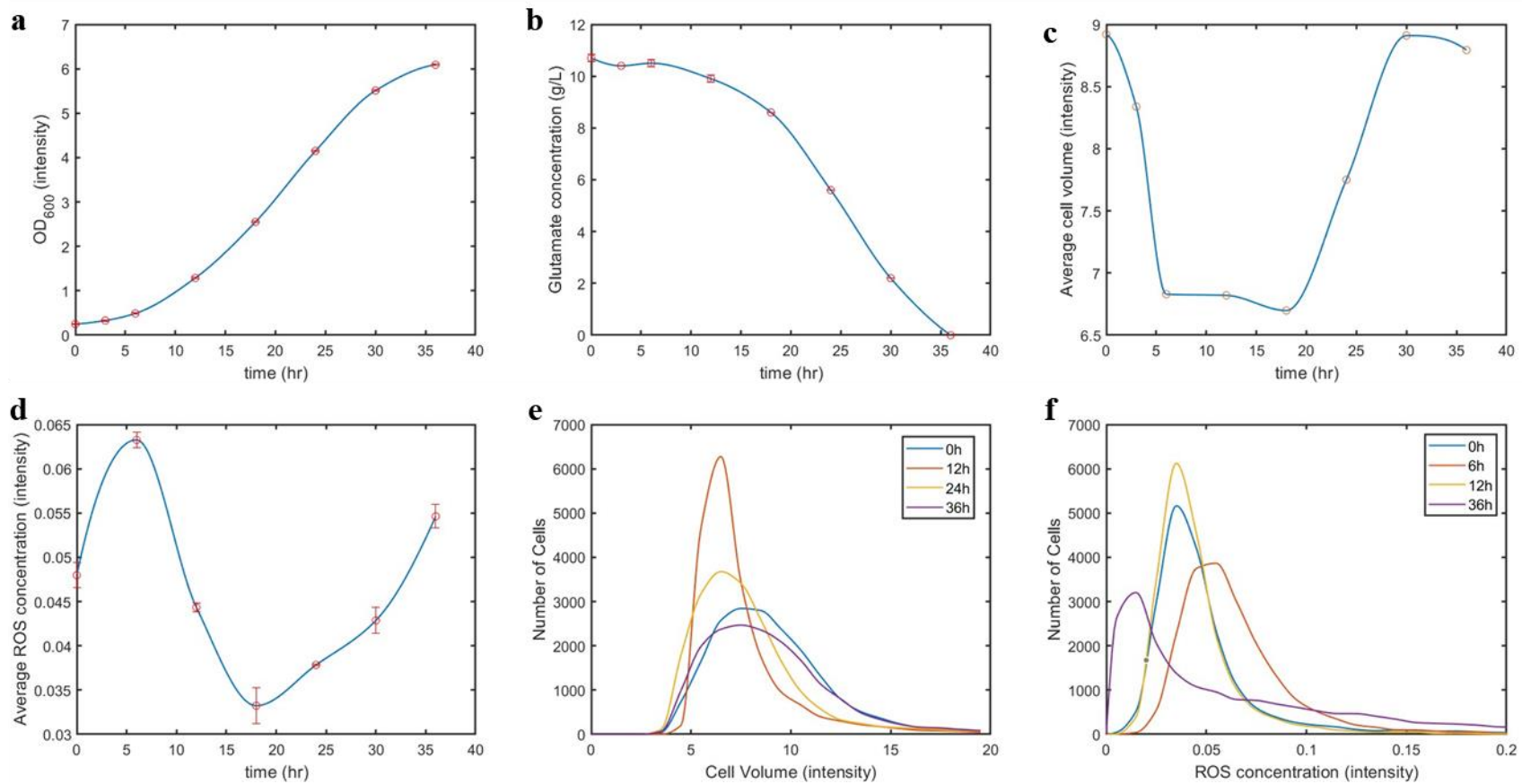


Figure 5-4: *B. pertussis* flask culture measurements: (a) optical density (OD), (b) extracellular glutamate concentration, (c) average cell volume, (d) average intracellular ROS concentration, (e) cell volume distributions and (f) extracellular ROS concentration distributions of a culture with an initial OD of 0.28 and initial glutamate concentration of 9.8 g/L.

5.4.2 Characteristic curves

The solution of the population balance model with the method of characteristics yields characteristic curves which can be used to determine the effect of different initial cell conditions on biomass growth and oxidative stress levels. Figure 5-5 depicts the characteristic curves for different initial cell sizes with the same starting intracellular ROS concentration. The curves in Figures 5-5b – d indicate that cell size has no effect on the intracellular concentrations of glutamate, NADPH and ROS. This is expected since the reaction rates given in Eqns (5.23) and (5.24) depend on the total biomass, but not on cell size, as explained above. On the other hand, the growth rate is proportional to cell size and so a larger initial size leads to faster growth (Figure 5-5a).

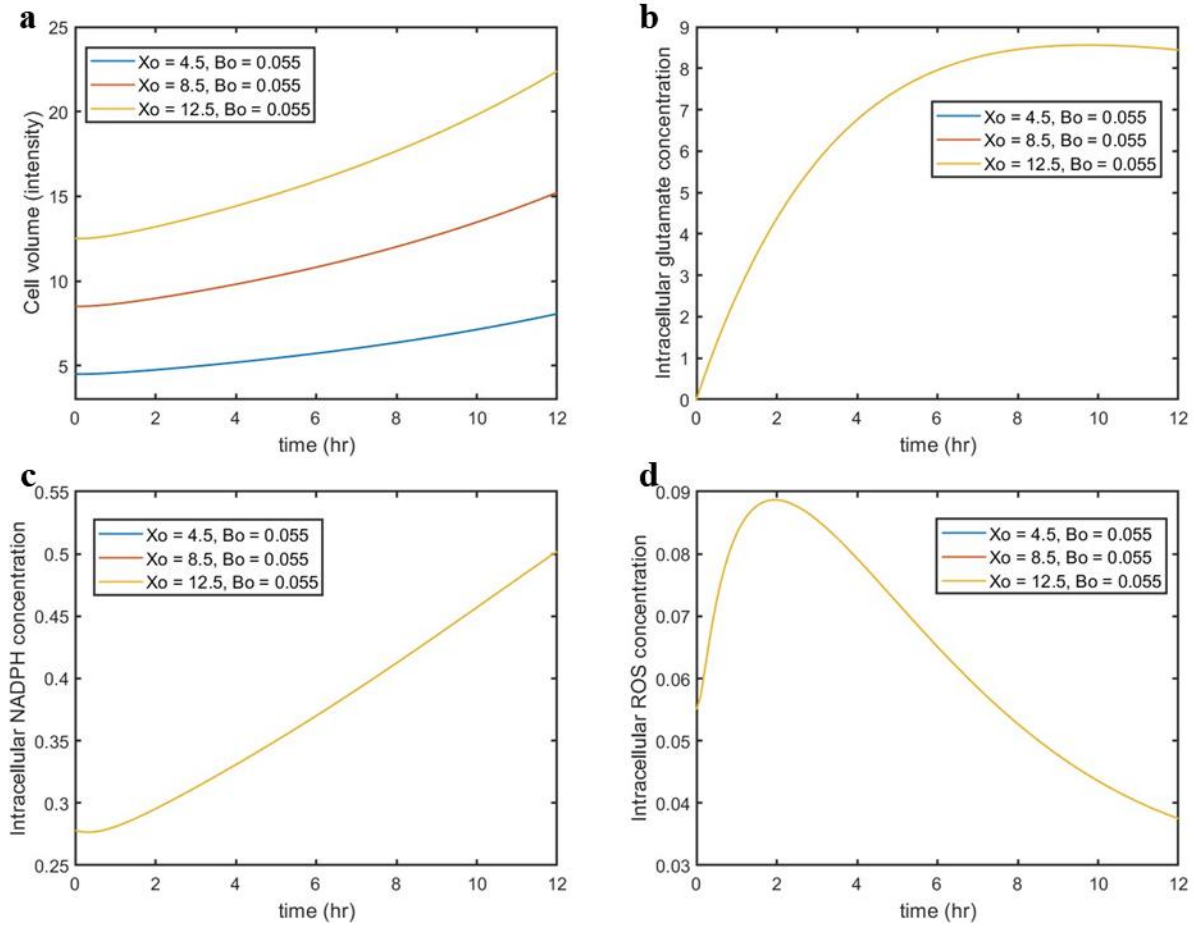


Figure 5-5: Characteristic curves: (a) cell volume, (b) intracellular glutamate concentration, (c) intracellular NADPH concentration and (d) intracellular ROS concentration for different initial cell sizes at the same initial ROS concentration intensity of 0.055.

Figure 5-6 shows the characteristic curves for cells with the same initial size (8.5) that are exposed to different initial intracellular ROS concentrations. In the mechanism, ROS indirectly inhibits cell growth since it reacts and consumes NADPH which would otherwise react to promote growth. This trend is observed in Figures 5-6a and c which show that an increase in the initial ROS concentration leads to a smaller cell size and less intracellular NADPH (Eqns (5.16) – (5.19)).

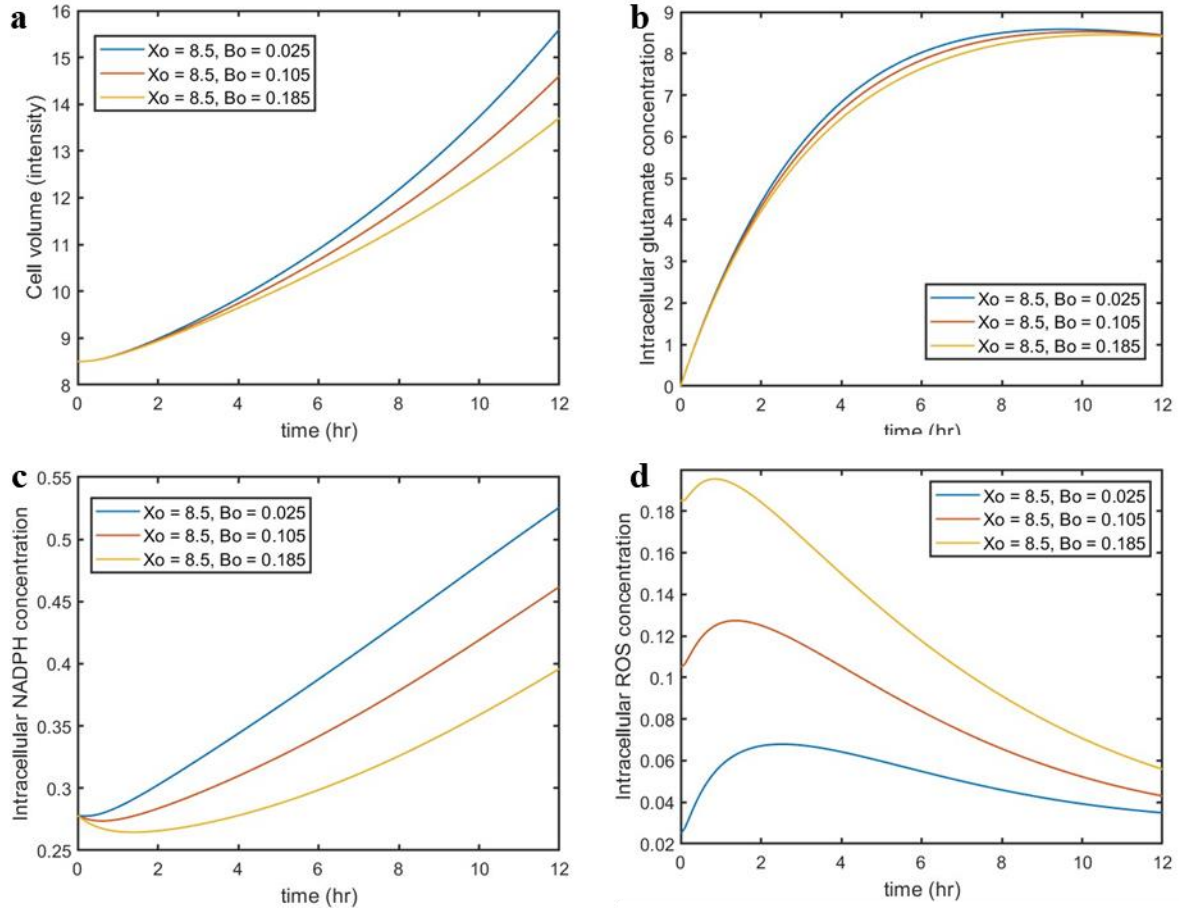


Figure 5-6: Characteristic curves: (a) cell volume, (b) intracellular glutamate concentration, (c) intracellular NADPH concentration and (d) intracellular ROS concentration for different initial ROS concentrations with cells with the same initial size (8.5).

The characteristic curves also play an important role in determining the partitioning rate of the cells since the volume of a mother cell must be twice that of the daughters into which it divides. Since the growth rate is proportional to cell size, the ratio between the cell sizes from any two characteristic curves at the same time remains constant (C_{x1}/C_{x2}) if the initial concentration of ROS, A , S and C are the same. For example, a characteristic curve corresponding to a cell size that is initially twice as large as another initial cell size will remain twice as large as the one for the smaller size over the entire duration of the culture for the same initial concentrations in ROS (B_o

= 0.055) and same initial concentrations of A , S and C (Figure 5-7b). The cell volume changes according to the product of specific growth rate μ and size X (Eqn (5.16)) and μ may change with time as a function of S , A , B and C . However, μ remains constant along the characteristic curve obtained with the same initial values of S , A , B and C since the kinetic rates depend on the total volume of cells (see N_b in the denominators of Eqns (5.23) and (5.24), but not on X).

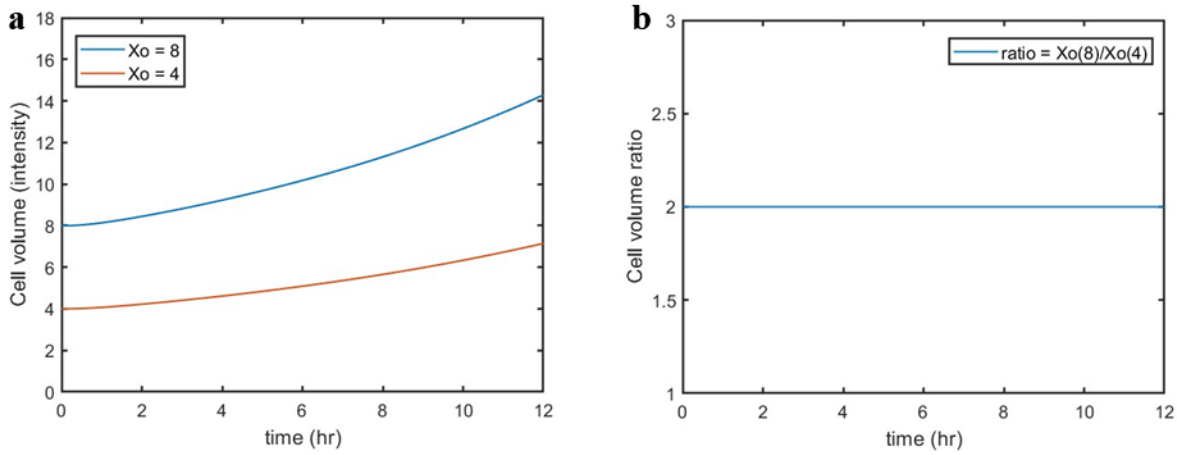


Figure 5-7: Characteristic curves for the same initial value of $B = 0.055$ and same initial values of A , S and C and for different initial values of cell size X : variation of (a) cell volume and (b) size ratio with time.

The volume of most *B. pertussis* cells falls in the range between 6-10 intensity units. Using the method of characteristics, it is possible to follow the trajectory of the growing cells over time. It is expected that the number of cells corresponding to a characteristic curve will be highest when the cell volume is within this range because of the highest probability assigned to this range of sizes in the partitioning rate probability function. Accordingly, larger cells have a higher chance of dividing and producing daughters within the intensity bins that contain cells with the common range for *B. pertussis* as shown in the experimental data (i.e., cell size range 4 - 14). Figure 5-8 depicts the characteristic curves for cell volume and the number of cells within that bin for the

same initial values of S , A , B and C . The bin with an initial cell volume of 3.5 does not contain any cells initially but cells divide into daughters that lie in this bin as the cell volume increases. The number of cells lying within this bin reaches a peak when the cell volume is approximately 8 intensity units (i.e., after ~16 hrs of incubation). This trend is common between the characteristic curves as the model is calibrated according to the measured average cell volume and standard deviation of the distribution.

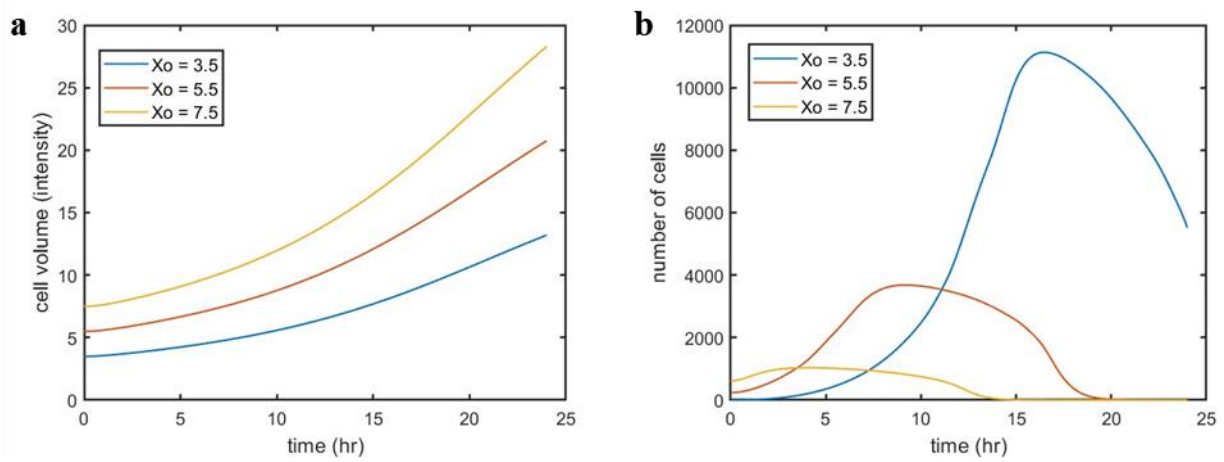


Figure 5-8: Characteristic curves with the same initial concentration of ROS ($B_o = 0.105$): (a) cell volume and (b) number of cells to examine the effect of the partitioning rate on cells in the characteristic curve.

5.4.3 Re-binning and projection grid

The solution using the method of characteristics poses specific numerical challenges related to the number of bins and the sizes to be used to describe the flow cytometry data. The numerical challenges arise from the fact that both the number of characteristic curves and the number of bins remain constant during the solution. Figure 5-8a shows that the differences between the characteristic curves at each time interval increase exponentially as time progresses. In view of this increase and the fact that the number of bins is kept constant, each bin must become

progressively wider as time continues to cover the entire range of the change of the state variable according to the method of characteristics. Accordingly, a smaller number of bins will be progressively assigned to the range of cell size that is relevant for this process (i.e. an cell volume intensity range of 4-12), thus leading to an ever lower numerical resolution over that range.

Possible approaches to address this problem are to: i) increase the number of bins to ensure enough discretization in the range of cell size of interest or ii) increase the number of bins in the lowest size range so that characteristic curves starting at the smallest sizes will contribute sufficient number of characteristic curves (i.e., enough bins) over the range of cell size of interest at all times. However, both of these alternatives require a significant increase in the computational effort which is already high.

An approximate solution to this problem that we propose involves re-binning the cells after every sampling interval (6 hours) calculation and then continuing to integrate the model from the initial conditions defined with respect to a fixed initial grid. Figure 5-9 qualitatively describes this re-binning procedure. The initial data points are found on all corners of the original grid (black open circles). As time passes, the cells follow the characteristic curve trajectories which end at the red open circles after a sampling interval of 6 hours.

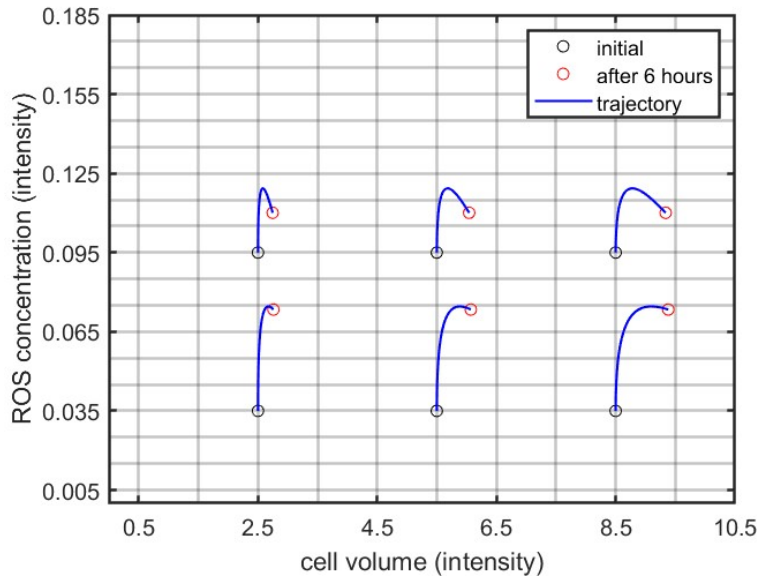


Figure 5-9: Projection window and moving window characteristic curves for cell volume and intracellular ROS concentration over 6 hours.

After the sample interval of 6 hours has elapsed, it is necessary to project the new results (open red circles) onto the original grid (open black circles) in order to continue use of the original grid corresponding to the black dots. The number of cells at the position of the newly calculated red dots must be assigned (“projected”) in the original grid. To achieve this projection, it is assumed that the number of cells at each red circle location is distributed in different proportions among the black circle locations (original grid) based on the idea that these proportions are inversely proportional to the distance between the red circle to each of the black circles. This projection operation is done by using Eqns (5.26) – (5.27) in conjunction with Figure 5-10. This operation preserves the total number of cells and total biomass available before the projection operation is applied. This also conveniently allows for the direct comparison of the model result distributions with the experimental distributions as the latter are defined in the original fixed grid given by the black circles. By projecting the model results at every sampling interval, numerical resolution is maintained in the cell volume intensity range of interest for *B. pertussis* even with the

exponentially increasing distance between each characteristic curves. The projection using Eqns (5.27) and (5.28) maintains the biomass concentration and the total number of cells before and after the projecting to the original grid.

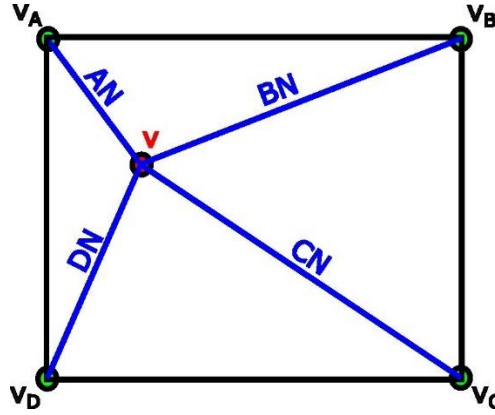


Figure 5-10: Re-binning characteristic curves after 6 hours onto the projection window.

$$\begin{aligned}
 dis_{total} = & \left(\frac{AN + BN + CN}{AN + BN + CN + DN} \right) + \left(\frac{AN + BN + DN}{AN + BN + CN + DN} \right) \\
 & + \left(\frac{AN + CN + DN}{AN + BN + CN + DN} \right) + \left(\frac{BN + CN + DN}{AN + BN + CN + DN} \right) = 3 \quad (5.27)
 \end{aligned}$$

$$v_A = \frac{\left(\frac{BN + CN + DN}{AN + BN + CN + DN} \right) v}{dis_{total}} \quad (5.28)$$

The intracellular concentrations of glutamate, NADPH and $NADP^+$ also have to be adjusted following the projection of the characteristic curves onto the original grid. These are adjusted through the MATLAB function *griddata* which uses a 3D linear interpolation of the intracellular concentrations of glutamate, NADPH and $NADP^+$ using the intracellular concentration of ROS and cell volume as the independent variables.

5.4.4 Calibration and validation of population balance model

The PBM calibration and validation plots for the evolution of biomass, extracellular glutamate and average intracellular ROS concentrations are found below in Figures 5-11 – 5-14. Since the oxidative stress model kinetics are used in the coupled PBM, the trends of the two models are similar and were already discussed in detail in Chpater 4.4.6.

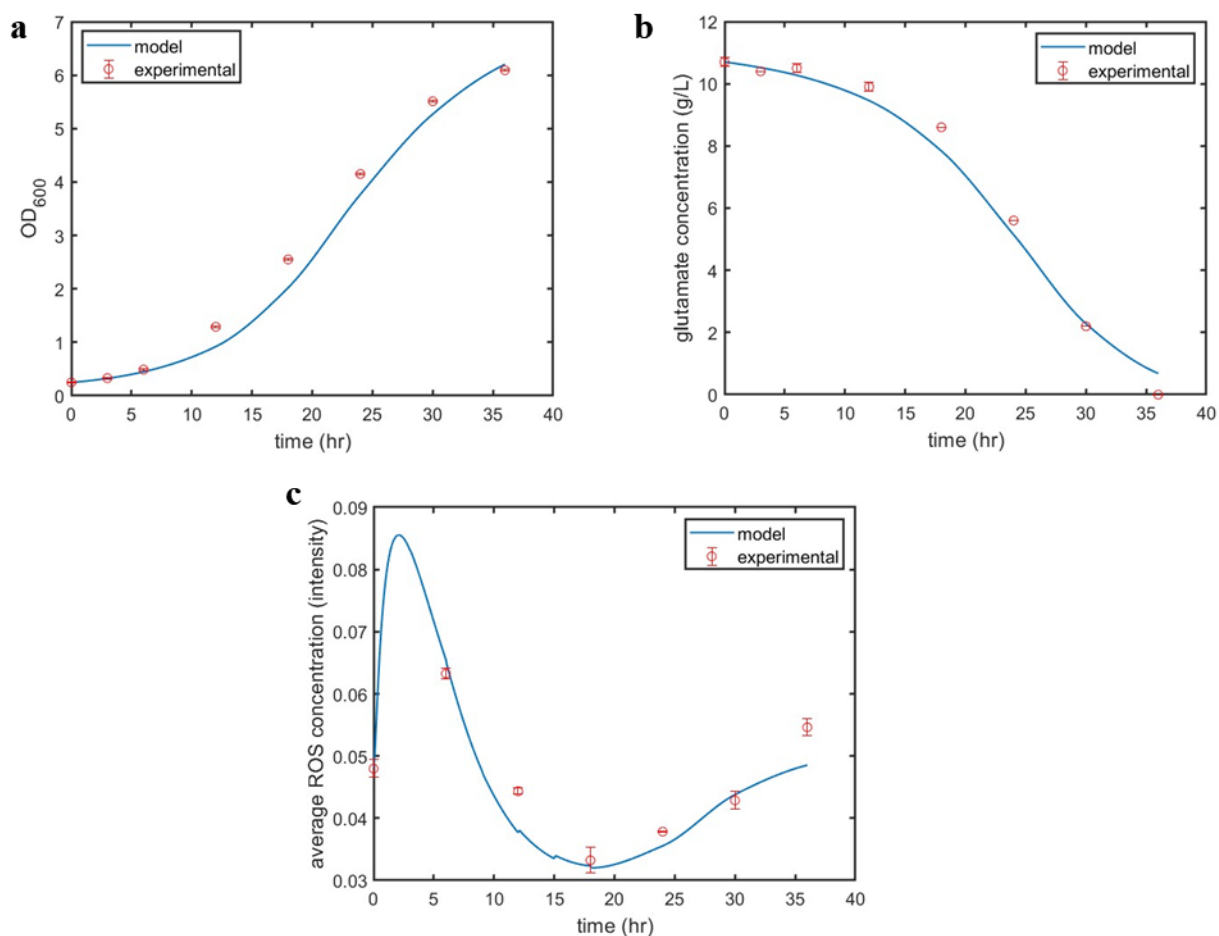


Figure 5-11: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration for a *B. pertussis* culture with an initial OD of 0.28 and initial glutamate concentration of 10.1 g/L.

