The effect of environmental conditions on growth and phenotype switching in *Mycobacterium smegmatis*

By

Davina-Nelson Apiyo

Thesis presented in partial fulfilment
of the requirements for the Degree

of

MASTER OF ENGINEERING
(CHEMICAL ENGINEERING)

in the Faculty of Engineering
at Stellenbosch University

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Supervisor:

Dr T. M. Louw

Co-Supervisors:

Dr J. M. Mouton
Prof. S. L. Sampson

March 2020

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: *March 2020*

PLAGIARISM DECLARATION

1. Plagiarism is the use of ideas, material and other intellectual property of another's work and to present is as my own.

2. I agree that plagiarism is a punishable offence because it constitutes theft.

3. I also understand that direct translations are plagiarism.

4. Accordingly, all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism.

5. I declare that the work contained in this assignment, except where otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

Initials and surname: D APIYO

Date: 18.02.2020

ABSTRACT

Mycobacterium tuberculosis can exist within a host in a seemingly dormant state, in which it can tolerate antibiotic challenge. This non-heritable survival mechanism is thought to be the cause of latent tuberculosis (TB) infection. The viable but non-replicating population exhibits a phenotype known as antibiotic tolerance, with these cells being referred to as 'persisters'. The intracellular environment of alveolar macrophages (a key habitat of *M. tuberculosis* bacilli) is detrimental to the survival of the bacteria, constituting antimicrobial effectors such as hypoxia, nutrient deficiency, nitrosative stress and acidic stress. Persistence arises when the bacilli can tolerate these host defence mechanisms. This study sought to investigate replication dynamics and phenotype switching of bacteria, but under *in vitro* environmental stresses — nutrient deficiency and acidic stress. *Mycobacterium smegmatis* (a non-pathogenic, fast growing *Mycobacterium* species) was used as a model for *M. tuberculosis*. To determine the response of *M. smegmatis* to the various growth stresses, two methods were utilized: mathematical modelling (for parameter estimation and prediction of cellular growth) and fluorescence dilution (FD — for examination of the persister population, using a dual-fluorescence replication reporter system).

Results from the experimental studies indicated that the fluorescence reporter was suitable for measuring bacterial replication dynamics for up to four generations, when compared to other conventional techniques such as optical density. Under acidic conditions (pH 4.6 media), the acute decline in bacterial growth, based on the calculated mean fluorescence intensity, was apparent. Under circumstances of nutrient deficiency, results from the reporter were inconclusive, since its minimum intensity had been reached before the cells in the culture could be influenced by the stresses (from t = 16 h). As for mathematical modelling, optimization of the relevant growth parameters was done through a weighted non-linear least squares approach. Quantitative comparison of the optimized model to the validation data — by calculating the normalized root mean squared error (NRMSE) — revealed a relatively good fit for the pH. For each of the five validation experiments (with varying environmental conditions), the NRMSE of the pH was 13.40%, 12.67%, 13.96%, 5.28% and 3.38%.

Based on these results, we conclude that the developed mathematical model was able to predict bacterial growth under diverse conditions, and that the reporter could accurately measure mycobacterial replication. Nonetheless, model predictability (more so for the biomass and ammonia variables) could be improved, by adding biochemical elements that influence the uptake and utilization of the substrates. It would also be beneficial to apply the model to slow-growing mycobacteria, to gauge its suitability in predicting *M. tuberculosis*

growth. Finally, FD results under nutrient-deficient conditions could be made more conclusive by withdrawing the inducer of the far-red fluorescent protein at a later timepoint during the experiment. This makes the comparison of replication dynamics to the normal case more perceptible.

OPSOMMING

Mycobacterium tuberculosis kan bestaan binne ’n gasheer in ’n oënskynlike dormante toestand, waar dit bestand is teen antibiotika. Daar word gedink dat hierdie onoordraaglike oorlewingsmeganisme die oorsaak van latente tuberkulose (TB) -infeksie is. Die uitvoerbare maar nie-repliserende populasie vertoon ’n fenotipe bekend as antibiotika toleransie, waar hierdie selle na verwys word as ‘vashoudendes’. Die intrasellulêre omgewing van alveolêre makrofagos (’n sleutelhabitat van *M. tuberculosis* basille) is nadelig tot die oorlewing van die bakterieë, wat antimikrobiale effektors bewerk soos hipoksie, nutriënttekort, stikstofstres en suurstress. Vashoudendheid kom voor wanneer die basille hierdie gasheer se verdedigingsmeganismes kan verdra. Hierdie studie het beoog om replikasiedinamieka en fenotipe wisseling van bakterieë te ondersoek, maar onder *in vitro* omgewingstres — nutriënttekort en suurstres. *Mycobacterium smegmatis* (’n nie-patogeniese, vinnig-groeiende *Mycobacterium*-spesie) is gebruik as ’n model vir *M. tuberculosis*. Om die respons van *M. smegmatis* op die verskillende groei stresse te bepaal, is twee metodes gebruik: wiskundige modellering (vir parameterberaming en voorspelling van sellulêre groei) en fluoressente verdunning (FD — vir eksaminering van die vashoudendes se populasie, deur ’n tweevoudige-fluoressente replikasie rapporteerderstelsel te gebruik).

Resultate van die eksperimentele studies het aangedui dat die fluoressente rapporteerder gepas was vir die afmeting van bakteriële replikasiedinamieka vir tot vier generasie, wanneer dit vergelyk word met ander konvensionele tegnieke soos optiese digtheid. Onder suur kondisies (pH 4.6 media) was die akute afneming in bakteriële groei, gebaseer op die berekende gemiddelde fluoressente intensiteit, duidelik. Onder omstandighede van nutriënttekort, was die resultate van die rapporteerder onbeslis, aangesien sy minimum intensiteit bereik is voor die selle in die kultuur beïnvloed kon word deur stres (van t = 16 h). Wat wiskundige modellering betref, is die optimering van die relevante groeiparameters gedoen deur ’n geweegde nie-liniêre kleinste kwadrate benadering. Kwantitatiewe vergelyking van die geoptimeerde model met die validasie data — deur die genormaliseerde wortel gemiddeld vierkantsfout (NRMSE) te bereken — het ’n relatiewe goeie passing vir die pH bekend gemaak. Vir elk van die vyf validasie-eksperimente (met variërende omgewingskondisies) was die NRMSE van die pH 13.40%, 12.67%, 13.96%, 5.28% en 3.38%.

Gebaseer op hierdie resultate het ons tot die gevolgtrekking gekom dat die ontwikkelde wiskundige model die bakteriese groei onder ’n verskeidenheid kondisies kon voorspel, en dat die rapporteerder die *Mycobacteria* se replikasie akkuraat kon meet. Nietemin,

modelvoorspelbaarheid (meer so vir die biomassa en ammoniaveranderlikes) kon verbeter word, deur biochemiese elemente wat die invloed van die opname en gebruik van die substrate beïnvloed, by te voeg. Dit sal ook voordelig wees om die model op stadiggroeiende *mycobacteria* toe te pas, om sy gepastheid te bepaal om *M. tuberculosis*-groei te voorspel. Laastens, FD-resultate onder nutriënttekortkondisies kan meer beslissend gemaak word deur die induseerder van die verste-rooi fluoressente proteïen by ’n later tydspunt tydens die eksperiment te onttrek. Dit maak die vergelyking van replikasiedinamieka met die normale gevallestudie meer merkbaar.

ACKNOWLEDGEMENT

I would like to offer my sincere gratitude to the following individuals and institutions, without whom the study would not have been a success:

- My supervisor, Dr T.M. Louw, who I have had the pleasure of working with through the final year of my undergraduate study, as well as my Masters studies. I am thankful for his valuable insight, more so in the mathematical modelling section of the study. Without his guidance, I would have never realised my interest in mathematical modelling of biological systems
- My co-supervisor, Dr J.M. Mouton, for her assistance in the wet-lab experiments at the Medical campus (Tygerberg), and familiarization with flow cytometry and fluorescence dilution
- My co-supervisor, Prof. S.L. Sampson, who accepted to take me on as a co-supervised Masters student, despite my unfamiliarity with tuberculosis research. I am also extremely grateful for her willingness to let me carry out my experiments in her laboratory, and for providing funding for the required equipment (especially in buying the numerous Erlenmeyer flasks)
- Ms Anja du Toit (Department of Viticulture and Oenology, Stellenbosch University) for the timely analysis of my glycerol and ammonia samples
- Ms Charnay Anderson-Small (Central Analytical Facilities, Stellenbosch University) for the timely analysis of my phosphorus samples, and the elemental analysis of *M. smegmatis*
- The Host-Pathogen Mycobactomics group at Tygerberg, who welcomed me into their space whole-heartedly. A special thanks goes to Trisha Parbhoo, for her assistance in the flow cytometry analysis with FlowJo software
- The National Research Foundation (NRF), the Department of Process Engineering and The Division of Molecular Biology and Human Genetics, for funding the study

DEDICATION

I dedicate this work to my mother, who helped me realize my passion for the sciences from a young age, and who has always inspired me to achieve greater heights.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

GENERAL	
ATP	Adenosine triphosphate
CDW	Cell dry weight
CFU	Colony forming unit
CO_2	Carbon dioxide
%DO	Dissolved oxygen
DAEs	Differential algebraic equations
DNA	Deoxyribonucleic acid
FD	Fluorescence dilution
FSC	Forward scatter
GFP	Green fluorescent protein
hip	High persistence
HCl	Hydrochloric acid
HNO_3	Nitric acid
Hyg	Hygromycin
HSR	Headspace ratio
ICP-OES	Inductively coupled plasma optical emission spectroscopy
INH	Isoniazid
Kan	Kanamycin
KatG	Catalase-peroxidase
KH_2PO_4	Monopotassium phosphate
LB	Luria-Bertani
MBC	Minimum bactericidal concentration
MDK	Minimum duration of killing
MDR-TB	Multidrug-resistant tuberculosis
MFI	Mean fluorescence intensity
$MgSO_4$	Magnesium sulphate
MIC	Minimal inhibitory concentration
MIRIAM	Minimum Information Requested in the Annotation of Models
MTBC	*Mycobacterium tuberculosis* complex
N	Number of generations
NADH	Nicotinamide adenine dinucleotide
NaOH	Sodium hydroxide
$(NH_4)_2SO_4$	Ammonium sulphate
nm	Nanometre
NO	Nitric oxide
NRMSE	Normalized root mean squared error
NRP	Nonreplicating persistence
NTM	Non-tuberculous mycobacteria
OD	Optical density
ODEs	Ordinary differential equations
OTR	Oxygen transfer rate

OUR	Oxygen uptake rate
PBS	Phosphate buffered saline
ppm	Parts per million
ppGpp	Guanosine tetraphosphate
QC	Quality control
R^2	Correlation coefficient
RIF	Rifampicin
RMSE	Root mean squared error
rpm	Rotations per minute
SBML	Systems Biology Markup Language
SSC	Side scatter
SSE	Sum of squared errors
TA	Toxin/antitoxin
TB	Tuberculosis
TCD	Thermal conductivity detection
μ	Dynamic viscosity
VBNR	Viable but non-replicating
XDR-TB	Extensively drug-resistant tuberculosis
$ZnSO_4$	Zinc sulphate
GROWTH DYNAMICS	
C	Glycerol
CD	Carbon dioxide
k_z	Saturation constant of component *z*
N	Ammonia
X	Biomass
$Y_{X/Z}$	Yield coefficient of biomass produced per mass of substrate *z* utilized or per mass of metabolite *z* produced
μ	Specific growth rate
μ_{max}	Maximum specific growth rate
pH MODELLING	
H^+	Hydrogen ions
HCO_3^-	Bicarbonate ions
Mg^{2+}	Magnesium ions
Na^+	Sodium ions
OH^-	Hydroxide ions
pK_a	Logarithmic acid dissociation constant
SO_4^{2-}	Sulphate ions
Zn^{2+}	Zinc ions

1 INTRODUCTION

1.1 Background to Research

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis;* the most common area of infection is the lungs (pulmonary TB), though other sites within a host can be affected (extrapulmonary TB). It is estimated that 1.7 billion people are latently infected with the disease, with a small population (5-10%) eventually developing active TB in their lifetime (World Health Organization, 2018). Thirty high TB burden countries account for 87% of global TB incidences, of which South Africa is a part (ranking 8th, with 3% incidence) (World Health Organization, 2018).

Tuberculosis is an epidemic which is ranked as one of the top ten causes of death worldwide. It accounts for 95% of the deaths that occur in low and middle-income countries. It is also the leading cause of death among individuals infected with the human immunodeficiency virus (HIV), accounting for approximately 300,000 deaths in 2017 (World Health Organization, 2018). Much as the disease is treatable, albeit with a long regimen of at least 6 months with first-line drugs, it has become more difficult to do so over the years, due to the emergence of multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB) stains. MDR-TB bacilli are not susceptible to either isoniazid (INH) or rifampicin (RIF), the two most-powerful first-line anti-TB drugs. The WHO estimates that 558,000 new cases of RIF resistance arose in 2017 (World Health Organization, 2018). XDR-TB bacilli are not only insensitive to first-line drugs, but are also resistant to second-line drugs (fluoroquinolones such as levofloxacin and moxifloxacin). More often than not, the cause of the emergence of such resistant strains is due to mismanagement of TB treatment, be it by the clinician or the patient – in the form of using poor-quality drugs, incorrect prescriptions or patient noncompliance (World Health Organization, 2018).

Much of the focus regarding improving TB treatment has been on the development of novel antimicrobial drugs for drug-resistant strains. However, the tolerance of bacteria to anti-TB drugs is often overlooked. Tolerance occurs when the physiological status or metabolism of a bacteria is changed in the presence of an antibiotic, such that it stops growing, but is still able to survive (da Silva and Palomino, 2011). The bacteria then resume growth once the antibiotic has been removed from the host. Thus, it can survive prolonged exposure to a drug, despite genetic susceptibility to the drug (Balaban *et al*., 2013). This is unlike resistance, which is attributed to heritable genetic mutations within the bacteria. The bacterial cells that survive in this dormant state are referred to as ‘persisters’.

The mechanisms leading to persistence, and the physiological state of these cells remain unclear. This is due, in part, to the difficulties associated with trying to examine this small drug-tolerant population. There has been little acknowledgement for the clinical significance of these persisters, and only recently has the research into antibiotic persistence increased. This is based on possible evidence that this viable but non-replicating (VBNR) population may be the cause of recalcitrance in biofilms associated with chronic respiratory infections, and that antibiotic persistence may be a stepping stone to antibiotic resistance, when repeated doses of antibiotics are used to treat an infection (Fisher, Gollan and Helaine, 2017). Nevertheless, recent research has been able to (partially) elucidate the persister phenotype, at a single-cell level, using high-resolution techniques such as microfluidic devices or flow cytometry-based fluorescence dilution (FD, using a fluorescent reporter plasmid) (Fisher, Gollan and Helaine, 2017; Helaine *et al*., 2010; Mouton *et al*., 2016). These techniques make it possible to single out persisters from the rest of an isogenic bacterial population, from which transcriptomic studies can be done, to comprehend the mechanisms leading to this phenotype.

Infections caused by various mycobacterial species are often associated with biofilm development. Indeed, one of the most fascinating findings regarding *M. tuberculosis* is that it is able to form biofilms *in vitro* under specific conditions (Ojha *et al*., 2008). However, it has yet to be determined if such growth occurs within a host. Biofilms, which are colonies of surface-attached cells, are resistant to harsh environmental stresses and antibiotics, in comparison to cells in their planktonic form. They make it harder for an infection to be treated, and for the infection to be resolved altogether, the biofilm must be physically eradicated. Several mechanisms have been suggested for the antimicrobial resistance of biofilms, one of which is the presence of persisters within this structured community (Ojha *et al*., 2008; Ojha and Hatfull, 2012). With respect to *M. tuberculosis*, mycobacterial biofilm resistance is still poorly understood.

Understanding the pathogenesis and immunology of TB is no longer restricted to laboratory experiments dealing with animal (mouse, guinea pig, rabbit or non-human primate) models or human samples; researchers are taking a more multidisciplinary approach, incorporating computational and mathematical modelling approaches to these wet-lab experiments. With systems biology, data across various systems can be integrated to develop a complex model reflecting human biology, from which various hypotheses can be tested and virtual experiments can be run. This approach has been useful in describing within-host dynamics during *M. tuberculosis* infection, offering new details on observed phenomena (Kirschner *et al*., 2017).

This study sought to explore the response of *Mycobacterium smegmatis* to environmental stresses *in vitro* using planktonic cultures; the stresses included nutrient limitation and acidity of the media. The techniques used to investigate the effect of the stresses on the growth of the microorganism included mathematical modelling, FD and optical density (OD) measurements. The OD established differences in bacterial growth as per the various environmental conditions used in the wet-lab experiments. The proposed mathematical model was used to estimate the growth parameters of the microbe through regression, while FD was used to investigate bacterial replication dynamics and possible persister formation under stressed growth conditions.

Rather than directly studying the effects of the growth stresses on *M. tuberculosis*, *M. smegmatis* was used as a model organism. In the first instance, *M. tuberculosis* is a Category three human pathogen, whose experimentation would require the use of a biosafety level three laboratory. Such stringent precaution is not required when working with *M. smegmatis*, it being a non-pathogenic organism. Furthermore, *M. tuberculosis* grows very slowly in liquid media, with a doubling time of 22 h, unlike that of *M. smegmatis*, which takes 3.5 h (Shiloh and Champion, 2010; Reyrat and Kahn, 2001).

1.2 Research Aims and Objectives

The aims of the project were: (a) to estimate the growth parameters of *M. smegmatis*, for the purpose of growth prediction under various environmental conditions and (b) to investigate persister formation under stressed environmental conditions, using FD. To achieve these aims, the following objectives were addressed:

1. **Assessing bacterial growth under normal conditions**: Trends in bacterial growth and nutrient consumption were examined, to determine the variables that were suitable for use in the mathematical model
2. **Assessing bacterial growth under stressed conditions**: Trends in bacterial growth and nutrient consumption, based on the environmental stressor in question, were examined. Dilution of a fluorescent reporter (TurboFP635) was used to further interrogate bacterial growth kinetics, as a measure of persister formation
3. **Model development and parameter estimation**: The Monod expression was chosen to model the growth of *M. smegmatis*, based on the variables identified in Objective 1. Growth parameters were estimated using ordinary least squares, with experimental data gathered from the first two objectives

4. **Model validation**: To ascertain the predictability of the mathematical model, the simulations using the optimized growth parameters (from Objective 3) were compared to the experimental data of the validation experiments

1.3 Research Hypothesis and Questions

It is hypothesized that a persister population will emerge under growth-limiting conditions (which include carbon deficiency, nitrogen deficiency and acidic stress). It is further hypothesized that the proposed mathematical model will be able to reproduce the observed effects of varying environmental stresses on bacterial growth.

The research questions that would address the aforementioned project aims and objectives are as follows:

1. What environmental stressors are likely to result in the formation of persisters?
2. Are the estimated parameters able to predict the growth dynamics of the bacteria under diverse environmental conditions?

1.4 Significance of the Research

Results from this study will be pertinent in inferring the effect of varying experimental conditions on growth and persister formation in a *Mycobacterium* biofilm. The estimated growth parameters from the proposed mathematical model can be used in an individual-based biofilm model, from which the development of the biofilm under various environmental stresses can be predicted. The fluorescent plasmid used in this study could also be applied to observe phenotypic heterogeneity (i.e. metabolically active and dormant cells) as a result of a generated spatially heterogeneous microenvironment within its matrix.

This study will also be key in developing a robust mathematical model that can predict *M. smegmatis* growth under diverse environmental conditions, hence, minimizing the need for wet-lab experiments. The model will be submitted to an open-source repository (BioModels Database: http://www.ebi.ac.uk/biomodels/) which stores numerous curated computational models of biochemical and cellular systems, making it available for use by other systems biology research groups. This would further the investigation into mycobacterial growth and phenotype switching from a mathematical perspective.

2 LITERATURE REVIEW

2.1 Introduction

The following literature review gives a brief summary of the *Mycobacterium* genus, categorizing mycobacterial species based on their pathogenic and non-pathogenic nature. The concept of antibiotic tolerance is also introduced, explicating the differences between persistence and antibiotic resistance, and the ways in which persisters are generated. Further, factors that influence the emergence of persisters are expounded on, as well as the mechanisms that enable this population to evade antibiotic challenge or environmental stresses. Also, the role that persisters play in the recalcitrance of biofilms is delved into.

Finally, efforts made in advancing our knowledge on TB infection and treatment, based on mathematical and computational modelling, are discussed.

2.2 Overview of the *Mycobacterium* Genus

Mycobacteria are rod-shaped obligate aerobes. Species belonging to the *Mycobacterium* genus have cell walls that are rich in lipids (specifically, mycolic acids) that resist conventional staining techniques (such as the Gram stain). Moreover, the cells resist decolourization with an acid or alcohol, when stained by basic dyes – hence, their 'acid-fast' nature (Brooks *et al*., 2013). Consequently, the Ziehl-Neelsen technique is used for identification of such acid-fast bacteria. This genus includes disease-causing microorganisms and environmental microorganisms that are broadly categorized into the following groups:

- ***Mycobacterium tuberculosis* complex (MTBC)**, which cause TB (or TB-like diseases) in humans and animals. Examples include *M. tuberculosis*, *Mycobacterium bovis* (including the vaccine strain *M. bovis* BCG), *Mycobacterium africanum* and *Mycobacterium canetti*. They are slow-growing microorganisms (doubling time between 18 and 24 hours), typically taking at least two weeks for the formation of visible colonies on laboratory media. Bacterial growth is enhanced with increased carbon dioxide (CO_2) tension (Simner, Woods and Wengenack, 2016)
- ***Mycobacterium leprae***, which causes leprosy. This bacterium grows very slowly, doubling after every 14 days. *In vitro* cultivation of *M. leprae* has generally not been successful, making its diagnosis difficult (Bhat and Prakash, 2012)
- **Nontuberculous mycobacteria (NTM)**, often found in the natural environment. Such species do not typically cause diseases, unless in cases where a patient is immunocompromised (such as people infected with HIV/AIDS or those with prior or recurrent pulmonary infections). NTM can be either slow-growing (for instance,

Mycobacterium avium complex, *Mycobacterium kansasii* and *Mycobacterium marinum*) or fast-growing (such as the *Mycobacterium abscessus* complex, *Mycobacterium chelonae* and *Mycobacterium fortuitum*). *M. smegmatis* falls under this category, but is a saprophyte rarely associated with human illness

Historically, NTM were identified based on the Runyon classification, which grouped them in accordance to their growth rate and pigment produced. Group I (photochromogens) constitute bacteria with slow growth and a pigment upon light exposure; Group II (scotochromogens) describe NTM that grow slowly, and produce a pigment in the dark; Group III (nonchromogens) have NTM that are nonpigmented slow-growers; Group IV (rapid growers) describe non-pigmented fast-growing mycobacteria (Simner, Woods and Wengenack, 2016). Such phenotypic methods of identification are now rarely used; instead, molecular methods – genetic sequencing, mass spectrometry and deoxyribonucleic acid (DNA) probes – are currently utilized.

2.3 Antibiotic Persistence

During drug exposure, a subpopulation of an isogenic microbial population may survive the bactericidal effects of a drug, while the rest of the microbes are killed off. The surviving subpopulation is thought to undergo metabolic quiescence; this serves as a survival mechanism, as inactive cells are less likely to be killed by antimicrobial agents. This phenotype switch, allowing a population subset to survive antibiotic exposure, is referred to as antibiotic persistence.

Persistence differs from antibiotic resistance in that persistence is a phenotype which is not transferred to subsequent progeny, unlike resistance, which is attributed to heritable genetic mutations. With resistance, cells can grow at high concentrations of an antibiotic, notwithstanding the duration of exposure. This makes resistance easily quantifiable, based on the minimum inhibitory concentration (MIC) of an antibiotic. Persistence, on the other hand, exhibits a cellular population that does not replicate in the presence of an antibiotic, and has no similar quantitative metric, though one study proposes quantifying persistence based on the minimum duration of killing (MDK) of an antibiotic (Brauner *et al*., 2016). The MDK involves measuring the duration at which a certain percentage of a bacterial population is killed, at antibiotic concentrations far exceeding the MIC, with a tolerant strain taking a longer duration than a susceptible strain (Brauner *et al*., 2016). Also, this study goes further to differentiate bacterial tolerance and persistence – two terms that are often used interchangeably in literature. As per the authors, tolerance is an attribute of an entire bacterial population which

describes survival to transient antibiotic exposure at concentrations that are higher than the MIC, while persistence, though similar, is specific to a subpopulation of bacteria.