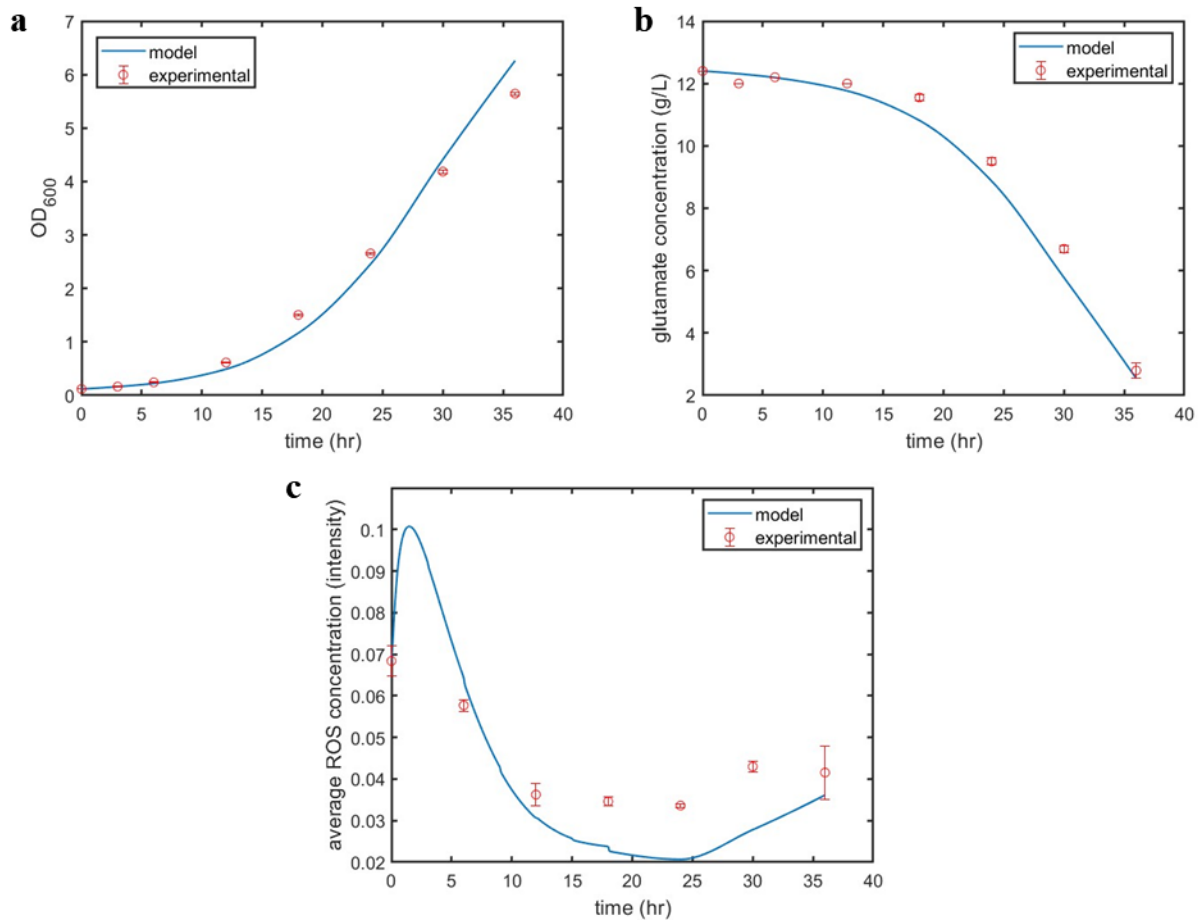


Figure 5-12: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration of a *B. pertussis* culture with an initial OD of 0.12 and initial glutamate concentration of 12.4 g/L.

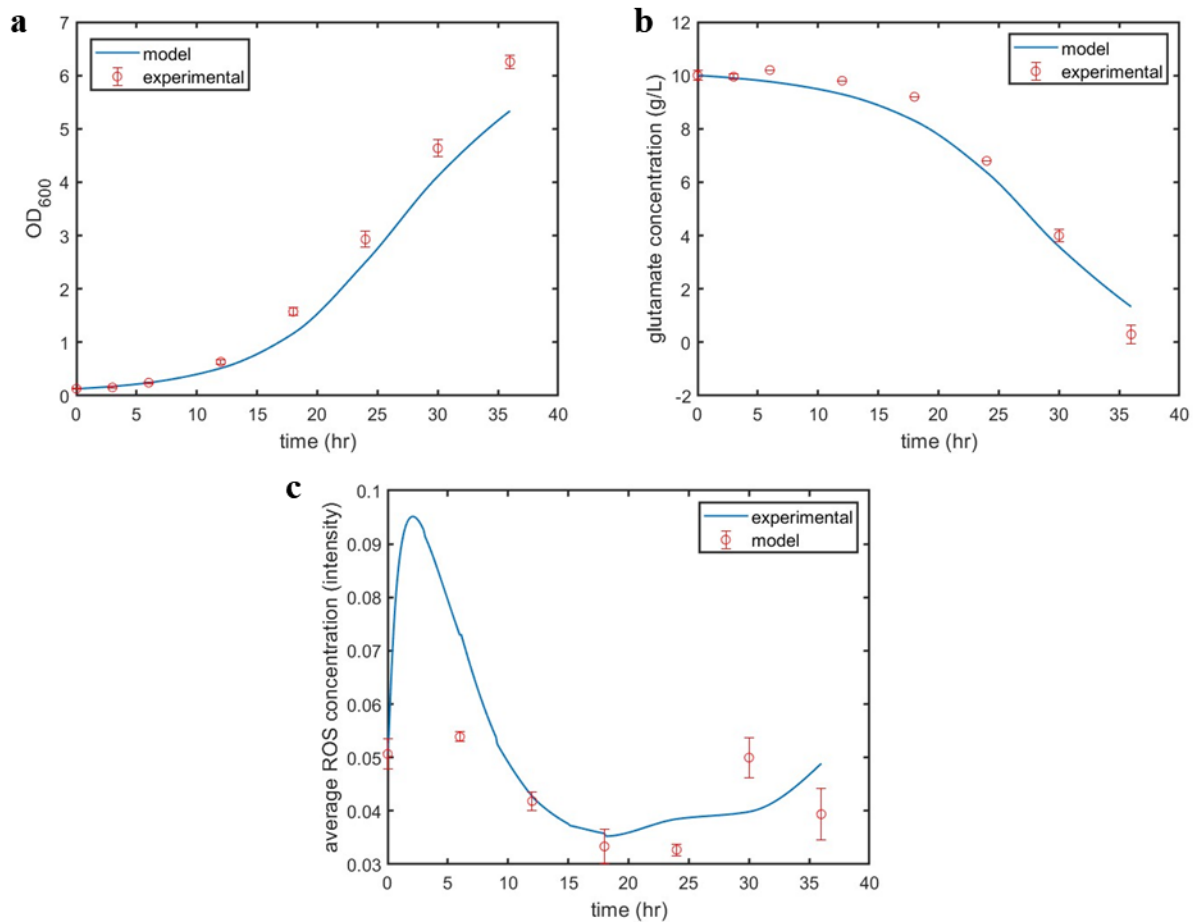


Figure 5-13: Model fitting of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration of a *B. pertussis* culture with an initial OD of 0.13 and initial glutamate concentration of 10.1 g/L.

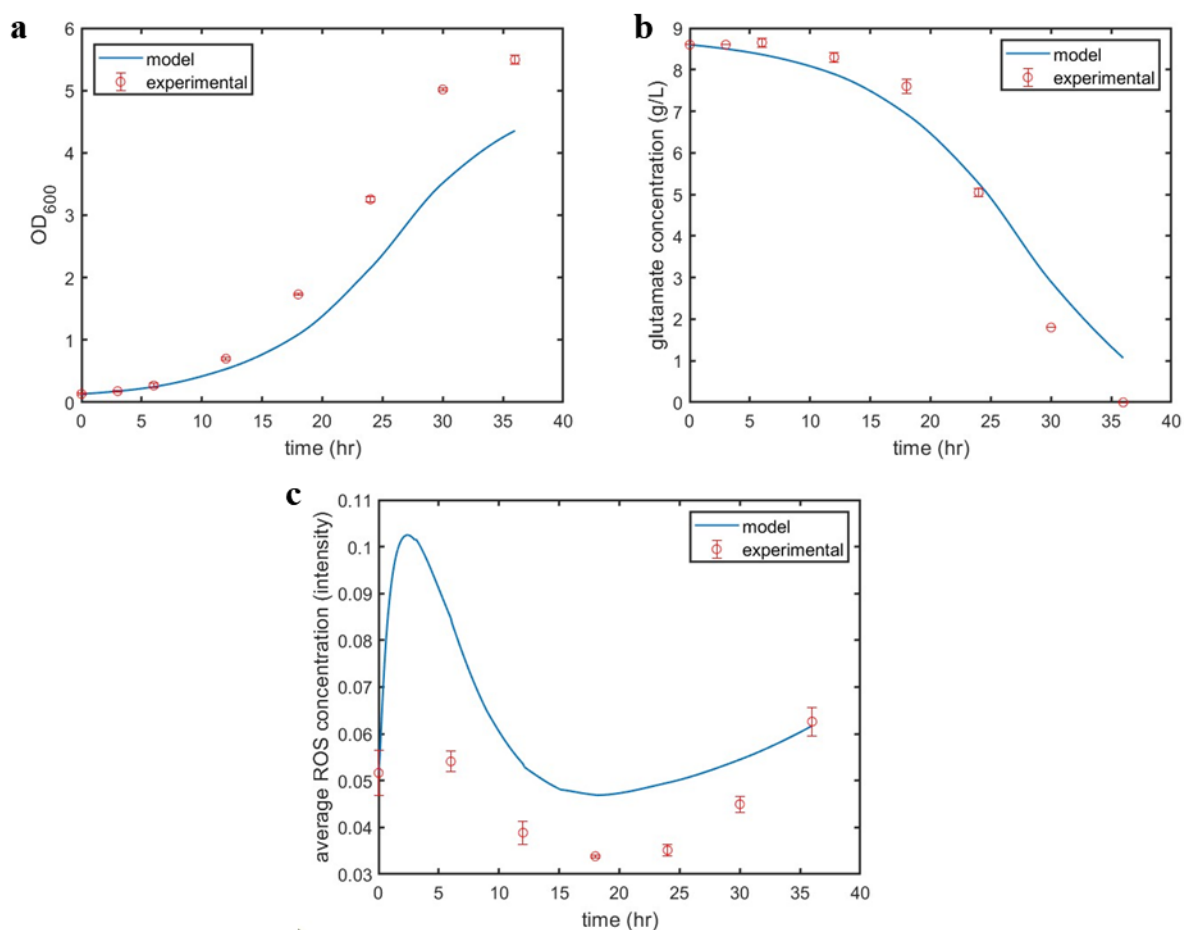


Figure 5-14: Model validation of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS concentration of a *B. pertussis* culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.

The root-mean squared errors (RMSEC) between the calibrated and measured values of OD, extracellular glutamate and intracellular ROS concentrations are presented in Table 5-5 below. The corresponding values for the validated model are included as well. The model provides very good fits for all the measured variables. The calibration and validation RMSEs are very similar to each other for both OD and extracellular glutamate indicating that the noise is not over-fitted in the calibration step. However, the average ROS concentration is not predicted as well in comparison to the calibration experiments, as reflected in the RMSE for validation being 1.88 times larger than the RMSE for calibration. This is shown in Figure 5-14c as the model over-

predicts the average concentration of ROS over the entire run of 36 hours. Including more experiments in the calibration set would provide more data and could potentially increase the predictability of the model for average intracellular ROS concentration.

Table 5-5: Root-mean-squared-error (RMSE) of measured variables for model calibration and validation. The range of each compound is also shown to compare with the RMSE.

	OD₆₀₀ (intensity)	Extracellular glutamate concentration (g/L)	Average Intracellular ROS concentration (intensity)
RMSEC	0.337	0.509	0.008
RMSEV	0.534	0.623	0.015
Range	[0-7]	[0-12.5]	[0-0.4]

The model was fitted to the distributions of the flow cytometry data for cell volume and intracellular ROS concentration based on minimization of the square-errors between the predictions and measurements in both the mean and standard deviations of the distributions. Table 5-6 presents the adjusted RMSE calculated based on the errors between the data and model distribution averages and standard deviations. The RMSE was adjusted by dividing by the mean measurement at each point as shown in Eqn (5.25) to compare the contribution of each error on the same level. Comparison of the results in Table 5-5 indicates that the model gives a better fit for the average cell volume and intracellular ROS concentration than the standard deviations of the same variables.

Table 5-6: Adjusted root mean-squared error (RMSE) of averages and standard deviations of the cell volume and intracellular ROS distributions for model calibration and validation and the range of values and error.

	Cell volume (intensity)		Intracellular ROS concentration (intensity)	
	Average	Standard deviation	Average	Standard deviation
RMSEC _{ad}	0.117	0.523	0.191	2.556
RMSEV _{ad}	0.113	0.333	0.336	1.087
Range	[6.45 – 9.06]	[0.0029 – 0.0109]	[0.0327 – 0.0633]	[2.91E-5 – 1.49E-4]
Error	0.8697	9.51E-4	0.0080	3.93E-5

Although two measured variable distributions (i.e., cell size and ROS) have been used in the development of the model, the measurements and fitted model data were compared along one dimension at a time to facilitate visualization of the results. Figures 5-15 – 5-22 show the single variable (cell volume and intracellular ROS concentration) distributions of the flow cytometry data and model. In the model, the cell volume is controlled by tuning the partitioning and growth rates. These figures show that the vast majority of cells have volumes within the intensity range from 4 to 13 according to both the measurements and fitted model. The cell volume measurements always fall along unimodal distributions, whereas the model yields multimodal distributions particularly over the earlier part of inoculation (Figure 5-15). Reducing the bin size exacerbated the problem of multimodal model distributions (data not shown). A much finer discretization both in terms of the cell size and time may be required to further smooth the distributions predicted by the model, but at the expense of additional computational effort which is already very high.

The intracellular ROS concentration distributions are much narrower than the cell volume distributions until the beginning of the stationary phase (Figure 5-16). However, the average values obtained from the experimental ROS distributions change over a larger range compared to the corresponding averages the average cell volume. The model is very accurate in terms of the

location of the peak intracellular ROS concentration but the widths of the predicted distributions are much narrower than those observed in the experiments. At the beginning of the stationary phase (~30-36 hours), the distribution of the measured intracellular ROS concentration becomes bimodal and generally exhibits a long tail on the right-hand side of the distribution (Figures 5-16f and g). The model does not fit the ROS distributions well in the stationary phase, since the error in the standard deviation of the distribution is 32.8% of the range of the measured standard deviation of the ROS distributions. In particular, the model is not able to predict the long tails in the measured ROS concentration distributions observed during the stationary phase that appear at higher ROS levels. Since the model distributions are so narrow, most cells lie within only a couple of bins. Increasing the number of bins would not significantly improve this result. A probable cause for the deviation between these measured and model-predicted distributions is that cells under starvation conditions may exhibit a different ROS generation rate that may depend on cell age. Thus, as the cells approach the stationary phase where glutamate is close to depletion, older cells may generate more ROS than younger cells and this may result in the long tail observed in the experimental distribution (e.g., Figure 5-16g). Since the individual cell ages were not calculated in our model, it was not possible to impose this age dependent ROS generation rate. An alternative possible adjustment widen the of distribution for intracellular ROS concentration would be to add a probability term for the occurrence of the reaction of glutamate to ROS that would increase as the substrate is depleted.

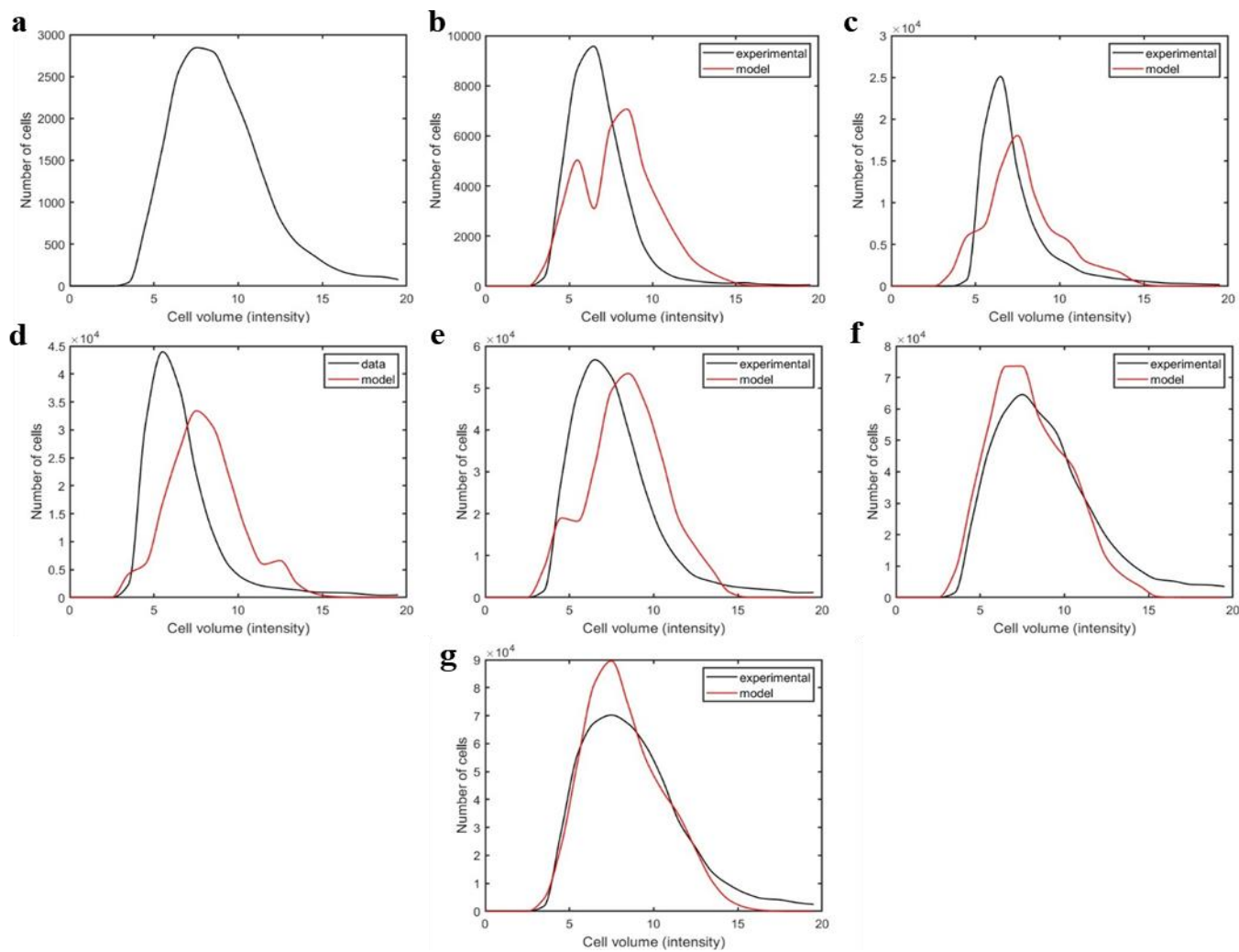


Figure 5-15: Model and experimental flow cytometry distributions of cell volume for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.28 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

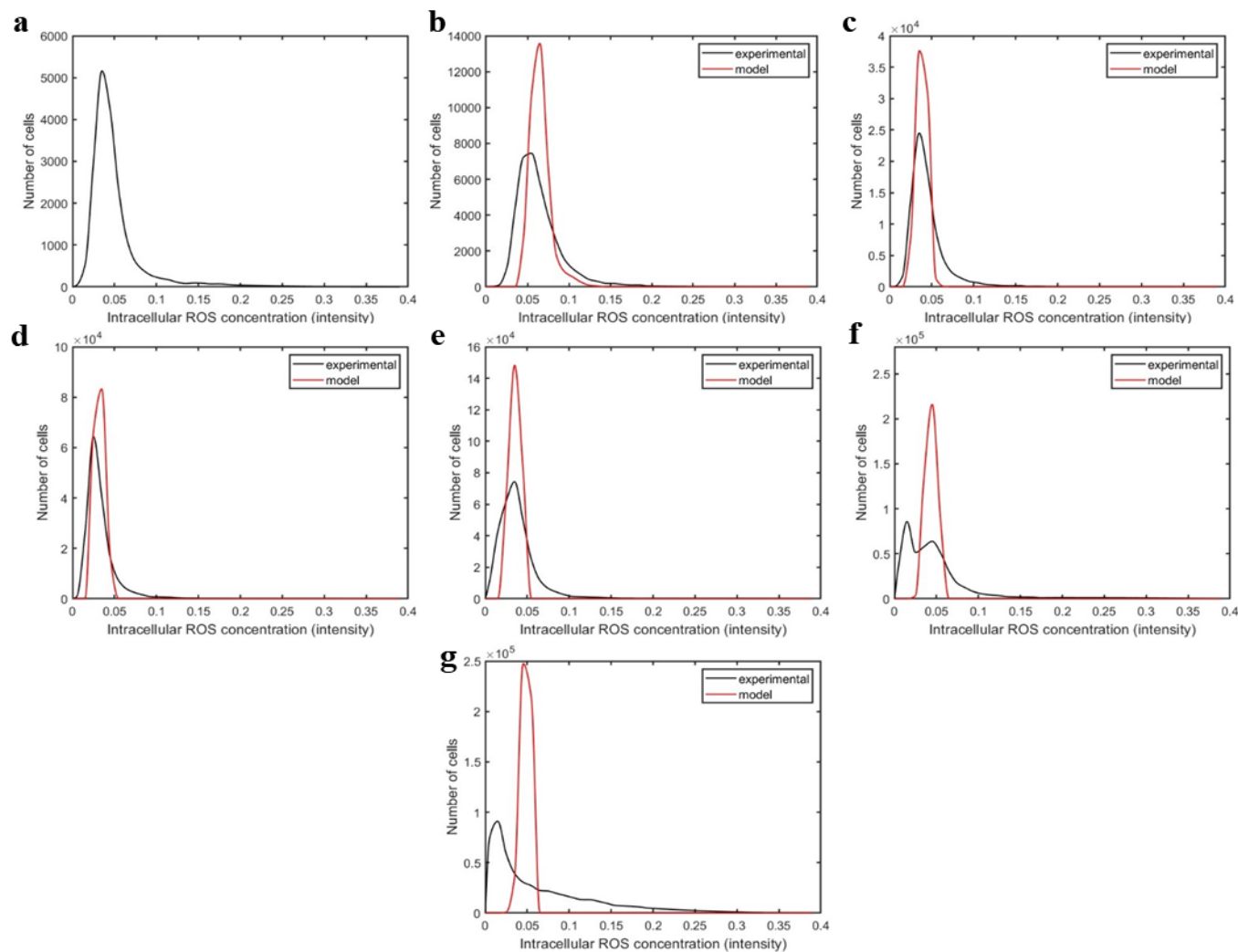


Figure 5-16: Model and experimental flow cytometry distributions of intracellular ROS concentration for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.28 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

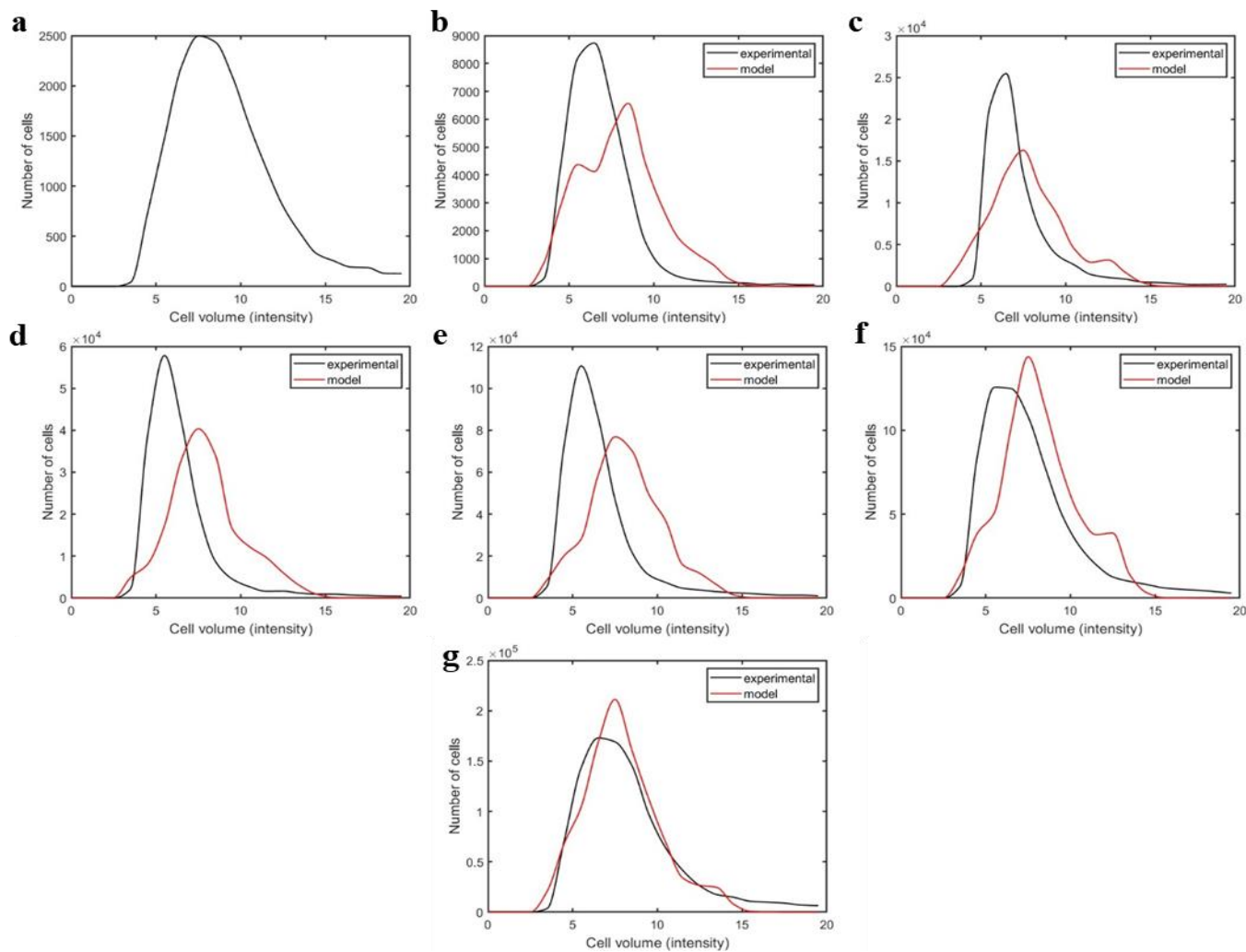


Figure 5-17: Model and experimental flow cytometry distributions of cell volume for a *B. pertussis* shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

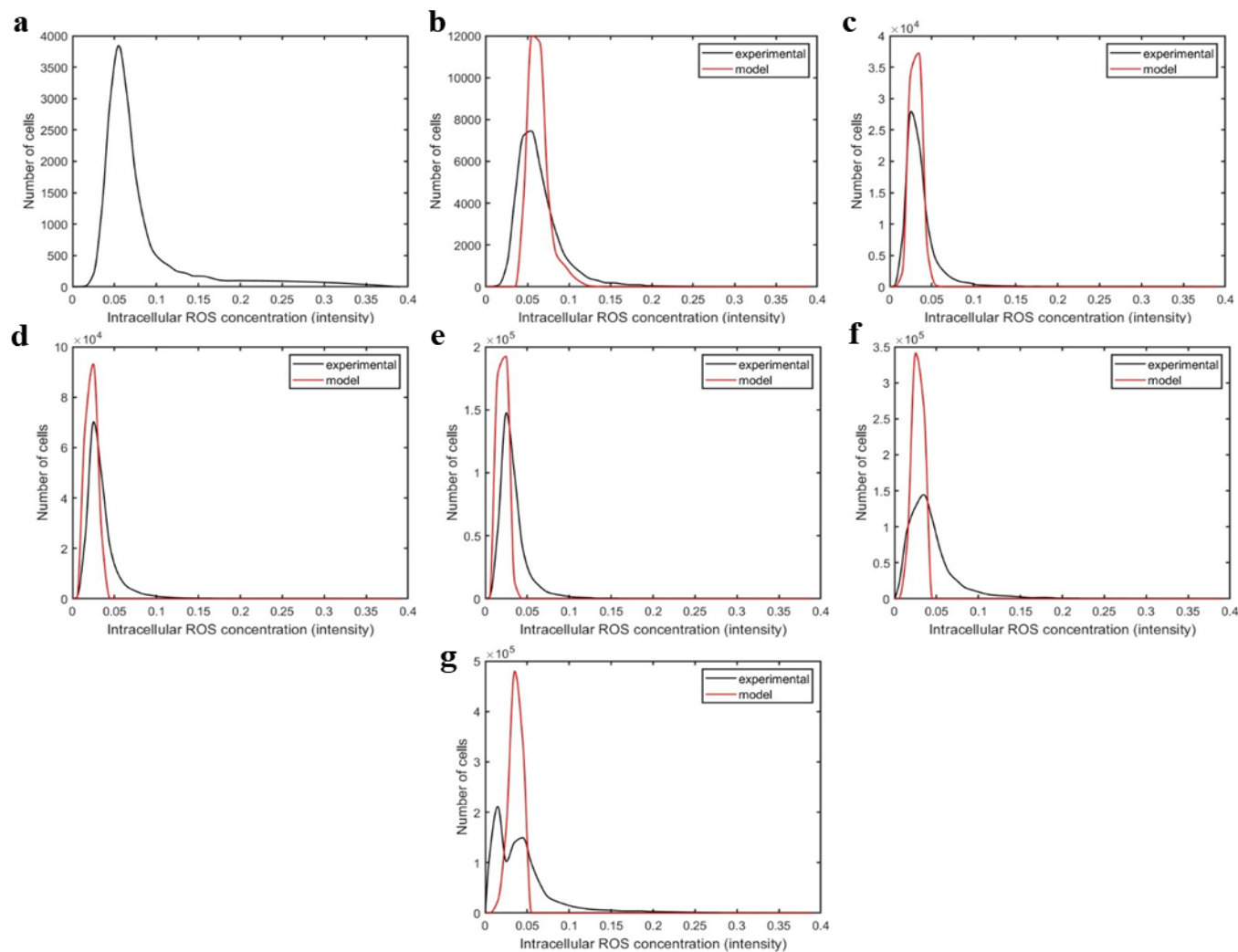


Figure 5-18: Model and experimental flow cytometry distributions of intracellular ROS concentration for a *B. pertussis* shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

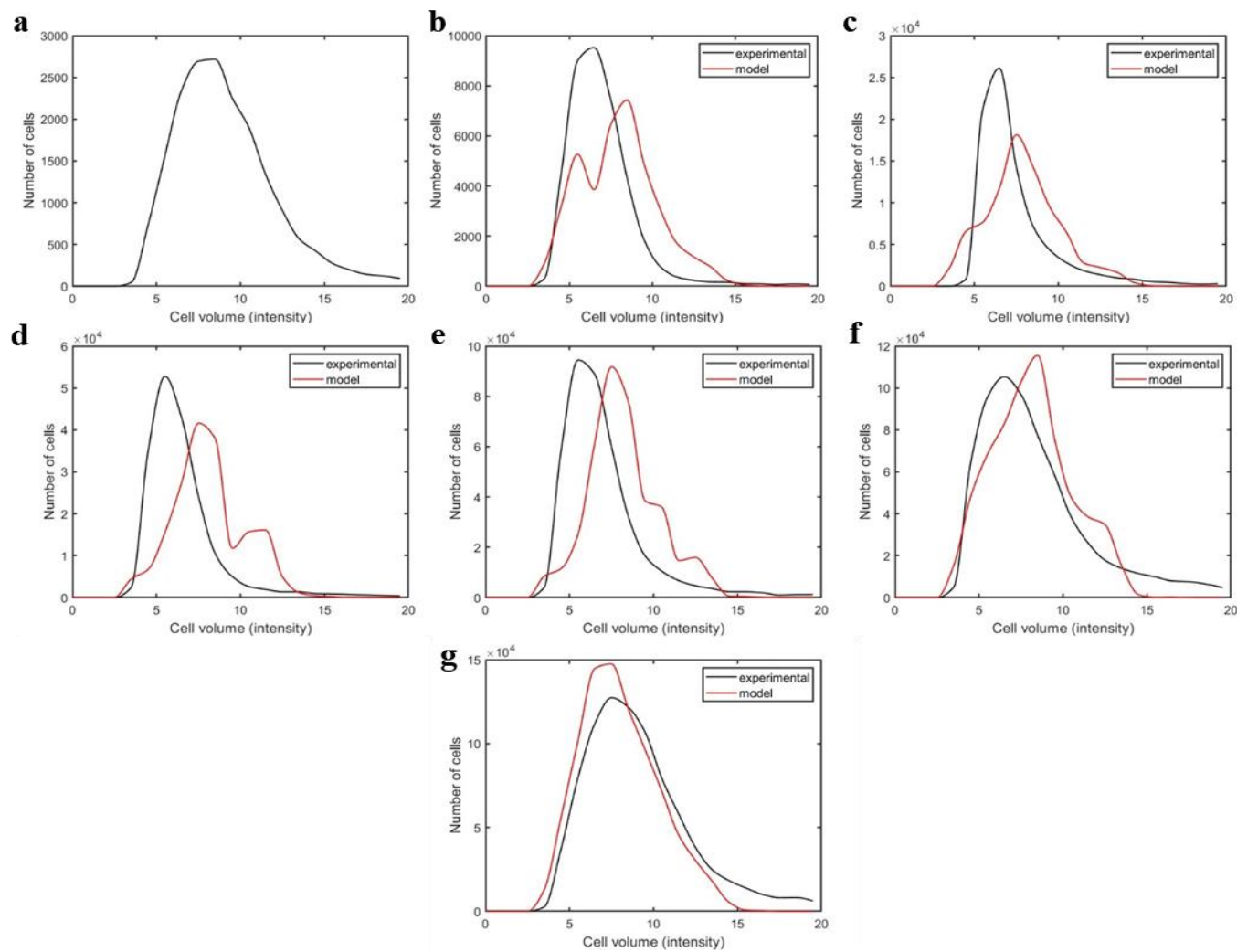


Figure 5-19: Model and experimental flow cytometry distributions of cell volume for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

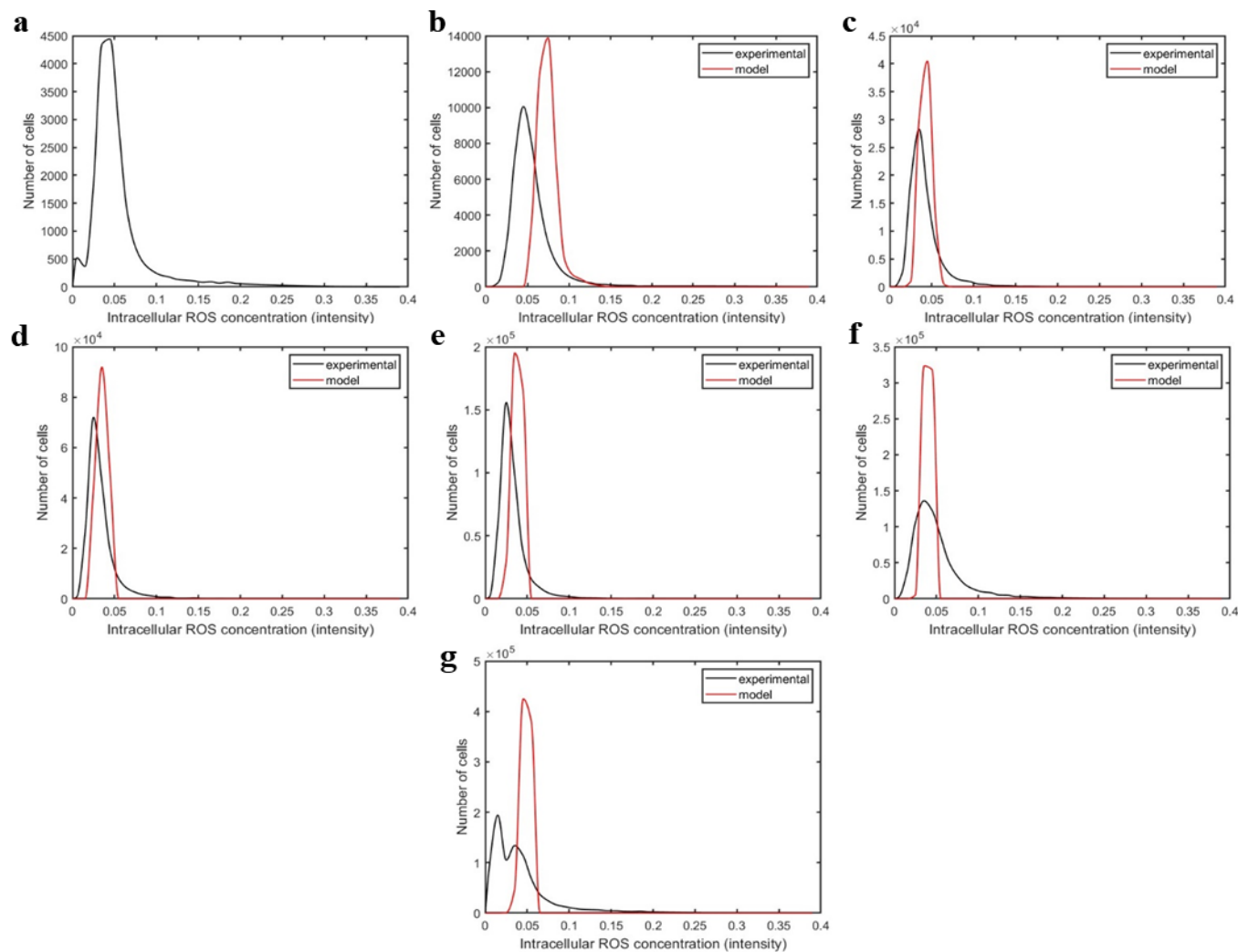


Figure 5-20: Model and experimental flow cytometry distributions of intracellular ROS concentration for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

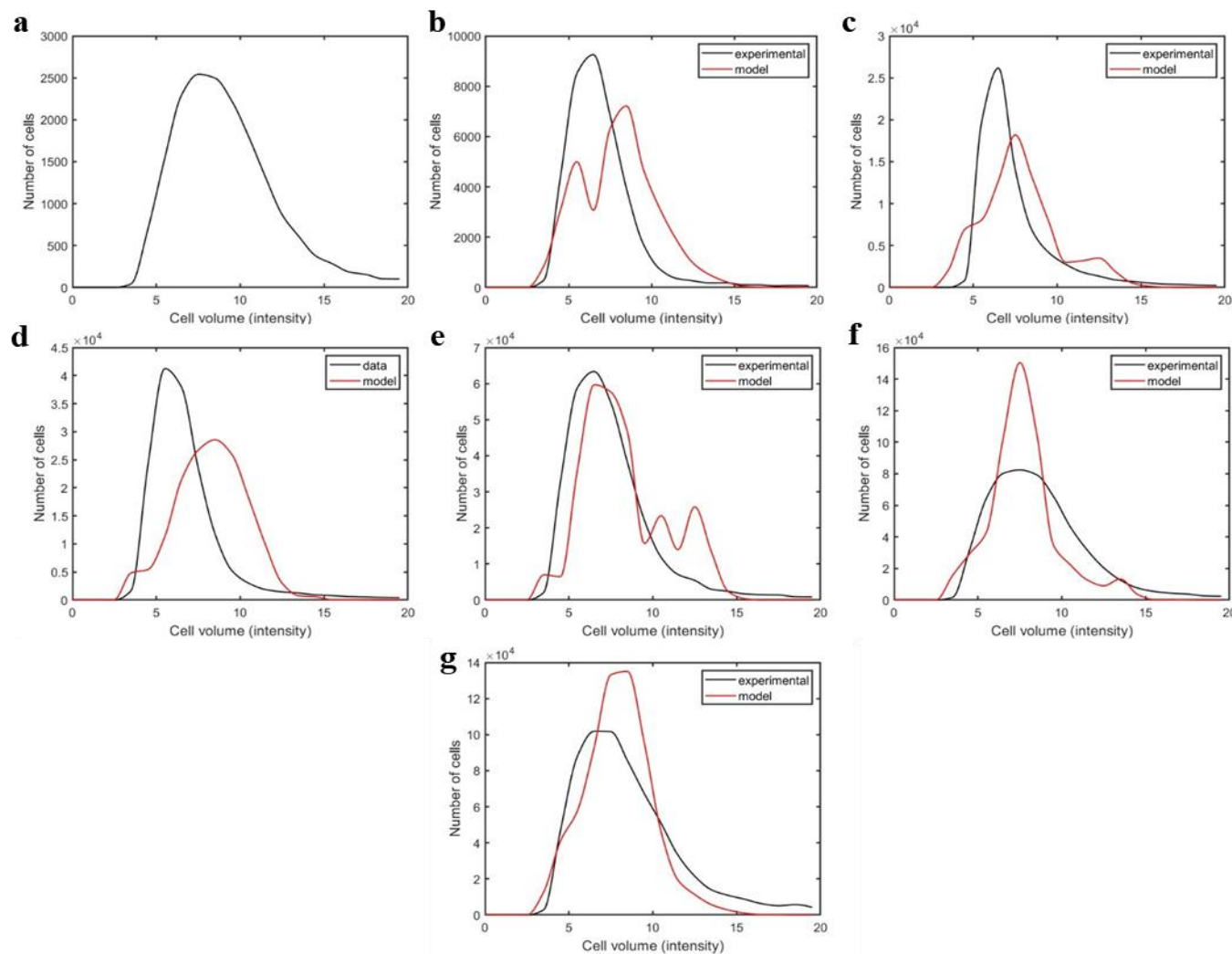


Figure 5-21: Model validation and experimental flow cytometry distributions of cell volume for a *B. pertussis* shake flask culture with 8.6 g/L glutamate in the media and starting optical density of 0.14 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

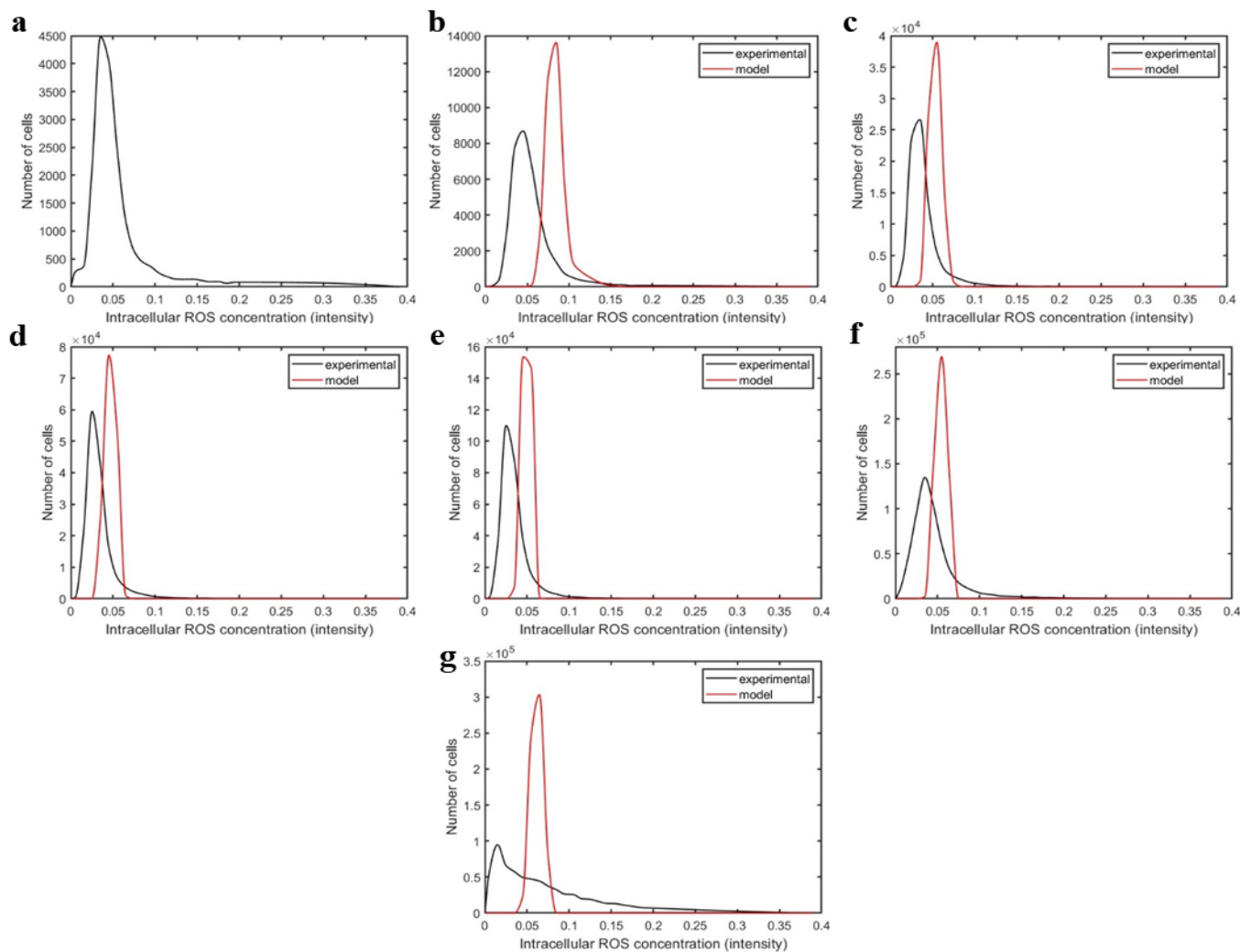


Figure 5-22: Model validation and experimental flow cytometry distributions of intracellular ROS concentration for a *B. pertussis* shake flask culture with 8.4 g/L glutamate in the media and starting optical density of 0.14 at (a) inoculation and after (b) 6, (c) 12, (d) 18, (e) 24, (f) 30 and (g) 36 hours of incubation.

In order to compare the 2-D experimental and model distributions, the Kullback-Liebler divergence was calculated using Eqn (5.27) at each time interval to yield the values shown in Table 5-7. The more similar are the predicted and measured distributions, the closer the Kullback-Liebler divergence will be to 0. According to these results, the fit of the model is better when the glutamate concentration is high (i.e., condition 2). Also, the Kullback-Liebler divergence tends to be highest when the cells reach the stationary phase. This observation is consistent with those that were made previously regarding the ROS concentration distributions (Figures 5-16, 5-18, 5-20 and 5-22). The contour distributions are presented in Figures 5-23 – 5-24. The contours for the remaining conditions can be found in Appendix B. It is also apparent from these figures that the model-predicted distributions of ROS concentration are much narrower than the experimental ones.

Table 5-7: Kullback-Liebler divergence based on the 2D cell volume and intracellular ROS distributions at all time points.

time (hr)	Kullback-Liebler Divergence			
	Condition 1	Condition 2	Condition 3	Condition 4
6	0.874	0.796	1.564	2.053
12	0.719	0.607	1.028	1.811
18	0.889	0.994	1.087	1.897
24	0.738	1.057	1.293	1.722
30	1.081	0.909	1.032	1.332
36	1.985	1.026	1.489	2.275

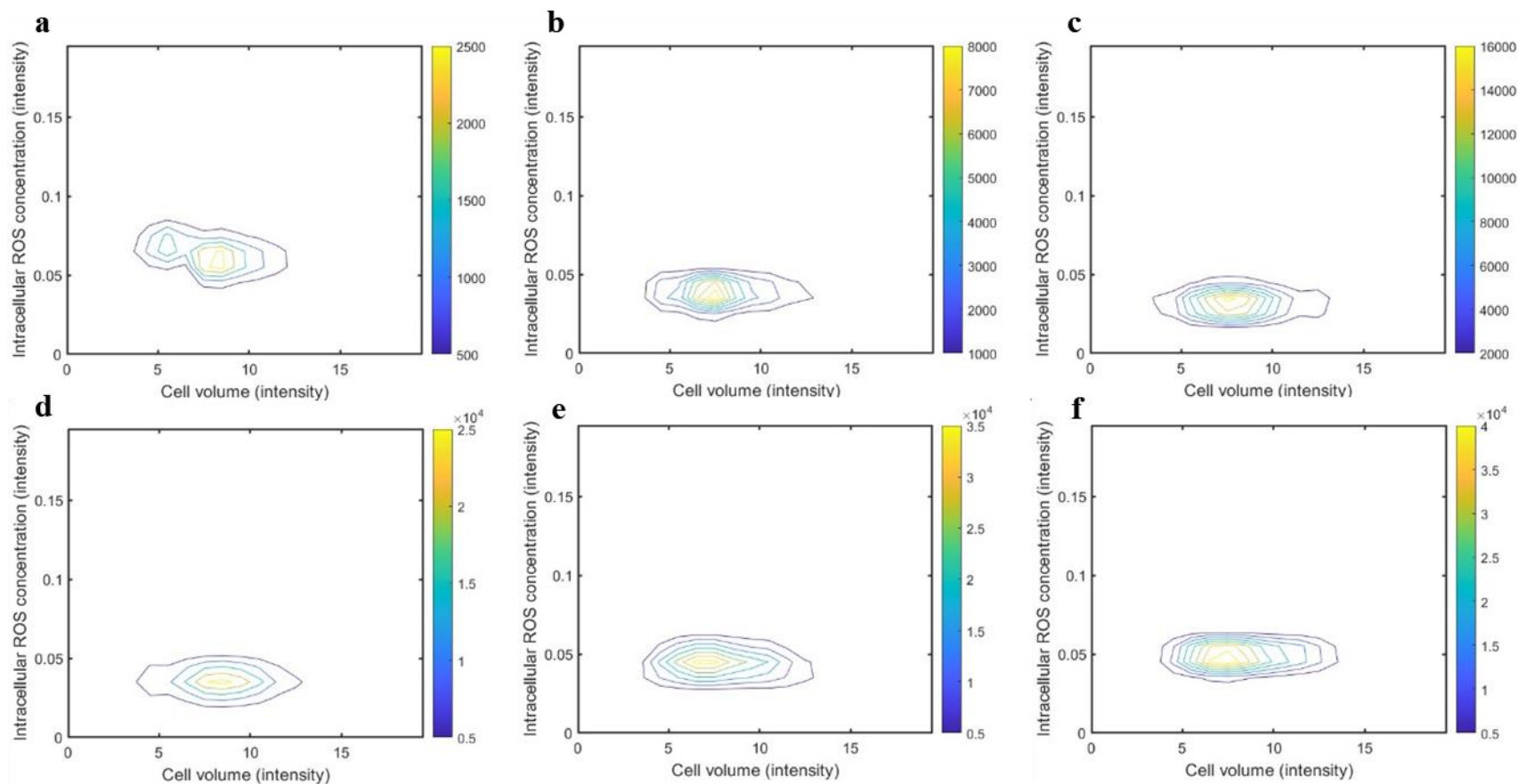


Figure 5-23: Model flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.25 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

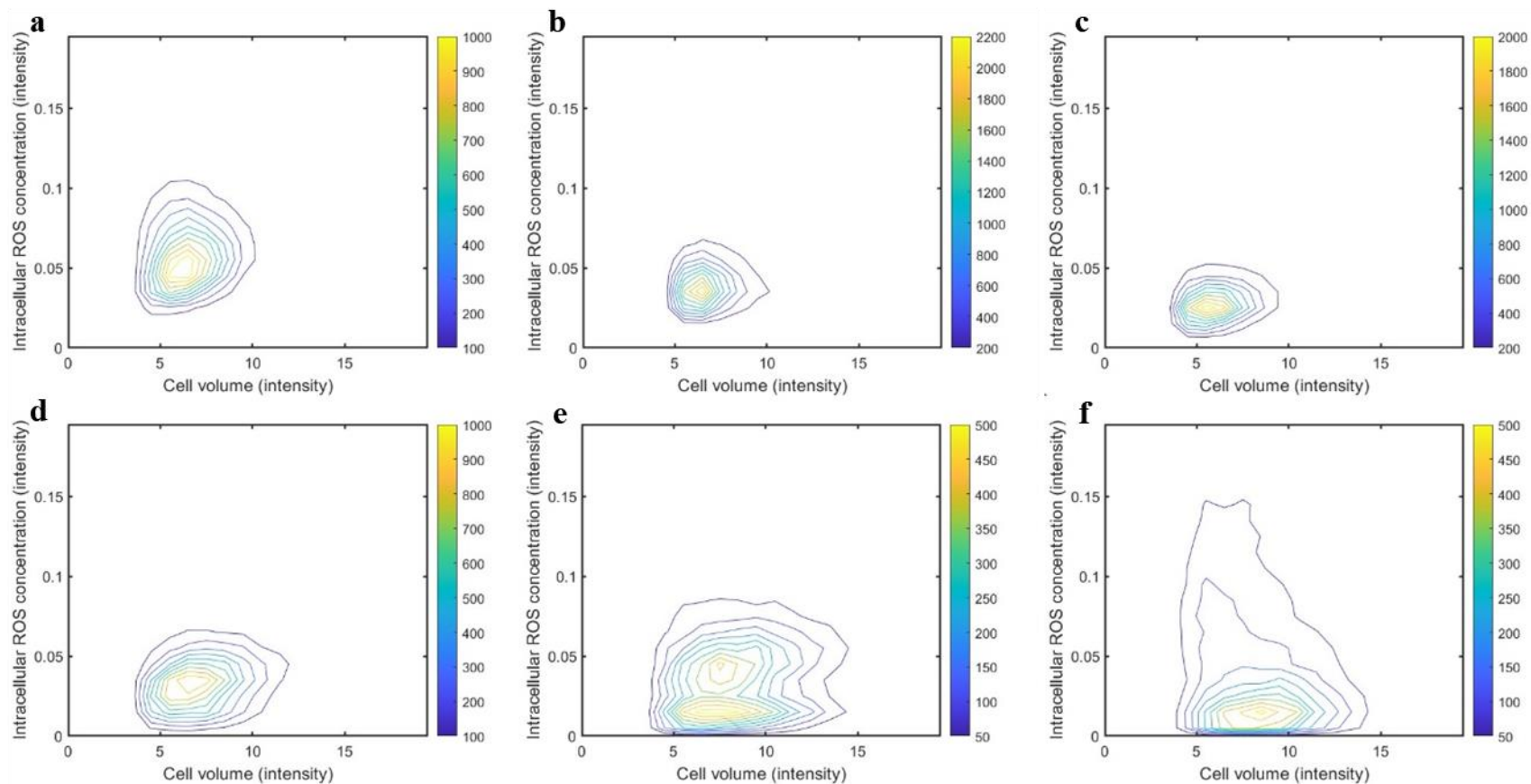


Figure 5-24: Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.25 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

5.4.5 Comparison of bulk and population balance models

The bulk oxidative stress model presented in Chapter 4 and the PBM are based on the same kinetics and therefore contain the same model parameters to be calibrated for fitting. Table 5-8 shows a comparison of the model parameter values obtained from fitting of the two models to the data. The additional parameters in the PBM are used to shape the probability of the partitioning rate. The partitioning rate controls the cell volume, but the cell volume has no effect on the concentration of intracellular species (Figure 5-5). Also, the assumption of equal partitioning means that the concentrations of the various species in the resulting daughter cells do not change from those in the dividing mother cells. Some of the parameter values are very similar in the two models, including K , k_m , $Y_{x/s}$, and d_B . The largest parameter change between the models is in the case of the maximum growth rate v_{max} . The PBM makes use of growth rates for individual cell volumes while the bulk model employs the growth rate for the entire biomass in the flask. The initial intracellular concentration of NADPH resulting from the fitting of the PBM to the data is much closer to the literature ranges of NADPH as compared to the value obtained for the bulk model (Goldbeck *et al.*, 2018; Shen *et al.*, 2021).

Table 5-8: Parameter estimates of bulk and PBM oxidative stress models

Parameter	Bulk	PBM
v_{max}	$16.0 \frac{mM}{hr}$	0.482
K	$1.72 mM$	$1.70 mM$
K_t	$1.73 mM$	$2.32 mM$
k_p	$0.536 \frac{1}{mM hr}$	$0.479 \frac{1}{mM hr}$
k_m	$0.001 \frac{1}{hr}$	$0.001 \frac{1}{hr}$

$Y_{x/s}$	$0.643 \frac{\text{intensity biomass}}{\text{g glutamate}}$	$0.620 \frac{\text{intensity biomass}}{\text{g glutamate}}$
K_{si}	4.99 mM	6.72 mM
d_B	$0.109 \frac{1}{\text{hr}}$	$0.109 \frac{1}{\text{hr}}$
A_o (secondary)	$[0.5\text{-}0.787] \text{ mM}$	0.278 mM

The major advantage of using a PBM is that it accounts for the distributions and can predict the heterogeneity of the cell population with respect to experimental conditions that are not used in bulk models. It can be shown mathematically that no differences can exist between the mean values calculated from a PBM and a lumped (bulk) model when a linear model is considered. On the other hand, since the kinetic equations of the model used in this study are highly nonlinear, the use of the distributions for cell size and ROS concentration will affect the mean intracellular ROS concentrations. Also, the distribution of ROS concentrations provides a stress level each cell has to combat. Cells that are more stressed will use NADPH to quench ROS while cells with a lower stress level will use NADPH to grow. Table 5-9 and Figures 5-25 – 5-28 compare the fitting and validation of the bulk and PB models. While a small difference between the models is observed in their abilities to predict the biomass and extracellular glutamate concentrations, the PBM provides much better predictions of the intracellular ROS concentration. The superior ability to predict intracellular ROS is clearly shown in Figures 5-31c – 5-34c. This result also confirms that the fitting of the predicted averages to experimental data is significantly improved by accounting for the distributed nature of a variable, given the nonlinearity of the model.

This result is significant since ROS is known to cause protein damage and affect synthesis rates. Thus, we hypothesize that similar improvement to the ones observed for intracellular ROS will be found when using distributed data to predict other intracellular quantities such as individual

cell antigen production which is of particular importance for manufacturing. At this point, the exact relation of intracellular ROS to protein production is not known. Future studies should identify this relation which will permit the proposed PBM to be extended to predict antigen productivity.