Despite the differences in the aforementioned mechanisms of bacterial survival to antibiotic exposure, it is theorised that tolerance (and persistence) may act as an intermediate step in the development of heritable drug resistance (Cohen, Lobritz and Collins, 2013). This association was proposed based on how bacteria respond to continuous cycles of intermediate antibiotic exposure. This was illustrated with *Escherichia coli*, intermittently exposed to sub-lethal doses of ampicillin (Levin-Reisman *et al*., 2017). The initial observation was that tolerance to ampicillin was achieved, after three or four cycles of antibiotic exposure, by the cells extending their lag phase. Subsequent exposures (between seven and seventeen cycles) led to the increase in the MIC of the cultures by at least sevenfold, making the cells resistant. Antibiotic resistance was further substantiated through whole-genome sequencing, which revealed a mutation that occurred in one of the genes coding for a β-lactamase that is responsible for ampicillin resistance when overexpressed (Levin-Reisman *et al*., 2017).

The phenomenon of bacterial persistence was discovered whilst growing *M. tuberculosis* in nutrient-limited media (Loebel, Shorr and Richardson, 1933a; b). Starvation – as a result of culturing cells in phosphate-buffered saline (PBS) – reduced the metabolism of the bacilli to minimal levels but was restored when the starved bacilli were transferred into nutrient-replete media. This phenotype was also observed in *Staphylococcus pyogenes*; penicillin could never fully sterilize *in vitro* cultures, resulting in the conclusion that there existed a small subpopulation of cocci insensitive to the drug due to the possibility of the cells being in a temporary, dormant, non-dividing phase. To distinguish these cells from resistant types, the ‘persister’ term was coined (Bigger, 1944). Other characterisics of this dormant subpopulation, as surmised by the author, were that the production of persisters was not due to exposure to a bactericidal agent and the progeny of the persister population was just as sensitive to the antibiotic as the normal cells. Moreover, some of the cocci in the inoculum were predestined to be persisters, though the phenotype could also be induced through contact with a new environment (Bigger, 1944).

The generation of persisters within a bacterial population can be either environmentally induced or spontaneous:

- **Environmentally induced persistence** occurs as a result of an external trigger – this could be in the form of heat shock, oxidative stress, acidic stress, hypoxia or nutrient starvation (Cohen, Lobritz and Collins, 2013). It appears the cells can sense an unfavourable environment, resulting in a small subset becoming dormant. This is a ‘bet-hedging’ strategy, in which some cells prepare for adverse conditions by becoming

inactive, while the rest grow, at the risk of being killed (Balaban *et al*., 2004). Environmentally-induced persistence is characterised by the generation of non-growing cells during the stationary phase of growth, and the negligible switching of cells from an active to a dormant state during exponential growth, meaning that the number of persisters directly correlates to the number of stationary phase cells inoculated into a batch culture (Balaban *et al*., 2004)

- **Spontaneous persistence** occurs in the absence of an external trigger (Balaban *et al*., 2004; Cohen, Lobritz and Collins, 2013), with stochastic variation in gene expression being proposed as the reason for pre-existing phenotypic heterogeneity (Balaban *et al*., 2004). This is exemplified in toxin/antitoxin (TA) modules, which are two-gene operons encoding a stable toxin (which inhibits one of many metabolic processes occurring within the cells) and an unstable antitoxin (which neutralises the toxin when a stress has been initiated) (Cohen, Lobritz and Collins, 2013; Fisher, Gollan and Helaine, 2017). Stressed growth leads to degradation of the antitoxin and overexpression of a toxin, thus, increasing the frequency of persisters in a culture. This is observed in *E. coli*, upon exposure to a fluoroquinolone antibiotic, where the overexpression of the TisB toxin disrupts the proton motive force, which in turn drops cellular adenosine triphosphate (ATP) levels, leading to a dormant state (Lewis, 2010). Aside from TA modules, other genes – such as global regulators and stress-response components – can also be variably expressed (Cohen, Lobritz and Collins, 2013). One such instance of this is in the case of *M. smegmatis*, when exposed to INH (Wakamoto *et al*., 2013). The bacteria was found to persist by dividing in the presence of the drug, because of the stochastic variation in the expression of the enzyme catalase-peroxidase (KatG), which is responsible for activating INH. Infrequent KatG pulsing by the persisters made them less vulnerable to INH killing. It could be argued, however, that this particular mechanism is an example of phenotypic resistance, rather than spontaneous persistence. Phenotypic resistance (a type of antibiotic resistance) involves cellular growth in the presence of a drug due to non-heritable changes affecting the drug target or the toxicity of the bactericidal agent, which contrasts persistence (associated with metabolic quiescence). Spontaneous persisters aren't generated at the stationary phase of growth, but rather, are continuously produced during steady-state exponential growth. Their fraction remains constant, so long as growth under this condition is upheld (Balaban *et al*., 2019)

A bacterial population containing persister cells is characterised by a biphasic killing curve (Balaban *et al*., 2019), based on the observation that not all cells within an isogenic bacterial population are killed at the same rate. Upon antibiotic exposure, the initial result is a fast killing

rate of majority of the population, followed by a much slower rate of cell death, as a result of survival by the persisters (Dhar and McKinney, 2007).

2.3.1 Isolation and Characterization of Persisters

There is a scarcity in information concerning the physiological state of persisters, owing to the difficulty associated with identifying, isolating and characterizing this bacterial subpopulation. This is due to the cells being found in very low numbers, and undergoing very little replication (if any) (Mouton *et al*., 2016). Also, this phenotype is a metastable state which is lost when cells are sub-cultured (Dhar and McKinney, 2007). Nonetheless, various high-resolution techniques have been developed over the recent years to be able to isolate and study this subpopulation at a single-cell level (Balaban *et al*., 2004; Shah *et al*., 2006; Helaine *et al*., 2010).

With a microfluidic device, the growth rate of individual cells can be tracked based on the length of the linear microcolonies formed from the resulting progeny of the inoculum, in conjunction with time-lapse microscopy (Balaban *et al*., 2004). Using a high persistence (*hip*) mutant of *E. coli*, the authors were able to identify the persister population after antibiotic exposure and observed that the persisters exhibited lower growth rates in comparison to majority of the cells, prior to antibiotic exposure. This result indicated that persistence in this *E. coli hip* population was associated with an inherent, pre-existing heterogeneity of growth rates (Balaban *et al*., 2004).

Flow cytometry has been used to separate persisters from metabolically active cells; a previous study applied cell sorting based on expression of fluorescence of a degradable green fluorescent protein (GFP), with the hypothesis that the rate of protein synthesis is low in dormant cells (Shah *et al*., 2006). Bright cells, constituting the majority, could be separated from those that had no detectable fluorescence. The dim cells were further confirmed to be dormant after exposure of the bacterial population (*E. coli*) to ofloxacin, which has no effect on persisters, but kills off metabolically active cells that are either growing or non-growing (Shah *et al*., 2006). On separation, the gene expression profile of the persisters was examined. The only limitation to this method is that it may lead to a decrease in the number of persisters during sorting, as dilution of the subpopulation resuscitates the dormant cells (Lewis, 2010). Use of flow cytometry has also extended into fluorescence dilution (FD), which can identify non-replicating cells, and allow the investigation of replication dynamics of an entire bacterial population at a single cell level (Helaine *et al*., 2010). Researchers were able to develop a dual fluorescence reporter for *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) – consisting of a far-red inducible fluorescent protein (DsRed, used to measure bacterial replication) and a green fluorescent protein that is either constitutive or inducible (EGFP, used

for bacterial detection if constitutive, or bacterial replication if induced). This reporter could accurately measure bacterial replication when compared to more conventional methods such as colony forming unit (CFU) counts, and revealed a non-replicating bacterial population within murine macrophages (Helaine *et al*., 2010). This approach was adapted for *M. tuberculosis*, developing a pTiGc plasmid with a constitutive GFP and inducible far-red fluorescent protein, TurboFP635, exhibiting similar results to that of *S.* Typhimurium (Mouton *et al*., 2016). The results of this study were generated using an *M. smegmatis* strain containing the pTiGc plasmid.

2.3.2 Emergence of Persisters

Infection of the host with *M. tuberculosis* bacteria is largely through the aerosol route; the pathogen makes its way into the lungs, where the cells are engulfed by alveolar macrophages – phagocytes which are the first line of defence against microbial pathogens (Magombedze and Mulder, 2012). The environment within these macrophages is detrimental to the survival of the bacteria; bactericidal stress factors include hypoxia, nutrient deficiency and nitrosative stress. Latency occurs when the pathogen can tolerate these anti-microbial effectors within the macrophage, thus, persisting within the host for an extended period. In this state, there is no characteristic manifestation of active TB. However, when the immune system of the host is impaired, the granuloma (formed from the infected macrophages) is disrupted, resulting in the spread of the infection.

To investigate the physiology of these persisters, several *in vitro* models have been developed, to mimic the environmental conditions within the granuloma. Most of these models focus on one environmental condition which would result in the emergence of persisters, though it can be argued that the combination of multiple stress factors would result in a phenotype closest to what is observed clinically. To date, only one study has so far attempted to imitate *in vivo* latency through combining several stress factors (Deb *et al*., 2009). Of all three conditions, hypoxia is the most widely studied factor.

2.3.2.1 Hypoxia

The most widely used model for oxygen limitation is the Wayne model (Wayne and Hayes, 1996), which mimics the gradual depletion of oxygen within a granuloma, resulting in tolerance to anaerobiosis. This is achieved by agitation of the media in a vessel with a consistent ratio of the air volume to culture volume – referred to as a headspace ratio (HSR). In the study, *M. tuberculosis* was cultured in a sealed screw-cap test-tube, with gentle stirring (which was enough to ensure homogeneity of the population, without agitating the surface of the media, thus limiting the oxygen transfer rate (OTR)) and an HSR of 0.5, which limits the amount of

oxygen that can be utilized by the cells. Continuous cellular growth caused oxygen depletion in the HSR, and the cells slowly adapted to reduced oxygen levels. From this point, two stages of non-replicating persistence were seen; the first stage occurred when the dissolved oxygen saturation reached 1%. This stage was deemed microaerophilic, and was characterised by a lack of replication or DNA synthesis, though ATP generation was still maintained (Wayne and Hayes, 1996). The second stage occurred at a saturation of 0.06% and was deemed anaerobic. Decolourization of methylene blue confirmed hypoxia within the media. Exposure of nonreplicating persistence (NRP) stage 2 bacilli to metronidazole showed bacterial sensitivity to the drug; metronidazole requires reduction under hypoxic conditions for activity (Betts *et al*., 2002), and this made it possible for the agent to kill off the dormant bacilli. This model was subsequently applied to *M. smegmatis*, which also showed similar results under oxygen-limited conditions (Dick, Lee and Murugasu-Oei, 1998).

Genetically, the DosR regulon, induced under anaerobic conditions, has been singled out as the genetic program responsible for the survival of *M. tuberculosis* (Leistikow *et al*., 2010). Induction of this regulon results in switching of the bacteria's metabolism (restricting aerobic respiration due to an insufficient amount of oxygen present) and maintaining ATP levels and redox balances. Experimental studies made it apparent that the wild type (H37Rv) had a clear advantage in survival under anaerobic conditions, as compared to the DosR mutant. Not only did the mutant rapidly lose its viability under oxygen deficiency, but also had a poor recovery, when switched from an anaerobic to an aerobic state of growth; the cells that could recover from this environmental switch ended up being more fragile than those of the wild type (Leistikow *et al*., 2010).

2.3.2.2 Nutrient Limitation

As previously mentioned, the *in vitro* nutrient-deprivation model by Loebel and colleagues (1933a; b) exhibited bacterial persistence, through the retention of viability of *M. tuberculosis* bacilli despite being moved from a nutrient-replete environment into PBS solution. In a different study, *M. tuberculosis* isolates from lung lesions were compared to nutrient-starved *in vitro* bacilli, and were seen to have similar morphologies (Nyka, 1974). The dormant bacilli lost their acid fastness but remained viable for two years. In addition to this, their staining capacity recovered when the chromophobic cells were transferred into nutrient-rich media. These observations made the author postulate that persistence of the isolates was as a result of nutrient deficiency within the lung environment.

Based on Loebel's study, a simple nutrient deprivation model was developed, which was used to test for antibiotics that were effective against persisters (Betts *et al*., 2002). Log-phase cultures of *M. tuberculosis* H37Rv were inoculated into sealed, standing flasks containing

PBS, and incubated at 37ºC. CFU levels were observed over a period of six weeks; the counts remained constant (≈10^7 CFU/mL), denoting that culture viability was not lost. Also, when methylene blue was added to the sealed, standing flasks, the dye was not decolourised. This indicated the presence of oxygen (though, this does not imply that cells were not respiring; respiration is expected to occur – albeit at a reduced rate prior to entering the NRP state – leading to a lower oxygen saturation level). This is the fundamental difference between the persisters formed based on the Wayne model versus the nutrient deprivation model. In addition to this, the starved cells were not sensitive to the effects of metronidazole (persisters in the Wayne model were), confirming that the nutrient-deficient cultures were not in an oxygen-deprived state (Betts *et al*., 2002).

This stress factor was also investigated using fast-growing mycobacteria; nutrient-deprived stationary phase *M. smegmatis* bacilli were shown to be able to survive for 650 days in either carbon, nitrogen or phosphorus-starved cultures (Smeulders *et al*., 1999). Furthermore, carbon-limited stationary phase cells could withstand aggressive environmental stresses, including osmotic and acidic stress. (Smeulders *et al*., 1999). A different study looked into the response of *M. smegmatis* to nitrogen limitation (Anuchin *et al*., 2009), with the authors observing the emergence of a morphologically distinct ovoid form with a low metabolic activity, increased resistance to antibiotics and high temperatures, and the inability to form colonies. Cells with this phenotype could revert back to the typical rod shape, after resuscitation in nutrient-rich media (Anuchin *et al*., 2009). This could imply that the ovoid cells were specialized dormant cells, with a morphology similar to bacilli isolated from people and animals with TB (Anuchin *et al*., 2009).

2.3.2.3 Nitric Oxide Stress

Exposure of *M. tuberculosis* to sub-toxic concentrations of nitric oxide (NO) causes the inhibition of bacterial respiration and replication (Voskuil *et al*., 2003). The physiological state brought about by exposure to NO is very similar to that resulting from the depletion of oxygen. In fact, in both cases, the DosR response regulator induces a common set of 48 genes, which enable *M. tuberculosis* to survive under anaerobic conditions (Voskuil *et al*., 2003; Leistikow *et al*., 2010). Most of these genes are also induced by the bacterium *in vivo*, when inoculated into activated murine macrophages (Schnappinger *et al*., 2003), leading to the presumption that latent TB infection in a host may be as a result of limited cellular respiration, due to the presence of NO and low oxygen concentration in a granuloma. That notwithstanding, the mechanisms of the regulon, facilitating adaptation under anaerobic conditions, is still unknown.

2.3.2.4 Other Stress Factors

The only *in vitro* model to date, which attempts to mimic *in vivo* dormancy through the induction of multiple stresses, is that of Deb *et al*., (2009). Aside from nutrient deficiency (10% Dubos medium without glycerol) and low atmospheric oxygen (5% saturation), *M. tuberculosis* bacilli were also subjected to acidic pH (pH 5) and high atmospheric CO_2 (10% saturation). Under these conditions, the wild type (H37Rv) was observed to accumulate storage lipids (specifically triacylglycerol), and antibiotic exposure showed phenotypic drug resistance to RIF and INH. A deletion mutant lacking the triacylglycerol synthase (*tgs1*) gene, when exposed to the same conditions, displayed different results. There was no accumulation of triacylglycerol, and antibiotic tolerance to RIF was lost. This led to the hypothesis that antibiotic tolerance is associated with the accumulation of storage lipids, which was strongly supported by the fact that complementation of the deletion mutant restored its tolerance to RIF (Deb *et al*., 2009).

Acid tolerance by mycobacteria has also been examined (O'Brien *et al*., 1996); exposure of *M. smegmatis* to a sub-lethal adaptive acidic pH for an extended period of time led to a higher survival rate of cells that were subsequently exposed to a lethal pH, as compared to un-adapted cells.

2.3.3 Survival Mechanisms

Dormancy has been previously speculated to be the main survival mechanism through which bacterial persistence emerges. Bigger (1944) hypothesized that staphylococci persisters could tolerate the presence of penicillin due to being in a dormant, non-dividing phase, a physiological state that was later proven by a landmark study using a *hip E. coli* population (Balaban *et al*., 2004). Nonetheless, there has been increasing evidence throughout the years which refutes this claim; granted, the *E. coli* persisters (Balaban *et al*., 2004) do demonstrate metabolic quiescence when exposed to ampicillin, but it cannot be assumed that this is a universal mechanism of persistence across various bacterial species. Rather, it seems that active cellular processes are occurring in tandem with the reduction of the bacterial growth rate, hence, promoting persistence (Cohen, Lobritz and Collins, 2013). These active processes include NO synthesis, upregulation of efflux pumps and the stringent response, which are subsequently discussed.

2.3.3.1 Detoxification

An instance of an active intracellular detoxification mechanism is the synthesis of endogenous NO from arginine during antibiotic exposure; this mechanism is specific to Gram-positive bacteria, which contain bacterial NO synthases that facilitate NO generation (Gusarov *et al*., 2009). Pathogens such as *Bacillus subtilis*, *Staphylococcus aureus* and *Bacillus anthracis* are

able to survive exposure to a broad spectrum of antibiotics, as NO not only chemically modifies the antibiotic to make it less potent, but also inhibits oxidative stress induced by the antibiotic (Gusarov *et al*., 2009). This method also makes it possible for *B. subtilis* to exist as a co-culture with *Pseudomonas aeruginosa* in the soil. *P. aeruginosa* produces a natural toxin – pyocyanin – which eliminates the bacilli through the induction of oxidative stress, but *B. subtilis* counters the stress by producing NO (Gusarov *et al*., 2009).

Another detoxification mechanism is through the upregulation of efflux pumps. In this instance, transport proteins are used to remove intracellular toxic substances (antibiotics) into the external environment (Webber and Piddock, 2003). This mechanism is induced following macrophage infection, as is the case when *M. marinum* and *M. tuberculosis* were used to infect mouse and human macrophages (Adams *et al*., 2011). Infection with *M. marinum* led to INH and RIF tolerance, while *M. tuberculosis* infection resulted in RIF tolerance. Efflux pump inhibitors (such as verapamil, reserpine and thioridazine) were able to ascertain the role of efflux pumps in macrophage-induced drug tolerance – addition of verapamil together with the relevant antibiotics made *M. marinum* more susceptible to killing by INH and RIF (15.6-fold reduction and 9.2-fold reduction in survival, respectively), while *M. tuberculosis* became more susceptible to RIF (1.9-fold reduction in survival, for 144-hr intracellular growth) (Adams *et al*., 2011).

In a different study, a space-confined bioreactor (microdialyser) was created to mimic intracellular confinement – similar to macrophage infection (Luthuli, Purdy and Balagaddé, 2015). Experimental results from bacterial growth in a 200 picolitre (pL) microdialyser culture chamber indicated tolerance of *M. smegmatis* to antibiotic challenge with RIF. Eleven out of twelve cultures grown under the same conditions exhibited significant growth in the presence of the drug (for comparison, with the same experimental setup, the bacterium was susceptible to INH, ofloxacin and hygromycin). Efflux activity was investigated by culturing the cells in the presence of RIF and verapamil; mycobacterial growth was severely inhibited in the presence of both substances (Luthuli, Purdy and Balagaddé, 2015). Therefore, *M. smegmatis* tolerance to RIF was mediated, for the most part, by efflux mechanisms.

2.3.3.2 Stringent Response

A stringent response is an adaptive response exhibited by bacteria subjected to nutrient-deficient conditions (Jain, Kumar and Chatterji, 2006), which is characterised by a decline in the synthesis of proteins and nucleic acid, with a converse upregulation in amino acid synthesis and protein degradation (Chatterji and Ojha, 2001). Bacterial survival is associated with the accumulation of guanosine tetraphosphate (ppGpp), a global transcription regulator whose concentration in the cell's cytosol is maintained by two enzymes, namely RelA and

SpoT (Chatterji and Ojha, 2001), both of which are expressed by Gram negative bacteria. A single homolog of these enzymes, Rel_{Mtb}, is found in *M. tuberculosis*, and has been proven to be pertinent in the long-term survival of the pathogen under starvation conditions (Primm *et al*., 2000). A mutant strain was developed by deleting the *rel* gene, and subsequent comparison of the long-term survival of the mutant to the wild type (under conditions of *in vitro* starvation and nutrient depletion in normal media) showed that the persistence of the mutant was severely affected (Primm *et al*., 2000). A more recent study conducted by Dutta *et al.* (2019) not only showed a similar outcome to the experimental results of Primm *et al.* (2000), but also determined that the metabolomics profile and ATP concentrations in the Δ*rel* mutant strain under nutrient starvation were akin to the exponentially-growing wild type in nutrient-replete media. Its growth rate was not slowed down despite the lack of nutrients, making the mutant susceptible to killing by INH in both *in vitro* and *in vivo* (mouse model) conditions (Dutta *et al*., 2019). The minimum bactericidal concentration (MBC) – described as the minimum concentration of an antibiotic required to eliminate 99% of the initial bacterial population – of INH during *in vitro* starvation conditions increased by 512-fold for the wild type (from 0.06 µg/mL to 30.72 µg/mL), while that of the mutant remained constant (at 0.06 µg/mL), thus being easily killed off by a low concentration of INH (Dutta *et al*., 2019).

This strategy has been seen in other bacterial species such as *P. aeruginosa*, in which the stationary-phase wild type was tolerant to ofloxacin exposure, but the stationary-phase mutant (formed by disrupting *relA* and *spoT*) was not (Nguyen *et al*., 2011).

2.3.4 Biofilm Resistance to Antimicrobials

Biofilms are an aggregate of cells growing on a living or inert surface, which are enclosed in a self-produced exopolymer matrix (Nayak, 2015). They are notoriously recalcitrant to antimicrobial therapy and are thought to be the cause for the difficulty in treating various chronic infections. Biofilm growth provides a survival mechanism, protecting cells from environmental aggressions and antimicrobials.

The significance of *Mycobacterium* biofilms in healthcare is becoming apparent. NTM, under normal circumstances, are (for the most part) environmental saprophytes found in various ecosystems without public health implications (Esteban and García-Coca, 2018). Nonetheless, on rare occurrences, these mycobacteria can cause human infections that are either associated with biomaterials or causing recurring infections in individuals with underlying diseases such as cystic fibrosis, bronchiectasis and pneumoconiosis (Esteban and García-Coca, 2018; Faria, Joao and Jordao, 2015). Biomaterials are used in the medical field in making medical implants and prosthetics. These opportunistic pathogens develop biofilms within the devices, causing adverse infections. There are limited ways to eradicate the biofilms

– either impregnating the implant with antibiotics as a prophylactic measure (at the risk of certain microbes forming resistance to the antimicrobial agent) or removing the device altogether (Esteban and García-Coca, 2018).

As for the pathogenic mycobacteria, what is intriguing is the discovery that *M. tuberculosis* can form biofilms *in vitro* (Ojha *et al*., 2008). This not only opens up a new angle in researching the pathogenesis of the disease based on possible *in vivo* biofilm development, but also proposes biofilm-forming mechanisms as a potential drug target in TB treatment (Esteban and García-Coca, 2018).

Many biofilm susceptibility studies test for the killing effectiveness of antimicrobials on a pre-formed biofilm, rather than growth inhibition during antimicrobial exposure, hence, the basis of biofilm 'resistance' to antibiotics (Lewis, 2001). Several factors are purported to contribute to the recalcitrant nature of biofilms:

1. **Restricted penetration:** During biofilm development, bacteria secrete a matrix which constitutes polymeric substances such as polysaccharides, lipids and nucleic acid (Faria, Joao and Jordao, 2015). This exopolymer matrix is theorized to be the a contributing factor towards the resistance of a biofilm, and the virulence of the bacteria in the biofilm (Faria, Joao and Jordao, 2015). The matrix may act as a barrier that inactivates an antibiotic or restricts the permeation of certain large molecules (such as antimicrobial proteins) through the biofilm. Smaller antibiotic molecules can equilibrate across the matrix, and such agents only serve to postpone cellular death. Nonetheless, the heterogeneity of a biofilm cannot be ignored, as this can influence the diffusion of molecules across the matrix – areas with more cells will certainly restrict diffusion across the biofilm (regardless of molecule size) unlike areas less densely packed with cellular matter (Lewis, 2001)
2. **Decreased growth rate:** The biofilm has a unique physiology, containing layered structures that result in a nutrient and oxygen concentration gradient across its matrix. This means that certain areas in the biofilm will be nutrient- and oxygen-deficient, hence, resulting in cells either being starved or growing at a slow rate (Costerton, Stewart and Greenberg, 1999). Slowly growing or nongrowing cells are not susceptible to most antimicrobials (which target metabolically active cells)
3. **Persister cells:** This is an intriguing factor thought to be the cause of biofilm resistance owing to the observations made when investigating the killing effectiveness of fluoroquinolones (which can equilibrate across a biofilm). *P. aeruginosa* biofilms were challenged with ofloxacin over a wide concentration range, with the results indicating a distinct biphasic killing curve - characteristic of the presence of persisters (Spoering and Lewis, 2001). Cells in the biofilm were eliminated at low ofloxacin concentrations, but a

further increase in the concentration of the antibiotic had no effect on cell death. Rather, a 'plateau' was observed, denoting the invulnerability of persisters to killing by a fluoroquinolone (Spoering and Lewis, 2001). This phenotypic variation in response to antibiotic exposure was confirmed with *M. tuberculosis* biofilms (Ojha *et al*., 2008), by comparing viability of one attenuated strain exhibiting biofilm growth ($mc^2$7000) to another which was defective in biofilm formation ($mc^2$7025). Like the *P. aeruginosa* biofilms, a biphasic killing curve was seen for the *M. tuberculosis* $mc^2$7000 strain (with a survival of 0.01% after five days of RIF exposure), while the cells in the biofilm-defective $mc^2$7025 strain were rapidly killed (bacterial survival in this case was 0.00001% after five days of RIF exposure) (Ojha *et al*., 2008)

A commonly accepted notion is that biofilms are more resistant to killing by antimicrobial agents than planktonic cells; this is the case when the contrast is made against log-phase planktonic cells. This concept, however, does not necessarily hold for stationary-phase planktonic cells. In fact, a previous study determined that stationary-phase *P. aeruginosa* planktonic cells were somewhat more tolerant to antibiotic exposure than its biofilms (Spoering and Lewis, 2001). The microbe was exposed to four different classes of antibiotics – a fluoroquinolone, aminoglycoside, β-lactam and an oxidant – and in all cases, the stationary-phase planktonic cells exhibited a higher survival rate than biofilms and log-phase planktonic cells (Spoering and Lewis, 2001). Whether this is a general phenomenon with other bacterial species has yet to be substantiated.

2.4 Computational Modelling in the Field of TB Research

There has been a relatively recent drive towards employing a multidisciplinary approach to advance our understanding of TB infection and treatment. One way to do so is through systems biology, an approach that uses mathematical and computational modelling to comprehend and predict the dynamics of complex biological systems. In the case of TB, systems biology has been used to investigate macrophage infection, formation of a granuloma within the lung and antibiotic and vaccine efficacy, among other aspects of the disease (Kirschner *et al*., 2017). These models are advantageous, as they not only assist in understanding proposed mechanisms, but also in explaining observed phenomena and running virtual experiments (under circumstances in which laboratory work with animal samples is deemed too expensive or logistically challenging) (Kirschner *et al*., 2017). They are also key in linking various forms of biological data measured across different scales – these can range from molecular interactions to interactions between cells, tissues, body organs and the host population (Young, Stark and Kirschner, 2008).

Many researchers within the systems biology community are interested in advancing scientific concepts and hypotheses through collaborative efforts with various research groups, which involves disseminating their own published models for use by other individuals. This, however, is often hampered because of different modelling environments – be it different programming languages, operating systems, mathematical frameworks or user interfaces – making it a challenge to reproduce these models and their associated data. Consequently, they are rendered of little benefit to the scientific community (Dräger and Palsson, 2014). For this reason, there has been a concerted effort to standardize computational models, enhancing their exchange between different simulation and analysis tools (Hucka *et al*., 2003). This has resulted in the development of several machine-readable representation formats, including CellML, NeuroML and Systems Biology Markup Language (SBML), which has been the most successful standard model exchange format (Li *et al*., 2010). SBML is based on a hierarchical eXtensible Markup Language (XML), with chemical reactions being broken down into components that describe the model, such as compartments (reaction containers of finite volume), species, reactions, parameters, unit definitions and rules (mathematical expressions) (Hucka *et al*., 2003).

These representation formats are advantageous, in that they not only support the simulation of a model across multiple analysis tools, but also ensure that a model survives beyond the lifetime of the software used to establish them (SBML, 2012). They also facilitate the distribution of models for use in different software environments; these models can be accessed through free, online repositories and databases. As an example, models standardized with the SBML format are submitted to the BioModels Database, most commonly used in the systems biology field. They undergo a thorough curation process, which involves checking that the model corresponds to its source article and ascertaining its compliance to the Minimum Information Requested in the Annotation of Models (MIRIAM) (Le Novere *et al*., 2006). Once the relevant criteria are met, the models are made accessible to the public.

Computational modelling in TB research encompasses three facets: epidemiology of the disease, intracellular/extracellular mycobacterium growth dynamics – influenced by antibiotic challenge and environmental stress factors (under which this study falls) – and host-pathogen interactions during infection (Kirschner *et al*., 2017). Examples of models falling under the third facet include granuloma models, signalling and gene regulatory models and two-compartment (lung/lymph node) models (Magombedze, Dowdy and Mulder, 2013). Despite their advantages, researchers may be faced with challenges when dealing with some of these computational models – be it their complexity in development, analysis and simulation or being computationally intensive (Magombedze, Dowdy and Mulder, 2013). It may also be difficult to

integrate disparate timescales, particularly with regard to persistent infections, where molecular and cellular interactions within the host can take seconds or minutes, but development of latent infection can continue for decades (Young, Stark and Kirschner, 2008).

Mycobacterium growth dynamics when investigating antibiotic persistence is not only done on a single-cell level; insight into the development of persister populations is also being investigated from a biofilm perspective. Aside from the basis of biofilm persistence, other key aspects that are investigated when modelling the life cycle of a biofilm include the biofilm structure (and how it influences the function of a biofilm) and the contribution of genetic and phenotypic heterogeneity to biofilm formation (Cogan, Gunn and Wozniak, 2011). Roberts and Stewart (2005) were able to illustrate the switch of active cells into persister cells in areas within the biofilm that were predicted to be substrate limited. Moreover, the study modelled the dependence of the biofilm thickness on the fraction of persisters formed, and the regrowth of a biofilm (to nearly its original thickness) after antibiotic exposure, due to the reactivated persisters (Roberts and Stewart, 2005). In another study, individual-based modelling was used to compare the switching dynamics of susceptible and persister cells. Production of persisters was tested using three switching strategies – constant switches, substrate-dependent switches and antibiotic-dependent switches – which influenced the ability of the biofilm to develop, survive and recover from antibiotic exposure (Carvalho *et al*., 2018).

Many of the developed infection models are geared towards antibiotic resistance. There has been minimal contribution in the elucidation of latent TB from a mathematical perspective, as the mechanisms leading to quiescence of the bacteria and the subsequent increase in replication of a bacterial subpopulation (leading to reactivation of the disease) are still poorly understood. This makes it difficult to discern the type of patients with the highest risk in developing active TB (Kirschner *et al*., 2017). Attempts have been made to understand the emergence of persistence and latent TB infection mathematically; Magombedze and Mulder (2012) did so through the combination of *M. tuberculosis* growth dynamics, gene expression data and environmental stresses. The authors were able to predict the dynamics of active, latent and dormant bacilli depending on the level of the environmental stress factors (concentration of oxygen, nutrients and NO, all of which influence gene expression dynamics) in question, and simulate possible scenarios of latent TB development and reactivation within the host (Magombedze and Mulder, 2012).

2.5 Conclusion

Chapter Two expounded on the concept of antibiotic persistence, detailing its discovery and the research breakthroughs encountered, to understand the physiological nature of these

drug-tolerant bacteria. Several hypotheses concerning this ‘persister’ phenotype have, by and large, been attributed to the assumptions made from the study of one Joseph Bigger (1944), and for a while, were held to be factual for most bacterial species. However, thanks to improved technological methods that enable the isolation and characterization of persisters from a clonal population, certain aspects of its physiology (such as its various survival mechanisms) have been elucidated, bringing researchers a step closer to understanding the implications of persisters in the treatment of chronic infections.

Having explored the environmental stresses resulting in the generation of a persister population, what follows is to investigate the growth of *M. smegmatis* under diverse environmental conditions that are likely to cause persister formation. This entails examining the presence of persisters using a dual-fluorescent reporter plasmid, pTiGc (Mouton *et al*., 2016), and predicting the growth dynamics of the bacteria by developing a relevant mathematical model.

3 RESEARCH METHODOLOGY

3.1 Introduction

This chapter introduces and expounds on the research methods that were used to achieve the project objectives, which included experimental studies and mathematical modelling.

The experimental studies encompassed generating calibration graphs for the determination of the cell dry weight of *M. smegmatis* (the wild-type strain, mc^2155 was used), and investigating the effect of various environmental stresses on its growth (*M. smegmatis::*pTiGc, was used). Once the relevant variables in the planktonic cell cultures were measured, data was collected and applied in a regression analysis, using the developed mathematical model, thus, estimating the significant growth parameters. Afterwards, a second round of experiments was carried out, from which data was collected, for the purpose of comparison to the optimized simulation of the mathematical model. This would authenticate its predictive power.

A summary of the methodology is depicted in Figure 3-1, in the following page:

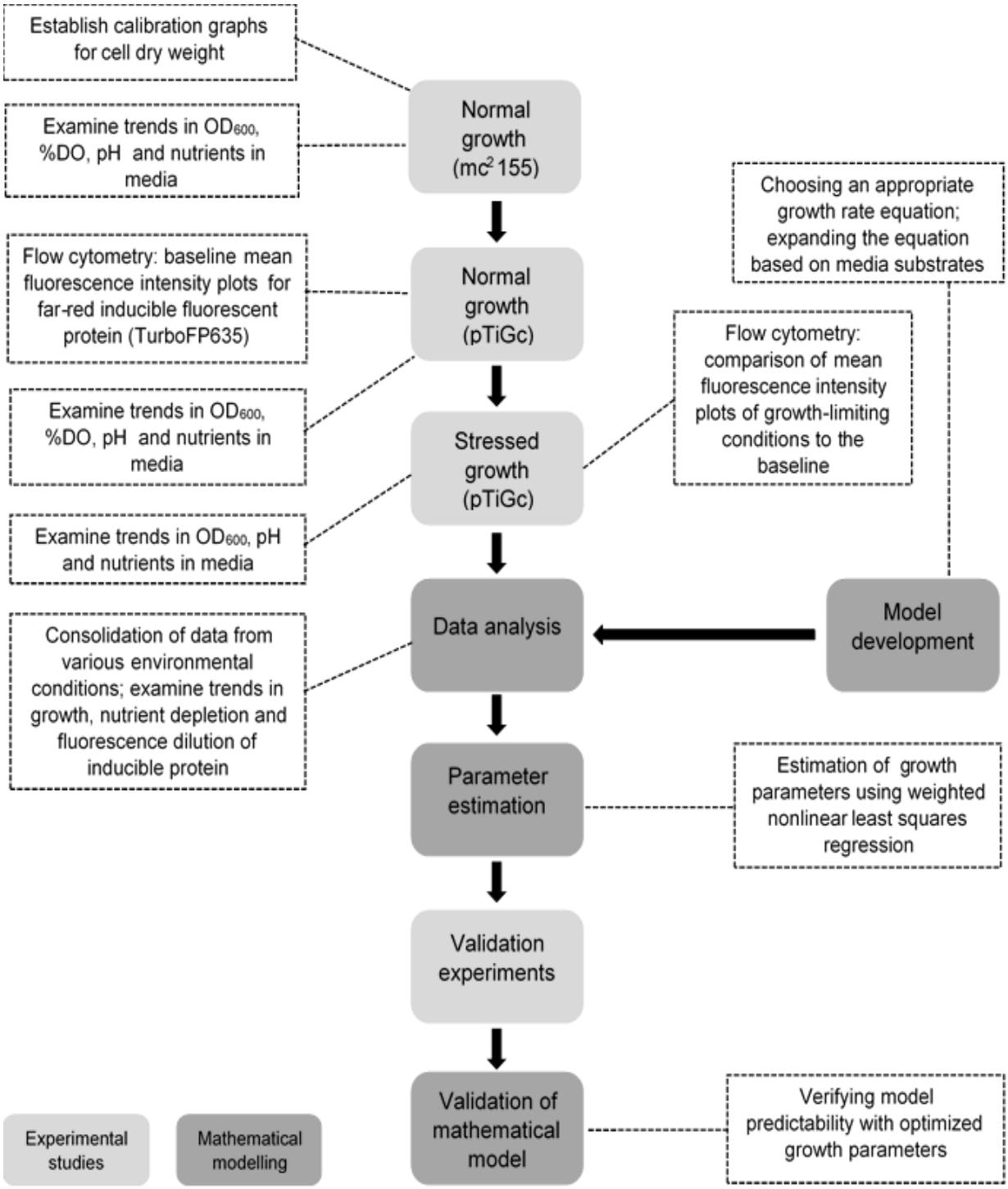

Figure 3-1: Flow diagram of project methodology

3.2 Experimental Studies

3.2.1 Bacterial Strains and Plasmids

The bacterial strains/plasmids used in this study are tabulated below:

Table 3-1: M. smegmatis plasmids/strains used

Strain/Plasmid	Description	Source	Use
mc^2155	• Non-tuberculous mycobacterium • Fast-growing organism	American Type Culture Collection (ATCC 700084)	Growth and calibration curves under normal conditions
pTiGc	• Carries an inducible far-red fluorescent protein (TurboFP635) and a constitutive GFP • A theophylline-inducible riboswitch promoter is used to control the expression of the inducible protein • Kanamycin (Kan)-resistant	(Mouton *et al*., 2016), Addgene plasmid number 78314	Growth curves and mean fluorescence intensity (MFI) plots under normal and stressed conditions
pST5552	• Carries inducible GFP • A theophylline-inducible riboswitch promoter is used to control the expression of the inducible protein • Kan-resistant	(Seeliger *et al*., 2012), Addgene plasmid number 36255	Single-colour control for flow cytometry
pCHARGE3	• Carries constitutive TurboFP635 • Expression of the red fluorescent protein is under control of a constitutive P_{smyc} promoter • Hygromycin (Hyg)-resistant	(Carroll *et al*., 2010), Addgene plasmid number 24658	

3.2.2 Growth Media

3.2.2.1 Modified Sauton's Solution

Bacterial growth under normal conditions was carried out using a Modified Sauton's solution, whose composition is described in Table 3-2 (all reagents were procured from Sigma-Aldrich®):

Table 3-2: Constituents of the modified Sauton's solution

Constituent	Amount (per litre of media)
Ammonium sulphate ($(NH_4)_2SO_4$)	3.96 g
Monopotassium phosphate (KH_2PO_4)	2.18 g
Citric acid	2 g
Magnesium sulphate ($MgSO_4$)	0.5 g
Ferric citrate	0.05 g
1% (w/v) zinc sulphate ($ZnSO_4$)	0.1 mL
10% (v/v) sterile tyloxapol	2.5 mL
50% (v/v) sterile glycerol	8 mL

To prevent bacterial clumping, tyloxapol was added to the media, rather than Tween-80. The latter may serve as a carbon source for the bacteria, which could influence the results of the carbon-starvation experiments (Tang *et al*., 2009).

To prepare growth media, $(NH_4)_2SO_4$, KH_2PO_4, $MgSO_4$, $ZnSO_4$, citric acid and ferric citrate were dissolved in 500 mL of moderately-heated deionised water (Milli-Q®) – the water was heated to dissolve the ferric citrate. The solution was left to cool to room temperature, at which its pH was set to 7 using 1 M sodium hydroxide (NaOH). The solution was topped up with more deionised water, up to 990 mL and transferred into a 1 L bottle. Finally, in a laminar-flow cabinet, the solution was filtered aseptically using a 1 L vacuum filtration unit (TPP®), with the respective amounts of sterile tyloxapol and glycerol being added to the filtered media.

For growth-limiting conditions, the media was adapted as follows (Table 3-3):

Table 3-3: Comparison: Normal media versus growth-limiting media

Stress Factor	Nutrient Source	Normal Growth Concentration and/or pH	Limiting Growth Concentration and/or pH	References
Carbon starvation	Glycerol	54.8 mM (0.4% v/v)[1]	11 mM (0.08% v/v)	(Smeulders *et al*., 1999)
Nitrogen starvation	$(NH_4)_2SO_4$	30 mM (3.96 g)	1 mM (0.132 g)	(Williams *et al*., 2015)
Acidic pH	-	pH 7	pH 4.6[2]	(Portaels and Pattyn, 1982)
Acidic pH	-	pH 7	pH 5.5[2]	(Portaels and Pattyn, 1982)

[1]In the given reference, the author used 0.2% (v/v) glycerol for normal growth of *M. smegmatis*. In this study, the concentration was doubled, to ensure that bacterial growth will not be impeded by a deficiency in glycerol at time points exceeding 24 h

[2]These values were chosen based on the pH range for growth of *M. smegmatis*; pH 4.6 is the lowest pH at which partial growth of the bacteria is observed, as per the experimental conditions of the author

3.2.2.2 Luria-Bertani (LB) Agar

This medium was used exclusively for the experiments pertaining to the generation of *M. smegmatis* calibration curves.

To prepare the agar, 10 g of LB Broth (Sigma-Aldrich®), 7.5 g of bacteriological agar (BIOLAB) and 500 mL of deionised water were added into a 500 mL bottle, which was then placed in an autoclave (121ºC, 20 minutes) to dissolve and sterilize the solution. The agar was allowed to cool to ≈ 60ºC, following which the plates were poured aseptically and left to set, before being sealed and transferred to a cold room for storage.

3.2.3 Inoculation and Sub-culturing

A 1 mL frozen stock culture of *M. smegmatis* mc^2155 cells was thawed and transferred aseptically into a 50 mL conical centrifuge tube (Falcon™) containing 9 mL of the modified Sauton's solution. A second 50 mL centrifuge tube (referred to as the control) was also prepared, which contained 10 mL of the media (no stock culture was present); this tube was to ascertain the media's sterility. Both tubes were placed overnight in a shaking incubator (200 rotations per minute (rpm), 37ºC).

On the following day, the starter culture was briefly sonicated in an ultrasonic bath (for 10 minutes, at room temperature) and filtered aseptically using a 40 µm strainer (Falcon™), to minimize bacterial clumps. The optical density (OD) of the starter culture was determined using a cell density meter (Biochrom Ltd), with a 1:5 dilution within the 1 mL cuvette. The OD_{600} of the starter culture was then adjusted to a value of 1 through pelleting and resuspension in an appropriate amount of fresh media.

The subsequent procedure was to sub-culture into a larger volume, to produce enough cells for the experiment. A volume of 2.5 mL of the starter culture (OD_{600} = 1) was transferred into a flask containing 47.5 mL of fresh growth media, establishing a starting OD_{600} of 0.05. This flask was left overnight in the shaking incubator.

The subculture was sonicated, filtered and had its OD_{600} readjusted to 1, before being transferred into the Erlenmeyer flasks (Thermo Scientific™ Nalgene™) that would be used for the experiment. Like the subculture, all growth curves for all experiments had an initial starting OD_{600} of 0.05.

3.2.4 Sequence of Experiments

To achieve a continuous growth curve, staggered culture start times were used; in this instance, two different sets of cultures were monitored over two different time frames. Both sets were inoculated at the same time, but as one set began growth, the other was stored at 4ºC, before being moved to a shaking incubator 8 hours later. This system is summarised in Table 3-4:

Table 3-4: Staggered growth times (normal and stressed conditions)

Sample No.	Culture	Growth time (h)	Beginning time of growth	Sampling time
1	a	0	08:00 day 1	08:00 day 1
2	a	3.5		11:30 day 1
3	a	7		15:00 day 1
4	b	16	16:00 day 1	08:00 day 2
5	b	19.5		11:30 day 2
6	b	23		15:00 day 2
7	a	24	08:00 day 1	08:00 day 2
8	a	27.5		11:30 day 2
9	a	31		15:00 day 2
10	b	40	16:00 day 1	08:00 day 3

3.2.4.1 Bacterial Growth under Normal Conditions

The experiments carried out in this sub-section were aimed at determining the trends in the concentration of the bacteria, oxygen, glycerol, phosphorus and ammonia, and the pH of the culture, with time. Two experimental runs were done – the first measured nutrient concentration and dissolved oxygen (%DO), whilst the second run measured pH. In both, bacterial inoculation (using *M. smegmatis* mc^2155) and sub-culturing was done as described in Section 3.2.3.

In the first experimental run, two groups of Erlenmeyer flasks were set up, the first group being used for nutrient sampling (which was done in triplicate), whilst the second group was used for measurement of %DO (single measurements were done at each time point). The Erlenmeyer flasks in each group had an initial starting volume of 50 mL, and both groups were sampled concurrently. For the first group, OD_{600} measurement and nutrient sampling was done aseptically, in a laminar flow cabinet. Four millilitres of the culture from each replicate was transferred into its respective 15 mL conical tube (Falcon™) and centrifuged (4000 rpm, 5 minutes). The supernatant was then transferred into three 2 mL Eppendorf tubes, with each tube containing a 1 mL volume of the supernatant – for the glycerol, ammonia and phosphorus

samples; the remainder of the supernatant in the conical tube was discarded. Samples were stored frozen at -80ºC, until the day of analysis. For the second group, 25 mL of the culture in the Erlenmeyer flask was transferred into a 50 mL conical tube, in which the %DO was measured – at an interval of five seconds – using a probe connected to a benchtop DO meter (Hanna® Instruments). The DO probe was polarised and calibrated prior to use. Recording of the %DO was done in a walk-in incubator, to maintain the culture at a constant temperature.

For the second run, 2 mL of the media was pipetted aseptically from the shake flask, into a 15 mL conical tube, and centrifuged (4000 rpm, 5 minutes). The pH of the supernatant was measured using a calibrated pH probe (Bante Instruments).

3.2.4.2 Bacterial Growth under Stressed Conditions

The subsequent experiments outlined in this section utilize *M. smegmatis* carrying the dual fluorescence reporter plasmid pTiGc (*M. smegmatis*::pTiGc). This was cultured with 2 mM theophylline for induction of the red fluorophore. The 10 mM theophylline stock solution was prepared by dissolving 0.09 g of theophylline into 50 mL of the modified Sauton's media. The stock solution was stored at 4ºC, in the dark, for no longer than one week.

3.2.4.2.1 Preparation of Single-Colour Controls

Inoculation and induction of the various strains and plasmids was done as per Table 3-5, using freshly prepared theophylline solution. All six tubes were placed in a shaking incubator, at 37ºC, and left overnight:

Table 3-5: Plasmids used for the colour controls

Strain	Stock culture (mL)	Induction	Antibiotics	Growth Media (mL)	Theophylline (mL)
M. smegmatis::pTiGc	1	2 mM	25 µg/mL Kan	7	2
M. smegmatis::pST5552	1	2 mM	25 µg/mL Kan	7	2
M. smegmatis::pCHARGE3	1	2 mM	50 µg/mL Hyg	7	2
M. smegmatis::pTiGc	1	-	25 µg/mL Kan	9	-
M. smegmatis mc^2155	1	-	-	9	-
Contamination Control	-	-	-	10	-

After sonication and filtering of the inocula (Section 3.2.3), 1 mL samples from each 50 mL conical tube were pipetted into their corresponding 2 mL microcentrifuge tubes and prepared for flow cytometry analysis (described in Section 3.2.5.1.2).

3.2.4.2.2 Normal Growth with *M. smegmatis*::pTiGc (Baseline)

Bacterial inoculation was done using three 50 mL conical tubes, each containing 7 mL of media, 2 mL of freshly prepared 10 mM theophylline solution, 1 mL of inoculum and an antibiotic concentration of 25 µg/mL Kan. The inocula grew overnight in a shaking incubator, and on the following day, they were sonicated, filtered, and had their OD_{600} measured. Thereafter, the samples were centrifuged (4000 rpm, 5 minutes) and the pellets washed to remove theophylline, and resuspended in fresh media (without theophylline) to an OD_{600} of 1. Growth cultures were started at an OD_{600} of 0.05 by pipetting 5 mL of the resuspended inoculum into a shake flask containing 95 mL media and 25 µg/mL Kan.

The collection and preparation of samples for flow cytometry, glycerol, ammonia, OD_{600} and pH measurements was done as described in Section 3.2.4.1, in **one** experimental run.

3.2.4.2.3 Stressed Growth with *M. smegmatis*::pTiGc

The environmental conditions in which stressed growth was initiated included nutrient starvation and acidic pH. Four experimental runs were carried out with the following initial conditions to facilitate stressed growth:

1. Carbon starvation – limiting the initial glycerol concentration to 0.08% (v/v)
2. Nitrogen starvation – limiting the initial $(NH_4)_2SO_4$ concentration to 1 mM
3. Acidic pH – setting the initial pH of the media to 4.6
4. Acidic pH – setting the initial pH of the media to 5.5

The procedure used for the baseline experiment (i.e. under normal growth conditions) was replicated across these four runs.

3.2.4.3 Model Validation Experiments

The objective of these experiments was to generate data which could be used for comparison with the optimized simulation of the proposed mathematical model. Similar trends would consequently authenticate the model. These experiments were run for a longer time frame (t = 0 h to t = 99.5 h) to examine the effect of a prolonged environmental stress on bacterial growth, which would give a higher chance of persister formation.

3.2.4.3.1 Combined Growth Stresses – Nutrient Starvation (*M. smegmatis*::pTiGc)

In this experimental run, the bacteria were grown in media that was glycerol **and** ammonia deficient. A baseline experimental run was conducted simultaneously, at normal conditions, for a proper comparison of the flow cytometry data.