Table 5-9: Comparison of RMSE between bulk oxidative stress and coupled PB models

Model	OD₆₀₀ (intensity)		Extracellular glutamate concentration (g/L)		Average intracellular ROS concentration (intensity)		AIC
	RMSEC	RMSEV	RMSEC	RMSEV	RMSEC	RMSEV	
Bulk	0.295	0.316	0.536	1.014	0.029	0.050	-260.2
PBM	0.337	0.534	0.509	0.623	0.008	0.015	-440.3

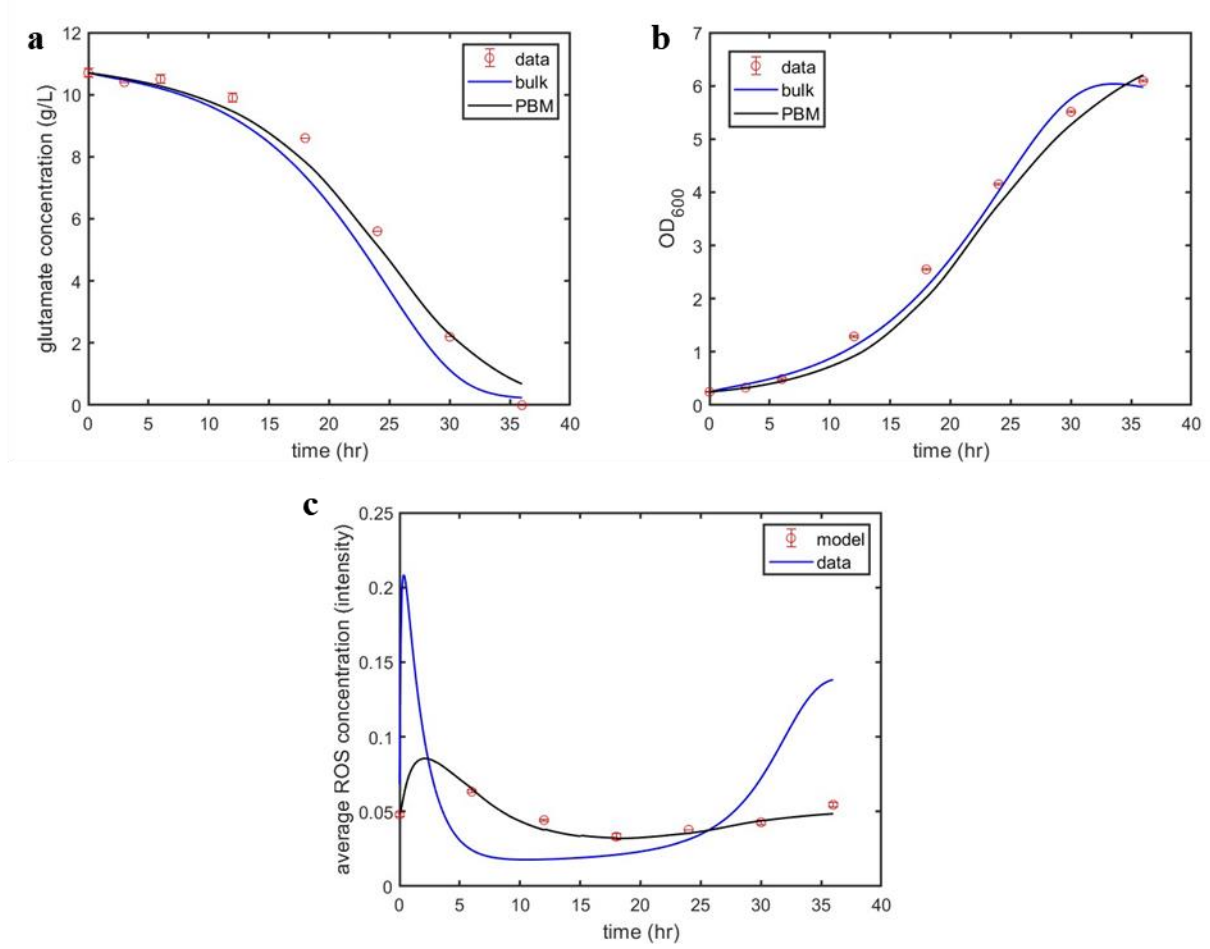


Figure 5-25: Comparison in model fitting between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a *B. pertussis* culture with an initial OD of 0.25 and initial glutamate concentration of 10.1 g/L.

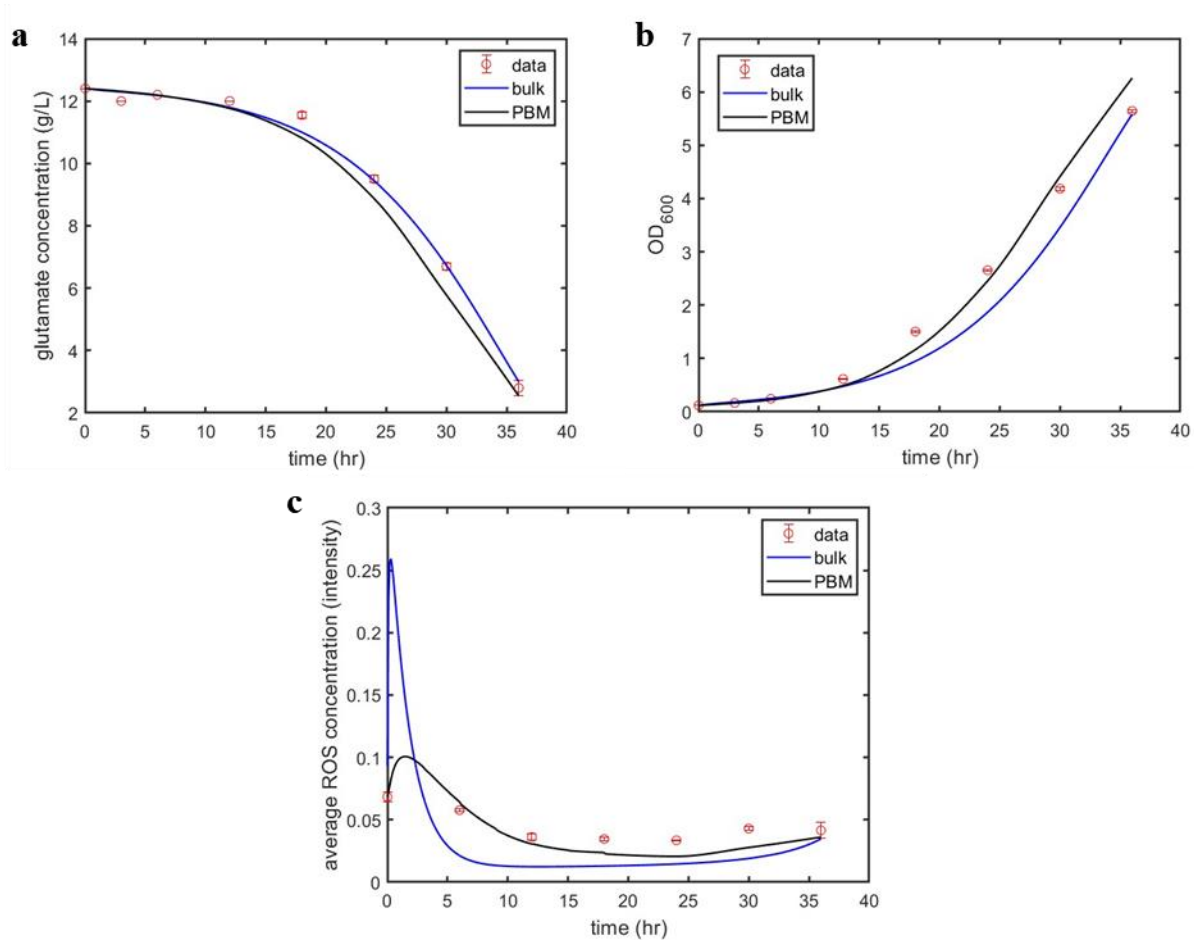


Figure 5-26: Comparison in model fitting between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a *B. pertussis* culture with an initial OD of 0.12 and initial glutamate concentration of 12.4 g/L.

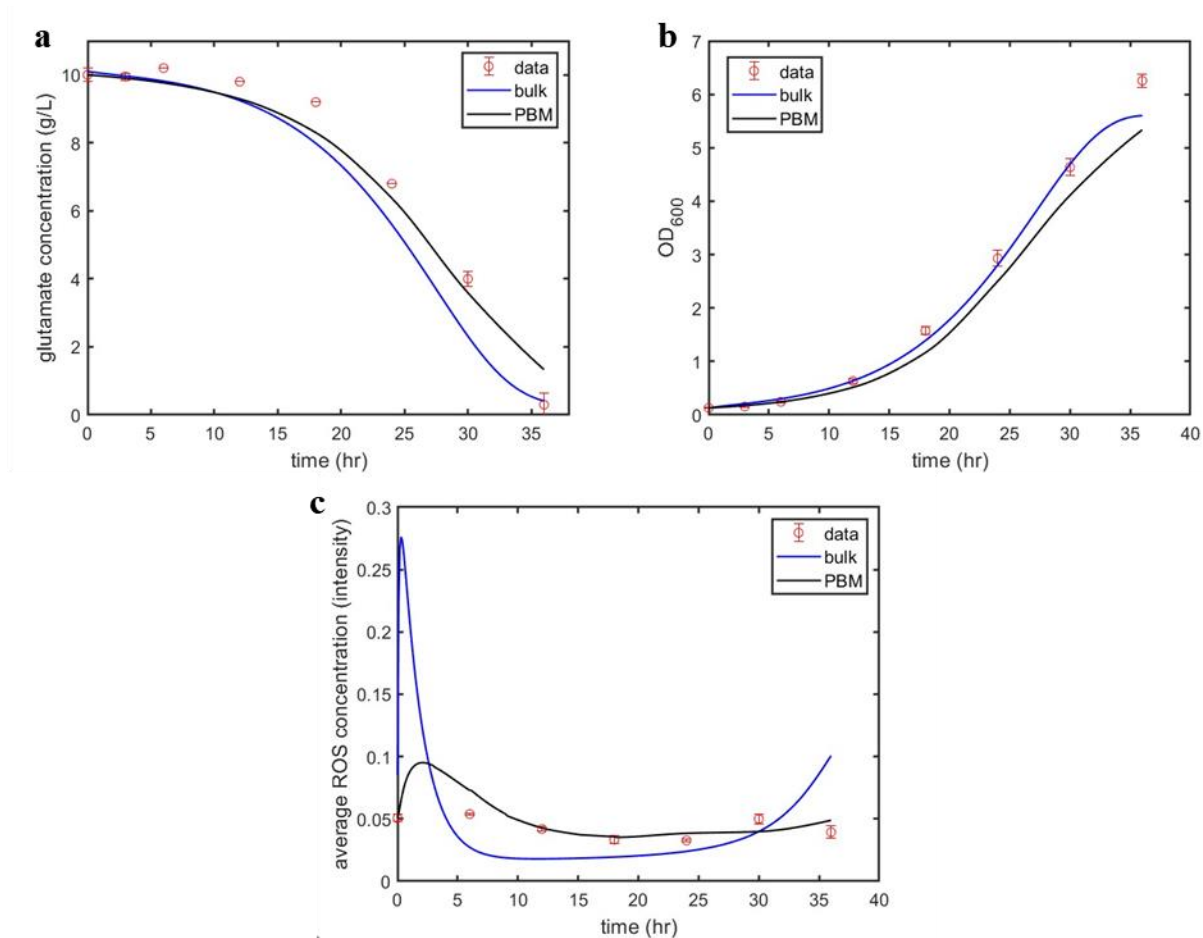


Figure 5-27: Comparison in model fitting between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a *B. pertussis* culture with an initial OD of 0.13 and initial glutamate concentration of 10.1 g/L.

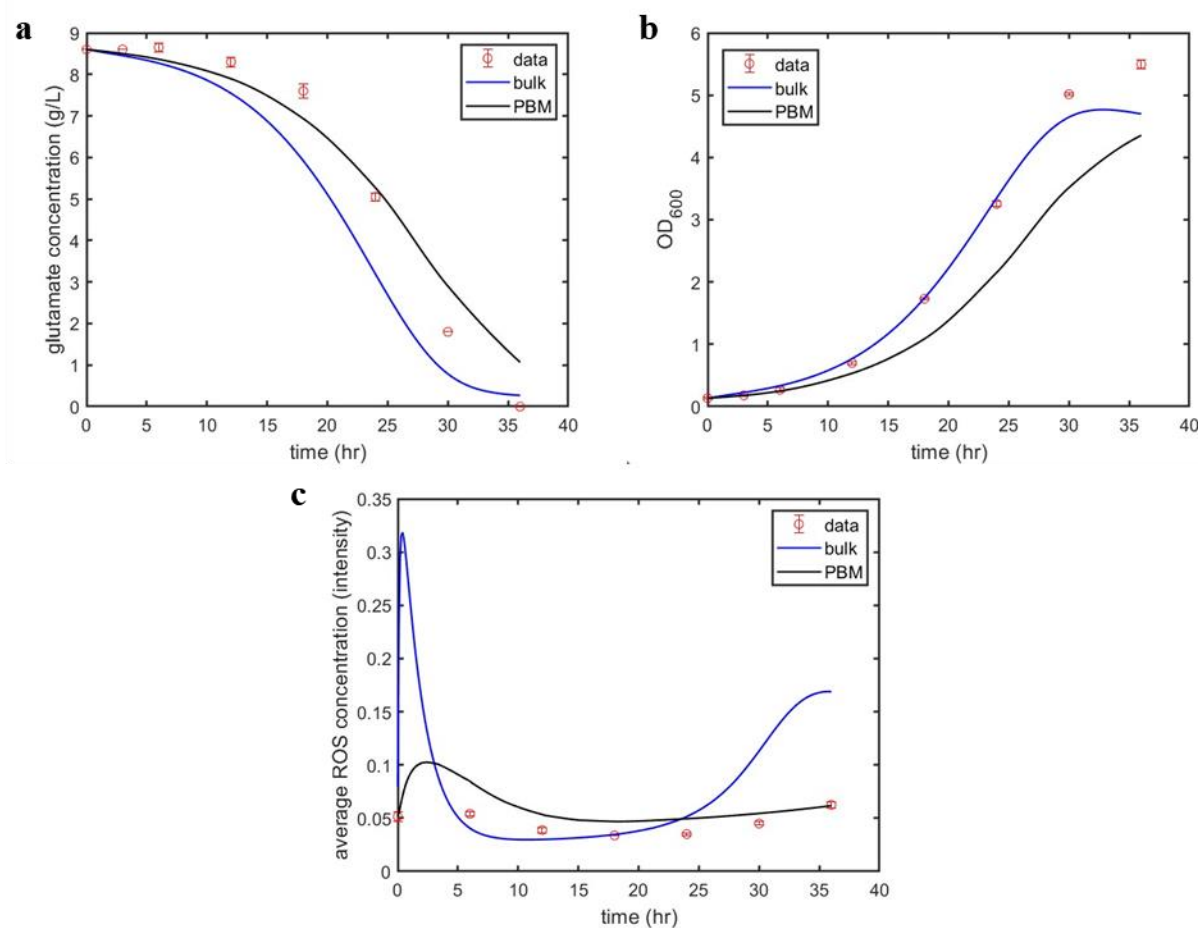


Figure 5-28: Comparison in model validation between bulk and PBM of (a) optical density (OD), (b) extracellular glutamate concentration and (c) average intracellular ROS for a *B. pertussis* culture with an initial OD of 0.14 and initial glutamate concentration of 8.6 g/L.

5.4.6 PBM and train of reactors

It should be recalled that phenomena observed in low pertactin-producing fermentations include an increase in the concentration of ROS at the beginning of the culture (Zavatti, 2017) and low OD₆₀₀ (Agarwal, 2022). One of the hypotheses behind the research of the current thesis is that the higher initial glutamate concentrations combined with the heterogeneity of the inoculum may cause adverse growth conditions that will propagate and be amplified along the train of bioreactors

used in the vaccine manufacturing process. Given that the population balance model can track the heterogeneity of the culture started with a heterogeneous inoculum, the proposed model can serve as a test of our original hypothesis. Similar to section 4.4.9 in Chapter 4, the coupled population balance model was used to replicate the effect that the initial glutamate concentration in the media would have on a train of reactors similar to the production process at Sanofi Pasteur. This is intended only as a qualitative study to gain some insight into the possible variability in the production process and the way in which a change in media composition propagates through the train of reactors.

The media composition can vary batch-to-batch, as discussed in Chapter 4.4.9. Accordingly, simulations are conducted for cases where the initial glutamate concentrations are 10 g/L and 10.5 g/L and the resulting biomass and glutamate concentrations in the train of reactors process are determined. As shown in Figure 5-29, the findings using the PBM are similar to those of the bulk model in section 4.4.9 in Chapter 4. When the media initially contains 10.5 g/L glutamate, the amount of biomass produced is lower in the first reactors of the train; however, by the time the discharge of the final 2000 L bioreactor is reached, the difference between the biomass levels produced under the two conditions is very small. Thus, the final biomass yield obtained by the train will be very similar for both media. This result is consistent with the observations at the Sanofi facility where the final OD is found not to vary significantly from batch-to-batch.

It should also be recalled that productivity is expected to be correlated to the area under the growth curve rather than the final biomass since the antigen is continuously secreted by the cells (Agarwal, 2022). The coupled PBM also enables a comparison of the distributions of cell volume and intracellular ROS concentration to be made when the train is fed by the two media (Figure 5-30). The ROS concentration at the end of the 200 and 2000 L reactors is higher in the train fed

with a media containing 10 g/L glutamate because less glutamate is available at the end of these fermentations and ROS has been found in our experiments to be generated under starvation conditions.

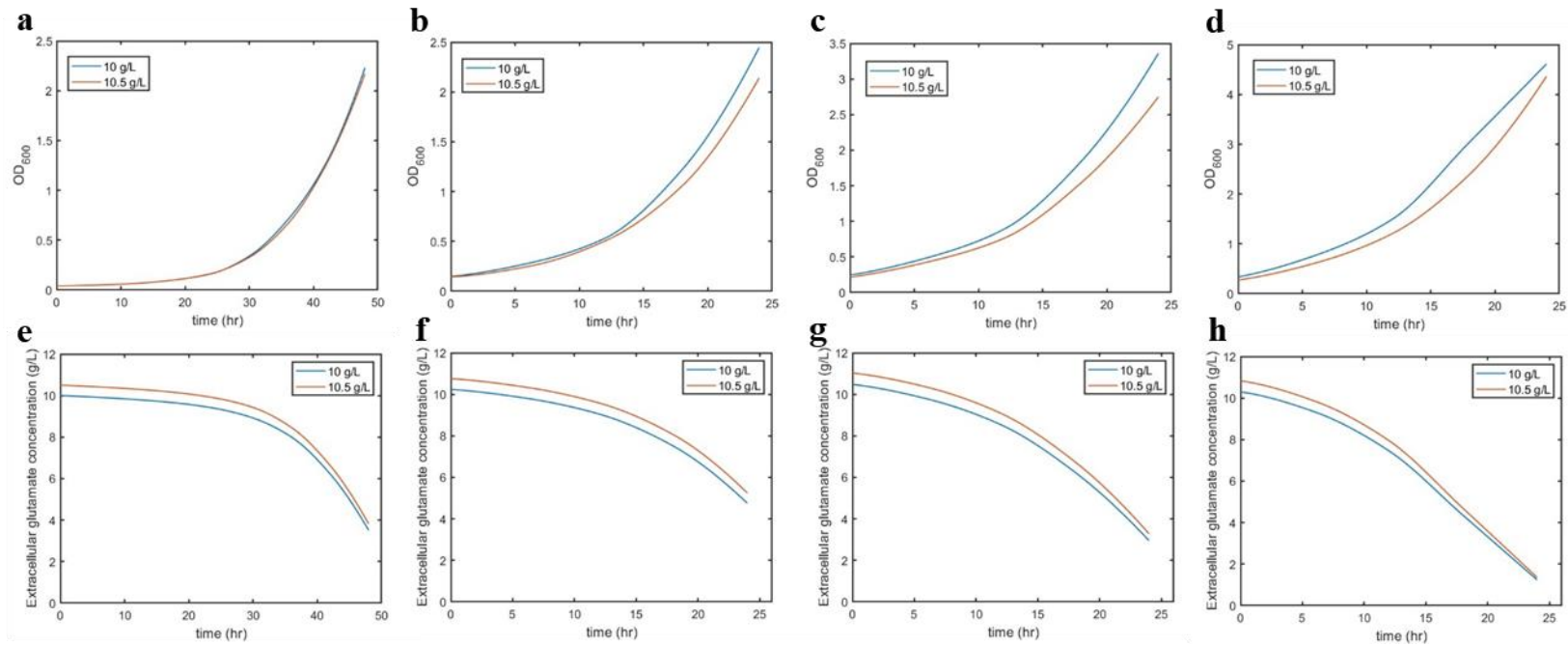


Figure 5-29: Population balance model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for *B. pertussis* cultures with media of 10 g/L and 10.5 g/L glutamate. Evolution of biomass coming from a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor. Evolution of glutamate concentration coming from e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor.

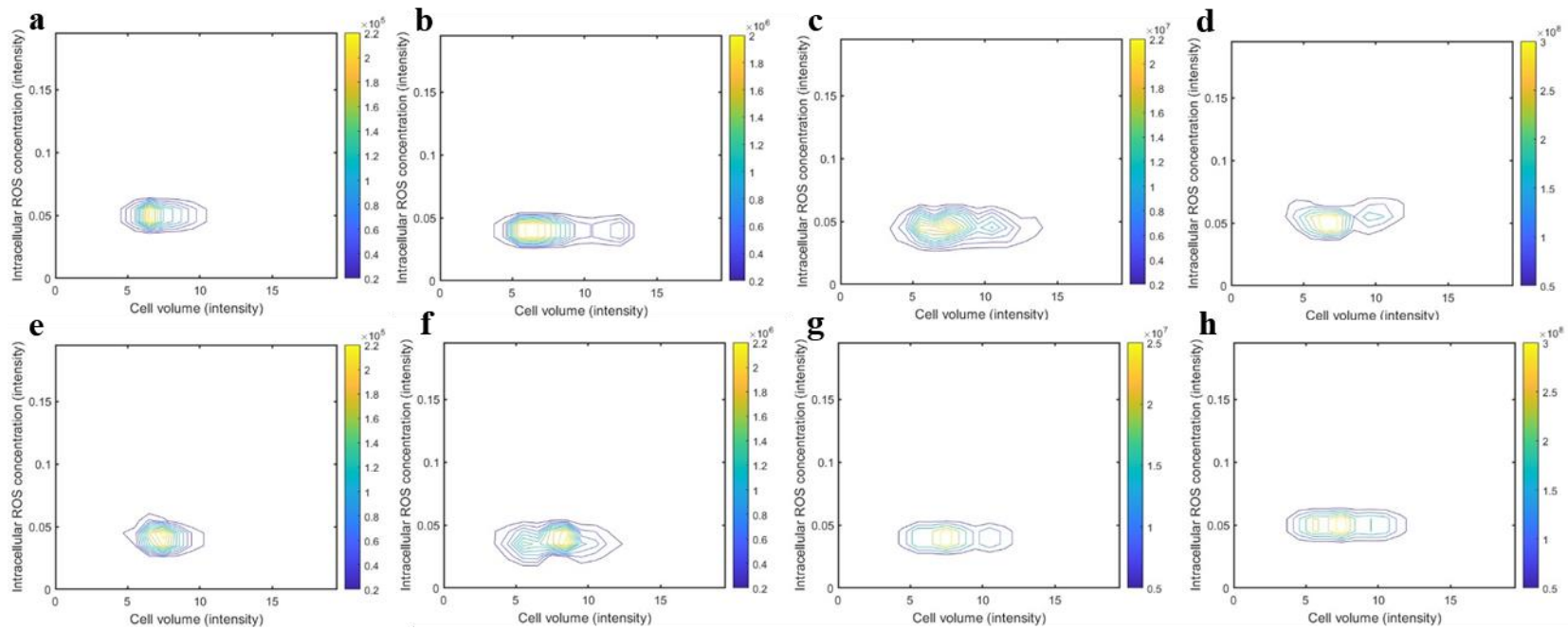


Figure 5-30: Population balance model prediction of Sanofi process (train of seed flask and three increasing sized reactors) for *B. pertussis* cultures with media of 10 g/L and 10.5 g/L glutamate. Distributions of cell volume and intracellular ROS concentration coming from a) seed flask, b) 20 L bioreactor, c) 200 L bioreactor and d) 2000 L bioreactor for a media with 10 g/L glutamate. Distributions of cell volume and intracellular ROS concentration coming from e) seed flask, f) 20 L bioreactor, g) 200 L bioreactor and h) 2000 L bioreactor for a media with 10.5 g/L glutamate.

5.6 Conclusions

A population balance model coupled to the metabolic model that uses oxidative stress presented in Chapter 4 was used to predict the growth of *B. pertussis*. The population balance model was solved with the method of characteristics and re-binning the distributions from the model after each time step was crucial to maintain numerical resolution over time in the range of cell volumes that was relevant for our experiments. The cell volumes were found to be lower in the exponential growth phase and higher in the lag and stationary phases. During the lag phase, cells increase in size and their metabolism is targeted to prepare the molecules necessary for replication; the increase in size during the stationary phase is due to an increase in pH because of the accumulation of secreted alkaline molecules. The model was calibrated and validated using measured OD, glutamate concentration and flow cytometry distributions for cell volume and intracellular ROS concentration to predict the outcomes of flask fermentations. The conditions analyzed in these experiments cover a range of initial glutamate and ROS concentrations. The coupled PBM model was able to outperform the bulk oxidative stress model with regard to predicting the intracellular ROS concentration. This is due to the nonlinearity of the kinetic equations which causes the distributions of intracellular properties to affect the evolution of the averages of the distributions. A population balance model may also be better suited to predict intracellular components such as antigen concentration which may also vary from cell to cell due to the correlation of the productivity to ROS or other intracellular properties.

Chapter 6 Flow Cytometry Sorting of *B. pertussis*

6.1 Introduction

Whooping cough is a highly contagious respiratory tract infection that can be prevented through vaccination at an early age. The vaccine produced by Sanofi Pasteur to treat whooping cough is an acellular vaccine which contains 5 antigens components: fimbriae (types 2 and 3), filamentous hemagglutinin, pertactin and pertussis toxin. The bottleneck in producing the vaccine is the antigen pertactin which is produced in small amounts compared to the other antigens. Also, the company is only able to collect a fraction of the pertactin generated by the cells since their patented process only allows for the recovery of antigen secreted into the supernatant, but none of the intracellular content. The yield of the antigens from the train of reactors in production can be highly variable. Based on flow cytometry based findings of a previous study conducted with samples from the manufacturing process (Zavatti *et al.*, 2019) the final yield of pertactin was found to be correlated mostly to the upstream part of the process. In particular, two factors measured via flow cytometry were found to contribute to the low pertactin yield batches: high oxidative stress and high cell complexity (side scattering) at beginning of the fermentations. Zavatti found the scattering to be correlated to oxidative stress and thus hypothesized that stress of the cell populations resulted in slower growth and lower productivity. The correlation between cell complexity (side scattering) and size (front scattering) with oxidative stress have been also reported for other bacteria (Zavatti *et al.*, 2020; Baatout *et al.*, 2006).

Flow cytometry is a sensitive and powerful tool that uses fluorescence and light scattering to simultaneously measure the physical and chemical properties of cells suspended in a fluid. Cytometry is a high throughput process that characterizes individual cells at very high rates and thus can detect rare events in a cell population. Flow cytometry sorting is an extension of flow cytometry that enables cells to be selected based on any property measured via flow cytometry. The actual sorting operation occurs downstream from the detectors that measure the light scattering and fluorescence energy.

One of the hypotheses of the current work was that productivity can be enhanced by sorting a population of highly producing cells followed by cultivation of the sorted population. Typically, to find high producers of cellular products, a search can be done for valuable mutants from an artificially induced subpopulation based on different methods such as random mutagenesis and nuclease genome editing (Copp *et al.*, 2014; Abuaita and Withey, 2011; Ramesh *et al.*, 2015; Lonowski *et al.*, 2017; Michener and Smolke, 2012). However, if a sorted *B. pertussis* cell was found to be a high producer of pertactin, a lengthy validation of the manufacturing process would have to be done in order for the corresponding cell line to be used in production. This process would include clinical trials where the manufacturing company would not see any profits from their vaccine for a long period of time. Instead, if a population of high producers could be effectively sorted from the current fermentation process, it could be used to inoculate the first fermenter in the train of bioreactors to enhance productivity. Naturally, the success of such strategy will strongly depend on whether the epigenetic changes are conserved after re-culturing the sorted cells.