Staggered culture start times were applied in the following manner (Table 3-6):

Table 3-6: Staggered growth times (validation experiments)

Sample No.	Culture	Growth time (h)	Beginning time of growth	Sampling time
1	a	0	08:00 day 1	08:00 day 1
2	a	7		15:00 day 1
3	b	16	16:00 day 1	08:00 day 2
4	a	27.5	08:00 day 1	11:30 day 2
5	b	40	16:00 day 1	08:00 day 3
6	a	51.5	08:00 day 1	11:30 day 3
7	b	64	16:00 day 1	08:00 day 4
8	a	75.5	08:00 day 1	11:30 day 4
9	b	88	16:00 day 1	08:00 day 5
10	a	99.5	08:00 day 1	11:30 day 5

During this experiment, the nitrogen starvation run indicated in Section 3.2.4.2.3 was repeated, but with the sampling times as shown in Table 3-6.

Inoculation and sampling, as per the protocol in Section 3.2.4.2.2, was used for this run.

3.2.4.3.2 Combined Growth Stresses – Nutrient Starvation and pH Stress (*M. smegmatis*::pTiGc)

The experimental setup involved initiating growth stress by growing the bacteria under nutrient limiting conditions, followed by spiking the media with a defined amount of the limiting nutrient, and finally, acidifying the media through addition of concentrated hydrochloric acid (HCl). Two runs were performed:

1. Carbon deficiency/acidic pH: Bacterial growth started off under glycerol-deficient conditions (t = 0 h); the media was then spiked with 1.5 mL of sterile 50% glycerol at t = 24 h, and was later acidified by adding four drops of concentrated HCl at t = 55 h
2. Nitrogen deficiency/acidic pH: The procedure is similar to the first run, but with bacteria growing under ammonia-deficient conditions, and the media being spiked with 3.4 mL of sterile 0.25 M $(NH_4)_2SO_4$ at t = 24 h

As in the previous section, a normal run was conducted simultaneously, to gather OD and flow cytometry data for comparison between the normal and stressed environmental states. Sampling times are summarised in Table 3-7, with inoculation and sampling being done as per Section 3.2.4.2.2. Staggered culture start times were not used in this case, as spiking needed to be done in the same Erlenmeyer flasks.

Table 3-7: Sampling and spiking times

Sample No.	Growth time (h)	Sampling time
1	0	08:00 day 1
2	7	15:00 day 1
3 (nutrient spike)	24	08:00 day 2
4	31	15:00 day 2
5	48	08:00 day 3
6 (addition of HCl)	55	15:00 day 3
7	72	08:00 day 4
8	79	15:00 day 4
9	96	08:00 day 5
10	99.5	11:30 day 5

3.2.5 Analytical Techniques

3.2.5.1 Assessment of Accuracy of Techniques used

The calibration curves of three different indirect analytical techniques were compared, to establish which method could be used to most accurately infer the cell dry weight (CDW) produced in the culture. These were OD measurement, CFU measurement and flow cytometric measurement of cell numbers.

For this experiment, bacteria were cultured for 40 hours, with samples taken at 10 time points (0 h, 3.5 h, 7 h, 16 h, 19.5 h, 23 h, 24 h, 27.5 h, 31 h and 40 h). For the setup, biological replicates were carried out – three baffled Erlenmeyer flasks were prepared for each time point. Bacterial inoculation and sub-culturing were done as previously described in Section 3.2.3. Samples for the OD, CFU and flow cytometry measurements were collected prior to determining the CDW produced at each time point.

3.2.5.1.1 Optical Density

One millilitre of the culture was taken out of each Erlenmeyer flask, and transferred to its respective 1 mL cuvette, after which its OD_{600} value was read using a cell density meter. One millilitre of fresh modified Sauton's media served as a blank for the reading. 1:5 dilutions were done for samples whose OD_{600} values exceeded 1.

3.2.5.1.2 Flow Cytometry

Sample preparation required the following solutions:

- 0.025% (v/v) phosphate-buffered saline/tyloxapol (PBS-T) – prepared aseptically by adding 1.25 mL of 10% tyloxapol to 500 mL of sterile PBS (Gibco™)

- 4% (v/v) formaldehyde – prepared aseptically by adding 1 mL of 40% (v/v) formaldehyde (Sigma-Aldrich®) to 9 mL of PBS-T

A 1 mL sample from each replicate was transferred to a 2 mL microcentrifuge tube (Eppendorf). The sample was pelleted (13000 rpm, 5 minutes) and resuspended in 0.2 mL of 4% (v/v) formaldehyde. This step, known as formaldehyde fixing, helps in not only retaining intracellular fluorescence and morphology, but also in decontaminating the sample, to avoid introducing live bacterial cells into the flow cytometer. After 30 minutes, 0.8 mL of PBS-T was added into the tube. The sample was pelleted and resuspended once more, in 0.2 mL of fresh PBS-T, and stored in the dark at 4ºC until further analysis at a later date.

Prior to running the samples in the flow cytometer, they were pelleted, resuspended in 0.5 mL of PBS and filtered through a 40 µm cell strainer cap, into their respective 5 mL round-bottom tubes (Falcon™).

Samples were analysed with the FACSJazz and LSRFortessa flow cytometers (Becton Dickinson). The FACSJazz was used to evaluate the samples pertaining to parameter estimation experiments, while the LSRFortessa was used for the validation experiments (at the time of analysis for the validation experiments, the FACSJazz was not available for use). The instruments captured the forward scatter (FSC; quantifies cell size) and side scatter (SSC; describes cell granularity) properties, as well as the fluorescence intensity of the reporters, under the following specifications (Table 3-8):

Table 3-8: Instrument Settings for the Flow Cytometers

	FACSJazz		**LSRFortessa**	
	GFP	**TurboPF635**	**GFP**	**TurboFP635**
Excitation/Emission (nm)	408/509	588/635	408/509	588/635
Laser (nm)	488	561	488	544
Filter (nm)	530/40	610/20	515/20	610/20
Voltage (volts)	57	78	557	566

Both flow cytometers captured 30,000 events for each sample and compensation (mathematical correction of spectral overlap of different fluorophores, due to their emission spectra being captured by all detectors of the instrument) was performed for each experiment, with the single-colour controls.

Flow cytometry data was evaluated using FlowJo 10.60.0 software. The procedure for analysing the fluorescence intensity of TurboFP635 in the bacterial population involved selecting a population subset of the collected events based on FSC/SSC properties (that is,

forming a primary gate), and then gating the GPG-positive (live) population (Figure 3-2, Q2 and Q3):

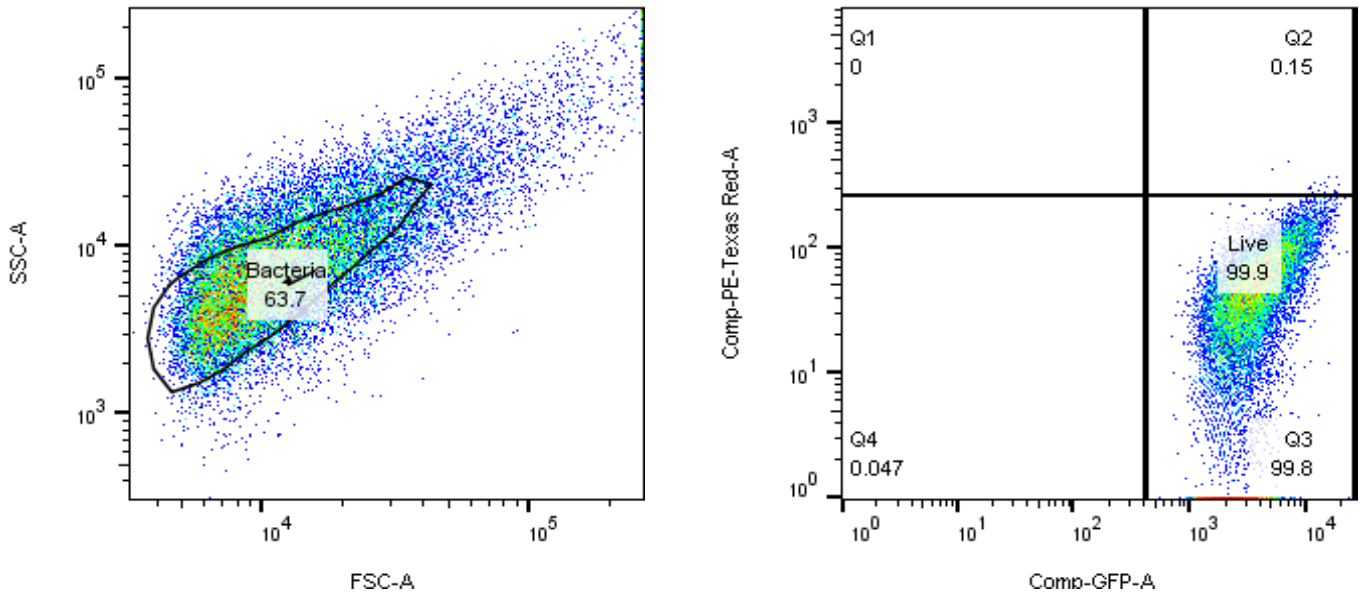

Figure 3-2: Capturing the TurboFP635+ and TurboFP635- population. Left: Setting a primary gate based on the FSC and SSC. Right: The bacteria within the primary gate are further sub-divided into four quadrants, based on the fluorescence intensity of the fluorophores. Q1: GFP- and TurboFP635+; Q2: GFP+ and TurboFP635+; Q3: GFP+ and TurboFP635-; Q4: GFP- and TurboFP635-

The second and third quadrants were used to determine the geometric mean of the TurboFP635 intensity. This data was used to calculate the number of bacterial generations (N) using the formula $\frac{Y_0}{Y_t} = 2^N$, where the ratio Y_0/Y_t represents the extent of bacterial replication (with Y being the geometric mean). In the case of optical density, N is calculated in a similar manner, but with an inverse of the ratio (i.e. Y_t/Y_0).

3.2.5.1.3 CFU Counts

Serial 10-fold dilutions were carried out in 48-well tissue culture plates (CELLSTAR®), using PBS-T as the diluent. Each well was filled with 0.9 mL PBS-T, which would dilute a 0.1 mL culture sample. The chosen dilution factors were based on the expected number of colonies, which were calculated from the anticipated OD_{600} values at each time point (Appendix B). They are shown in Table 3-9:

Table 3-9: Chosen dilution factors

Time Point	Dilution Factor
0 h – 3.5 h	10^{-3}, 10^{-4} and 10^{-5}
7 h	10^{-4}, 10^{-5} and 10^{-6}
16 h – 40 h	10^{-5}, 10^{-6} and 10^{-7}

After plating, the petri dishes were incubated at 37ºC for three days, before counting the colonies that formed.

3.2.5.1.4 Cell Dry Weight

Two millilitre microcentrifuge tubes needed for the biomass samples were prepared beforehand by perforating the caps of the tubes – to facilitate the evaporation of liquid from the samples during the drying process – and labelling said tubes. Following this, the tubes were individually weighed, to determine their masses prior to sample addition.

The remaining volume of the culture within the Erlenmeyer flask was transferred into a 50 mL conical centrifuge tube and pelleted (4000 rpm, 5 minutes). With the supernatant discarded, the pellet was resuspended in 0.5 mL PBS-T, transferred into its respective microcentrifuge tube and placed in a centrifugal vacuum concentrator, where all samples were dried overnight. Finally, the tubes were re-weighed, and the CDW was calculated by subtracting the mass of the tube without the sample from the mass of the same tube with the sample.

3.2.5.2 Calibration Curves

From the previously mentioned techniques, the calibration curves shown in Figure 3-3 were generated:

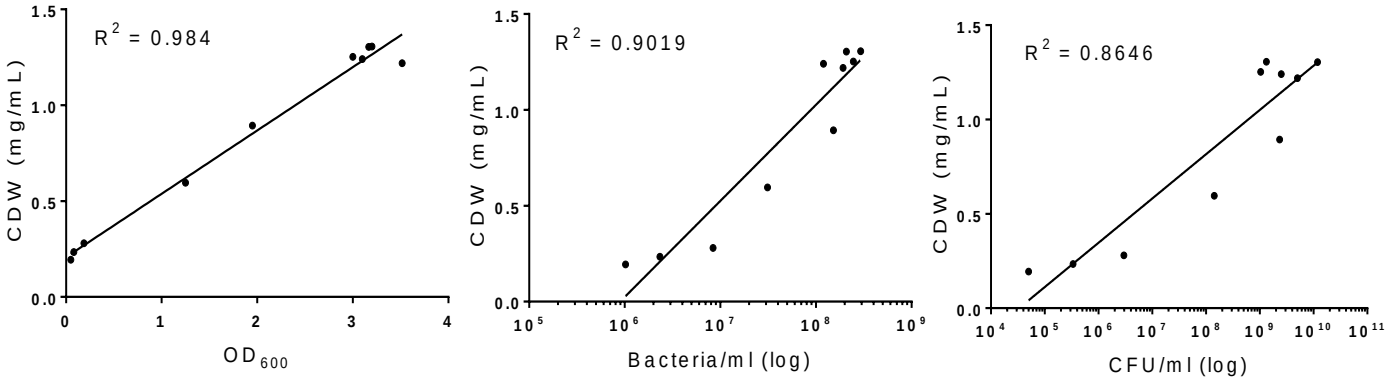

Figure 3-3: Calibration curves for OD_{600}, flow cytometry and CFU (respectively)

Based on the values of the correlation coefficient (R^2) in each graph, the OD_{600} is the most accurate technique in the determination of the CDW, hence, was used for the rest of the experimental runs.

3.2.5.3 Glycerol Concentration

The ENZYTEC™ *fluid* Glycerol method was used for the photometric assay of the carbon source. The glycerol concentration is proportional to the colorimetric product formed from coupled enzymatic reactions involving glycerokinase, adenosine diphosphate-dependent hexokinase and glucose-6-phosphate dehydrogenase (Thermo Fisher Scientific, 2010). The assay was performed automatically, using a Thermo Scientific Arena™ 20XT biochemistry

analyser. Sample preparation involved centrifugation, followed by manual dilution of samples whose concentration fell outside the calibration range of the instrument. Assay specifications are denoted in Table 3-10.

3.2.5.4 Ammonia Concentration

The ENZYTEC™ *fluid* Ammonia method was used for the photometric assay of the nitrogen source. In this test, the ammonia concentration is proportional to the decrease in absorbance due to oxidized nicotinamide adenine dinucleotide (NADH), in the presence of glutamate dehydrogenase (Thermo Fisher Scientific, 2016). As with glycerol, this assay was performed automatically, with the samples being centrifuged before being run in the biochemistry analyser. Manual dilutions were done for the samples whose ammonia concentration fell out of the calibration range. Assay specifications are denoted in Table 3-10:

Table 3-10: Assay specifications for glycerol and ammonia analysis

	Glycerol	**Ammonia**
Wavelength	340 nanometres (nm)	
Temperature	37ºC	
Reaction time	8 minutes	10 minutes
Measurement	against water	
Calibration range	0.0183 – 0.365 g/L	10-50 mg/L

3.2.5.5 Phosphorus Concentration

Inductively coupled plasma optical emission spectroscopy (ICP-OES) was used for the determination of the phosphorus concentration in the samples. Liquid sample preparation involved acidification using ultra-pure nitric acid (HNO_3), to a final concentration of 2% (the acid prevents the precipitation of elemental components from solution). The Thermo iCAP 6000 series (Thermo Fisher Scientific) was used for the elemental analysis, under the following instrument conditions (Table 3-11):

Table 3-11: Spectrometer specifications

Radio frequency power (W)	1350
Volumetric flow rate of carrier gas (L/min)	0.65 (Argon)
Volumetric flow rate of auxiliary gas (L/min)	1.00 (Argon)
Nebuliser	2 mL/min micro mist
Internal standard used	1 part per million (ppm) Yttrium

The optimum emission wavelength of phosphorus, as per the instrument settings, is 177.4 nm.

3.2.5.6 Elemental Analysis of M. smegmatis

The elemental composition of *M. smegmatis* is pertinent in estimation of certain yield coefficients of the substrates, based on a balance of its metabolic equation. As its molecular formula is not available in literature, an elemental analysis was done, through inductively coupled plasma mass spectrometry (ICP-MS) and scanning electron microscopy (SEM). Preparation of dried *M. smegmatis* mc^2155 samples was done as per Section 3.2.5.1.4.

Samples were firstly analysed using ICP-MS, which determined its carbon, hydrogen, nitrogen and sulphur content. The equipment used in this process was the Vario EL cube elemental analyser (Elementar), under the operating conditions as specified in Table 3-12:

Table 3-12: Operating conditions of the elemental analyser

Carrier gas used	Argon (± 1 bar)
Dosing gas for combustion	Oxygen (2 bar)
Combustion temperature	1050ºC – 1150ºC
Reduction column temperature	850ºC
Thermal conductivity detection (TCD) temperature	60ºC
TCD Argon flow	220 mL/min
TCD background signal	<100 (usually after blanks and washout)

The oxygen content in the sample was analysed using SEM. Samples were firstly mounted onto double-sided carbon tape on aluminium SEM stubs, and sputter-coated with a thin layer of gold, using a Leica EM ACE200 sputter-coater (Leica Microsystems). This was done to enhance its conductivity. Next, the coated sample was loaded onto a Zeiss MERLIN Field Emission Scanning Electron Microscope (Carl Zeiss Microscopy). Images from the microscope were generated using a Zeiss backscattered electron detector and Zeiss Smart

SEM software, whereas sample quantification was done via energy dispersive x-ray spectrometry. Gold and aluminium were automatically excluded from the analysis.

3.3 Mathematical Modelling

3.3.1 Growth Kinetics of Cell Cultures

In this study, extracellular factors influencing bacterial growth kinetics were assessed, therefore, an unstructured model was used to simulate bacterial growth. The Monod model (Monod, 1949), which has served as the basis of microbial growth dynamics under various engineering and environmental applications, was chosen:

$$\mu = \mu_{max} \frac{S}{K_s + S} \quad [1]$$

where μ represents the specific growth rate, μ_{max} represents the maximum specific growth rate, S is the substrate concentration and K_s is the saturation constant, which represents the substrate concentration when $\mu = 0.5\mu_{max}$. This constant gives an indication of the concentration at which the limiting substrates start to negatively influence the rate of bacterial growth. The kinetics take substrate limitation into account in an aqueous environment.

Nonetheless, the main limitation of this model is that it does not account for the initial lag phase after inoculation of the bacteria, and as such, is only valid for the exponential and decelerating growth phases.

3.3.2 Conceptual Dynamic Mathematical Model

Implementing the Monod model, with glycerol and ammonia serving as substrates for bacterial growth, resulted in the following expression:

$$\mu = \mu_{max} \left(1 - e^{\frac{-t}{t_{lag}}}\right) \left(\frac{C}{C + k_C}\right) \left(\frac{N}{N + k_N}\right) I_{pH} \quad [2]$$

where C and N represent the concentrations of glycerol and ammonia, respectively. The following extensions were added to Equation [2], for greater accuracy of the simulation:

- An inhibition constant, I_{pH}, to account for the decrease in the bacterial growth rate as a result of an increasingly acidic pH environment. This constant is denoted by the following equation (IWA Task group for Mathematical Modelling of Anaerobic Digestion Processes, 2002):

$$I_{pH} = e^{\left(I_{val} \frac{pH - pH_{UL}}{pH_{UL} - pH_{LL}}\right)^2} \quad [3]$$

- An expression incorporating the initial lag phase after inoculation $\left(1 - e^{\frac{-t}{t_{lag}}}\right)$, where t_{lag} is the length of the lag phase, as observed from the experimental data (Bergter and Knorre, 1972)

Expounding the model, to simulate the concentrations of the substrates utilized and the metabolites produced, resulted in the following ordinary differential equations (ODEs) listed in Table 3-13:

Table 3-13: Batch model equations for bacterial growth

Variables	Equation	
Biomass (X)	$\frac{dX}{dt} = (\mu - k_d)f(X)$	*[4]*
Glycerol (C)	$\frac{dC}{dt} = -\left(\frac{\mu}{Y_{\frac{X}{C}}}\right)f(X)$	*[5]*
Ammonia (N)	$\frac{dN}{dt} = -\left(\frac{\mu}{Y_{\frac{X}{N}}}\right)f(X)$	*[6]*
CO_2 (CD)	$\frac{dCD}{dt} = \left(\frac{\mu}{Y_{\frac{X}{CD}}}\right)f(X)$	*[7]*

The rates of substrate utilization and metabolite production are directly proportional to the biomass produced by *M. smegmatis* (Equation [4]). This is accounted for in the above equations, but with an additional modification, which limits further biomass production once its carrying capacity has been reached:

$$f(X) = X - \frac{e^{\alpha X} - 1}{e^{\alpha X_{max}} - 1} \qquad [8]$$

The purpose of the carrying capacity term (X_{max}) in f(X) is to restrict bacterial growth under environmental conditions in which the concentrations of glycerol and ammonia are in abundance.

Table 3-14 gives a summary of the parameters used in Equations [2] to [8]. Note that the shaded cells represent the growth parameters that underwent optimization; parameters in the non-shaded cells remained constant throughout the simulation.

Table 3-14: Growth parameters of model

Parameter	Description	Units
μ_{max}	Maximum specific growth rate of bacteria	h^{-1}
k_C	Saturation constant of glycerol	g/L
k_N	Saturation constant of ammonia	g/L
k_d	Microbial death rate	h^{-1}
pH_{LL}	Lower limit of pH inhibition	-
α	Carrying capacity coefficient	-
I_{val}	Empirical lower inhibition coefficient	-
$Y_{X/CD}$	Yield coefficient of CO_2	-
pH_{UL}	Upper limit of pH inhibition	-
X_{max}	Biomass carrying capacity	g/L
$Y_{X/C}$	Yield coefficient of glycerol	-
$Y_{X/N}$	Yield coefficient of ammonia	-

The yield coefficients of glycerol and ammonia were determined based on the data of their respective nutrient-deficient experiments (refer to Appendix B for calculations).

3.3.3 Inclusion of Oxygen Dynamics

The dissolved oxygen (%DO) in the growth media is dependent upon two factors:

1. The oxygen transfer rate (OTR) from the gaseous phase to the aqueous phase, due to agitation of the shake flask (with the assumption of complete mixing of the gas in the aqueous phase)
2. The utilization of the gas by the organism, due to aerobic respiration

Therefore, oxygen dynamics are represented by the following ODE:

$$\frac{dO}{dt} = (k_L a(O^* - O)) - \left(\left(\frac{\mu}{Y_{\frac{X}{O}}}\right) * X\right) \qquad [9]$$

where:

- O^* is the saturated oxygen concentration in the aqueous phase
- k_La is the overall volumetric mass transfer coefficient of oxygen
- $Y_{X/O}$ is the yield coefficient of oxygen
- O is the oxygen concentration (in g/L)

When added to μ in Equation [2], the resulting growth rate equation is as follows:

$$\mu = \mu_{max}\left(1 - e^{\frac{-t}{t_{lag}}}\right)\left(\frac{C}{C+k_C}\right)\left(\frac{N}{N+k_N}\right)\left(\frac{O}{O+k_O}\right)I_{pH} \qquad [10]$$

with k_O acting as the saturation constant of oxygen.

3.3.4 Modelling pH Change in Growth Media

The steps followed in establishing the charge balance equations were as follows:

1. **Identification of media components and metabolic products that influence the pH**

This process was based on the value of the logarithmic acid dissociation constant (pK_a) of each component and product. Those that had pK_a values exceeding 9 were not incorporated into the model, as they would undergo very little dissociation, thus, have little to no effect on the pH.

The pK_a values, with accompanying dissociation equations, are listed in Table 3-15:

Table 3-15: pK_a values and acid dissociation equations of media components

Constituent	**pK_a**	**Acid Dissociation Equations**	**References**
KH_2PO_4	6.86 (pK_{a1})[1]	$KH_2PO_4 \leftrightarrow H^+ + KHPO_4^-$	(Mathews *et al*., 1990)
Citric acid	3.13 (pK_{a2})	$C_6H_8O_7 \leftrightarrow H^+ + C_6H_7O_7^-$	(Haynes, 2016)
	4.76 (pK_{a3})	$C_6H_7O_7^- \leftrightarrow H^+ + C_6H_6O_7^{2-}$	(Haynes, 2016)
	6.40 (pK_{a4})	$C_6H_6O_7^{2-} \leftrightarrow H^+ + C_6H_5O_7^{3-}$	(Haynes, 2016)
CO_2	6.35 (pK_{a5})	$CO_2 + H_2O \leftrightarrow H^+ + HCO_3^-$	(Wang *et al*., 2008)
	10.33	$HCO_3^- \leftrightarrow H^+ + CO_3^{2-}$	(Wang *et al*., 2008)
Glycerol	14.15	$C_3H_8O_3 \leftrightarrow H^+ + C_3H_7O_3^-$	(Serjeant *et al*., 1979)
$(NH_4)_2SO_4$	9.25	$NH_4^+ \leftrightarrow H^+ + NH_3$	(Haynes, 2016)

[1]The K_a for each species is calculated using the pKa as follows: $10^{-pK_a} = K_a$

Referring to Table 3-15, it is seen that the pK_a values of bicarbonate (HCO^-_3), glycerol and ammonium are greater than 9; therefore, dissociation of these components was neglected.