Epigenetics changes include any process that alters gene activity without changing the DNA sequence and leads to modifications that can be transmitted to daughter cells (Weinhold,

2006). In *B. pertussis*, the BvgAS sensory transduction system mediates an alteration between the Bvg⁺ phase, characterized by expression of adhesins and toxins, and the Bvg⁻ phase, where virulence factors are not expressed (Martinez de Tejada *et al.*, 1998; Moon *et al.*, 2017). In the manufacturing process at Sanofi, the last reactor in the train (2000 L reactor) is first run as a batch process and then changed to fed-batch when a rise in the dissolved oxygen level is observed followed by a decrease in airflow and agitation (representing low glutamate concentration in the media). In a study at Sanofi Pasteur, it was determined that the surface antibody concentration of pertactin is much higher during the batch phase than during the fed-batch phase due to a change from a Bvg⁺ to Bvg⁻ phase (data not shown). However, not all *B. pertussis* cells in the population express a low surface concentration of pertactin in the fed-batch stage of production which is hypothesized to be caused by epigenetics.

This chapter presents a flow cytometry protocol to sort *B. pertussis* based on side scattering and surface concentration of pertactin. To assess whether the epigenetic phenotype is maintained in the daughter cells of the sorted populations, the sorted cells were re-cultured and monitored to verify the phenotype. A flask study was completed to measure biomass, ROS, and surface antigen concentration by spectrophotometry and flow cytometry to compare the sorted cultures against an unsorted control. The main goal of the sorting protocol is to gain insight into whether the properties enhanced by sorting, e.g., productivity changes due to switch of Bvg⁺ and Bvg⁻ phases, can be maintained post-sorting and be used for re-culturing.

The remainder of this chapter is organized as follows. Section 2 presents the flow cytometry sorting protocol and experimental procedures used. Section 3 presents the experimental results including the sorting of high/ low side scattering and pertactin surface concentration and the shake flask experiments of the re-cultured sorted populations. Section 4 discusses the potential

improvements of sorting high pertactin producers of *B. pertussis* and other possible applications of using flow cytometry sorting in the manufacturing process. Section 5 presents conclusions.

6.2 Materials and Methods

6.2.1 Streptomycin resistant *B. pertussis* seed

Since the cell sorting process is essentially open to the environment, this places an inherent risk of contamination of the sorted population. Initial experiments in our research showed a high level of contamination. Thus, to avoid contamination during the sorting protocol, it is common practice to administer an antibiotic to the seed population (Abuaita and Withey, 2011; Li *et al.*, 2018; Sun *et al.*, 2016). Based on this, we mutated *B. pertussis* to have streptomycin resistance. It should be noted that the seed at Sanofi Pasteur previously exhibited this antibiotic resistance, but the mutation led to a loss in this resistance.

The following protocol was developed to create a streptomycin-resistant seed:

- 1) *B. pertussis* was plated on a Bordet-Gengou agar plate with 50 µg/mL of streptomycin. The plate was then inserted in an incubator at 36 °C for 4 days. Visible colonies appeared on the plate on the fourth day which are resistant to streptomycin while those that did not developed resistance were killed.
- 2) A surviving colony from step 1 was cloned (re-plated) onto another Bordet-Gengou agar plate with 50 µg/mL streptomycin. The plate was inserted in an incubator at 36 °C for 2 days.

- 3) Clones from step 2 were inoculated in a 250 mL flask with 50 mL of culture media (Table 3-1) containing streptomycin (50 µg/mL). The contents of the flask were then grown for 36 hours.
- 4) 1.5 mL of the culture obtained from the flask were added to 1 mL of a 40% glycerol solution to generate the seeds. The seeds were then placed in a freezer for storage at -80 °C.

6.2.2 Flow cytometry sorting protocol

B. pertussis cells were sorted into cultured media based on the side scattering and pertactin surface concentration using an S3e Cell Sorter (Bio Rad). A total of 100,000 cells were sorted in per sample in purity mode at a rate of 500 events per second.

6.2.3 Re-Culture Protocol

B. pertussis is sensitive to the sorting procedure and does not readily recover to grow in liquid media. Since *B. pertussis* grows better on solid media, we have devised the following procedure to re-culture the cells:

- 1) Sorted cells were plated on a Bordet-Gengou agar plate with 50 µg/mL streptomycin with an inoculation needle. The plates were placed in an incubator at 36 °C until colonies were observed after approximately 2 days.
- 2) Colonies were inoculated in a 50 mL flask with working volume of 5 mL containing streptomycin (50 µg/mL). After the contents of the flask were grown for 48 hours, samples were collected for OD and flow cytometry measurements.
- 3) Secondary flasks were prepared using the culture in step 2 to measure growth at various times along the culture. *B. pertussis* fermentations were carried out in 250 mL flasks

(polycarbonate, sterile, with a vented cap having 0.22 μm pore-size PTFE membrane from VWR, Canada) with a working volume of 60 mL by incubation at 36°C and 200 rpm and a starting optical density of 0.25 – 0.3.

6.2.4 Sample Measurements

A detailed method for these measurements can be found in Chapter 3. All measurements (OD, ROS and pertactin surface concentration) were completed in duplicate. Below are summaries of the experimental methods used.

6.2.4.1 OD₆₀₀

To measure OD, the sample was diluted by a factor of 20 in 0.9% saline solution and OD was obtained using optical spectrophotometry by measurement of the absorbance at 600 nm.

6.2.4.2 ROS Concentration

Oxidative stress was measured using the fluorescein derivative carboxy-2',7'-dichlorodihydrofluorescein diacetate (carboxy-H₂DCFDA) from Sigma-Aldrich (Canada). Samples with an OD of 0.05 were prepared in PBS and carboxy-H₂DCFDA was added to the samples to obtain a final concentration of 50 μM . Flow cytometry was conducted using an S3e Cell Sorter (Bio Rad). To measure the ROS distribution, the FL1 diode was used for DCF emission (green fluorescence) which fluoresces at an excitation/ emission pair of 503/523 nm. A total of 20,000 events per sample were collected at a rate of 1000 events per second.

6.2.4.3 Pertactin surface concentration

To measure the surface concentration of pertactin, an antibody provided by Sanofi (PRN 3-16) was conjugated to fluorescein (FITC) using a FITC conjugation kit (Abcam, Canada). 1 mL

samples with an OD of 0.05 were prepared and centrifuged at 4,000 G for 4 minutes and the supernatant discarded. 25 µL of the antibody solution (containing 1µg/mL pertactin antibody) was then added to each tube and incubated for 1 hour in the dark at 4°C. Then, 1 mL 1xPBS was added to each tube and centrifuged at 4,000 G for 4 minutes to wash the sample. The supernatant was discarded and the pellet re-suspended in 1 mL PBS. Along with the PRN 3-16 antibody, a negative control was also conjugated to FITC using the same method. This antibody does not attach to *B. pertussis* and is used to check for nonspecific binding of the conjugation. To measure the surface antibody concentration distribution across the cell population (of both negative control and pertactin antibodies), the FL1 diode was used for FITC emission (green fluorescence) which fluoresces at an excitation/ emission pair of 491/516 nm. A total of 20,000 events per sample were collected at a rate of 1000 events per second.

6.3 Results

6.3.1 Oxidative stress of antibiotic resistant strain of *B. pertussis*

Streptomycin resistance was necessary to eliminate the risk of contamination in the sorted population of *B. pertussis*. However, the addition of streptomycin in the media increases the oxidative stress on the culture at inoculation as shown in Figure 6-1a. The average ROS concentration intensity is increased by 324% in the streptomycin resistant (strepR) strain with 50 ug/mL of streptomycin in the media compared to the control seeds without antibiotic in the media. Antibiotics have been shown to cause oxidative stress (Guillouzo and Guguen-Guillouzo, 2020; Grant and Hung, 2013) and specifically streptomycin has been shown to cause oxidative stress in *E. coli* (Goswami *et al.*, 2007). In the presence of streptomycin, the strepR strain is able to recover

from the oxidative stress but exhibits a longer lag phase compared to the control strain (Figure 6-1b). Our original plan was to re-culture a cell population that had been sorted according to its low ROS concentration. However, because of the observed increase in ROS concentration at inoculation caused by the presence of streptomycin in the media, ROS concentration-based cell sorting was not conducted as initially planned.

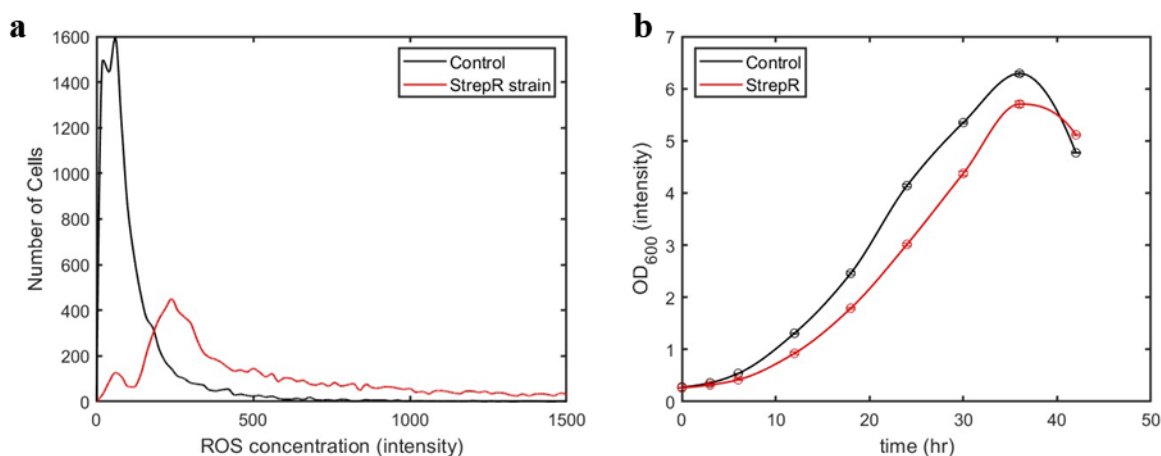


Figure 6-1: Comparison between control and streptomycin-resistant (strepR) strains of *B. pertussis*: (a) number distribution of initial ROS concentration and (b) evolution of growth of secondary cultures.

6.3.2 Flow cytometry sorting

A primary flask of *B. pertussis* was grown until the stationary phase where the population evolved into a bimodal distribution according to PRN surface concentration (Figure 6-2b). Flow cytometry sorting was used to sort *B. pertussis* cells with high and low intensity of side scattering (Figure 6-2a) and PRN antibody surface concentration (Figure 6-2b). The ranges and averages of each gate are found in Figure 6-2 and shown in Table 6-1.

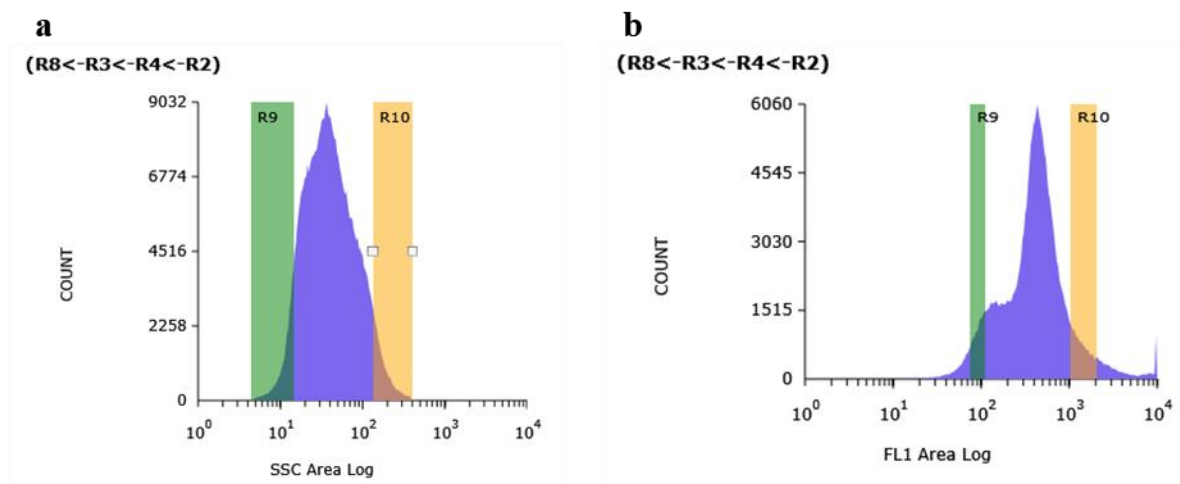


Figure 6-2: Sorting gates of *B. pertussis* for (a) side scattering (R9 – low SSC, R10 – high SSC) and (b) PRN antibody concentration (R9 – low PRN, R10 – high PRN).

Table 6-1: Flow cytometry sorting population averages and ranges for side scattering and PRN surface concentration.

Sort Property	Low Intensity Population		High Intensity Population	
	Average	Range	Average	Range
Side scattering	11.49	[3.2 – 10.6]	183.70	[138 – 406]
PRN surface concentration	94.28	[68 – 108]	1418.74	[1015 – 2022]

After sorting the populations according to side scattering and PRN surface concentration, the sorted populations were plated on TSA and BG agar (with 50 ug/mL streptomycin) plates to check for contamination. Figure 6-3 depicts the TSA plate of the sorted population of low PRN surface concentration cells grown for 7 days. The TSA plate shows no growth of colonies which corroborates that the sorted cells are *B. pertussis* only and that none of the sorted populations has been contaminated by other type of cells.



Figure 6-3: Tryptic soy agar plated with low PRN surface concentration population.

6.3.3 Epigenetic effects of re-cultured cells

The goal of sorting cells based on side scattering and PRN surface concentration was to determine whether the changes in the property that served as the basis for sorting would be maintained when the sorted cells were cultured in agar plates. After approximately 2 days on the BG agar plate, the sorted cells formed colonies. Colonies were taken from the plate and tested for side scattering and PRN surface concentration to determine whether the sorted properties (i.e., side scattering and PRN surface concentration) were maintained in culturing the sorted cells. Figure 6-4 shows the distributions of the re-cultured populations for side scattering and PRN surface concentration and the averages for these distributions are found in Table 6-2.

No observable differences in the side scattering distributions are evident in all the sorted populations (low SSC, high SSC, low PRN and high PRN) as shown in Figure 6-4a. Side scattering is a measure of the complexity and granularity of cells and as changes in this property have been generally associated with the status of the cell along the cell cycle (Papa *et al.*, 1991; Gant *et al.*, 1993). The observed small differences in the distributions of side scattering for the sorted populations can be explained since the selected cells were plated at the same time and exposed to media for the same duration.

The PRN surface concentration is also expected to be somewhat related to the growth phase in which *B. pertussis* is found. It has been reported that cell expresses PRN in the Bvg+ phase since it is used in adhesion while PRN is not expressed in the Bvg- phase (Martinez de Tejada *et al.*, 1998; Moon *et al.*, 2017). It has been shown that spontaneous mutations of the Bvg system are the cause of Bvg- mutants which out-compete other bacteria in the culture, showing that the Bvg- phase *B. pertussis* has a growth advantage (Belcher, 2017). Through phenotypic assays it has been shown that *B. pertussis* in the Bvg- phase grows and divides quicker, although Bvg+ growth consumes more glutamate per gram of biomass (Belcher, 2017).

Interestingly, the sorted population with low SSC exhibits the highest surface pertactin concentration (Figure 6-4b). The explanation for this behaviour is not clear and more testing is required to better explain this result. Also, the sorted population with low PRN resulted in a higher surface PRN concentration on the plated cells after 2 days than the sorted population with low PRN. This shows that while cells can be sorted based on PRN surface concentration, the changes are not propagated to the daughter cells.

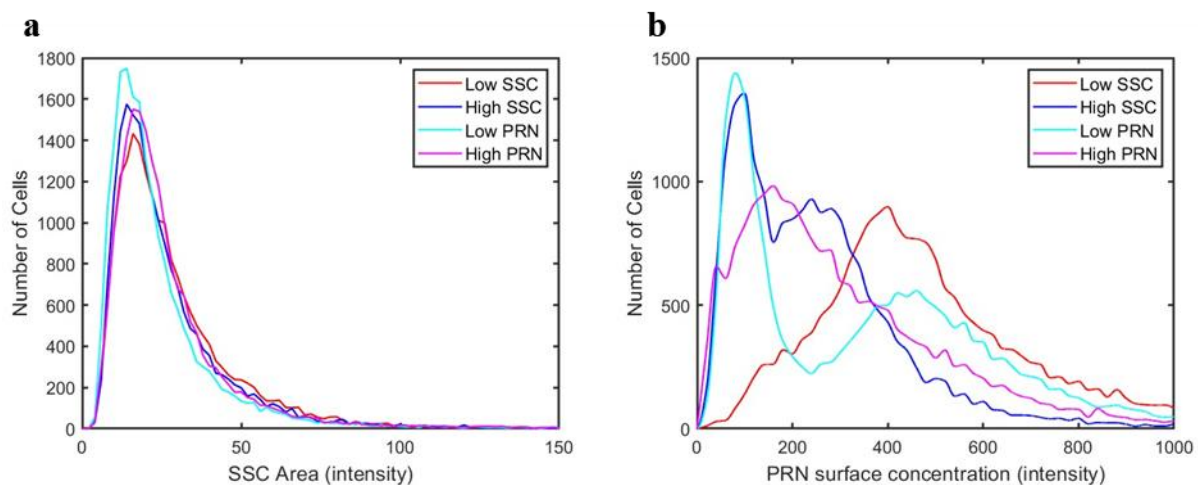


Figure 6-4: Plate properties of (a) side scattering distribution and (b) surface PRN concentration of sorted cell colonies.

Table 6-2: Average plate properties of sorted cell colonies 2 days after plating the sorted populations

Property	Low SSC	High SSC	Low PRN	High PRN
Side Scattering	28.16	26.17	23.53	26.54
PRN surface concentration	527.54	252.18	373.90	313.62

After colonies were found on the BG agar plates from the sorted populations, individual colonies were grown in flasks. For each sorted property, two colonies were selected and grown in separate primary flasks. These primary flasks contained a working volume of 5 mL and were inoculated and grown for 48 hours. Then a secondary flask was inoculated with the culture from the primary flask with a starting OD of 0.3. The growth profiles of the secondary flasks from the sorted populations (low SSC, high SSC, low PRN and high PRN) were compared to a control (Sanofi seed stock), all with the same media outlined in section 3.1 in Chapter 3. While the sorted cell populations did not maintain their epigenetic properties in the daughter cells, re-culturing of the sorted cells did result in faster growing *B. pertussis* cells, as shown in Figure 6-5. It can be noticed from this figure that re-culturing of all sorted populations, regardless of the property and range used for sorting, leads to higher growth compared to the control culture. This suggests a correlation between the sorting and re-culturing procedures and the observed enhancement of growth. Specifically, *B. pertussis* appears to be sensitive to sorting since the sorted populations do not grow in liquid media and have to be plated first. Then, it appears that plating of the sorted cells can repair the inherent damage resulting from the sorting operation and the colonies recovered from the plate grow faster. The sorted and plated populations for side scattering (low SSC and high SSC) are faster growing than the sorted and plated populations for PRN surface concentration (low PRN and high PRN). This may be due to the potentially detrimental effect of the dye used for surface PRN-based sorting operation since the PRN surface concentration during the latter step is

measured with an antibody conjugated to FITC. Another possible reason for the faster growth of the sorted populations would be that they contain Bvg- phase cells but this has to be verified by testing the antigen concentrations at the end of the fermentation.

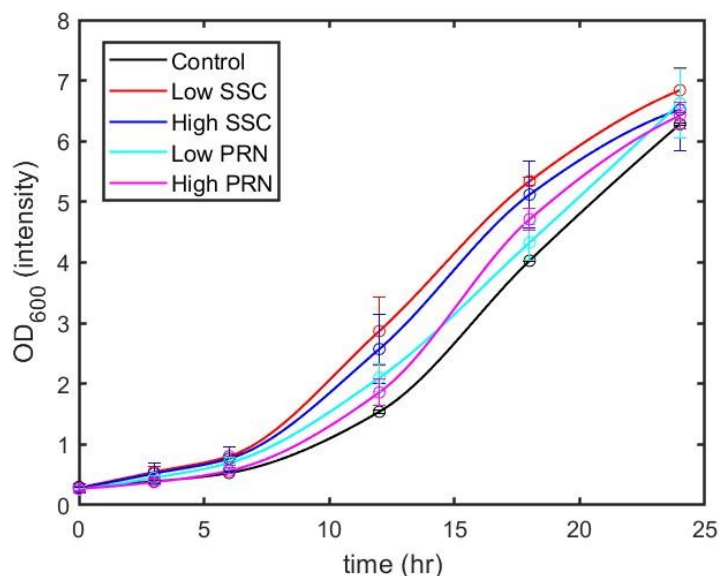


Figure 6-5: Growth of secondary cultures for control and sorted *B. pertussis*.

6.4 Discussion

The productivity of pertactin is the major bottleneck to manufacture the vaccine for whooping cough at Sanofi. After fermentation, the downstream procedure begins with centrifugation to separate the fermentation broth into the centrate and cells. Pertactin is further isolated from the centrate by a series of filtrations and separation steps. The fermentation operation up to the centrifugation steps are often referred to as the upstream phase of the process and the purification/separation steps as the downstream phase. ELISA measurements are taken at the final stages of the upstream phase before and after centrifugation and at different points along the

downstream phase of the process to determine the intermediate PRN concentrations. In the current operation no correlation has been found between any of the ELISA measurements conducted along the process and the final yield of pertactin obtained by the Kjeldahl method for a fully purified sample. In general, it is known that a high error is associated with the ELISA measurements (data not shown) in the fermentation step because of the lack of specificity of the antibody used for this test. In the experiments presented in this study we have implicitly assumed that the surface concentration of PRN is associated with high secretion of pertactin. This may not always be the case since pertactin is cleaved to the surface of the cell to be secreted.

On the other hand, in view of the lack of precision of ELISA, we hypothesize that droplet flow cytometry sorting can instead be used to determine the amount of secreted pertactin during the fermentation stage. This process involves encapsulating single cells in a water-oil-water (w/o/w) double-emulsion where the inner water-phase contains the media and secreted products (118-119). The major advantage of double emulsion encapsulation is that it prevents the secreted metabolites or products from being transported away from the cells and traps them within the droplet for detection. To use this technique, a suitable marker would have to be found. The conjugated antibody used in this work would not be an option because the antibody is unable to cross the oil boundary to enter the droplet and cannot be added before the droplet is formed because even an unbound conjugated antibody would generate a signal. The options would be to design a marker that can travel through the hydrophobic barrier (i.e., Ma *et al.*, 2016) or be placed in the media. Both options would have to provide no fluorescent signal until it is bound to pertactin where the attachment allows it to become fluorescent (i.e., Filonov *et al.*, 2014; Zou *et al.*, 2015; Kellermann *et al.*, 2017). Although the use of droplet flow cytometry sorting to find *B. pertussis* cells that

secrete larger amounts of pertactin procedure can be promising, this technique may still be inaccurate for quantifying productivity since it measures both surface and secreted pertactin.

Another development in flow cytometry sorting is the use of miniaturized devices, some of which carry out FACS by taking advantage of a range of physical properties. These smaller devices are termed microfluidic FACS (μ FACS) and fall within the category of *lab-on-a-chip* technologies. The use of μ FACS has several advantages over that of conventional FACS in that it eliminates potentially biohazardous aerosols, sample loss and cross-contamination due to the closed-platform designs and disposable chips (Shields *et al.*, 2015; Shen *et al.*, 2019). Using microfluidic chips may allow for sorting of *B. pertussis* without the need for streptomycin resistance that had to be used in the current work to avoid contamination. The presence of streptomycin in the media resulted in an increase in oxidative stress and a longer lag phase (Figure 6-1b).

6.5 Conclusions

Protocols involving sorting followed by re-culturing were developed for *B. pertussis* which involved obtaining streptomycin resistance to avoid contamination, sorting the cells according to a pre-defined property of interest and plating the sorted population on BG agar plates. The main goals of the sorting protocol were to investigate whether: i- cells sorted with respect to their scattering properties resulted in advantageous cell growth and ii- cells sorted based on surface PRN maintained the property post-sorting and continued to be used for re-culturing. To this end, flow cytometry sorting was completed on the basis of low and high side scattering as well as low and high PRN surface concentration. The sorted populations of cells did not maintain the selected

properties but the sorting protocol did manage to capture *B. pertussis* cells which grew at a faster rate than the control seed provided by Sanofi. It is believed that the combined sorting and plating procedures led to natural selection of cell populations that are more fit for growth.

Chapter 7 Conclusions and Future Work

7.1 Conclusions

7.1.1 Oxidative stress in flask experiments

A main objective of the current research was to identify sources of variability in the manufacturing process of the whooping cough vaccine operated by Sanofi Pasteur. Variability in medium or inoculum were the expected causes of this variability. In terms of growth media, glutamate was a main potential source for batch-to-batch variation since it is the main nutrient in the fermentation process and thus small changes in its concentration may have a large impact on the process. Starvation was also targeted since the operation switches from batch to fed-batch mode during the course of the manufacturing process and the culture may become partially or completely depleted of nutrients for a duration of time during this switch-over. Measurement error and variability in seed preparation related to the size and/or its origin of the inoculum were also considered to be sources of variability. Based on previous studies which found that oxidative stress is correlated to productivity, we hypothesized that changes in glutamate concentration, starvation conditions or variability in inoculum can separately or simultaneously contribute to oxidative stress and thus may affect growth and productivity. In order to determine the effect of glutamate concentration, initial OD, starvation and seed batches on oxidative stress, a control using standard operating conditions and media from Sanofi was monitored to understand the pattern of growth and oxidative stress in *B. pertussis*. For this control, the ROS concentration was found to rise sharply during the lag phase, then drop and maintain low levels during the exponential growth

phase. Then, later during the culture, the ROS level was again found to increase rapidly during the stationary phase into the death phase.

When the glutamate concentration in the media was increased, the intracellular ROS concentration during the lag and early exponential phases was found to be higher than in the control culture and to increase rapidly at the beginning of the batch. A possible mechanism for the fast increase in ROS level involves the fast generation of NADH in the TCA cycle in response to the high glutamate concentration, which exceeds the consumption rate capacity of the electron chain cycle. It has been reported for other microorganisms that ROS production increases and can lead to oxidative stress when nutrients are present in excess (Wellen and Thompson, 2011). In this work we also try to elucidate whether oxidative stress is correlated, or it is the direct cause for the observed slower growth rate found due to variations in glutamate concentration or inoculum biomass concentration. The results from the experimental data show that the growth rate of conditions with a high concentration of glutamate are related to substrate inhibition. For the case of low biomass concentration in the inoculum, after a lag period is over, the growth rate is faster than the control even with an observed initially higher concentration of ROS. The oxidative stress model also verifies this observation because the fitting accuracy of the model decreased when it was assumed that oxidative stress had a direct impact on the growth rate. Therefore, we conclude that oxidative stress is observed in batches with low biomass levels but it is not the root cause for low growth rates. In other words, low biomass levels occur in batches with high initial glutamate concentration or small inoculum as compared to the control, but the growth rates values during these cultures at each time does not correlate with the corresponding oxidative stress concentrations at that time.

Experiments carried out with lower initial density cultures of *B. pertussis* (i.e., small size inoculum) also exhibited higher oxidative stress during the lag phase and early portion of the exponential growth phase; however, this condition had no discernable effect on the biomass growth rate. In this case, the changes in glutamate concentration for different inoculum sizes were not significant enough to explain the changes in ROS levels. Quorum sensing can provide a possible explanation for the increased stress in low inoculum cultures. Other bacteria have been shown to use quorum sensing in response to environmental conditions that result in higher oxidative stress.

The duration of the lag phase in a flask (taken as the time required for the initial biomass population to double) that is inoculated with cells that were subjected to starvation conditions is approximately 36 hours more than in the case of the control culture. This effect is presumably due to the fact that the percentage of viable cells at the beginning of the secondary culture conducted post-starvation is only 20.4%. Nevertheless, the cells were able to recover and eventually accumulate biomass. The comparison of the two seed batches (one made in 2015 and the other in 2018) followed a similar trend to the starved cells. Both seeds had low viability at inoculation – the 2015 and 2018 seeds contained 55.3% and 80.2% viable cells, respectively. This indicates that cryogenic storage does not conserve cells entirely and the viability of the seed depends strongly on the length of time over which it is stored (Walters *et al.*, 2004).

7.1.2 Oxidative stress model

A model was developed to describe the interplay between the main nutrient (glutamate), ROS and NADPH by modifying a general model of Himeoka and Kaneko to include a relation between the extents of starvation and oxidative stress as well as the role NADPH plays in the synthesis of major cellular components and quenching of ROS. To assess the validity of this model

for this specific group of *B. pertussis* pathways, we conducted an experimental flask study in which the ROS, NADPH, glutamate and biomass concentrations were measured under conditions of varied oxidative stress. The model provided very good fits for the OD, glutamate and NADPH concentrations for all experimental conditions, but a less accurate fit for ROS concentration based on a comparison of the RMSE over the range of experimental measurements. Nevertheless, the fit of the oxidative stress model was considerably better than that of a Contois-based model based on the RMSE and the AIC. The AIC is a particularly useful measure of the trade-off between model dimensionality and predictability in this case since the oxidative stress model involves a greater number of parameters than the Contois model.