2. Balancing the cations and anions in solution

Based on the dissociation equations of the media components, the following charge balance results:

$$[H^+] = [OH^-] + [KHPO_4^-] + 3[C_6H_5O_7^{3-}] + 2[C_6H_6O_7^{2-}] + [C_6H_7O_7^-] + [HCO_3^-] \qquad [11]$$

It is at this point that certain ionic species, whose concentration remains constant in the media, are accounted for. They include magnesium (Mg^{2+}), zinc (Zn^{2+}), sodium (Na^+) and sulphate (SO_4^{2-}) ions. This brings about the subsequent charge balance:

$$[H^+] + [Na^+] + 2[Zn^{2+}] + 2[Mg^{2+}] = [OH^-] + [KHPO_4^-] + 3[C_6H_5O_7^{3-}] + 2[C_6H_6O_7^{2-}] + [C_6H_7O_7^-] + [HCO_3^-] + 2[SO_4^{2-}] \quad [12]$$

Because the concentration of the Mg^{2+}, Zn^{2+}, Na^+ and SO_4^{2-} ions do not change, the charge balance was simplified by taking the first three ionic species to the right side of the equation, and grouping them together with the SO_4^{2-} ions, resulting in an expression which denotes the total concentration of all four ions in solution (I_{cst}), throughout each experimental run:

$$[H^+] = [OH^-] + [KHPO_4^-] + 3[C_6H_5O_7^{3-}] + 2[C_6H_6O_7^{2-}] + [C_6H_7O_7^-] + [HCO_3^-] - [I_{cst}] \quad [13]$$

When simulating the cellular system with MATLAB software, different values of I_{cst} were used, to ensure that the initial pH of the simulation under each environmental condition would correspond to the initial pH of the relevant experimental data.

3. **Expression of ionic species in the charge balance as a function of hydrogen ions (H^+)**

K_a is an equilibrium constant denoting dissociation in acid-base reactions. Using HB as an arbitrary acid:

$$HB \leftrightarrow H^+ + B^- \quad [14]$$

$$K_a = \frac{[H^+][B^-]}{[HB]} \quad [15]$$

The total concentration of HB in solution is:

$$HB_{tot} = [B^-] + [HB] \quad [16]$$

Assuming HB_{tot} is known, [HB] can be substituted using the K_a expression in Equation [15]:

$$HB_{tot} = [B^-] + \frac{[H^+][B^-]}{K_a} = [B^-]\left(1 + \frac{[H^+]}{K_a}\right) \quad [17]$$

Therefore:

$$[B^-] = \frac{HB_{tot}}{\left(1+\frac{[H^+]}{K_a}\right)} \quad [18]$$

Applying this method to the anions in the charge balance, the following expressions were derived (Table 3-16):

Table 3-16: Anion expression as a function of H^+

Anion	Substitution	Equation
[$KHPO_4^-$]	$\frac{KH_2PO_{4,tot}}{\left(1+\frac{[H^+]}{K_{a1}}\right)}$	*[19]*
[$C_6H_5O_7^{3-}$]	$\frac{C_6H_8O_{7,tot}}{\frac{[H^+]^3}{K_{a2}K_{a3}K_{a4}}+\frac{[H^+]^2}{K_{a3}K_{a4}}+\frac{[H^+]}{K_{a4}}+1}$	*[20]*
[$C_6H_6O_7^{2-}$]	$\left(\frac{[H^+]}{K_{a4}}\right)*[C_6H_5O_7^{3-}]$	*[21]*
[$C_6H_7O_7^-$]	$\left(\frac{[H^+]}{K_{a3}}\right)*[C_6H_6O_7^{2-}]$	*[22]*
[HCO_3^-]	$\frac{CO_{2,tot}}{\left(\frac{[H^+]}{K_{a5}}+1\right)}$	*[23]*
[OH^-]	$\frac{K_{a6}}{[H^+]}$	*[24]*

For the hydroxide [OH^-] ions, the substituted expression is different from the rest of the equations of the other ionic species. This is due to the autoionization of water, in which its chemical activity – the effective concentration of a species in solution – is taken to be unity, as water is a pure liquid. Therefore, [H_2O] is not added to the dissociation expression:

$$K_{a6} = [H^+][OH^-] \qquad [25]$$

4. Calculation of the concentration of [H^+] ions

The only unknown in the charge balance is [H^+]. Rearranging Equation [13]:

$$Z = [OH^-] + [KHPO_4^-] + 3[C_6H_5O_7^{3-}] + 2[C_6H_6O_7^{2-}] + [C_6H_7O_7^-] + [HCO_3^-] - [I_{cst}] - [H^+] \quad [26]$$

[H^+] can be solved for by minimizing Z, such that Z([H^+]) = 0 (the *fzero* function in MATLAB is appropriate in this case). To calculate the pH of the media:

$$pH = -log_{10}[H^+] \qquad [27]$$

Solving the ODEs in Section 3.3.2 and the H^+ equations in Section 3.3.4 renders this problem a system of differential algebraic equations (DAEs). As such, the solver used in MATLAB was *ode15s*, suitable for DAE systems. *ode15s* is a variable-step, variable-order solver that implements numerical differential formulas of order 1 to 5 (Shampine and Reichelt, 1997) and is used in instances where a system exhibits stiffness (difficulty in evaluation).

3.3.5 Parameter Estimation

An initial attempt at estimation of the growth parameters was done, with the assumption that the magnitudes of the residuals of the model response variables were equal. This assumption resulted in the use of a nonlinear least squares regression approach. However, as the optimized simulation had a poor fit, the residuals for biomass, glycerol and ammonia were examined once more, and it was realized that their magnitudes were different. This called for standardization of these residuals, and as such, a **weighted** nonlinear least squares regression approach was utilized (NIST, 2012):

$$R = \sum_{i=1}^{n} w_i \left(\frac{y_i - y_{i,min}}{y_{i,max} - y_{i,min}} - \frac{f\left(\vec{x_i}, \vec{\beta}\right) - y_{i,min}}{y_{i,max} - y_{i,min}} \right)^2 \qquad [28]$$

where y_i represents the experimental data of the variables in question, β represents the model parameters to be estimated, and $f(x_i;\beta)$ represents the corresponding model predictions, which are a function of the independent variable in the model (time, x_i) and the relevant growth parameters. $y_{i,min}$ and $y_{i,max}$ denote the minimum and maximum values (respectively) of the data for the relevant response variable; they are used in the normalization of the model predictions and the raw data.

w_i represents the weights to be applied to the data points, such that each is given a proper amount of influence over the estimated parameters. Granted, this method assumes that the exact values of the weights are known beforehand, which is a rare occurrence (NIST, 2012), therefore, w_i has to be estimated. Consequently, since sample triplicates were taken for each variable, w_i can be assumed to be the inverse of the sample variance at each time point:

$$w_{ij} \; \alpha \; \frac{1}{\hat{\sigma}_i^2} \qquad [29]$$

When implementing the model in MATLAB, a user-defined function was created, which calculated the weighted residuals of the response variables for each environmental condition. Two different optimization algorithms were used – the first solved the least squares problem with a local solver, whose solution changed depending on the initial conditions used in the model. The second approach used a global optimization algorithm, which worked together with the local solver, to find multiple local minima. The minimum with the lowest residual norm would be the global minimum, giving the most optimal fit. The local solver used in minimizing the residuals was *lsqnonlin*, with the default 'trust-region-reflective' algorithm. As for the global optimization approach, *MultiStart* was utilized, with the algorithm being run for fifty instances.

3.3.6 Statistical Analysis

To quantify the predictive accuracy of the proposed mathematical model, the root mean squared error (RMSE), defined as the square root of the mean of the squared differences between observed values and their corresponding predicted values, was calculated:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - f\left(\vec{x_i};\vec{\beta}\right)\right)^2}{n}} \qquad [30]$$

Generally, a lower RMSE would indicate a better fit of the model to the observed data. However, given that the RMSE is scale-dependent, it would be incorrect to compare this statistical value across different response variables (i.e. biomass, glycerol, ammonia and pH). Therefore, to account for this, the normalized RMSE (NRMSE) for each response variable, under each environmental condition, was determined by dividing the RMSE by the range of the measured data:

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \qquad [31]$$

The NRMSE is expressed as a percentage, and as with the RMSE, a lower value indicates a better fit.

3.4 Conclusion

This chapter detailed the procedures used in carrying out the wet laboratory experiments and in developing the dynamic mathematical model. Having obtained the raw experimental data, the mathematical equations in Sections 3.3.2 to 3.3.4 were implemented and simulated with MATLAB software, with comparisons being drawn between the actual growth data versus simulation predictions based on the estimated growth parameters. Possible persister formation was also explored, by analysing the bacterial generation plots that were created using MFI data. An in-depth analysis of results obtained is given in the next chapter.

4 RESULTS AND DISCUSSIONS

4.1 Introduction

The results for bacterial growth under normal and stressed conditions, the flow cytometry analysis and the computational simulations are discussed herein. The flow cytometry analysis explores bacterial replication dynamics through FD and compares bacterial generation plots of the various growth-limiting conditions to that of the normal conditions, to establish the possible presence of the VBNR population. The computational simulations consider parameter estimation, with the estimated parameters being used to investigate model predictability, by contrasting the optimized simulation with data from the validation experiments. The results are discussed within the framework of the project objectives.

4.2 Objective 1: Bacterial Growth under Normal Conditions

The trends in bacterial growth, pH and nutrient consumption of *M. smegmatis* mc^2155 are illustrated in Figure 4-1, in the following page:

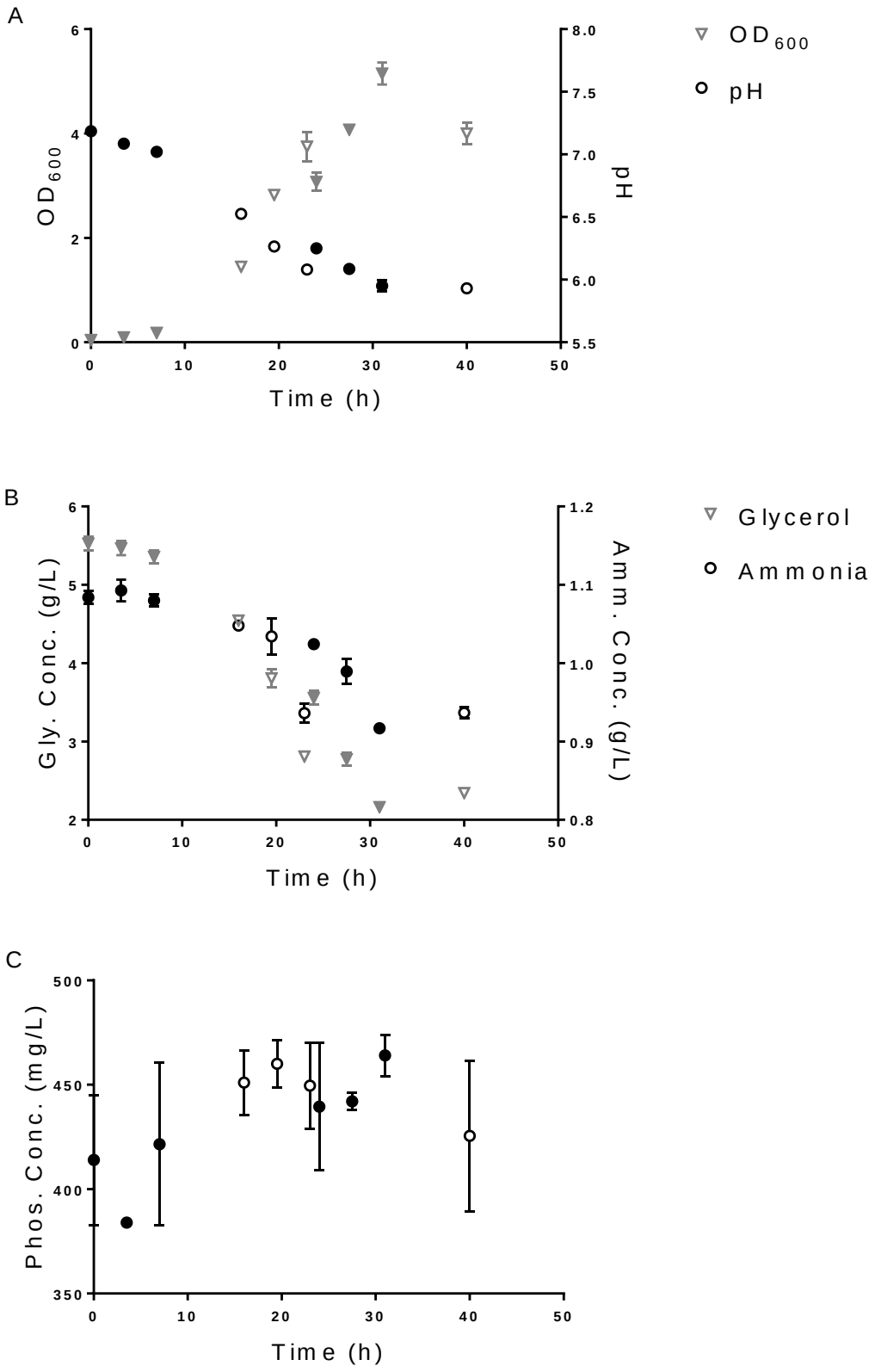

Figure 4-1: Growth trends of mc²155. A: OD (triangles) and pH (circles); B: glycerol (triangles) and ammonia (circles); C: phosphorus. Staggered culture start times were used to measure the continuous growth of the bacteria. Shaded symbols represent culture a; open symbols represent culture b

The results for OD, pH, glycerol and ammonia consumption are as expected (Figure 4-1 A and B). Glycerol is broken down aerobically by the bacteria through glycolysis, which results in the increase in its OD, and a decrease in the pH of the media, due to the production of CO_2 as a metabolic product.

Likewise, nitrogen uptake by the cells resulted in the depletion of ammonia in the media over time. Its assimilation follows two main pathways, depending on the extracellular concentration of ammonia that is available: the glutamine synthetase (GS)/ glutamine oxoglutarate aminotransferase (GOGAT) pathway or the glutamate dehydrogenase (GDH) pathway (Figure 4-2) (Harper *et al*., 2010):

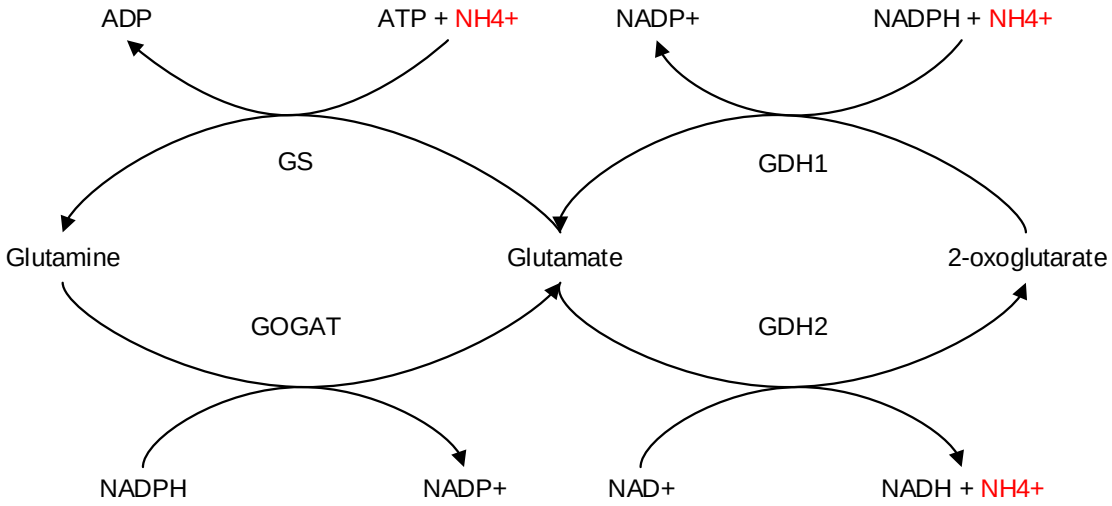

Figure 4-2: Pathways for nitrogen assimilation in M. smegmatis - adapted from (Kirsten, 2011)

When the nitrogen source is in excess, the GDH pathway is in effect. This enzyme, which has a low affinity for ammonium, catalyses the reversible amination of 2-oxoglutarate, forming glutamate (Harper *et al*., 2010). The GS/GOGAT pathway, however, is primarily active under nitrogen-limited conditions; GS, which has a high affinity for ammonium, catalyses the production of glutamine from glutamate and ammonium, while GOGAT catalyses the synthesis of glutamate from glutamine (Harper *et al*., 2010). The GS/GOGAT pathway is also more energy-intensive than the GDH pathway, as one mole of ATP is consumed for every mole of glutamine produced (Harper *et al*., 2010). In this instance, as the nitrogen concentration in the media was in excess, it can be assumed that the GDH pathway was in effect.

The phosphorus concentration (Figure 4-1 C) did not appear to change over time. Based on the previously observed concentration profiles of ammonia and glycerol, one would assume that the extracellular phosphorus concentration would also decrease with time, but this is not evident in the figure. To examine this anomaly, the theoretical amount of phosphorus that would have been taken up by the bacteria within the entire duration of the experiment was

calculated. Using a maximal uptake velocity of 0.4 nmol/min/mg cells, derived from Michaelis-Menten dynamics (Wolschendorf, Mahfoud and Niederweis, 2007), it was estimated that 45.45 mg phosphorus/L would have been taken up by the cells within 40 hours (see Appendix B for calculations). This value is very small, compared to the initial phosphorus concentration in the media (496.12 mg phosphorus/L). This indicates that the phosphorus uptake is not as significant as that of glycerol or ammonia, thus, was excluded from subsequent experimental analysis and mathematical modelling.

Normal growth with *M. smegmatis*::pTiGc produced similar growth trends (Figure 4-3), confirming that carriage of the plasmid did not affect the observed variables.

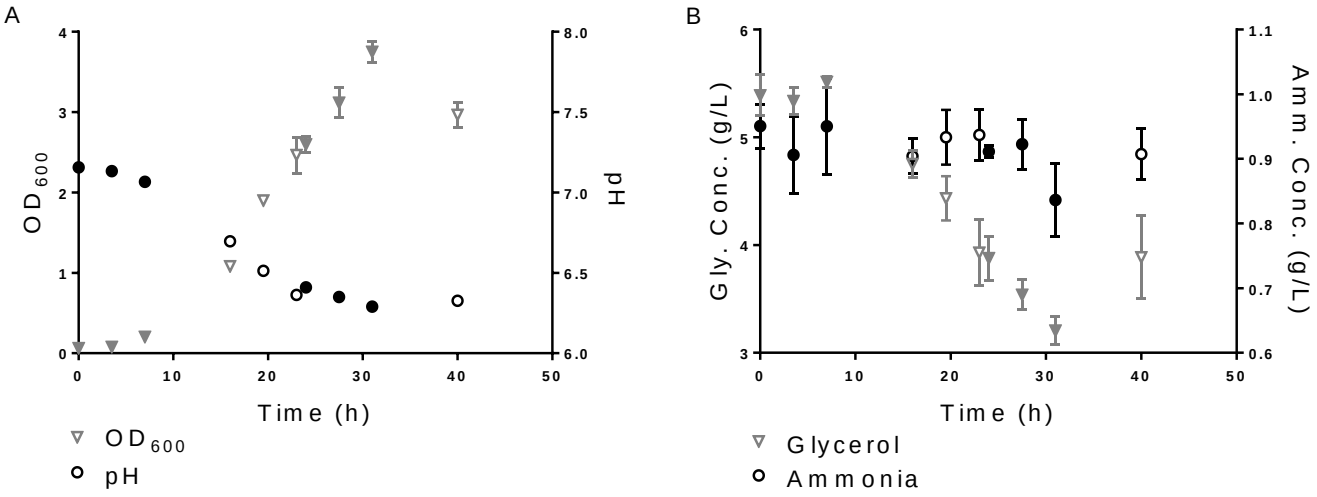

Figure 4-3: Growth trends of M. smegmatis::pTiGc. A: OD (triangles) and pH (circles); B: glycerol (triangles) and ammonia (circles). Shaded symbols represent culture a; open symbols represent culture b (staggered start times)

4.3 Objective 2: Bacterial Growth under Stressed Conditions

4.3.1 Growth Trends

Under growth-limiting conditions, *M. smegmatis*::pTiGc experienced an overall reduction of the OD (and in turn, its biomass) in the growth media (Figures 4-4 and 4-5):

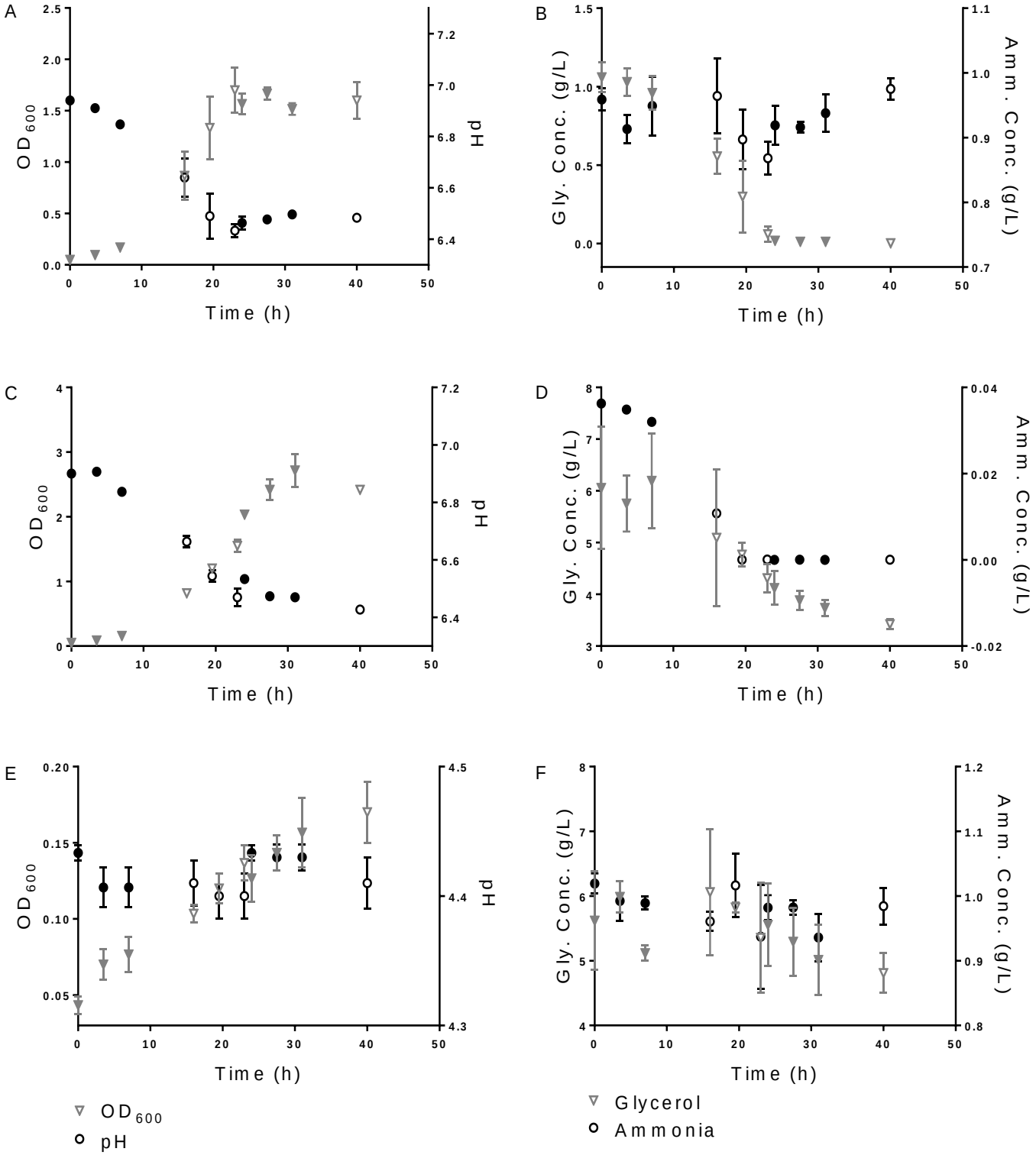

Figure 4-4: Growth limitation due to nutrient deficiency and acidic media. Left: OD (triangles) and pH (circles). Right: glycerol (triangles) and ammonia (circles). Shaded symbols represent culture a; open symbols represent culture b (staggered start times). A and B: glycerol-deficient; C and D: ammonia-deficient; E and F: pH 4.6 media

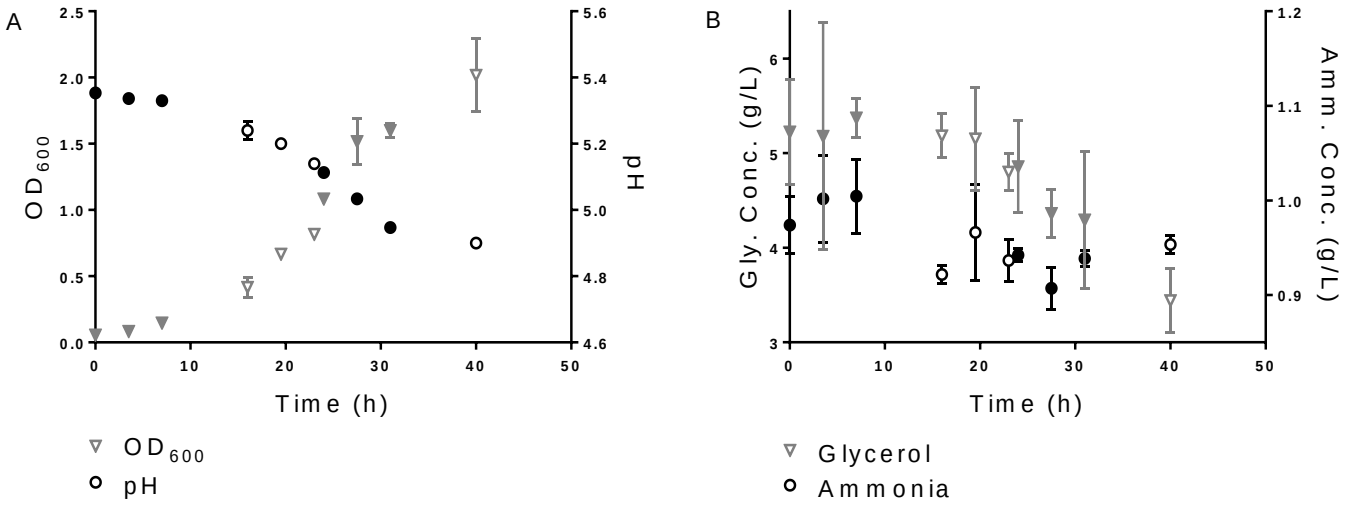

Figure 4-5: Growth limitation due to acidic media. Left: OD (triangles) and pH (circles). Right: glycerol (triangles) and ammonia (circles). Shaded symbols represent culture a; open symbols represent culture b (staggered start times). A and B: pH 5.5 media

Under carbon-deficient conditions, the exhaustion of glycerol at t = 23 h results in the bacterial population entering a stationary phase of growth (Figure 4-4 A). This is anticipated, as glycerol is the sole energy source in the growth media; its eventual depletion would severely restrict bacterial growth. This, however, is not seen in the case of ammonia deficiency (Figure 4-4 C and D). Rather than bacterial growth stagnating due to the apparent absence of the sole nitrogen source in the media (from t = 19.5 h to t = 40 h), *M. smegmatis* still continued growing, albeit at a lower growth rate – an observation corroborated in literature (Williams *et al*., 2013). A possible suggestion for this occurrence may be explained by use of the organism's genomics; previous studies have explored the nitrogen-response mechanism in *M. smegmatis*, highlighting the orphan response regulator, GInR, as the transcriptomic factor responsible for gene expression under nitrogen limitation. Aside from the fact that GInR upregulates genes that are associated with nitrogen uptake and assimilation from nitrogen sources other than ammonium (i.e. nitrates, nitrites, urea, amino acids and peptides (Jenkins *et al*., 2013)), it was also discovered that GInR increases the expression of genes which encode enzymes that are related to the breakdown of cellular components into ammonium (Jenkins *et al*., 2013). This suggests that, under nitrogen deficiency, the response of the bacteria would be to break down cellular components, thus providing ammonium to be used for bacterial growth and survival. Another hypothesis could be that ammonia is still present within the media, but at a concentration that cannot be detected by the biochemistry analyser (the detection limit being 10 mg/L). Due to nitrogen deficiency, the cells adapt by slowing down the uptake rate of the remaining ammonia – this would involve switching of nitrogen

assimilation pathways (from GDH to GS/GOGAT, Figure 4-2). This suggests that *M. smegmatis* can sense nitrogen austerity.

Very acidic growth conditions inhibit mycobacterial growth; as such, it is expected that minimal growth will be realized at a pH of 4.6, as was observed here (Figure 4-4 E and F). At pH 5.5 (Figure 4-5), bacterial growth is at a higher rate than in the previous case, though not to the extent that is seen under normal conditions. No stationary growth phase is attained under these acidic conditions.

4.3.2 Flow Cytometry

Previous studies have been able to successfully demonstrate the application of FD in monitoring *in vitro* bacterial replication (Mouton *et al*., 2016; Helaine *et al*., 2010) . To ascertain that this is the same case for these experiments, the number of generations of the bacteria, using OD values, was calculated and compared against that of the MFI data (Figure 4-6).

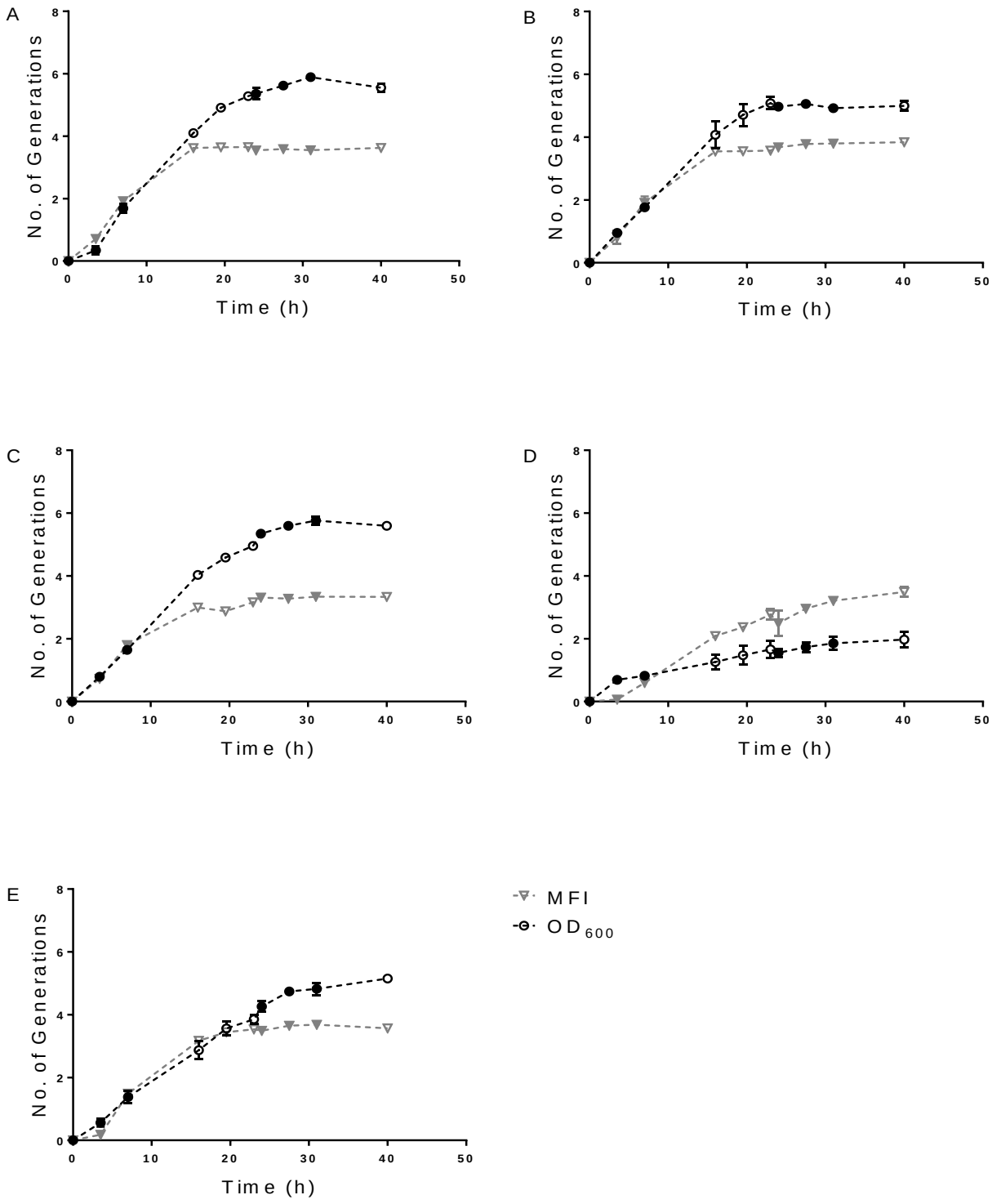

Figure 4-6: Bacterial generation numbers, calculated using the OD (black dotted line, circles) and fluorescence intensity (grey dotted line, triangles). A: normal growth; B: glycerol deficiency; C: ammonia deficiency; D: pH 4.6; E: pH 5.5. Shaded symbols represent culture a; open symbols represent culture b (staggered start times)

Examination of the previous illustration shows that there is a fairly good correlation between the two methods, for up to four generations. Beyond this, the number of generations using the MFI plateaus, as its minimum intensity has been reached (at approximately t = 16 h). This result closely relates to what is observed in the study done by Mouton *et al*. (2016), where the same reporter was able to accurately measure mycobacterial replication for up to five generations.

Having established that FD is appropriate in investigating mycobacterial replication dynamics, the differences in bacterial growth numbers – based on the OD and number of generations using MFI data, was considered (Figure 4-7):

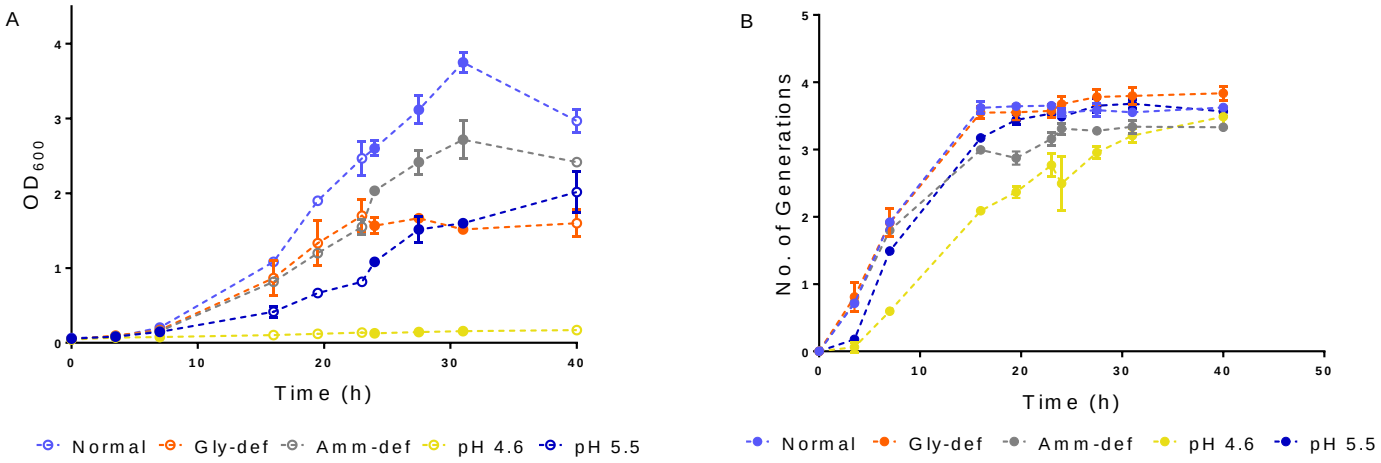

Figure 4-7: Comparison of bacterial replication using OD (A) and MFI (B)

Bacterial growth based on the OD shows acute differences among the various growth conditions; growth under normal conditions exhibits the highest turbidity, whilst growth using pH 4.6 media exhibits the lowest turbidity. This result, however, does not translate to the MFI data. As per Figure 4-7 B, the results indicate that despite the application of growth-limiting conditions, bacterial replication in four of the five conditions reaches the point at which the fluorescence intensity is at its minimum (≈ 4 generations at t = 16 h). This makes it difficult to discern the possible presence of a persister population beyond the 16-hour mark, where the fluorescence of the induced protein is not detectable. The exception of this is the cells grown in pH 4.6 media, which show promising results; the replication rate at the near-lethal pH was relatively low, indicating that there was a slow dilution of the fluorescence intensity. One would, therefore, surmise, that a persister population might be emerging under this growth condition.

4.4 Objective 3: Model Development and Parameter Estimation

4.4.1 Computational Simulations

To determine the values of the parameters that would result in the best fit of the model, two different optimization methods were carried out in MATLAB. The first utilized a local solver, *lsqnonlin*, in which three different simulations were run, each with a different set of initial values. The second method used the same solver, but with a global optimization algorithm, *MultiStart*. The results are as seen in Table 4-1:

Table 4-1: Initial and optimized values of the growth parameters

			Local Solver						**Global Solver**	
			Run 1		**Run 2**		**Run 3**			
Param.	**LB**[1]	**UB**[2]	**IC**[3]	**OV**[4]	**IC**	**OV**	**IC**	**OV**	**IC**	**OV**
μ_{max}	0.005	0.25	0.15	0.1440	0.01	0.1454	0.2	0.1563	0.15	0.1455
k_C	0.01	1	0.08	0.3021	0.9	0.3813	0.02	0.3629	0.08	0.3828
k_N	0.0001	0.05	0.01	0.0001	0.03	0.0001	0.0005	0.0001	0.01	0.0001
k_d	0.00001	0.001	0.0001	0.0010	0.0005	0.0010	0.001	0.0010	0.0001	0.0010
pH_{LL}	2	8	4.6	4.2083	3	4.3292	5	4.3210	4.6	3.4687
α	1	13	1.5	8.7898	5	12.9999	8	6.7025	1.5	12.7492
I_{val}	1	5	3	2.7062	1.5	2.4645	4	3.3633	3	4.0413
$Y_{X/CD}$	0.1	1	0.35	0.3813	0.6	0.3692	0.2	0.3232	0.35	0.3687
Residual Norm				2891.26		2865.07		3643.57		2864.97

[1]Lower bound; [2]Upper bound; [3]Initial conditions; [4]Optimized values

Given that this system is non-linear, it is bound to have several local minima. Therefore, finding the solution of the set of ODEs with a local solver will yield parameters that may not relate to the global minimum of the model. This is seen in the table above, whereby different solutions are generated by the optimization algorithm, depending on the initial conditions. As such, determining the global minimum would require the use of a global optimization algorithm. In this case, using *MultiStart* (which ran *lsqnonlin* for fifty instances) yielded the lowest objective function value (the residual norm, highlighted in dark grey). Its respective optimized values were used in subsequent simulations, as well as being contrasted to the experimental data (Figures 4-8 and 4-9). The corresponding NRMSE values of each environmental condition can be referred to in Appendix B (Table B 8).

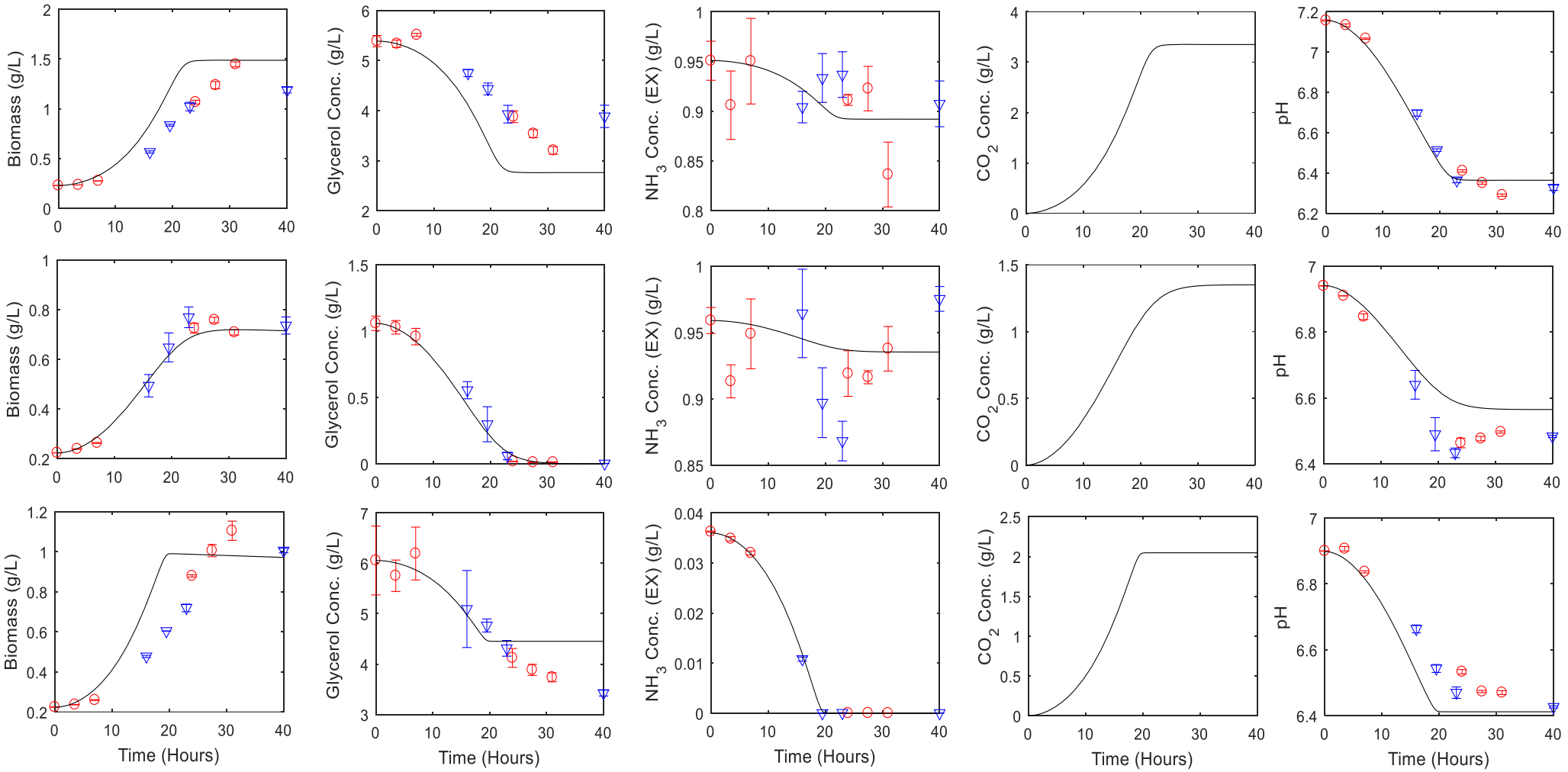

Figure 4-8: Simulation of the mathematical model with optimized growth parameters for normal growth and nutrient-deficient conditions. Normal conditions, glycerol deficiency and ammonia deficiency are represented by the first, second and third row (respectively). Solid line: optimized simulation; circles: culture a; triangles: culture b

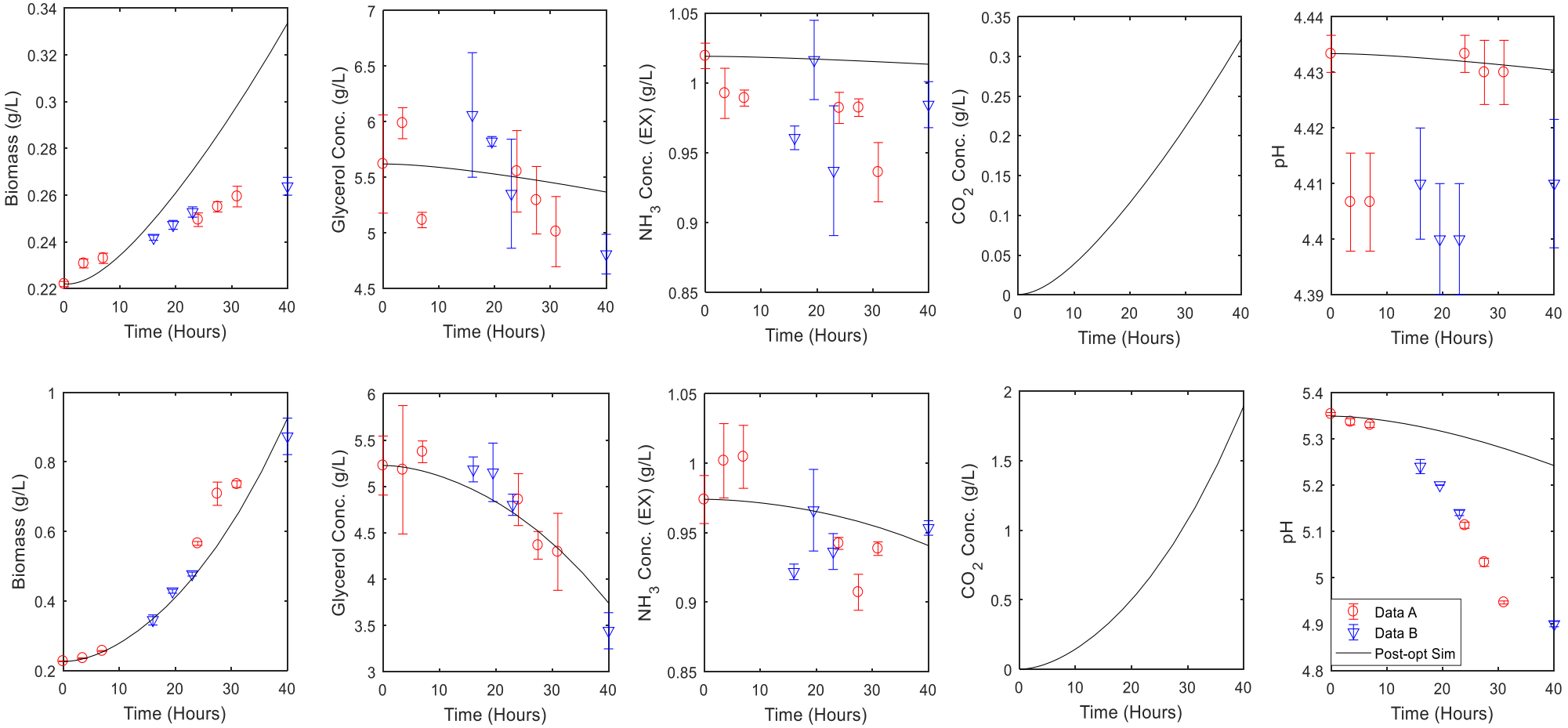

Figure 4-9: Simulation of the mathematical model with optimized growth parameters for pH-stressed conditions. pH 4.6 and pH 5.5 are represented by the first and second row (respectively). Solid line: optimized simulation; circles: culture a; triangles: culture b

Under normal conditions, the optimized simulation depicts the onset of a stationary phase occurring at approximately t = 23 h, which is contrary to what is seen in the experimental data. This supposed stationary phase is as a result of the carrying capacity term within the model, which restricts growth under normal conditions past a certain biomass concentration. Going back to the basis of the Monod model, this equation relates microbial growth under conditions in which the nutrient concentration is *limiting***.** Therefore, under normal conditions, lack of a carrying capacity term would result in continued growth until one of the nutrients is exhausted, with the predicted biomass concentration for the proposed model far exceeding what is observed in the experimental data. In this case, the response variable with the best fit is the pH, with an NRMSE of 3.53%.

Glycerol deficiency exhibits the best fit in the optimized model; the carbon-limited stationary phase is properly predicted for the biomass and glycerol response variables (NRMSE of 6.06% and 4.29%, respectively).

Under ammonia deficiency, the comparison of the experimental data to the optimized simulation indicates that the ammonia concentration has the best fit (NRMSE of 2.12%). Due to the ammonia being depleted at t = 19.5 h, the optimized simulation portrays the onset of a nitrogen-limited stationary phase, an observation which contradicts what is seen in the experimental data (continued growth in the absence of ammonium in the media).

The optimized simulation under pH 5.5 displays a good fit for the biomass and glycerol concentrations (NRMSE of 11.07% and 10.13%, respectively). This is unlike the simulation of bacterial growth in pH 4.6 media, where the fit of the model is relatively poor. This may be, in part, due to the incorporation of the lag phase to the Monod expression – while its inclusion does improve the overall fit of the proposed model, this phase is not observed when growing cells at this near-lethal pH. In fact, bacterial growth appears to be linear.