The model was also applied to simulate a train of fermenters in order to assess whether the variability in the process can propagate and be further amplified over this sequence. This simulation showed that even a small increase in the concentration of glutamate (10 g/L versus 10.5 g/L) yielded less biomass in the early reactors in the train (i.e., 20 L and 200 L reactors), but the difference vanished by the end of the last reactor (2000 L reactor). However, while the final OD may be similar, the final pertactin productivity may be lower. This can be explained by the fact that the final yield of pertactin was found to be correlated with the area under the growth curve (Agarwal, 2022) rather than the final OD. Therefore, a difference in glutamate concentration may be a possible cause in the variability found in production because it results in a smaller area under the curve of biomass concentration over time.

7.1.3 Coupled population balance – oxidative stress model

Due to the heterogeneity of the cell population, a control action that on only the average culture properties may not necessarily guarantee consistent process outcomes. The solution of a coupled population balance-oxidative stress model using the method of characteristics was

developed to predict distributions in cell size and intracellular glutamate, ROS (reactive oxidative species), NADPH and NADP^+ concentrations in shake flask cultures of *B. pertussis*. The model was calibrated with glutamate, NADPH, OD and flow cytometry data of the distributions of the cell population with respect to cell size and ROS levels. The model was fitted to the measured extracellular glutamate concentrations, OD_{600} for biomass and the mean and standard deviation for the flow cytometric distributions of cell volume and intracellular ROS concentration. It was found that the cell volume measurements always fall along unimodal distributions, whereas the model yields multimodal distributions particularly over the earlier part of inoculation. A much finer discretization both in terms of the cell size and time may be required to further smooth the distributions predicted by the model, but at the expense of additional computational effort which is already very high. The intracellular ROS concentration distributions predicted by the model were much narrower than the distributions from the experimental data. However, the peak or average intracellular ROS concentration was very accurate as shown by the low RMSE for this variable. A probable cause for the deviation between the widths of the measured and model-predicted distributions for the intracellular ROS concentration is that cells may exhibit a different ROS generation rate that depends on cell age.

The major advantage of using a PBM is that it accounts for the distributions and can predict the heterogeneity of the cell population with respect to experimental conditions that are not used in bulk models. When comparing the coupled population balance – oxidative stress model to the bulk oxidative stress model in Chapter 4, it is apparent that the PBM provides much better predictions of the intracellular ROS concentration. It can be shown mathematically that no differences can exist between the mean values calculated from a PBM and a lumped (bulk) model when the model is linear. On the other hand, since the kinetic equations of the model used in this

study are highly nonlinear, the use of the distributions for cell size and ROS concentration will affect the mean intracellular ROS concentrations.

7.1.4 Flow cytometry sorting of epigenetic populations

Epigenetics changes include any process that alters gene activity without changing the DNA sequence and leads to modifications that can be transmitted to daughter cells. In particular, *B. pertussis* contains a regulon that involves a Bvg⁺ phase, characterized by expression of adhesins and toxins, and a Bvg⁻ phase where virulence factors are not expressed. The bottleneck of the production of the vaccine for whooping cough is the antigen pertactin. A sorting and re-culture protocol was developed for *B. pertussis* which involved obtaining streptomycin resistance to avoid contamination, sorting into cultured media and plating on BG agar plates. The main goal of the sorting protocol is to gain insight into whether Bvg⁺ and Bvg⁻ phases can be maintained post-sorting and continued for re-culturing. To this end, flow cytometry sorting was completed on basis of low and high side scattering (a factor from a previous study in Zavatti *et al.* (2019) that showed a correlation with pertactin yield) as well as low and high PRN surface concentrations. The sorted populations of cells did not maintain the properties sorted but the sorting protocol did manage to select *B. pertussis* cells which grew at a faster rate than the control seed provided by Sanofi. It is believed that the cells that are re-cultured following the sorting protocol and plating may become more resilient to survive the stress from sorting and grow faster.

7.2 Next Steps

Following the conclusions of this research, this section outlines future work both in terms of development and improvement of the oxidative stress and coupled population balance –

oxidative stress models and the implementation of the flow cytometry sorting protocol for *B. pertussis*.

The following research investigations could be pursued:

1. The predictability of the oxidative stress model can be improved by adding additional parameters or conducting additional measurements (intracellular NADPH to include in the model or metabolomic analysis). For example, the inclusion of a maintenance term or an age-related term for the production of ROS is recommended since the current population balance model cannot accurately predict the outcome of starved cells and cells in the stationary and death phases.
2. Intracellular NADPH concentrations can be measured with flow cytometry using commercially available dyes or via autofluorescence if excited using the correct laser wavelength. The use of such dyes may enable the simultaneous monitoring of ROS and NADPH concentrations which could help corroborate the relation between these two species.
3. The pathway for the generation of oxidative stress can be included within a dynamic metabolic flux model. The identification of this pathway may be a key factor in better understanding and modeling the metabolic pathways of *B. pertussis*. A dynamic flux model that incorporates the oxidative stress pathways may help explain the relation between oxidative stress and all amino acids included in the media.
4. The ribosome levels during periods of high oxidative stress should be measured to gain a better understanding of the possible indirect effect of stress on growth rate.
5. A population balance model incorporating the surface concentration of pertactin measured with flow cytometry should be developed to predict the final yield of pertactin from a

fermentation. We hypothesize that intracellular and cell surface quantities can be better modeled with population balance models for such applications as individual cell antigen production which is of particular importance for manufacturing. The challenge of this model would be developing the metabolic pathways for pertactin production and the release mechanism as the cells enter the Bvg- phase.

6. Multi-parameteric flow cytometry can be used to measure the relation between cell surface antigen level and ROS concentration which would help elucidate a potential direct effect of ROS level on the antigen concentration.
7. A population balance model that includes the mixing profile within the bioreactors should be developed. While heterogeneity of the cell population can originate in the inocula, additional factors contributing to heterogeneity stem from the surrounding media conditions and mass transfer effects.
8. The flow cytometry sorting protocol and random mutagenesis can be used to sort high producers of the antigens in the whooping cough vaccine. Instead of re-culturing these populations, gene expression or PCR analysis of the factors leading to high production mutations should lead to better understanding of the metabolic pathways of *B. pertussis* and more optimal process control of the operations at Sanofi to increase antigen production.

References

- Abuaita BH, Withey JH. Genetic screening for bacterial mutants in liquid growth media by fluorescence-activated cell sorting. *Journal of microbiological methods* 2011;84(1):109-113.
- Abujarour, R., Valamehr, B., Robinson, M., Rezner, B., Vranceanu, F. and Flynn, P., 2013. Optimized surface markers for the prospective isolation of high-quality hiPSCs using flow cytometry selection. *Scientific reports*, 3(1), pp.1-11.
- Agarwal, P., 2022. Application of Deep Learning in Chemical Processes: Explainability, Monitoring and Observability.
- Ahn, R.S., Taravati, K., Lai, K., Lee, K.M., Nititham, J., Gupta, R., Chang, D.S., Arron, S.T., Rosenblum, M. and Liao, W., 2017. Transcriptional landscape of epithelial and immune cell populations revealed through FACS-seq of healthy human skin. *Scientific reports*, 7(1), pp.1-9.
- Akaike, H., 1998. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199-213). Springer, New York, NY.
- Andrä, I., Ulrich, H., Dürr, S., Soll, D., Henkel, L., Angerpointner, C., Ritter, J., Przibilla, S., Stadler, H., Effenberger, M. and Busch, D.H., 2020. An evaluation of T-cell functionality after flow cytometry sorting revealed p38 MAPK activation. *Cytometry Part A*, 97(2), pp.171-183.
- Arias-Fuenzalida, J., Jarazo, J., Qing, X., Walter, J., Gomez-Giro, G., Nickels, S.L., Zaehres, H., Schöler, H.R. and Schwamborn, J.C., 2017. FACS-assisted CRISPR-Cas9 genome editing facilitates Parkinson's disease modeling. *Stem Cell Reports*, 9(5), pp.1423-1431.
- Arné, E.S. and Holmgren, A., 2000. Physiological functions of thioredoxin and thioredoxin reductase. *European journal of biochemistry*, 267(20), pp.6102-6109.
- Arts, I.S., Gennaris, A. and Collet, J.F., 2015. Reducing systems protecting the bacterial cell envelope from oxidative damage. *FEBS letters*, 589(14), pp.1559-1568.

- Åstrand, M., Nilvebrant, J., Björnmalm, M., Lindbo, S., Hober, S. and Löfblom, J., 2016. Investigating affinity-maturation strategies and reproducibility of fluorescence-activated cell sorting using a recombinant ADAPT library displayed on staphylococci. *Protein Engineering, Design and Selection*, 29(5), pp.187-195.
- Aubron, C., Glodt, J., Matar, C. *et al.* Variation in endogenous oxidative stress in *Escherichia coli* natural isolates during growth in urine. *BMC Microbiol* **12**, 120 (2012).
- Baatout, S., De Boever, P. and Mergeay, M., 2006. Physiological changes induced in four bacterial strains following oxidative stress. *Applied Biochemistry and Microbiology*, 42(4), pp.369-377.
- Babu, M.M., Bhargavi, J., Saund, R.S. and Singh, S.K., 2001. Virulence factors of *Bordetella pertussis*. *Current science*, pp.1512-1522.
- Baret, J.C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., Rick, C., Samuels, M.L., Hutchison, J.B., Agresti, J.J. and Link, D.R., 2009. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab on a Chip*, 9(13), pp.1850-1858.
- Belcher, T. 2017. Investigating the growth and metabolic difference of Bvg⁺ and Bvg⁻ phase *Bordetella pertussis*.
- Bertrand, R.L., 2019. Lag phase is a dynamic, organized, adaptive, and evolvable period that prepares bacteria for cell division. *Journal of bacteriology*, 201(7), pp.e00697-18.
- Boder, E.T. and Wittrup, K.D., 1997. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology*, 15(6), pp.553-557.
- Boone, C.H., Grove, R.A., Adamcova, D., Seravalli, J. and Adamec, J., 2017. Oxidative stress, metabolomics profiling, and mechanism of local anesthetic induced cell death in yeast. *Redox biology*, 12, pp.139-149.
- Bordet, J. and Gengou, O., 1906. Le microbe de la coqueluche. *Ann. l'Institut Pasteur* 20, 731–74.

- Brooks, D.L. and Seagroves, T.N., 2018. Fluorescence-activated cell sorting of murine mammary cancer stem-like cell subpopulations with HIF activity. In *Hypoxia* (pp. 247-263). Humana Press, New York, NY.
- Brynildsen, M.P., Winkler, J.A., Spina, C.S., MacDonald, I.C. and Collins, J.J., 2013. Potentiating antibacterial activity by predictably enhancing endogenous microbial ROS production. *Nature biotechnology*, 31(2), pp.160-165.
- Buehler, D.P., Wiese, C.B., Skelton, S.B. and Southard-Smith, E.M., 2012. An optimized procedure for fluorescence-activated cell sorting (FACS) isolation of autonomic neural progenitors from visceral organs of fetal mice. *JoVE (Journal of Visualized Experiments)*, (66), p.e4188.
- Butz, M., Kast, P. and Hilvert, D., 2014. Affinity maturation of a computationally designed binding protein affords a functional but disordered polypeptide. *Journal of Structural Biology*, 185(2), pp.168-177.
- Caen, O., Schütz, S., Jammalamadaka, M.S., Vrignon, J., Nizard, P., Schneider, T.M., Baret, J.C. and Taly, V., 2018. High-throughput multiplexed fluorescence-activated droplet sorting. *Microsystems & nanoengineering*, 4(1), pp.1-10.
- Cashel, M. and Potrykus, K., 2001. Stringent Response.
- Cashel, M., 1996. The stringent response. *Escherichia coli and Salmonella typhimulium: cellular and molecular biology*, 2, pp.1458-1496.
- Cavanaugh, J.E. and Neath, A.A., 2019. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*, 11(3), p.e1460.
- Chatterji, D. and Ojha, A.K., 2001. Revisiting the stringent response, ppGpp and starvation signaling. *Current opinion in microbiology*, 4(2), pp.160-165.
- Chen, I., Dorr, B.M. and Liu, D.R., 2011. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences*, 108(28), pp.11399-11404.

- Chen, J., Vestergaard, M., Jensen, T.G., Shen, J., Dufva, M., Solem, C. and Jensen, P.R., 2017. Finding the needle in the haystack—the use of microfluidic droplet technology to identify vitamin-secreting lactic acid bacteria. *MBio*, 8(3), pp.e00526-17.
- Chen, Z., Wang, X., Zhao, N., Han, L., Wang, F., Li, H., Cui, Y. and Zhao, X., 2018. Improving the immunogenicity and protective efficacy of the EtMIC2 protein against *Eimeria tenella* infection through random mutagenesis. *Vaccine*, 36(18), pp.2435-2441.
- Cheng, Z., Wu, X., Cheng, J. and Liu, P., 2017. Microfluidic fluorescence-activated cell sorting (μ FACS) chip with integrated piezoelectric actuators for low-cost mammalian cell enrichment. *Microfluidics and Nanofluidics*, 21(1), pp.1-11.
- Cho, S.H., Chen, C.H., Tsai, F.S., Godin, J.M. and Lo, Y.H., 2010. Human mammalian cell sorting using a highly integrated micro-fabricated fluorescence-activated cell sorter (μ FACS). *Lab on a Chip*, 10(12), pp.1567-1573.
- Christodoulou, D., Link, H., Fuhrer, T., Kochanowski, K., Gerosa, L. and Sauer, U., 2018. Reserve flux capacity in the pentose phosphate pathway enables *Escherichia coli*'s rapid response to oxidative stress. *Cell systems*, 6(5), pp.569-578.
- Contois, D.E., 1959. Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures. *Microbiology*, 21(1), pp.40-50.
- Copp JN, Williams EM, Rich MH, Patterson AV, Smaill JB, Ackerley DF. Toward a high-throughput screening platform for directed evolution of enzymes that activate genotoxic prodrugs. *Protein Engineering, Design & Selection* 2014;27(10):399-403.
- Copp, J.N., Mowday, A.M., Williams, E.M., Guise, C.P., Ashoorzadeh, A., Sharrock, A.V., Flanagan, J.U., Smaill, J.B., Patterson, A.V. and Ackerley, D.F., 2017. Engineering a multifunctional nitroreductase for improved activation of prodrugs and PET probes for cancer gene therapy. *Cell Chemical Biology*, 24(3), pp.391-403.
- Cossarizza, A., Chang, H.D., Radbruch, A., Acs, A., Adam, D., Adam-Klages, S., Agace, W.W., Aghaeepour, N., Akdis, M., Allez, M. and Almeida, L.N., 2019. Guidelines for the use of flow cytometry and cell sorting in immunological studies. *European journal of immunology*, 49(10), pp.1457-1973.

- Croce, A.C. and Bottiroli, G., 2014. Autofluorescence spectroscopy and imaging: a tool for biomedical research and diagnosis. *European journal of histochemistry: EJH*, 58(4).
- Crouch, E.E. and Doetsch, F., 2018. FACS isolation of endothelial cells and pericytes from mouse brain microregions. *Nature Protocols*, 13(4), pp.738-751.
- Cuny, C., Lesbats, M. and Dukan, S., 2007. Induction of a global stress response during the first step of *Escherichia coli* plate growth. *Applied and environmental microbiology*, 73(3), pp.885-889.
- Denny, W.A., 2003. Prodrugs for gene-directed enzyme-prodrug therapy (suicide gene therapy). *Journal of Biomedicine and Biotechnology*, 2003(1), p.48.
- Derfus, G.E., Abramzon, D., Tung, M., Chang, D., Kiss, R. and Amanullah, A., 2010. Cell culture monitoring via an auto-sampler and an integrated multi-functional off-line analyzer. *Biotechnology progress*, 26(1), pp.284-292.
- Deweid, L., Neureiter, L., Englert, S., Schneider, H., Deweid, J., Yanakieva, D., Sturm, J., Bitsch, S., Christmann, A., Avrutina, O. and Fuchsbauer, H.L., 2018. Directed evolution of a bond-forming enzyme: ultrahigh-throughput screening of microbial transglutaminase using yeast surface display. *Chemistry—A European Journal*, 24(57), pp.15195-15200.
- Dippong, M., Carl, P., Lenz, C., Schenk, J.A., Hoffmann, K., Schwaar, T., Schneider, R.J. and Kuhne, M., 2017. Hapten-specific single-cell selection of hybridoma clones by fluorescence-activated cell sorting for the generation of monoclonal antibodies. *Analytical chemistry*, 89(7), pp.4007-4012.
- Domínguez, Á., Ciruela, P., Hernández, S., García-García, J.J., Soldevila, N., Izquierdo, C., Moraga-Llop, F., Díaz, A., de Sevilla, M.F., González-Peris, S. and Campins, M., 2017. Effectiveness of the 13-valent pneumococcal conjugate vaccine in preventing invasive pneumococcal disease in children aged 7-59 months. A matched case-control study. *PloS one*, 12(8), p.e0183191.

- Dukan, S. and Nyström, T., 1998. Bacterial senescence: stasis results in increased and differential oxidation of cytoplasmic proteins leading to developmental induction of the heat shock regulon. *Genes & development*, 12(21), pp.3431-3441.
- Dukan, S. and Nyström, T., 1999. Oxidative stress defense and deterioration of growth-arrested *Escherichia coli* cells. *Journal of Biological Chemistry*, 274(37), pp.26027-26032.
- Dürr, R., Duvigneau, S., Laske, T., Bachmann, M. and Kienle, A., 2016. Analyzing the impact of heterogeneity in genetically engineered cell lines for influenza vaccine production using population balance modeling. *IFAC-PapersOnLine*, 49(26), pp.225-230.
- Dürr, R., Franz, A. and Kienle, A., 2015. Combination of limited measurement information and multidimensional population balance models. *IFAC-PapersOnLine*, 48(20), pp.261-266.
- Dürr, R., Franz, A. and Kienle, A., 2015. Combination of limited measurement information and multidimensional population balance models. *IFAC-PapersOnLine*, 48(20), pp.261-266.
- Eakman, J.M., Fredrickson, A.G. and Tsuchiya, H.M., 1966. Statistics and dynamics of microbial cell populations. /*AMERICAN INST. OF CHEMICAL ENGINEERS*.
- Enfors, S.O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jürgen, B., Krüger, E., Schweder, T., Hamer, G., O'beirne, D. and Noisommit-Rizzi, N., 2001. Physiological responses to mixing in large scale bioreactors. *Journal of biotechnology*, 85(2), pp.175-185.
- Eruslanov, E. and Kusmartsev, S., 2010. Identification of ROS using oxidized DCFDA and flowcytometry. In *Advanced protocols in oxidative stress II* (pp. 57-72). Humana Press, Totowa, NJ.
- Farha, M.A. and Brown, E.D., 2013. Discovery of antibiotic adjuvants. *Nature biotechnology*, 31(2), pp.120-122.
- Farr, S.B. and Kogoma, T., 1991. Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. *Microbiological reviews*, 55(4), pp.561-585.
- Fasnacht, M. and Polacek, N., 2021. Oxidative Stress in Bacteria and the Central Dogma of Molecular Biology. *Frontiers in Molecular Biosciences*, 8, p.392.

- Ferullo, D.J., Cooper, D.L., Moore, H.R. and Lovett, S.T., 2009. Cell cycle synchronization of *Escherichia coli* using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods*, 48(1), pp.8-13.
- Filonov, G.S., Moon, J.D., Svensen, N. and Jaffrey, S.R., 2014. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *Journal of the American Chemical Society*, 136(46), pp.16299-16308.
- Finegersh, A. and Homanics, G.E., 2016. Chromatin immunoprecipitation and gene expression analysis of neuronal subtypes after fluorescence activated cell sorting. *Journal of neuroscience methods*, 263, pp.81-88.
- Frazier, T.P., Bowles, A., Lee, S., Abbott, R., Tucker, H.A., Kaplan, D., Wang, M., Strong, A., Brown, Q., He, J. and Bunnell, B.A., 2016. Serially transplanted nonpericytic CD146–adipose stromal/stem cells in silk bioscaffolds regenerate adipose tissue in vivo. *Stem Cells*, 34(4), pp.1097-1111.
- Fredrickson, A.G., Ramkrishna, D. and Tsuchiya, H.M., 1967. Statistics and dynamics of procaryotic cell populations. *Mathematical Biosciences*, 1(3), pp.327-374.
- Fried, J., Doblin, J., Takamoto, S., Perez, A., Hansen, H. and Clarkson, B., 1982. Effects of Hoechst 33342 on survival and growth of two tumor cell lines and on hematopoietically normal bone marrow cells. *Cytometry: The Journal of the International Society for Analytical Cytology*, 3(1), pp.42-47.
- Gaj, T., Gersbach, C.A. and Barbas III, C.F., 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in biotechnology*, 31(7), pp.397-405.
- Gant, V.A., Warnes, G., Phillips, I. and Savidge, G.F., 1993. The application of flow cytometry to the study of bacterial responses to antibiotics. *Journal of Medical Microbiology*, 39(2), pp.147-154.
- Gentry, D.R., Hernandez, V.J., Nguyen, L.H., Jensen, D.B. and Cashel, M., 1993. Synthesis of the stationary-phase sigma factor sigma^s is positively regulated by ppGpp. *Journal of bacteriology*, 175(24), pp.7982-7989.

- George, S.E., Hrubesch, J., Breuing, I., Vetter, N., Korn, N., Hennemann, K., Bleul, L., Willmann, M., Ebner, P., Götz, F. and Wolz, C., 2019. Oxidative stress drives the selection of quorum sensing mutants in the *Staphylococcus aureus* population. *Proceedings of the National Academy of Sciences*, 116(38), pp.19145-19154.
- Geslewitz, W.E., Percopo, C.M. and Rosenberg, H.F., 2018. FACS isolation of live mouse eosinophils at high purity via a protocol that does not target Siglec F. *Journal of immunological methods*, 454, pp.27-31.
- Ghatnekar, M.V., 1964. Primary effects of different mutagens and the disturbances induced in the meiosis of X1 and X2 of *Vicia faba*. *Caryologia*, 17(1), pp.219-244.
- Goldbeck, O., Eck, A.W. and Seibold, G.M., 2018. Real time monitoring of NADPH concentrations in *Corynebacterium glutamicum* and *Escherichia coli* via the genetically encoded sensor mBFP. *Frontiers in microbiology*, p.2564.
- González-Flecha, B. and Demple, B., 1995. Metabolic Sources of Hydrogen Peroxide in Aerobically Growing *Escherichia coli**. *Journal of Biological Chemistry*, 270(23), pp.13681-13687.
- Goswami, M., Mangoli, S.H. and Jawali, N., 2007. Effects of glutathione and ascorbic acid on streptomycin sensitivity of *Escherichia coli*. *Antimicrobial agents and chemotherapy*, 51(3), pp.1119-1122.
- Grant, S.S. and Hung, D.T., 2013. Persistent bacterial infections, antibiotic tolerance, and the oxidative stress response. *Virulence*, 4(4), pp.273-283.
- Grose, J.H., Joss, L., Velick, S.F. and Roth, J.R., 2006. Evidence that feedback inhibition of NAD kinase controls responses to oxidative stress. *Proceedings of the National Academy of Sciences*, 103(20), pp.7601-7606.
- Guillouzo, A. and Guguen-Guillouzo, C., 2020. Antibiotics-induced oxidative stress. *Current Opinion in Toxicology*, 20, pp.23-28.
- Hai, M. and Magdassi, S., 2004. Investigation on the release of fluorescent markers from w/o/w emulsions by fluorescence-activated cell sorter. *Journal of controlled release*, 96(3), pp.393-402.

- Halliwell, B. and Whiteman, M., 2004. Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean?. *British journal of pharmacology*, 142(2), pp.231-255.
- He, L., Su, J., Ming, M., Bernardo, L., Chen, T., Gisonni-Lex, L. and Gajewska, B., 2018. Flow cytometry: an efficient method for antigenicity measurement and particle characterization on an adjuvanted vaccine candidate H4-IC31 for tuberculosis. *Journal of Immunological Methods*, 452, pp.39-45.
- Helmrich, U., Marsano, A., Melly, L., Wolff, T., Christ, L., Heberer, M., Scherberich, A., Martin, I. and Banfi, A., 2012. Generation of human adult mesenchymal stromal/stem cells expressing defined xenogenic vascular endothelial growth factor levels by optimized transduction and flow cytometry purification. *Tissue Engineering Part C: Methods*, 18(4), pp.283-292.
- Henge-Aronis, R., 2000. The general stress response in *Escherichia coli*. *Bacterial stress responses*.
- Himeoka, Y. and Kaneko, K., 2017. Theory for transitions between exponential and stationary phases: universal laws for lag time. *Physical Review X*, 7(2), p.021049.
- Hirai, Y., Takagi, D., Anai, S., Chihara, Y., Tsuchiya, T., Fujimoto, K., Hirao, Y. and Tabata, O., 2015. ALA-induced fluorescence detection with photoresist-based microfluidic cell sorter for bladder cancer diagnosis. *Sensors and Actuators B: Chemical*, 213, pp.547-557.
- Hölzenspies, J., Cruz, G.D. and Brickman, J.M., 2015. Resolving heterogeneity: fluorescence-activated cell sorting of dynamic cell populations from feeder-free mouse embryonic stem cell culture. In *Embryonic Stem Cell Protocols* (pp. 25-40). Humana Press, New York, NY.
- Hu, F.J., Volk, A.L., Persson, H., Säll, A., Borrebaeck, C., Uhlen, M. and Rockberg, J., 2018. Combination of phage and Gram-positive bacterial display of human antibody repertoires enables isolation of functional high affinity binders. *New biotechnology*, 45, pp.80-88.
- Hubalek, Z., 2003. Protectants used in the cryopreservation of microorganisms. *Cryobiology*, 46(3), pp.205-229.

- Huizer, K., Mustafa, D.A., Spelt, J.C., Kros, J.M. and Sacchetti, A., 2017. Improving the characterization of endothelial progenitor cell subsets by an optimized FACS protocol. *PLoS One*, 12(9), p.e0184895.
- Immanuel, S.R.C., Ghanate, A.D., Parmar, D.S., Marriage, F., Panchagnula, V., Day, P.J. and Raghunathan, A., 2018. Integrative analysis of rewired central metabolism in temozolomide resistant cells. *Biochemical and Biophysical Research Communications*, 495(2), pp.2010-2016.
- Ishii, K., Suzuki, N., Mabuchi, Y., Sekiya, I. and Akazawa, C., 2017. Technical advantage of recombinant collagenase for isolation of muscle stem cells. *Regenerative therapy*, 7, pp.1-7.
- Jager, S.B., Pallesen, L.T. and Vaegter, C.B., 2018. Isolation of satellite glial cells for high-quality RNA purification. *Journal of neuroscience methods*, 297, pp.1-8.
- Jahan-Tigh, R.R., Ryan, C., Obermoser, G. and Schwarzenberger, K., 2012. Flow cytometry. *The Journal of investigative dermatology*, 132(10), p.e1.
- Jerono, P., Schaum, A. and Meurer, T., 2021. Parameter identification of a yeast batch cell population balance model. *IFAC-PapersOnLine*, 54(7), pp.144-149.
- Kanehisa, M. and Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), pp.27-30.
- Kanehisa, M. and Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), pp.27-30.
- Kashmiri, Z.N. and Mankar, S.A., 2014. Free radicals and oxidative stress in bacteria. *Int J Curr Microbiol App Sci*, 3(9), pp.34-40.
- Kellermann, S.J. and Rentmeister, A., 2017. A FACS-based screening strategy to assess sequence-specific RNA-binding of Pumilio protein variants in E. coli. *Biological chemistry*, 398(1), pp.69-75.

- Khelef, N., DeShazer, D., Friedman, R.L. and Guiso, N., 1996. In vivo and in vitro analysis of Bordetella pertussis catalase and Fe-superoxide dismutase mutants. *FEMS microbiology letters*, 142(2-3), pp.231-235.
- Kirkman, H.N., Rolfo, M., Ferraris, A.M. and Gaetani, G.F., 1999. Mechanisms of protection of catalase by NADPH: kinetics and stoichiometry. *Journal of Biological Chemistry*, 274(20), pp.13908-13914.
- Kodym, A. and Afza, R., 2003. Physical and chemical mutagenesis. In *Plant functional genomics* (pp. 189-203). Humana Press.
- Koh, H.J., Lee, S.M., Son, B.G., Lee, S.H., Ryoo, Z.Y., Chang, K.T., Park, J.W., Park, D.C., Song, B.J., Veech, R.L. and Song, H., 2004. Cytosolic NADP⁺-dependent isocitrate dehydrogenase plays a key role in lipid metabolism. *Journal of Biological Chemistry*, 279(38), pp.39968-39974.
- Kohanski, M.A., Dwyer, D.J. and Collins, J.J., 2010. How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology*, 8(6), pp.423-435.
- Kolter, R., Siegele, D.A. and Tormo, A., 1993. The stationary phase of the bacterial life cycle. *Annual review of microbiology*, 47, pp.855-875.
- Kullback, S. and Leibler, R.A., 1951. On information and sufficiency. *The annals of mathematical statistics*, 22(1), pp.79-86.
- Kunnath-Velayudhan, S. and Porcelli, S.A., 2018. Isolation of intact RNA from murine CD4⁺ T cells after intracellular cytokine staining and fluorescence-activated cell sorting. *Journal of immunological methods*, 456, pp.77-80.
- Labrou, N.E., 2010. Random mutagenesis methods for in vitro directed enzyme evolution. *Current Protein and Peptide Science*, 11(1), pp.91-100.
- Lakowicz, J.R. ed., 2006. *Principles of fluorescence spectroscopy*. Boston, MA: Springer US.
- Lange, H., Taillandier, P. and Riba, J.P., 2001. Effect of high shear stress on microbial viability. *Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology*, 76(5), pp.501-505.