4.4.2 Consideration of Oxygen as a Variable in the Mathematical Model

M. smegmatis is an aerobic microorganism, making oxygen a necessary factor for its growth. Previous *in vitro* studies have proven that oxygen depletion is a signal that can trigger dormancy in mycobacteria (Wayne and Hayes, 1996; Dick, Lee and Murugasu-Oei, 1998). The study done by Wayne and Hayes (1996) was able to establish a setup exhibiting the increased tolerance of homogenous populations of *M. tuberculosis* bacilli to various stages of oxygen depletion, leading to the observation of at least two stages of nonreplicating persistence. Nonetheless, oxygen was not added as a response variable for the following reasons:

1. Differences in the experimental setup: Wayne and Hayes (1996) created an oxygen-limited environment by using an appropriate agitation rate (120 rpm) and HSR (0.5) which would minimize the OTR into the media (note that these cultures were also sealed). In this project, however, rapid agitation of the culture flasks was utilized (with the Erlenmeyer flasks containing baffles which would increase the OTR into the media) and HSRs were not considered, making it unlikely that oxygen limitation would occur. This experimental setup was required to facilitate culture uniformity, large numbers of culture vessels, repeated sampling and for some measures, sampling of large volumes, and was not compatible with a system such as that employed by Wayne and Hayes
2. The measured %DO values: the recorded values are illustrated in Figure 4-10. Based on the figure, it seems as though the system is oversaturated with oxygen, as most of the time points have saturation values above 100% (except at t = 0 h and t = 3.5 h). Under normal circumstances, the oxygen saturation decreases as the cells grow in the media. This observed trend may, therefore, have been due to instrument error whilst using the %DO meter.

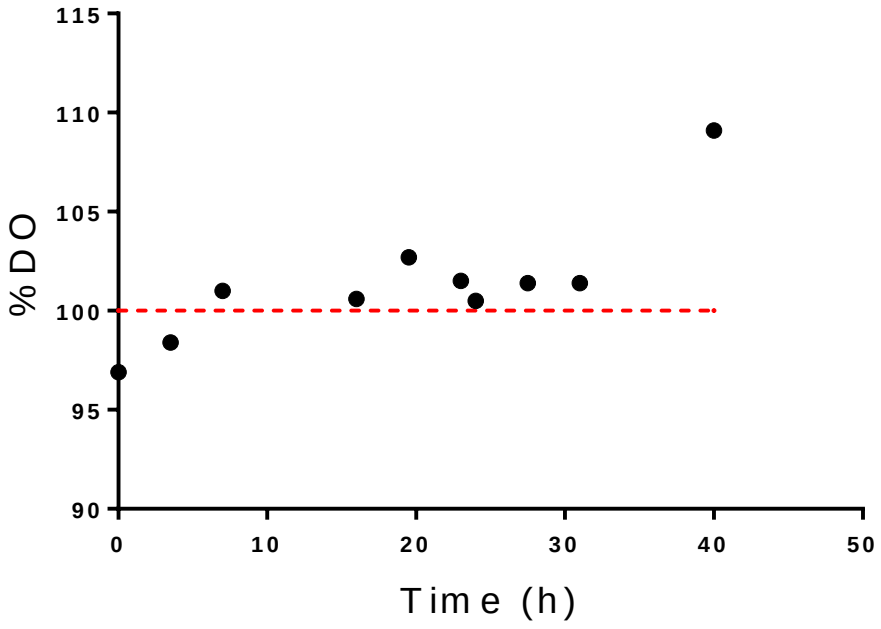

Figure 4-10: Initial recorded values of %DO at each time point. Values above the dashed line indicate oversaturation of oxygen in the media

If oxygen were added as an additional response variable to the mathematical model, simulation of the system would yield results as depicted in Figures 4-11 and 4-12. Estimation of additional constant model parameters – k_La and $Y_{X/O}$ – is described in Appendix B.

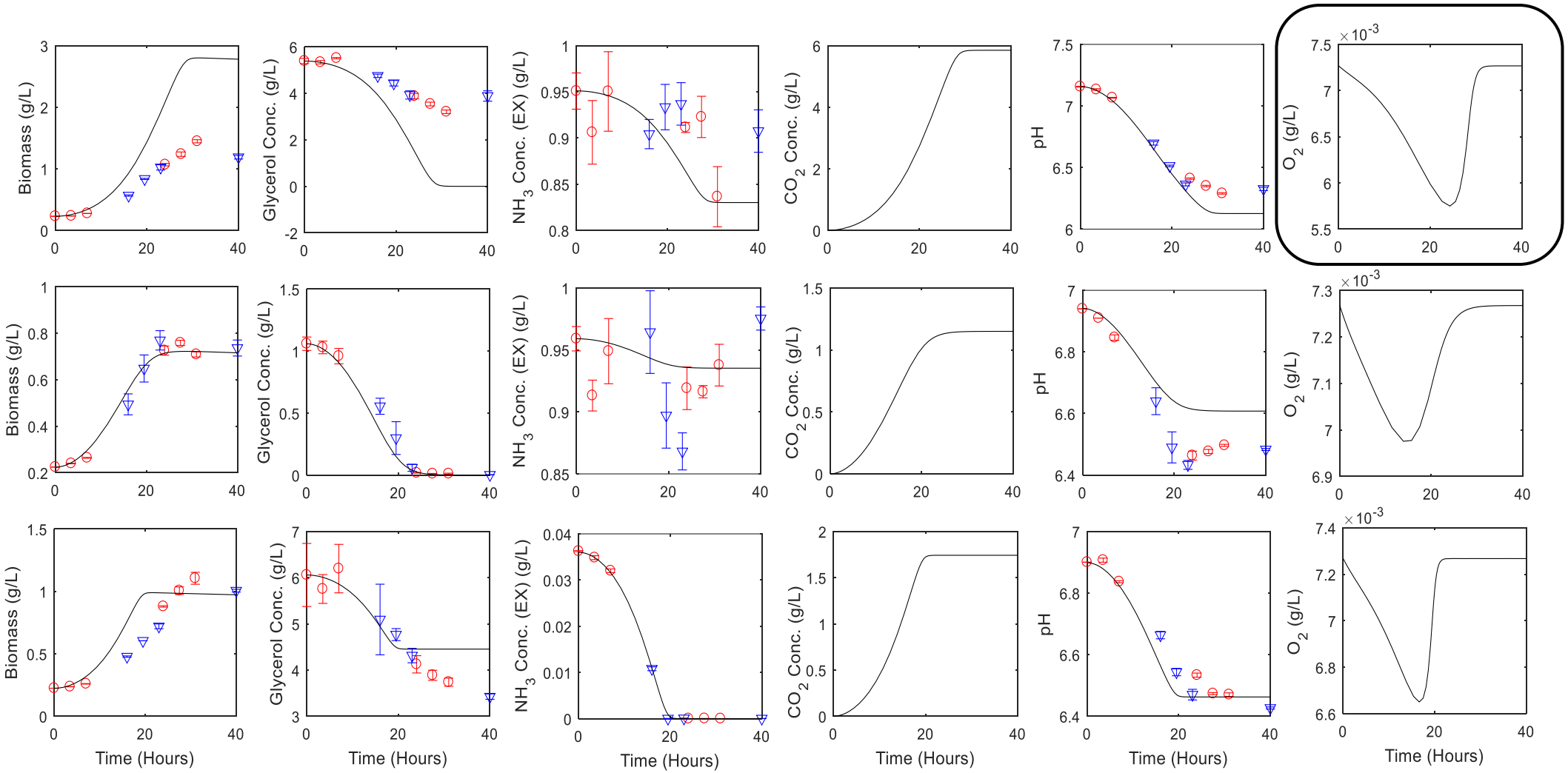

Figure 4-11: Simulation of the mathematical model with oxygen as an added variable for normal growth and nutrient-deficient conditions. Normal conditions, glycerol deficiency and ammonia deficiency are represented by the first, second and third row (respectively). Solid line: optimized simulation; circles: culture a; triangles: culture b

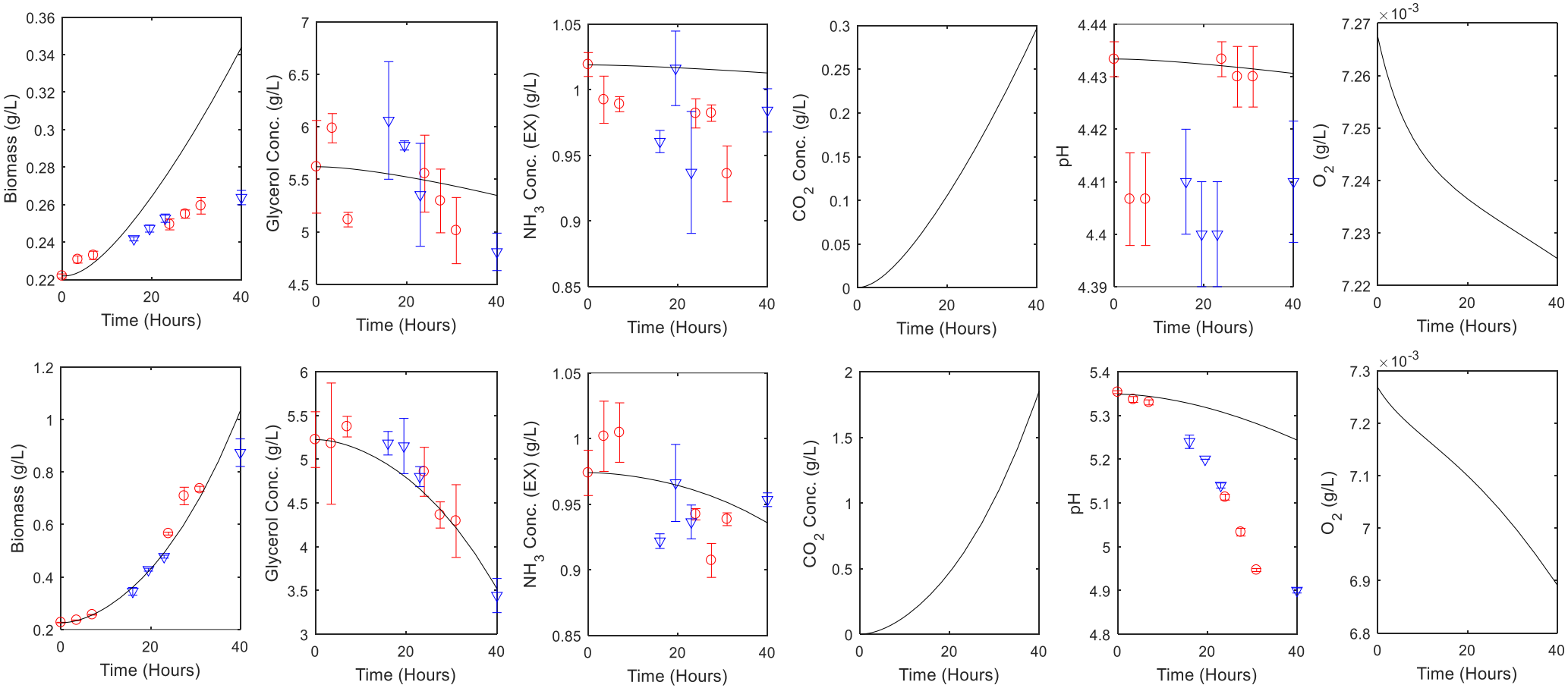

Figure 4-12: Simulation of the mathematical model with oxygen as an added variable for pH-stressed conditions. pH 4.6 and pH 5.5 are represented by the first and second row (respectively). Solid line: optimized simulation; circles: culture a; triangles: culture b

For this scenario, the simulations were run without a carrying capacity coefficient, to investigate what effect the concentration of oxygen in the system would have on the growth of the bacteria. Under normal conditions (Figure 4-11, row 1), without α or X_{max}, excessive growth is predicted by the mathematical model, up until glycerol is depleted, hence, initiating a stationary phase at t = 31 h. From this figure, it is evident that the variable that limits growth under normal conditions is glycerol, and not the oxygen saturation. The concentration of oxygen (circled graph in Figure 4-11) recovers during the stationary phase that is predicted by the optimized simulation.

Therefore, much as oxygen limitation is a factor that could contribute to mycobacterial persistence, its concentration (as per the experimental setup of this study) has a minimal effect on the decline in bacterial growth, if at all.

4.5 Objective 4: Model Validation

The procedure for carrying out the validation experiments was like that of the parameter-estimation experiments, albeit with a longer duration. In this case, a combination of growth stresses was applied to various experimental runs, with the data being used to explore the predictive power of the proposed mathematical model. This was done by comparing the experimental data to the optimized computational simulation (Figures 4-13, 4-14 and 4-15).

Under normal conditions, the growth rate is relatively over-estimated at the beginning of the experiment (due to the substrates being in excess), up to the point at which the carrying capacity of the model is reached. Nonetheless, the model gives a relatively good prediction of the pH (NRMSE = 13.40%, Figure 4-13).

Ammonia deficiency was repeated, under a longer time frame. Much as the model was able to predict the exhaustion of extracellular ammonia (NRMSE = 2.99%), it was expected that the model would not capture the increase in biomass, and the decrease in glycerol concentration, after ammonia depletion (Figure 4-13).

Regarding the glycerol- and ammonia-deficient experiment (Figure 4-13), the model does offer a better prediction for three of the response variables (NRMSE of 13.32%, 6.01% and 13.96%, for biomass, glycerol and pH, respectively). It, however, is not able to properly predict the depletion of the nitrogen source. Glycerol (as per the simulation) is utilized faster than ammonia, leading to exhaustion of the carbon source before the nitrogen source. This contradicts experimental data, where both substrates are depleted by 27.5 h.

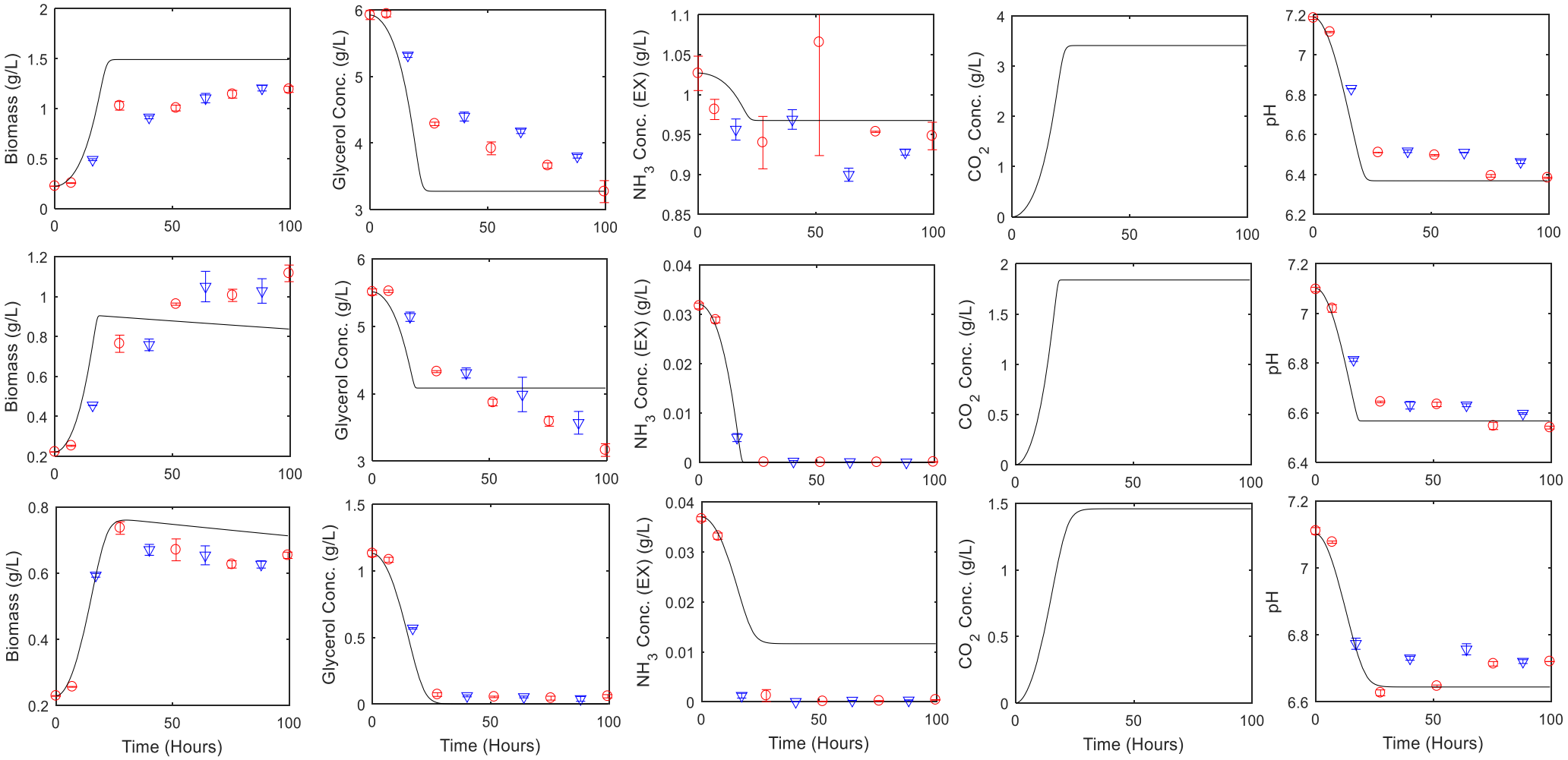

Figure 4-13: Comparison of the optimized simulation (solid line) to experimental data for normal growth and nutrient-limited conditions. Normal growth, ammonia deficiency and ammonia- and glycerol-deficiency are represented by the first, second and third row, respectively. Circles: culture a; triangles: culture b

For the spiking experiments (Figures 4-14 and 4-15), stress factors were applied sequentially; growth started off with nutrient-deficiency (before the media was spiked with the deficient substrate at t = 24 h), followed by acidic stress (at t = 55 h). The aim of this approach was to test the robustness of the model in predicting the effect of consecutive stresses on the microbial system.

An interesting aspect to note from the experimental data is the bacterial response to nutrient spiking – glycerol spiking elicits a higher OD in the media than ammonia spiking, despite neither nutrient being limiting after the 24-hour mark in both runs (Figures 4-14 and 4-15).

Comparison of the simulation to the data of the ammonia spiking experiment (Figure 4-14) denotes a very good fit for the prediction of the ammonia concentration and the pH (NRMSE of 4.39% and 5.28%, respectively). Due to the system no longer being ammonia-deficient, excessive bacterial growth is envisaged (by the model) after spiking, up until the carrying capacity is reached. Furthermore, adding concentrated HCl to the media has a minimal effect on the viability of the bacteria – *M. smegmatis* can grow across a wide range of pH values, therefore, bacterial growth is not severely affected on addition of HCl, so long as the pH is kept above the range in which partial growth would occur (pH 4.6 to 4.9) (Portaels and Pattyn, 1982). Likewise, the model does not capture the influence of the acid on bacterial growth.

A similar result is also observed in the glycerol spiking experiment (Figure 4-15), in which the simulation of the glycerol concentration and the pH exhibit very similar trends to the observed values (NRMSE of 5.80% and 3.38%, respectively). Addition of HCl, also, does not negatively influence bacterial growth.

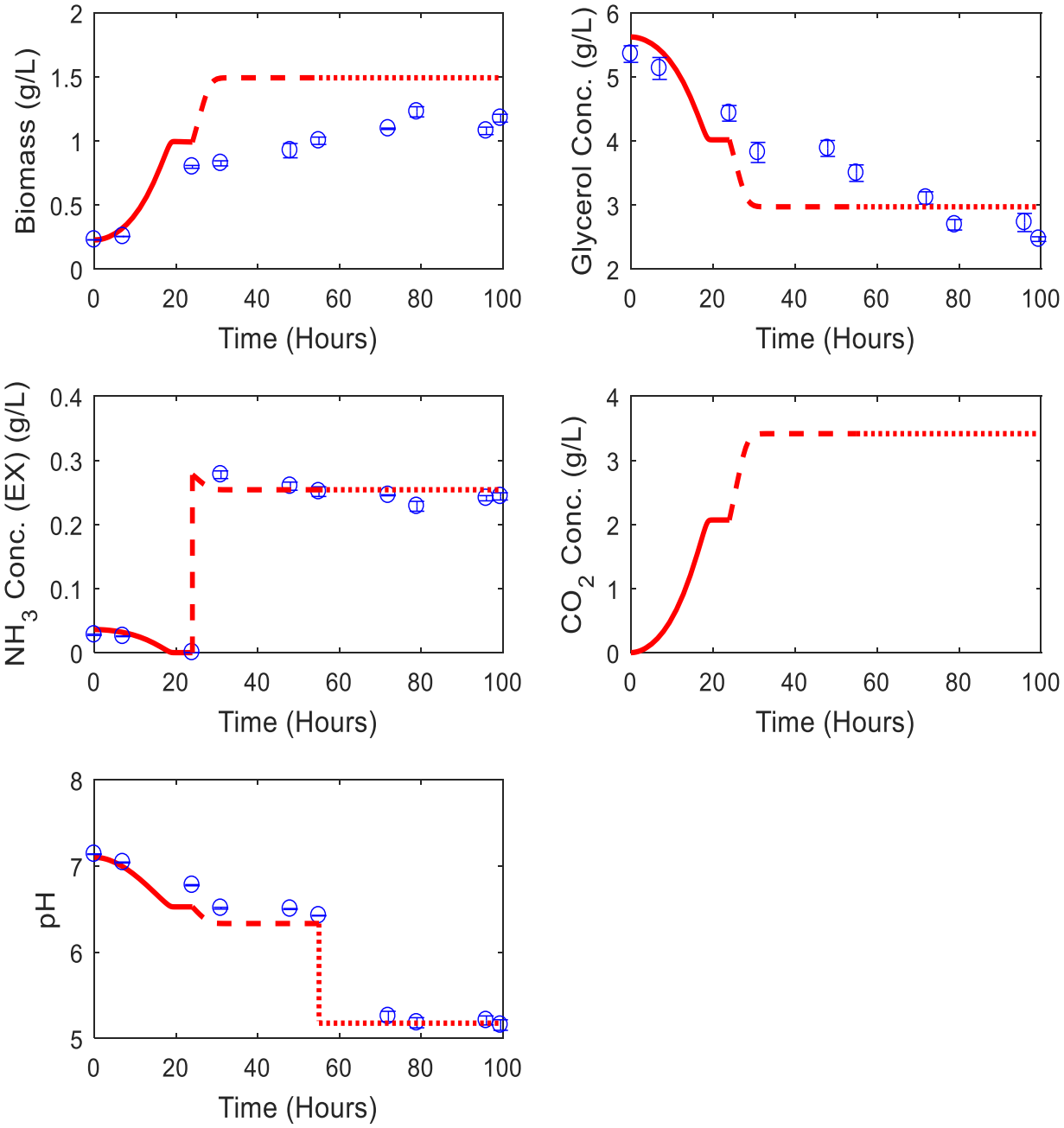

Figure 4-14: Comparison of the simulation to experimental data, with spiking. The solid line represents the simulation under ammonia deficiency; the dashed line represents the simulation after ammonia spiking; the dotted line represents the simulation after acidifying the media with concentrated HCl. Ammonia spiking was done at t = 24 h with 0.25 M $(NH_4)_2SO_4$; acid spiking was done at t = 55 h

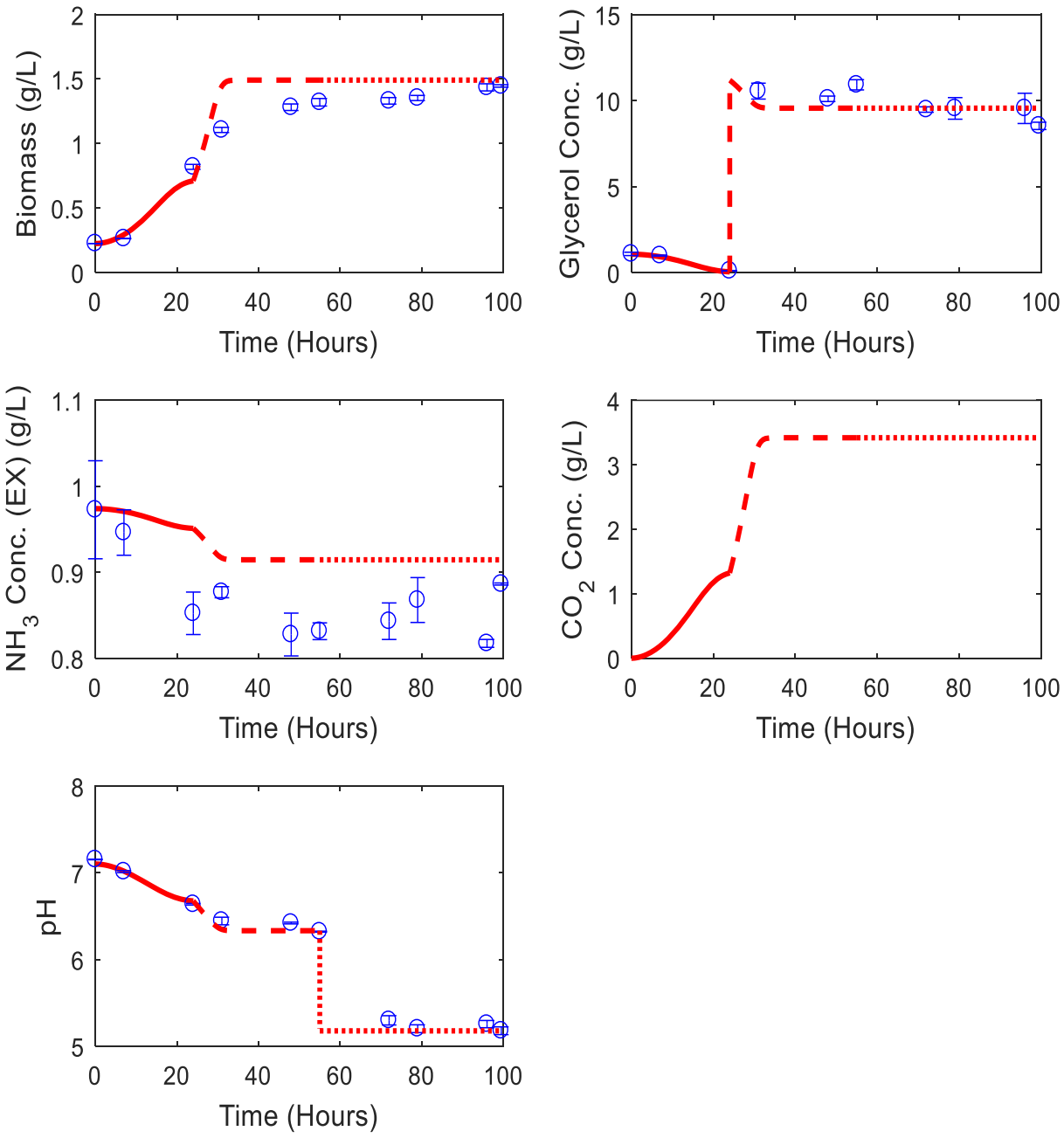

Figure 4-15: Comparison of the simulation to experimental data, with spiking. The solid line represents the simulation under glycerol deficiency; the dashed line represents the simulation after glycerol spiking; the dotted line represents the simulation after acidifying the media with concentrated HCl. Glycerol spiking was done at t = 24 h with 50% glycerol; acid spiking was done at t = 55 h

4.6 Conclusion

In this chapter, bacterial replication dynamics was explored, with the aim of investigating the bacterial response to growth-limiting conditions. In this regard, two techniques were used: mathematical modelling (for prediction of bacterial growth and substrate utilization under various environmental conditions) and FD (for examining the possible emergence of a persister population under stressed growth conditions).