- Lara, A.R., Galindo, E., Ramírez, O.T. and Palomares, L.A., 2006. Living with heterogeneities in bioreactors. *Molecular biotechnology*, 34(3), pp.355-381.
- Leo, J.C., Grin, I. and Linke, D., 2012. Type V secretion: mechanism (s) of autotransport through the bacterial outer membrane. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592), pp.1088-1101.
- Li, N., Huang, X., Zou, J., Chen, G., Liu, G., Li, M., Dong, J., Du, F., Cui, X. and Tang, Z., 2018. Evolution of microbial biosensor based on functional RNA through fluorescence-activated cell sorting. *Sensors and Actuators B: Chemical*, 258, pp.550-557.
- Li, N., Huang, X., Zou, J., Chen, G., Liu, G., Li, M., Dong, J., Du, F., Cui, X. and Tang, Z., 2018. Evolution of microbial biosensor based on functional RNA through fluorescence-activated cell sorting. *Sensors and Actuators B: Chemical*, 258, pp.550-557.
- Li, P., Ma, Z., Zhou, Y., Collins, D.J., Wang, Z. and Ai, Y., 2019. Detachable acoustophoretic system for fluorescence-activated sorting at the single-droplet level. *Analytical chemistry*, 91(15), pp.9970-9977.
- Lim, Y., 2005. Modeling and prediction of cell population dynamics. In *Computer Aided Chemical Engineering* (Vol. 20, pp. 517-522). Elsevier.
- Lindberg, H., Härd, T., Löfblom, J. and Ståhl, S., 2015. A truncated and dimeric format of an Affibody library on bacteria enables FACS-mediated isolation of amyloid-beta aggregation inhibitors with subnanomolar affinity. *Biotechnology journal*, 10(11), pp.1707-1718.
- Liu, Y., Xue, Z.L., Chen, S.P., Wang, Z., Zhang, Y., Gong, W.L. and Zheng, Z.M., 2016. A high-throughput screening strategy for accurate quantification of menaquinone based on fluorescence-activated cell sorting. *Journal of Industrial Microbiology and Biotechnology*, 43(6), pp.751-760.
- Llufrio, E.M., Wang, L., Naser, F.J. and Patti, G.J., 2018. Sorting cells alters their redox state and cellular metabolome. *Redox biology*, 16, pp.381-387.
- Locht, C., 1999. Molecular aspects of *Bordetella pertussis* pathogenesis. *International Microbiology*, 2(3), pp.137-144.

- Lonowski LA, Narimatsu Y, Riaz A, Delay CE, Yang Z, Niola F, Duda K, Ober EA, Clausen H, Wandall HH, Hansen SH. Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis. *Nature protocols* 2017;12(3):581.
- López, J.E., Sharma, J., Avila, J., Wood, T.S., VanDyke, J.E., McLaughlin, B., Abbey, C.K., Wong, A., Myagmar, B.E., Swigart, P.M. and Simpson, P.C., 2017. Novel large-particle FACS purification of adult ventricular myocytes reveals accumulation of myosin and actin disproportionate to cell size and proteome in normal post-weaning development. *Journal of molecular and cellular cardiology*, 111, pp.114-122.
- Luftig, M., Nikitin, P., Yan, C., Forte, E., Tourigny, J., Price, A. and Dave, S., 2013. G101 studies at the oncogenic virus/host interface: dynamic regulation of Epstein-Barr virus-mediated B cell immortalization. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 62, p.58. = PRICE
- Lumjiaktase, P., Diggle, S.P., Loprasert, S., Tungpradabkul, S., Daykin, M., Camara, M., Williams, P. and Kunakorn, M., 2006. Quorum sensing regulates dpsA and the oxidative stress response in *Burkholderia pseudomallei*. *Microbiology*, 152(12), pp.3651-3659.
- Lunder, M., Bratkovič, T., Doljak, B., Kreft, S., Urleb, U., Štrukelj, B. and Plazar, N., 2005. Comparison of bacterial and phage display peptide libraries in search of target-binding motif. *Applied biochemistry and biotechnology*, 127(2), pp.125-131.
- Lushchak, V.I., 2001. Oxidative stress and mechanisms of protection against it in bacteria. *Biochemistry (Moscow)*, 66(5), pp.476-489.
- Lv, D., Ma, Q.H., Duan, J.J., Wu, H.B., Zhao, X.L., Yu, S.C. and Bian, X.W., 2016. Optimized dissociation protocol for isolating human glioma stem cells from tumorspheres via fluorescence-activated cell sorting. *Cancer Letters*, 377(1), pp.105-115.
- Ma, F., Fischer, M., Han, Y., Withers, S.G., Feng, Y. and Yang, G.Y., 2016. Substrate engineering enabling fluorescence droplet entrapment for IVC-FACS-based ultrahigh-throughput screening. *Analytical Chemistry*, 88(17), pp.8587-8595.

- Ma, N., Koelling, K.W. and Chalmers, J.J., 2002. Fabrication and use of a transient contractional flow device to quantify the sensitivity of mammalian and insect cells to hydrodynamic forces. *Biotechnology and bioengineering*, 80(4), pp.428-437.
- Maesner, C.C., Almada, A.E. and Wagers, A.J., 2016. Established cell surface markers efficiently isolate highly overlapping populations of skeletal muscle satellite cells by fluorescence-activated cell sorting. *Skeletal Muscle*, 6(1), pp.1-10.
- Maitra, A. and Dill, K.A., 2015. Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proceedings of the National Academy of Sciences*, 112(2), pp.406-411.
- Maki, Y., Yoshida, H. and Wada, A., 2000. Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase Escherichia coli. *Genes to cells*, 5(12), pp.965-974.
- Mantzaris, N.V., Daoutidis, P. and Sreenc, F., 2001a. Numerical solution of multi-variable cell population balance models: I. Finite difference methods. *Computers & Chemical Engineering*, 25(11-12), pp.1411-1440.
- Mantzaris, N.V., Daoutidis, P. and Sreenc, F., 2001b. Numerical solution of multi-variable cell population balance models. II. Spectral methods. *Computers & Chemical Engineering*, 25(11-12), pp.1441-1462.
- Mantzaris, N.V., Liou, J.J., Daoutidis, P. and Sreenc, F., 1999. Numerical solution of a mass structured cell population balance model in an environment of changing substrate concentration. *Journal of Biotechnology*, 71(1-3), pp.157-174.
- Martinez de Tejada, G., Cotter, P.A., Heininger, U., Camilli, A., Akerley, B.J., Mekalanos, J.J. and Miller, J.F., 1998. Neither the Bvg[−] phase nor the vrg6 locus of Bordetella pertussis is required for respiratory infection in mice. *Infection and immunity*, 66(6), pp.2762-2768.
- McCullum, E.O., Williams, B.A., Zhang, J. and Chaput, J.C., 2010. Random mutagenesis by error-prone PCR. In *In vitro mutagenesis protocols* (pp. 103-109). Humana Press, Totowa, NJ.
- McDougald, D., Gong, L., Srinivasan, S., Hild, E., Thompson, L., Takayama, K., Rice, S.A. and Kjelleberg, S., 2002. Defences against oxidative stress during starvation in bacteria. *Antonie Van Leeuwenhoek*, 81(1), pp.3-13.

- Michener JK, Smolke CD. High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch. *Metabolic engineering* 2012;14(4):306-316.
- Miller, M.B. and Bassler, B.L., 2001. Quorum sensing in bacteria. *Annual Reviews in Microbiology*, 55(1), pp.165-199.
- Mollet, M., Godoy-Silva, R., Berdugo, C. and Chalmers, J.J., 2007. Acute hydrodynamic forces and apoptosis: a complex question. *Biotechnology and bioengineering*, 98(4), pp.772-788.
- Mollet, M., Godoy-Silva, R., Berdugo, C. and Chalmers, J.J., 2008. Computer simulations of the energy dissipation rate in a fluorescence-activated cell sorter: Implications to cells. *Biotechnology and bioengineering*, 100(2), pp.260-272.
- Moon, K., Bonocora, R.P., Kim, D.D., Chen, Q., Wade, J.T., Stibitz, S. and Hinton, D.M., 2017. The BvgAS regulon of *Bordetella pertussis*. *MBio*, 8(5), pp.e01526-17.
- Morchain, J., Pigou, M. and Lebaz, N., 2017. A population balance model for bioreactors combining interdivision time distributions and micromixing concepts. *Biochemical engineering journal*, 126, pp.135-145.
- Mueller, E.A., Westfall, C.S. and Levin, P.A., 2020. pH-dependent activation of cytokinesis modulates *Escherichia coli* cell size. *PLoS genetics*, 16(3), p.e1008685.
- Mullarky, E. and Cantley, L.C., 2015. Diverting glycolysis to combat oxidative stress. *Innovative medicine*, pp.3-23.
- Mutafooulos, K., Spink, P., Lofstrom, C.D., Lu, P.J., Lu, H., Sharpe, J.C., Franke, T. and Weitz, D.A., 2019. Traveling surface acoustic wave (TSAW) microfluidic fluorescence activated cell sorter (μ FACS). *Lab on a Chip*, 19(14), pp.2435-2443.
- Nakamura, T. and Omasa, T., 2015. Optimization of cell line development in the GS-CHO expression system using a high-throughput, single cell-based clone selection system. *Journal of bioscience and bioengineering*, 120(3), pp.323-329.
- Nguyen, R., Perfetto, S., Mahnke, Y.D., Chattopadhyay, P. and Roederer, M., 2013. Quantifying spillover spreading for comparing instrument performance and aiding in multicolor panel design. *Cytometry Part A*, 83(3), pp.306-315.

- Nilsson, H., Krawczyk, K.M. and Johansson, M.E., 2014. High salt buffer improves integrity of RNA after fluorescence-activated cell sorting of intracellular labeled cells. *Journal of biotechnology*, 192, pp.62-65.
- Noofeli, M., 2008. *Genetic analysis and characterisation of the BapC autotransporter of bordetella pertussis* (Doctoral dissertation, University of Glasgow).
- Ostafe, R., Prodanovic, R., Commandeur, U. and Fischer, R., 2013. Flow cytometry-based ultra-high-throughput screening assay for cellulase activity. *Analytical biochemistry*, 435(1), pp.93-98.
- Pannala, V.R., Bazil, J.N., Camara, A.K. and Dash, R.K., 2013. A biophysically based mathematical model for the catalytic mechanism of glutathione reductase. *Free Radical Biology and Medicine*, 65, pp.1385-1397.
- Papa, S., Zamai, L., Cecchini, T., Del Grande, P. and Vitale, M., 1991. Cell cycle analysis in flow cytometry: use of BrdU labelling and side scatter for the detection of the different cell cycle phases. *Cytotechnology*, 5(1), pp.103-106.
- Parkhill, J., Sebahia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T., Churcher, C.M., Bentley, S.D., Mungall, K.L. and Cerdeño-Tárraga, A.M., 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nature genetics*, 35(1), pp.32-40.
- Pigou, M. and Morchain, J., 2015. Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. *Chemical Engineering Science*, 126, pp.267-282.
- Póvoa, E.E.F., Ebbing, A.L., Betist, M.C., Van Der Veen, C. and Korswagen, H.C., 2020. An optimized dissociation protocol for FACS-based isolation of rare cell types from *Caenorhabditis elegans* L1 larvae. *MethodsX*, 7, p.100922.
- Quedeville, V., Ouazait, H., Polizzi, B., Fox, R.O., Villedieu, P., Fede, P., Létisse, F. and Morchain, J., 2018. A two-dimensional population balance model for cell growth including multiple uptake systems. *Chemical Engineering Research and Design*, 132, pp.966-981.

- Ramesh B, Frei CS, Cirino PC, Varadarajan N. Functional enrichment by direct plasmid recovery after fluorescence activated cell sorting. *BioTechniques* 2015;59(3):157-161.
- Ramkrishna, D., 1979. Statistical models of cell populations. In *Advances in Biochemical Engineering, Volume 11* (pp. 1-47). Springer, Berlin, Heidelberg.
- Ren, L., Yang, S., Zhang, P., Qu, Z., Mao, Z., Huang, P.H., Chen, Y., Wu, M., Wang, L., Li, P. and Huang, T.J., 2018. Standing surface acoustic wave (SSAW)-based fluorescence-activated cell sorter. *Small*, 14(40), p.1801996.
- Richardson, G.M., Lannigan, J. and Macara, I.G., 2015. Does FACS perturb gene expression?. *Cytometry Part A*, 87(2), pp.166-175.
- Rinnan, Å. and Andersen, C.M., 2005. Handling of first-order Rayleigh scatter in PARAFAC modelling of fluorescence excitation–emission data. *Chemometrics and intelligent laboratory systems*, 76(1), pp.91-99.
- Rolfe, M.D., Rice, C.J., Lucchini, S., Pin, C., Thompson, A., Cameron, A.D., Alston, M., Stringer, M.F., Betts, R.P., Baranyi, J. and Peck, M.W., 2012. Lag phase is a distinct growth phase that prepares bacteria for exponential growth and involves transient metal accumulation. *Journal of bacteriology*, 194(3), pp.686-701.
- Rover Jr, L., Fernandes, J.C., de Oliveira Neto, G., Kubota, L.T., Katekawa, E. and Serrano, S.H., 1998. Study of NADH stability using ultraviolet–visible spectrophotometric analysis and factorial design. *Analytical biochemistry*, 260(1), pp.50-55.
- Sadreddini, S., Jadidi-Niaragh, F., Younesi, V., Pournak, T., Afkham, A., Shokri, F. and Yousefi, M., 2016. Evaluation of EBV transformation of human memory B-cells isolated by FACS and MACS techniques. *Journal of Immunotoxicology*, 13(4), pp.490-497.
- Sánchez-Luengo, M.Á., Rovira, M., Serrano, M., Fernandez-Marcos, P.J. and Martinez, L., 2017. Analysis of the advantages of cis reporters in optimized FACS-G al. *Cytometry Part A*, 91(7), pp.721-729.
- Sandkvist, M. and Bagdasarian, M., 1996. Secretion of recombinant proteins by Gram-negative bacteria. *Current opinion in biotechnology*, 7(5), pp.505-511.

- Santos, J.M., Freire, P., Vicente, M. and Arraiano, C.M., 1999. The stationary-phase morphogene *bolA* from *Escherichia coli* is induced by stress during early stages of growth. *Molecular microbiology*, 32(4), pp.789-798.
- Schurig-Briccio, L.A., Farías, R.N., Rodríguez-Montelongo, L., Rintoul, M.R. and Rapisarda, V.A., 2009. Protection against oxidative stress in *Escherichia coli* stationary phase by a phosphate concentration-dependent genes expression. *Archives of biochemistry and biophysics*, 483(1), pp.106-110.
- Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z. and Hwa, T., 2010. Interdependence of cell growth and gene expression: origins and consequences. *Science*, 330(6007), pp.1099-1102.
- Seaver, L.C. and Imlay, J.A., 2001. Hydrogen peroxide fluxes and compartmentalization inside growing *Escherichia coli*. *Journal of bacteriology*, 183(24), pp.7182-7189.
- Serra, D.O., Lücking, G., Weiland, F., Schulz, S., Görg, A., Yantorno, O.M. and Ehling-Schulz, M., 2008. Proteome approaches combined with Fourier transform infrared spectroscopy revealed a distinctive biofilm physiology in *Bordetella pertussis*. *Proteomics*, 8(23-24), pp.4995-5010.
- Sharma, P., Jha, A.B., Dubey, R.S. and Pessarakli, M., 2012. Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. *Journal of botany*, 2012.
- Shen, Y., Yalikun, Y. and Tanaka, Y., 2019. Recent advances in microfluidic cell sorting systems. *Sensors and Actuators B: Chemical*, 282, pp.268-281.
- Shen, Y.P., Liao, Y.L., Lu, Q., He, X., Yan, Z.B. and Liu, J.Z., 2021. ATP and NADPH engineering of *Escherichia coli* to improve the production of 4-hydroxyphenylacetic acid using CRISPRi. *Biotechnology for biofuels*, 14(1), pp.1-10.
- Shields IV, C.W., Reyes, C.D. and López, G.P., 2015. Microfluidic cell sorting: a review of the advances in the separation of cells from debulking to rare cell isolation. *Lab on a Chip*, 15(5), pp.1230-1249.

- Siloto, R.M. and Weselake, R.J., 2012. Site saturation mutagenesis: Methods and applications in protein engineering. *Biocatalysis and Agricultural Biotechnology*, 1(3), pp.181-189.
- Simione, F., 1992. Key issues relating to the genetic stability and preservation of cells and cell banks. *PDA Journal of Pharmaceutical Science and Technology*, 46(6), pp.226-232.
- Simon, R.U.P.A.P., Priefer, U. and Pühler, A., 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/technology*, 1(9), pp.784-791.
- Singh, R., Lemire, J., Mailloux, R.J. and Appanna, V.D., 2008. A novel strategy involved anti-oxidative defense: the conversion of NADH into NADPH by a metabolic network. *PLoS One*, 3(7), p.e2682.
- Singh, R., Mailloux, R.J., Puiseux-Dao, S. and Appanna, V.D., 2007. Oxidative stress evokes a metabolic adaptation that favors increased NADPH synthesis and decreased NADH production in *Pseudomonas fluorescens*. *Journal of bacteriology*, 189(18), pp.6665-6675.
- Singh, R., Mailloux, R.J., Puiseux-Dao, S. and Appanna, V.D., 2007. Oxidative stress evokes a metabolic adaptation that favors increased NADPH synthesis and decreased NADH production in *Pseudomonas fluorescens*. *Journal of bacteriology*, 189(18), pp.6665-6675.
- Smith, G.P., 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, 228(4705), pp.1315-1317.
- Stadinski, B.D. and Huseby, E.S., 2020. How to prevent yourself from seeing double. *Cytometry. Part A: the journal of the International Society for Analytical Cytology*, 97(11), p.1102.
- Stainer, D. and Scholte, M.J., 1970. A simple chemically defined medium for the production of phase I *Bordetella pertussis*. *Microbiology*, 63(2), pp.211-220.
- Storz, G. and Imlay, J.A., 1999. Oxidative stress. *Current opinion in microbiology*, 2(2), pp.188-194.
- Sun, Y., Ban, B., Bradbury, A., Ansari, G.S. and Blake, D.A., 2016. Combining yeast display and competitive FACS to select rare hapten-specific clones from recombinant antibody libraries. *Analytical chemistry*, 88(18), pp.9181-9189.

- Sundararaghavan, A., Mukherjee, A., Sahoo, S. and Suraishkumar, G.K., 2020. Mechanism of the oxidative stress-mediated increase in lipid accumulation by the bacterium, *R. opacus* PD630: Experimental analysis and genome-scale metabolic modeling. *Biotechnology and bioengineering*, 117(6), pp.1779-1788.
- Szalóki, G. and Goda, K., 2015. Compensation in multicolor flow cytometry. *Cytometry Part A*, 87(11), pp.982-985.
- Ter Huurne, M., Chappell, J., Dalton, S. and Stunnenberg, H.G., 2017. Distinct cell-cycle control in two different states of mouse pluripotency. *Cell stem cell*, 21(4), pp.449-455.
- Terekhov, S.S., Smirnov, I.V., Stepanova, A.V., Bobik, T.V., Mokrushina, Y.A., Ponomarenko, N.A., Belogurov Jr, A.A., Rubtsova, M.P., Kartseva, O.V., Gomzikova, M.O. and Moskovtsev, A.A., 2017. Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity. *Proceedings of the National Academy of Sciences*, 114(10), pp.2550-2555.
- Toli, D., Buttigieg, D., Blanchard, S., Lemonnier, T., d'Incamps, B.L., Bellouze, S., Baillat, G., Bohl, D. and Haase, G., 2015. Modeling amyotrophic lateral sclerosis in pure human iPSc-derived motor neurons isolated by a novel FACS double selection technique. *Neurobiology of disease*, 82, pp.269-280.
- Tomlinson, J.E., Wagner, B., Felipe, M.J.B. and Van de Walle, G.R., 2018. Multispectral fluorescence-activated cell sorting of B and T cell subpopulations from equine peripheral blood. *Veterinary immunology and immunopathology*, 199, pp.22-31.
- Tsuchiya, H.M., Fredrickson, A.G. and Aris, R., 1966. Dynamics of microbial cell populations. In *Advances in chemical engineering* (Vol. 6, pp. 125-206). Academic Press.
- Ueta, M., Ohniwa, R.L., Yoshida, H., Maki, Y., Wada, C. and Wada, A., 2008. Role of HPF (hibernation promoting factor) in translational activity in *Escherichia coli*. *Journal of biochemistry*, 143(3), pp.425-433.
- Van Beek, L.F., de Gouw, D., Eleveld, M.J., Bootsma, H.J., De Jonge, M.I., Mooi, F.R., Zomer, A. and Diavatopoulos, D.A., 2018. Adaptation of *Bordetella pertussis* to the respiratory tract. *The Journal of infectious diseases*, 217(12), pp.1987-1996.

- Van Brussel, I., Ammi, R., Rombouts, M., Cools, N., Vercauteren, S.R., De Roover, D., Hendriks, J.M., Lauwers, P., Van Schil, P.E. and Schrijvers, D.M., 2015. Fluorescent activated cell sorting: an effective approach to study dendritic cell subsets in human atherosclerotic plaques. *Journal of Immunological Methods*, 417, pp.76-85.
- Varma, S. and Voldman, J., 2015. A cell-based sensor of fluid shear stress for microfluidics. *Lab on a Chip*, 15(6), pp.1563-1573.
- Vemuri, G. N., Altman, E., Sangurdekar, D. P., Khodursky, A. B., & Eiteman, M. A. (2006). Overflow metabolism in *Escherichia coli* during steady-state growth: transcriptional regulation and effect of the redox ratio. *Applied and environmental microbiology*, 72(5), 3653–3661.
- Vitelli, M., Budman, H., Pritzker, M. and Tamer, M., 2021. Applications of flow cytometry sorting in the pharmaceutical industry: A review. *Biotechnology Progress*, 37(4), p.e3146.
- von Recum-Knepper, J., Sadewasser, A., Weinheimer, V.K. and Wolff, T., 2015. Fluorescence-activated cell sorting-based analysis reveals an asymmetric induction of interferon-stimulated genes in response to seasonal influenza A virus. *Journal of Virology*, 89(14), pp.6982-6993.
- Wagner, J.M., Liu, L., Yuan, S.F., Venkataraman, M.V., Abate, A.R. and Alper, H.S., 2018. A comparative analysis of single cell and droplet-based FACS for improving production phenotypes: riboflavin overproduction in *Yarrowia lipolytica*. *Metabolic engineering*, 47, pp.346-356.
- Walters, C., Wheeler, L. and Stanwood, P.C., 2004. Longevity of cryogenically stored seeds. *Cryobiology*, 48(3), pp.229-244.
- Wan, B., Zhang, Q., Ni, J., Li, S., Wen, D., Li, J., Xiao, H., He, P., Ou, H.Y., Tao, J. and Teng, Q., 2017. Type VI secretion system contributes to Enterohemorrhagic *Escherichia coli* virulence by secreting catalase against host reactive oxygen species (ROS). *PLoS pathogens*, 13(3), p.e1006246.

- Wan, H.M., Chang, B.Y. and Lin, S.C., 2002. Anchorage of cyclodextrin glucanotransferase on the outer membrane of *Escherichia coli*. *Biotechnology and bioengineering*, 79(4), pp.457-464.
- Wang, B.L., Ghaderi, A., Zhou, H., Agresti, J., Weitz, D.A., Fink, G.R. and Stephanopoulos, G., 2014. Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. *Nature biotechnology*, 32(5), pp.473-478.
- Wang, J.D. and Levin, P.A., 2009. Metabolism, cell growth and the bacterial cell cycle. *Nature Reviews Microbiology*, 7(11), p.822.
- Wang, L., Flanagan, L.A., Jeon, N.L., Monuki, E. and Lee, A.P., 2007. Dielectrophoresis switching with vertical sidewall electrodes for microfluidic flow cytometry. *Lab on a Chip*, 7(9), pp.1114-1120.
- Wang, Y., Pati, S. and Schreiber, M., 2018. Cellular therapies and stem cell applications in trauma. *The American Journal of Surgery*, 215(5), pp.963-972.
- Waters, C.M. and Bassler, B.L., 2005. Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.*, 21, pp.319-346.
- Weinhold, B., 2006. Epigenetics: the science of change.
- Wellen, K.E. and Thompson, C.B., 2010. Cellular metabolic stress: considering how cells respond to nutrient excess. *Molecular cell*, 40(2), pp.323-332.
- Wilson, D.S. and Keefe, A.D., 2000. Random mutagenesis by PCR. *Current protocols in molecular biology*, 51(1), pp.8-3.
- Wilson, K., Webster, S.P., Iredale, J.P., Zheng, X., Homer, N.Z., Pham, N.T., Auer, M. and Mole, D.J., 2017. Detecting drug-target binding in cells using fluorescence-activated cell sorting coupled with mass spectrometry analysis. *Methods and Applications in Fluorescence*, 6(1), p.015002.

- Witvliet, M.H., Burns, D.L., Brennan, M.J., Poolman, J.T. and Manclark, C.R., 1989. Binding of pertussis toxin to eucaryotic cells and glycoproteins. *Infection and immunity*, 57(11), pp.3324-3330.
- Wu, T.H., Chen, Y., Park, S.Y., Hong, J., Teslaa, T., Zhong, J.F., Di Carlo, D., Teitell, M.A. and Chiou, P.Y., 2012. Pulsed laser triggered high speed microfluidic fluorescence activated cell sorter. *Lab on a Chip*, 12(7), pp.1378-1383.
- Wylot, B., Konarzewska, K., Bugajski, L., Piwocka, K. and Zawadzka, M., 2015. Isolation of vascular endothelial cells from intact and injured murine brain cortex—technical issues and pitfalls in FACS analysis of the nervous tissue. *Cytometry Part A*, 87(10), pp.908-920.
- Xu, L.M., Zhao, J.Z., Liu, M., Yin, J.S., Liu, H.B. and Lu, T., 2016. Recombinant scFv antibodies against infectious pancreatic necrosis virus isolated by flow cytometry. *Journal of virological methods*, 237, pp.204-209.
- Yang, J.H., Wright, S.N., Hamblin, M., McCloskey, D., Alcantar, M.A., Schrübbbers, L., Lopatkin, A.J., Satish, S., Nili, A., Palsson, B.O. and Walker, G.C., 2019. A white-box machine learning approach for revealing antibiotic mechanisms of action. *Cell*, 177(6), pp.1649-1661.
- Yang, X.S., 2019. *Introduction to algorithms for data mining and machine learning*. Academic press.
- Yim, S.S., Bang, H.B., Kim, Y.H., Lee, Y.J., Jeong, G.M. and Jeong, K.J., 2014. Rapid isolation of antibody from a synthetic human antibody library by repeated fluorescence-activated cell sorting (FACS). *PLoS one*, 9(10), p.e108225.
- Yuan, D., Chen, L., Li, M., Xia, H., Zhang, Y., Chen, T., Xia, R., Tang, Q., Gao, F., Mo, X. and Liu, M., 2015. Isolation and characterization of circulating tumor cells from human gastric cancer patients. *Journal of cancer research and clinical oncology*, 141(4), pp.647-660.
- Yumlu, S., Stumm, J., Bashir, S., Dreyer, A.K., Lisowski, P., Danner, E. and Kühn, R., 2017. Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9. *Methods*, 121, pp.29-44.

- Zavatti, V., 2014. Monitoring of an Antigen Manufacturing Process Using Fluorescence (Master's thesis, University of Waterloo).
- Zavatti, V., 2019. Application of Flow Cytometry and Fluorescence Spectroscopy to Monitor and Predict the Fermentation Activity in a Vaccine Manufacturing Process.
- Zavatti, V., Budman, H., Legge, R.L. and Tamer, M., 2020. Investigation of the effects of oxidative stress-inducing factors on culturing and productivity of *Bordetella pertussis*. *Biotechnology Progress*, 36(1), p.e2899.
- Zeng, W., Guo, L., Xu, S., Chen, J. and Zhou, J., 2020. High-throughput screening technology in industrial biotechnology. *Trends in biotechnology*, 38(8), pp.888-906.
- Zhao, R., Jiang, S., Zhang, L., & Yu, Z. (2019). Mitochondrial electron transport chain, ROS generation and uncoupling (Review). *International Journal of Molecular Medicine*, 44, 3-15.
- Zheng, W., Zhao, Y., Zhang, M., Wei, Z., Miao, K. and Sun, W., 2009. Oxidative stress response of *Inonotus obliquus* induced by hydrogen peroxide. *Medical mycology*, 47(8), pp.814-823.
- Zhou, M., Diwu, Z., Panchuk-Voloshina, N. and Haugland, R.P., 1997. A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. *Analytical biochemistry*, 253(2), pp.162-168.
- Zimmermann, N., McBride, M.L., Yamada, Y., Hudson, S.A., Jones, C., Cromie, K.D., Crocker, P.R., Rothenberg, M.E. and Bochner, B.S., 2008. Siglec-F antibody administration to mice selectively reduces blood and tissue eosinophils. *Allergy*, 63(9), pp.1156-1163.
- Zou, J., Huang, X., Wu, L., Chen, G., Dong, J., Cui, X. and Tang, Z., 2015. Selection of intracellularly functional RNA mimics of green fluorescent protein using fluorescence-activated cell sorting. *Journal of molecular evolution*, 81(5), pp.172-178.

Appendices

Appendix A Method of Characteristics

After applying the assumptions for the system of interest in this work (see Chapter 5.2), the population balance model is shown below:

$$\frac{\delta N(\mathbf{x}, t)}{\delta t} + \nabla_{\mathbf{x}}(\mathbf{r}(\mathbf{x}, S)N(\mathbf{x}, t)) + \gamma(\mathbf{x}, t)N(\mathbf{x}, t) - 4\gamma(2\mathbf{x}, t)N(2\mathbf{x}, t) = 0 \quad (\text{A.1})$$

This first order PDE has initial conditions $N(\mathbf{x}, 0) = N_o(\mathbf{x})$ at $t = 0$. Eqn (A.1) can be expanded ($\nabla_{\mathbf{x}}$) into the form:

$$\begin{aligned} \frac{\delta N(\mathbf{x}, t)}{\delta t} + \mathbf{r}(\mathbf{x}, S) \frac{\delta N(\mathbf{x}, t)}{\delta \mathbf{x}} + \frac{d\mathbf{r}(\mathbf{x}, S)}{d\mathbf{x}} N(\mathbf{x}, t) + \gamma(\mathbf{x}, t)N(\mathbf{x}, t) \\ - 4\gamma(2\mathbf{x}, t)N(2\mathbf{x}, t) = 0 \end{aligned} \quad (\text{A.2})$$

It can be observed that $\frac{\delta N(\mathbf{x}, t)}{\delta t} + \mathbf{r}(\mathbf{x}, S) \frac{\delta N(\mathbf{x}, t)}{\delta \mathbf{x}}$ is a directional derivative in the direction of the vector $(\mathbf{r}(\mathbf{x}, t), 1)$ in the (\mathbf{x}, t) plane. Therefore, all the direction vectors in the (\mathbf{x}, t) plane can be plotted to obtain a direction field. In the field, curves $\mathbf{x} = \mathbf{C}_x(t)$ can be found by fitting the curves $\mathbf{C}_x(t)$ into this direction field obtained by solving the following system of ODEs:

$$\frac{d\mathbf{C}_x(t)}{dt} = \mathbf{r}(\mathbf{C}_x(t), t) \quad (\text{A.3})$$

with initial conditions of $\mathbf{C}_x(0) = \mathbf{x}_0$. The curves defined by $\mathbf{C}_x(t)$ are called characteristic curves which each curve representing how \mathbf{x} changes over time given \mathbf{x}_0 . Each property in \mathbf{x}_0 will have its own characteristic curve. Now, consider a solution $N(\mathbf{x}, t)$ of the initial value problem of (A.1), (A.2) and the characteristic curve $\mathbf{x} = \mathbf{C}_x(t)$ with $\mathbf{C}_x(t) = \mathbf{x}_0$. Restricting $N(\mathbf{x}, t)$ to the characteristic curve gives the function:

$$v(t) = N(\mathbf{C}_x(t), t) \quad (\text{A.4})$$

Using the chain rule, $\frac{dv}{dt}$ can be expanded as:

$$\frac{dv}{dt} = \frac{\delta N(\mathbf{C}_x(t), t)}{\delta \mathbf{C}_x(t)} \frac{d\mathbf{C}_x(t)}{dt} + \frac{\delta N(\mathbf{C}_x(t), t)}{\delta t} \quad (\text{A.5})$$

Eqn (A.5) can be substituted into Eqn (A.2) to give the following equation:

$$\frac{dv(t)}{dt} + \frac{d\mathbf{r}(\mathbf{C}_x(t), S)}{d\mathbf{C}_x(t)} v(t) + \gamma(\mathbf{C}_x(t), t)v(t) - 4\gamma(2\mathbf{C}_x(t), t)v_2(t) = 0 \quad (\text{A.6})$$

with initial conditions of $v_o = N_o(\mathbf{x}_0)$. Since $N(2\mathbf{x}, t)$ would also be subjected to the method of characteristics, $v_2(t) = N(2\mathbf{C}_x(t), t)$. The first order PDE can be solved for using a family of first order ODEs in the form of Eqn (A.6). To consolidate Eqn (A.6) into Eqn (5.9), $\frac{d\mathbf{r}(\mathbf{C}_x(t), S)}{d\mathbf{C}_x(t)} v(t) = 0$ based on the reasoning provided in section 5.2.2 in Chapter 5.

Appendix B Population Balance Model and Experimental Contours

The contour distributions for the remaining conditions presented in Chapter 5 are shown in Figures B-1 to B-6.

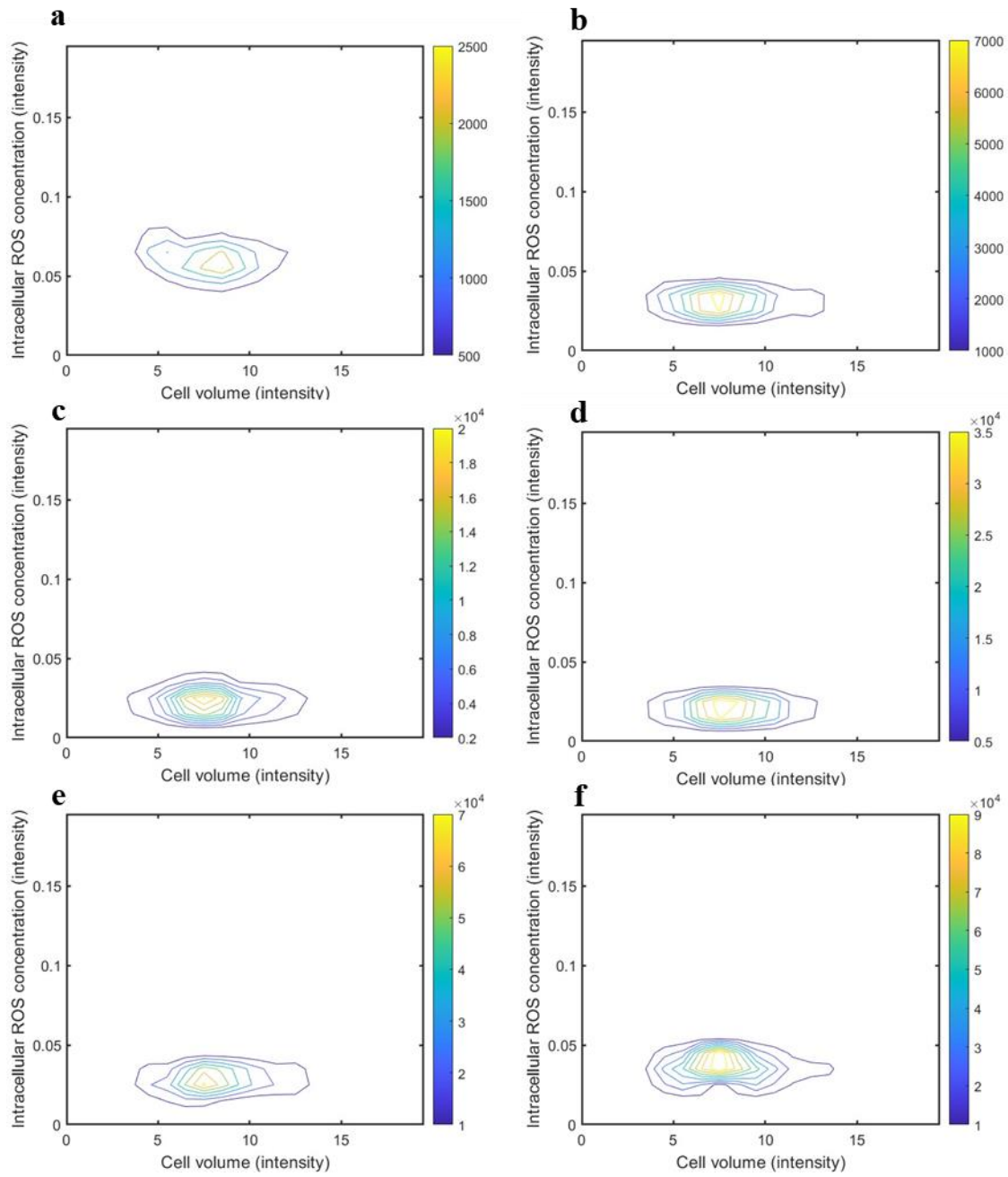


Figure B-1: Model flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

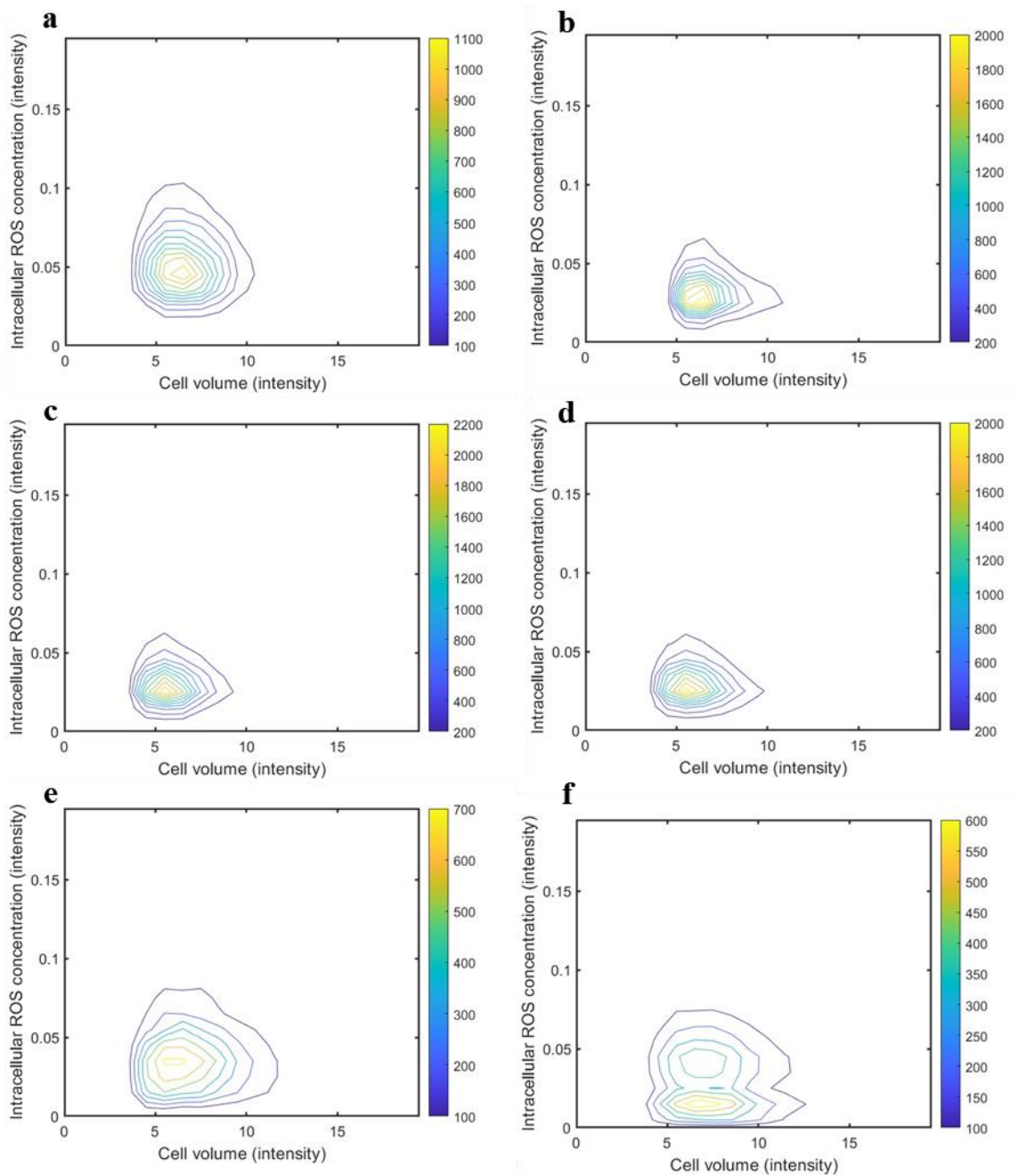


Figure B-2: Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 12.4 g/L glutamate in the media and starting optical density of 0.12 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

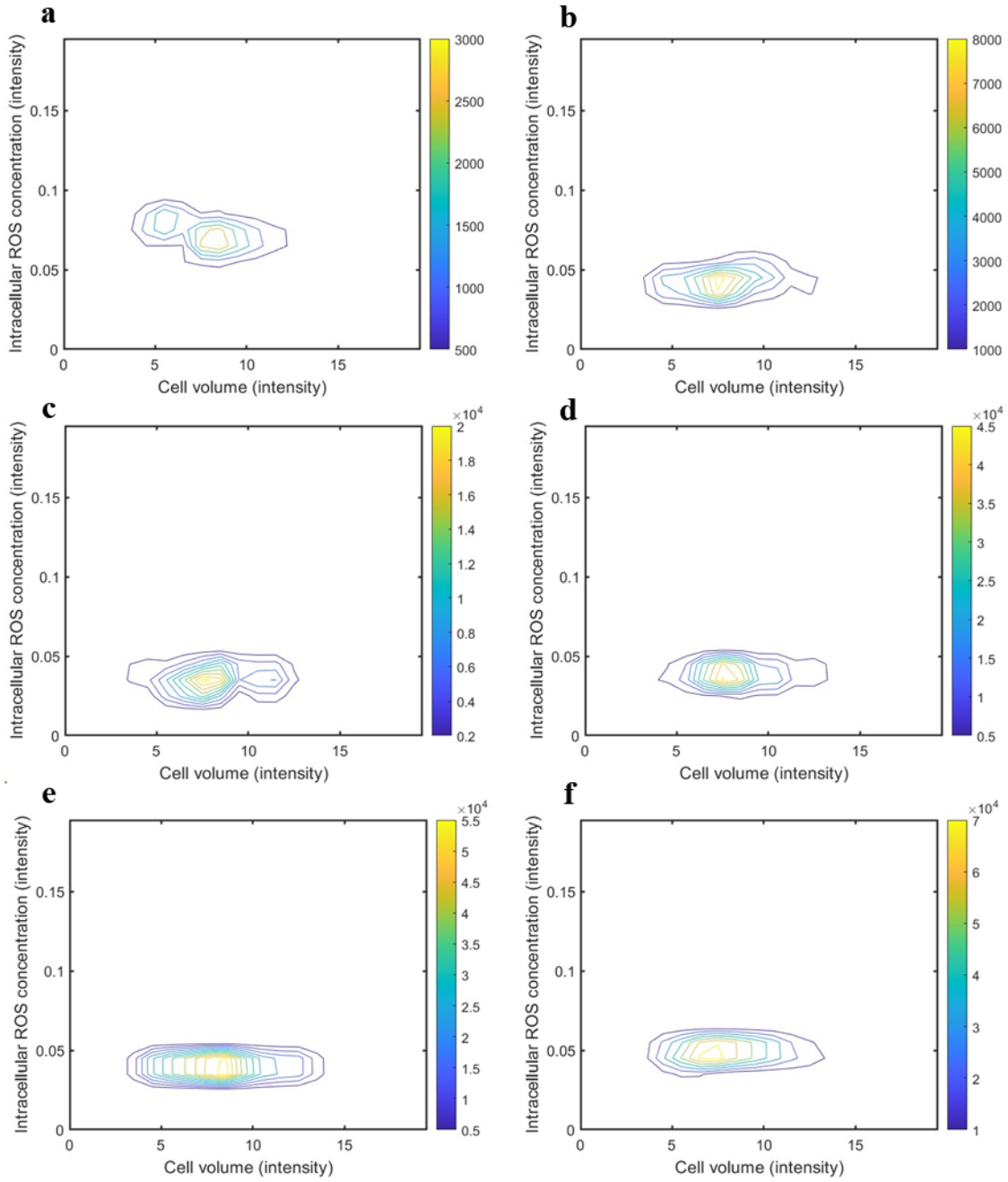


Figure B-3: Model flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

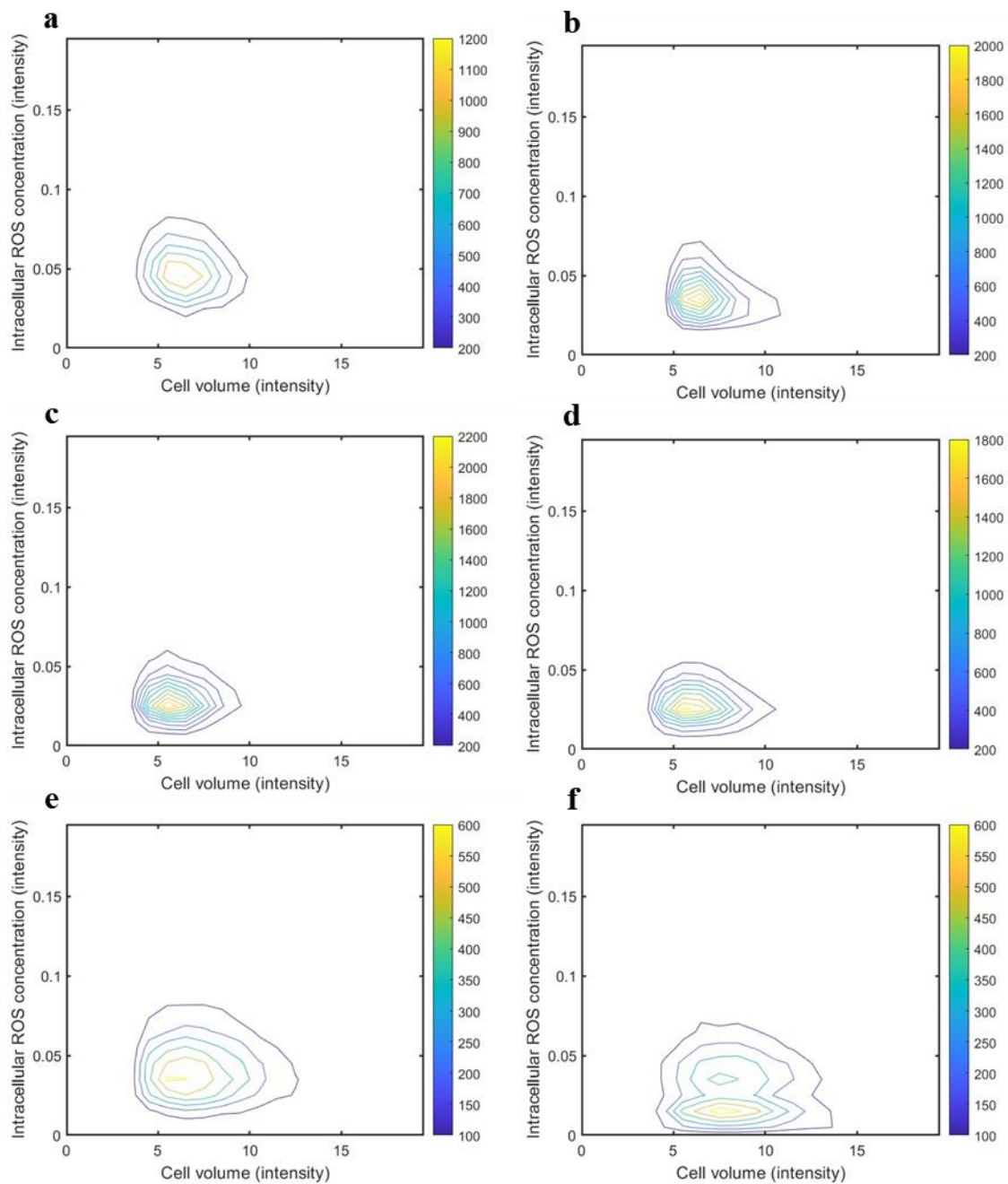


Figure B-4: Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 10.1 g/L glutamate in the media and starting optical density of 0.13 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

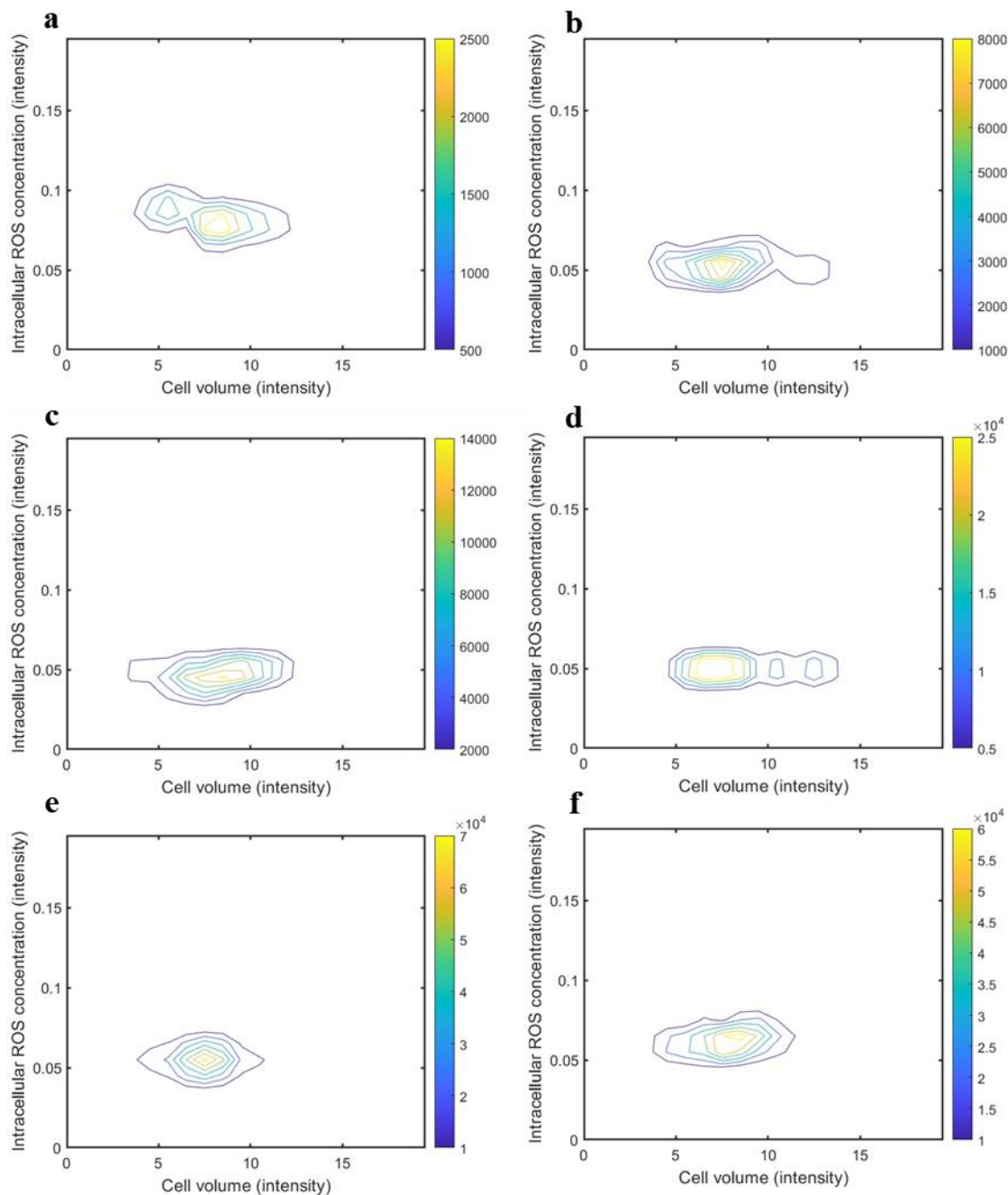


Figure B-5: Model validation of flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 8.6 g/L glutamate in the media and starting optical density of 0.14 at (a) 6, (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.

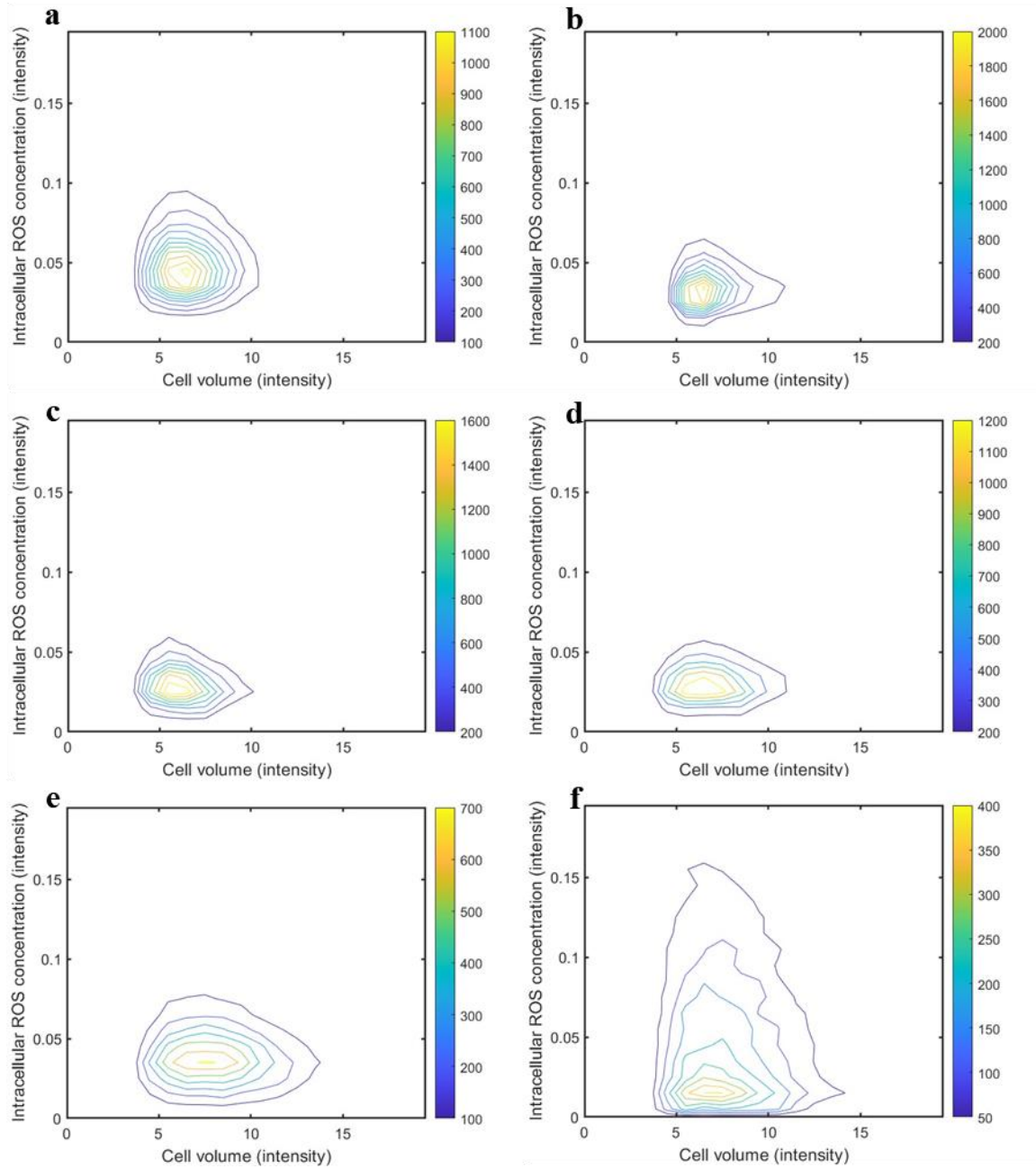


Figure B-6: Experimental flow cytometry distribution contours of cell volume intracellular ROS concentration for a *B. pertussis* shake flask culture with 8.6 g/L glutamate in the media and starting optical density of 0.14 at (a) 6 , (b) 12, (c) 18, (d) 24, (e) 30 and (f) 36 hours.