The proposed mathematical model was used to estimate the relevant growth parameters of *M. smegmatis*, with the predictive power of the optimized simulation being assessed through comparison with experimental data. Additionally, wet-lab experiments were carried out, to not only come up with data which would be used for model regression and validation, but to also analyse the dilution of the induced red fluorescence under the various growth conditions.

The Monod expression, which formed the basis of the proposed mathematical model, was used to predict the production of biomass, substrate utilization and pH of the microbial system. A global optimization algorithm, which utilized a weighted non-linear least squares regression approach, was implemented to estimate the growth parameters. The optimized simulation gave a relatively good prediction of bacterial growth, with the pH showing the best fit (based on the calculated NRMSE values). Further modification of the mathematical model, nevertheless, is necessary, to improve its fit to the experimental data.

Bacterial replication dynamics under growth-limiting conditions was explored, with the use of a dual-fluorescence reporter. Comparison of the MFI plots to the normal case showed that the cells growing in pH 4.6 media retained the red fluorescence for a longer period of time, indicating that they were growing at a very slow rate. It is probable that this environmental condition could encourage the switching of cells from an active to a dormant state.

The results in this chapter will be tied back to the research questions and project objectives, to establish whether the aims of the study were attained. In addition to this, recommendations for future research will be discussed (in the following chapter).

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The aims of the study were to establish a mathematical model that could predict the growth of *M. smegmatis* under various environmental conditions, and to explore possible persister formation under conditions that would limit mycobacterial growth; the former was achieved. Based on the project objectives, the following conclusions can be drawn:

1. Through conducting growth experiments under normal conditions, using both *M. smegmatis* mc^2155 and *M. smegmatis*::pTiGc, the appropriate variables affecting growth were identified, and were consequently added to the mathematical model
2. FD can indeed be used as a measure of bacterial replication dynamics, as the number of generations calculated using MFI data closely correlated to that which was calculated using OD data, for up to four generations. However, the differences in bacterial growth under the various environmental conditions, based on the shift in the intensity of the red fluorescent protein, was not as evident, except in the case of cells grown in pH 4.6 media. The slow growth of the cells in this highly acidic environment (leading to a higher retention of the red fluorescence) may indicate the possible emergence of a persister population. The MFI profiles under the other growth-limiting conditions, on the other hand, closely mimicked the dilution trend under normal conditions, where the minimum intensity was reached at around t = 16 h. This made it difficult to discern how these stresses affected *M. smegmatis* growth, given that their effects would only be felt at the later time points.
3. The proposed mathematical model was developed based on the variables influencing the growth of *M. smegmatis* in the liquid media. The Monod expression was chosen to describe the growth rate of the bacteria, with relevant modifications to incorporate the lag phase after inoculation, and the inhibitory effect of an increasingly acidic pH environment on bacterial growth. Using a global optimization algorithm – which incorporated a weighted non-linear least squares regression technique – the relevant growth parameters were estimated. At this juncture, the justification for omitting oxygen from the proposed model was reinforced, by running the simulation with oxygen as a response variable. The results revealed that under normal conditions, the variable resulting in the decline of bacterial growth is the glycerol concentration, and not the oxygen saturation
4. The predictive power of the proposed mathematical model, when the simulation was contrasted with data from the validation experiments, was relatively good. Nonetheless, it is evident that the model needs to be modified further, to incorporate the following observations seen in the experimental data:

- The slow decline in bacterial growth under normal conditions – none of the substrates in this case are limiting, so it is apparent that another variable is at play, resulting in the decrease of the growth rate of the bacteria. Doing so would eliminate the need for the carrying capacity parameter
- The continued growth of the organism after depletion of the nitrogen source – the current model predicts the onset of a stationary phase once ammonium is exhausted, but experimental data shows otherwise

5.2 Recommendations

1. **Bacterial replication dynamics with FD**: The MFI results from the second project objective indicated similar trends in the dilution of the intensity of the red fluorescent protein (when comparing the normal case to that of the growth-limiting conditions), even though a smaller shift in the MFI is expected under cases that limit bacterial growth. For that reason, it would be recommended that the induced cells be cultured into the nutrient-deficient media, *without* removing the theophylline, and only have the inducer withdrawn after t = 16 h. Knowing that the effect of the stresses on bacterial growth will be more apparent at the later time points, this approach ensures that differences in dilution of the red fluorescence across various environmental conditions are more ostensible. Furthermore, given that the MFI cannot wholly ascertain the presence of a VBNR population, further analysis would be necessary to positively identify persisters under the various stresses. A possible strategy would be to sort the live bacterial population based on the expression of the red fluorophore, then treat both groups with an antibiotic (which would have no effect on the persisters). The TurboFP635+ subpopulation should have a higher survival rate than the TurboFP635- majority, hence, proving that persisters are present within the culture.
2. **Oxygen modelling**: Even though oxygen was not modelled, due to the experimental setup used in this study, it is still an important variable that can influence bacterial growth, as well as phenotypic changes to the cells, particularly in instances of hypoxia. Thus, it is recommended that oxygen be added to the mathematical model but based on a setup that would result in hypoxic conditions being formed as the experiment progresses with time. For instance, using a rubber septum to seal the Erlenmeyer flask (to prevent entry of oxygen into the headspace volume), sampling by use of a syringe (to prevent having to open the top of the Erlenmeyer flask, which would reintroduce oxygen into the headspace volume) and agitating the liquid culture at an rpm that would ensure homogeneity within the culture, but not disturb its surface (minimizing the OTR into the media).

3. **Microbial growth under normal conditions**: The Monod equation relates microbial growth under conditions in which the substrate is limiting, and as such, is an aspect which affects the simulation of microbial growth under normal conditions, where the substrates are in abundance (thus, the need for a carrying capacity coefficient to limit bacterial growth up to a certain biomass concentration). Model predictability could be greatly improved if the variable resulting in a decreasing bacterial growth rate under normal conditions would be captured, rather than forcing bacterial growth to stop at a certain point, through the introduction of a carrying capacity parameter. Hence, it would be of benefit to look into the metabolites (other than CO_2, which has been modelled) produced by the microorganism during growth, to examine those that may result in growth inhibition. This may be done through a high-performance liquid chromatography analysis of the growth media.
4. **Use of antibiotics**: Having explored the response of *M. smegmatis* to nutrient deficiency and acidic pH, it would be of interest to determine if the model would be able to properly simulate growth dynamics of the bacteria upon exposure to a bactericidal agent.
5. **Modelling of ammonia uptake**: Having seen that the proposed model is not able to properly predict bacterial growth after depletion of a nitrogen source, future studies could involve adding more structure to the mathematical model, particularly regarding ammonia uptake and utilization. Rather than exclusively modelling bacterial growth based on the concentration of ammonia, the metabolites and enzymes associated with its uptake and utilization can also be accounted for, thereby improving model predictability.
6. **Application of the model to slow-growing mycobacteria**: When examining mycobacterial growth, *M. smegmatis* is often used as a model for *M. tuberculosis*, as it is a non-pathogenic species with a faster generation time. Having established a mathematical model describing its growth kinetics, it would be interesting to investigate if this proposed model would be suitable in predicting the growth of *M. tuberculosis*, by varying the μ_{max} parameter whilst keeping the rest constant.
7. **Phenotype switching in a biofilm**: Having seen that one of the factors influencing the recalcitrance of biofilms to environmental aggressions and bactericidal agents is the presence of a persister population, a recommendation for future studies would be to substitute the estimated growth parameters into an individual-based biofilm model, and investigate phenotypic heterogeneity that would arise during simulated biofilm development. The same dual-fluorescence reporter would also be used, alongside confocal microscopy, to be able to identify the VBNR population.

References

Adams, K.N., Takaki, K., Connolly, L.E., Wiedenhoft, H., Winglee, K., Humbert, O., Edelstein, P.H., Cosma, C.L. and Ramakrishnan, L., 2011. Drug tolerance in replicating mycobacteria mediated by a macrophage-induced efflux mechanism. *Cell*, 145(1), pp.39–53.

Anuchin, A.M., Mulyukin, A.L., Suzina, N.E., Duda, V.I., El-Registan, G.I. and Kaprelyants, A.S., 2009. Dormant forms of Mycobacterium smegmatis with distinct morphology. *Microbiology*, 155(4), pp.1071–1079.

Balaban, N.Q., Gerdes, K., Lewis, K. and McKinney, J.D., 2013. A problem of persistence: still more questions than answers? *Nature reviews. Microbiology*, 11(8), pp.587–91.

Balaban, N.Q., Helaine, S., Lewis, K., Ackermann, M., Aldridge, B., Andersson, D.I., Brynildsen, M.P., Bumann, D., Camilli, A., Collins, J.J., Dehio, C., Fortune, S., Ghigo, J.M., Hardt, W.D., Harms, A., Heinemann, M., Hung, D.T., Jenal, U., Levin, B.R., Michiels, J., Storz, G., Tan, M.W., Tenson, T., Van Melderen, L. and Zinkernagel, A., 2019. Definitions and guidelines for research on antibiotic persistence. *Nature Reviews Microbiology*, 17(July), pp.441–448.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L. and Leibler, S., 2004. Bacterial persistence as a phenotypic switch. *Science*, 305(5690), pp.1622–1625.

Bergter, F. and Knorre, W.A., 1972. Computer simulation of growth and product formation in Saccharomyces cerevisiae. *Journal of Basic Microbiology*, 12(8), pp.613–629.

Betts, J.C., Lukey, P.T., Robb, L.C., McAdam, R.A. and Duncan, K., 2002. Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. 43, pp.1–15.

Bhat, R.M. and Prakash, C., 2012. Leprosy: An overview of pathophysiology. *Interdisciplinary Perspectives on Infectious Diseases*, 2012, pp.1–6.

Bigger, J.W., 1944. Treatment of Staphylococcal infections by intermittent sterilisation. *The Lancet*, pp.497–500.

Brauner, A., Fridman, O., Gefen, O. and Balaban, N.Q., 2016. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nature Reviews Microbiology*, 14(5), pp.320–330.

Brooks, G.F., Carroll, K.C., Butel, J.S., Morse, S.M. and Mietzner, T.A., 2013. *Jawetz, Melnick & Adelberg's Medical Microbiology*. 26th ed. Mc Graw Hill.

Carroll, P., Schreuder, L.J., Muwanguzi-Karugaba, J., Wiles, S., Robertson, B.D., Ripoll, J.,

Ward, T.H., Bancroft, G.J., Schaible, U.E. and Parish, T., 2010. Sensitive detection of gene expression in mycobacteria under replicating and non-replicating conditions using optimized far-red reporters. *PLoS ONE*, 5(3), pp.1–17.

Carvalho, G., Balestrino, D., Forestier, C. and Mathias, J.D., 2018. How do environment-dependent switching rates between susceptible and persister cells affect the dynamics of biofilms faced with antibiotics? *npj Biofilms and Microbiomes*, 4(1), pp.2–9.

Chatterji, D. and Ojha, A.K., 2001. Revisiting the stringent response, ppGpp and starvation signaling. *Current Opinion in Microbiology*, 4(2), pp.160–165.

Cogan, N.G., Gunn, J.S. and Wozniak, D.J., 2011. Biofilms and infectious diseases: Biology to mathematics and back again. *FEMS Microbiology Letters*, 322(1), pp.1–7.

Cohen, N.R., Lobritz, M.A. and Collins, J.J., 2013. Microbial persistence and the road to drug resistance. *Cell Host and Microbe*, 13(6), pp.632–642.

Costerton, J.W., Stewart, P.S. and Greenberg, E.P., 1999. Bacterial Biofilms: A Common Cause of Persistent Infections. *Science*, 284(May), pp.1318–1322.

Deb, C., Lee, C.M., Dubey, V.S., Daniel, J., Abomoelak, B., Sirakova, T.D., Pawar, S., Rogers, L. and Kolattukudy, P.E., 2009. A novel in vitro multiple-stress dormancy model for mycobacterium tuberculosis generates a lipid-loaded, drug-tolerant, dormant pathogen. *PLoS ONE*, 4(6), pp.1–15.

Dhar, N. and McKinney, J.D., 2007. Microbial phenotypic heterogeneity and antibiotic tolerance. *Current Opinion in Microbiology*, 10(1), pp.30–38.

Dick, T., Lee, B.H. and Murugasu-Oei, B., 1998. Oxygen depletion induced dormancy in Mycobacterium smegmatis. *FEMS microbiology letters*, 163(2), pp.159–164.

Dräger, A. and Palsson, B., 2014. Improving collaboration by standardization efforts in systems biology. *Frontiers in Bioengineering and Biotechnology*, 2(DEC), pp.1–20.

Dutta, N.K., Klinkenberg, L.G., Vazquez, M.J., Segura-Carro, D., Colmenarejo, G., Ramon, F., Rodriguez-Miquel, B., Mata-Cantero, L., Francisco, E.P. De, Chuang, Y.M., Rubin, H., Lee, J.J., Eoh, H., Bader, J.S., Perez-Herran, E., Mendoza-Losana, A. and Karakousis, P.C., 2019. Inhibiting the stringent response blocks Mycobacterium tuberculosis entry into quiescence and reduces persistence. *Science Advances*, 5(3), pp.1–13.

Engineers Edge, 2019. *Water-Density, Viscosity, Specific Weight*. [online] Available at: <https://www.engineersedge.com/physics/water__density_viscosity_specific_weight_13146.htm> [Accessed 13 Aug. 2019].

Esteban, J. and García-Coca, M., 2018. Mycobacterium biofilms. *Frontiers in Microbiology*, 8(JAN), pp.1–8.

Faria, S., Joao, I. and Jordao, L., 2015. General Overview on Nontuberculous Mycobacteria, Biofilms, and Human Infection. *Journal of Pathogens*, 2015, pp.1–10.

Fisher, R.A., Gollan, B. and Helaine, S., 2017. Persistent bacterial infections and persister cells. *Nature Reviews Microbiology*, 15(8), pp.453–464.

Gusarov, I., Shatalin, K., Starodubtseva, M. and Nudler, E., 2009. Endogenous Nitric Oxide Protects Bacteria Against a Wide Spectrum of Antibiotics. *Science*, 325(5946), pp.1380–1384.

Han, P. and Bartels, D.M., 1996. Temperature Dependence of Oxygen Diffusion in H2O and D2O. *The Journal of Physical Chemistry*, 100(13), pp.5597–5602.

Harper, C.J., Hayward, D., Kidd, M., Wiid, I. and van Helden, P., 2010. Glutamate dehydrogenase and glutamine synthetase are regulated in response to nitrogen availability in Mycobacterium smegmatis. *BMC microbiology*, 10(138), pp.1–12.

Haynes, W.M., 2016. *CRC Handbook of Chemistry and Physics*. 97th ed. CRC Press.

Helaine, S., Thompson, J.A., Watson, K.G., Liu, M., Boyle, C. and Holden, D.W., 2010. Dynamics of intracellular bacterial replication at the single cell level. *Proceedings of the National Academy of Sciences*, 107(8), pp.3746–3751.

Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., Cuellar, A.A., Dronov, S., Gilles, E.D., Ginkel, M., Gor, V., Goryanin, I.I., Hedley, W.J., Hodgman, T.C., Hofmeyr, J.H., Hunter, P.J., Juty, N.S., Kasberger, J.L., Kremling, A., Kummer, U., Le Novère, N., Loew, L.M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E.D., Nakayama, Y., Nelson, M.R., Nielsen, P.F., Sakurada, T., Schaff, J.C., Shapiro, B.E., Shimizu, T.S., Spence, H.D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J. and Wang, J., 2003. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4), pp.524–531.

IWA Task group for Mathematical Modelling of Anaerobic Digestion Processes, 2002. *Anaerobic Digestion Model No. 1 (ADM1)*. London: IWA Publishing.

Jain, V., Kumar, M. and Chatterji, D., 2006. ppGpp: Stringent response and survival. *Journal of Microbiology*, 44(1), pp.1–10.

Jenkins, V.A., Barton, G.R., Robertson, B.D. and Williams, K.J., 2013. Genome wide analysis of the complete GlnR nitrogen-response regulon in Mycobacterium smegmatis. *BMC*

Genomics, 14(1), pp.1–15.

Kirschner, D., Pienaar, E., Marino, S. and Linderman, J.J., 2017. A review of computational and mathematical modeling contributions to our understanding of Mycobacterium tuberculosis within-host infection and treatment. *Current Opinion in Systems Biology*, 3, pp.170–185.

Kirsten, C.J., 2011. *Nitrogen metabolism and the regulation thereof in Mycobacterium smegmatis*. Stellenbosch University.

Klöckner, W. and Büchs, J., 2012. Advances in shaking technologies. *Trends in Biotechnology*, 30(6), pp.307–314.

Leistikow, R.L., Morton, R.A., Bartek, I.L., Frimpong, I., Wagner, K. and Voskuil, M.I., 2010. The Mycobacterium tuberculosis DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *Journal of Bacteriology*, 192(6), pp.1662–1670.

Levin-Reisman, I., Ronin, I., Gefen, O., Braniss, I., Shoresh, N. and Balaban, N.Q., 2017. Antibiotic tolerance facilitates the evolution of resistance. *Science*, 355(6327), pp.826–830.

Lewis, K., 2001. Riddle of Biofilm Resistance. *Antimicrobial Agents and Chemotherapy*, 45(4), pp.999–1007.

Lewis, K., 2010. Persister Cells. *Annual Review of Microbiology*, 64(1), pp.357–372.

Li, C., Donizelli, M., Rodriguez, N., Dharuri, H., Endler, L., Chelliah, V., Li, L., He, E., Henry, A., Stefan, M.I., Snoep, J.L., Hucka, M., Le Novère, N. and Laibe, C., 2010. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4(92), pp.1–14.

Loebel, R.O., Shorr, E. and Richardson, H.B., 1933a. The influence of adverse conditions upon the respiratory metabolism and growth of human tubercle bacilli. *Journal of Bacteriology*, XXVI(2), pp.167–200.

Loebel, R.O., Shorr, E. and Richardson, H.B., 1933b. The influence of foodstuffs upon the respiratory metabolism and growth of human tubercle bacilli. *Journal of Bacteriology*, XXVI(2), pp.139–166.

Luthuli, B.B., Purdy, G.E. and Balagaddé, F.K., 2015. Confinement-induced drug-tolerance in mycobacteria mediated by an efflux mechanism. *PLoS ONE*, 10(8), pp.1–16.

Magombedze, G., Dowdy, D. and Mulder, N., 2013. Latent Tuberculosis: Models, Computational Efforts and the Pathogen's Regulatory Mechanisms during Dormancy. *Frontiers in Bioengineering and Biotechnology*, 1(August), pp.1–15.

Magombedze, G. and Mulder, N., 2012. A mathematical representation of the development of

Mycobacterium tuberculosis active, latent and dormant stages. *Journal of Theoretical Biology*, 292, pp.44–59.

Mathews, C.K., van Holde, K.E., Appaling, E.R. and Anthony-Cahill, S.J., 1990. *Biochemistry*. Redwood City: Benjamin/Cummings.

Molecular Probes Inc, 2004. *LIVE/DEAD® BacLight™ Bacterial Viability and Counting Kit*.

Monod, J., 1949. The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1), pp.371–394.

Mouton, J.M., Helaine, S., Holden, D.W. and Sampson, S.L., 2016. Elucidating population-wide mycobacterial replication dynamics at the single-cell level. *Microbiology*, 162, pp.966–978.

Nayak, N., 2015. Mycobacterium Tuberculosis Biofilm – A new perspective. *Indian Journal of Tuberculosis*, 62(1), pp.4–6.

Nguyen, D., Joshi-datar, A., Lepine, F., Bauerle, E., Olakanmi, O., Beer, K., McKay, G., Siehnel, R., Schafhauser, J., Wang, Y., Britigan, B.E. and Singh, P.K., 2011. Active Starvation Responses Mediate Antibiotic Tolerance in Biofilms and Nutrient-Limited Bacteria. *Science (New York, N.Y.)*, 334(6058), pp.982–986.

NIST, 2012. Accounting for Non-Constant Variation Across the Data. In: *NIST/SEMATECH e-Handbook of Statistical Methods*. [online] NIST/SEMATECH. Available at: <http://www.itl.nist.gov/div898/handbook/>.

Le Novere, N., Bornstein, B., Broicher, A., Courtot, M., Donizelli, M., Dharuri, H., Li, L., Sauro, H., Schilstra, M., Shapiro, B., Snoep, J.L. and Hucka, M., 2006. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34(90001), pp.D689–D691.

Nyka, W., 1974. Studies on the effect of starvation on mycobacteria. *Infection and Immunity*, 9(5), pp.843–850.

O'Brien, L.M., Gordon, S. V, Roberts, I.S. and Andrew, P.W., 1996. Response of Mycobacterium smegmatis to acid stress. *FEMS microbiology letters*, 139, pp.11–17.

Ojha, A.K., Baughn, A.D., Sambandan, D., Hsu, T., Trivelli, X., Guerardel, Y., Alahari, A., Kremer, L., Jacobs, W.R. and Hatfull, G.F., 2008. Growth of Mycobacterium tuberculosis biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Molecular Microbiology*, 69(1), pp.164–174.

Ojha, A.K. and Hatfull, G.F., 2012. Biofilms of Mycobacterium tuberculosis: New Perspectives

of an Old Pathogen. In: P.-J. Cardona, ed. *Understanding Tuberculosis - Deciphering the Secret Life of the Bacilli*. IntechOpen.pp.181–192.

Portaels, F. and Pattyn, S.R., 1982. Growth of mycobacteria in relation to the pH of the medium. *Annales de microbiologie*, 133(2), pp.213–21.

Primm, T.P., Andersen, S.J., Mizrahi, V., Avarbock, D., Rubin, H. and Barry, C.E., 2000. The stringent response of Mycobacterium tuberculosis is required for long-term survival. *Journal of Bacteriology*, 182(17), pp.4889–4898.

Reyrat, J.-M. and Kahn, D., 2001. Mycobacterium smegmatis: An absurd model for tuberculosis? *Trends in Microbiology*, 9(10), pp.472–473.

Roberts, M.E. and Stewart, P.S., 2005. Modelling protection from antimicrobial agents in biofilms through the formation of persister cells. *Microbiology*, 151(1), pp.75–80.

SBML, 2012. *Basic Introduction to SBML*. [online] Available at: <http://sbml.org/Basic_Introduction_to_SBML> [Accessed 7 Nov. 2019].

Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., Dolganov, G., Efron, B., Butcher, P.D., Nathan, C. and Schoolnik, G.K., 2003. Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: Insights into the phagosomal environment. *Journal of Experimental Medicine*, 198(5), pp.693–704.

Seeliger, J.C., Topp, S., Sogi, K.M., Previti, M.L., Gallivan, J.P. and Bertozzi, C.R., 2012. A Riboswitch-Based inducible gene expression system for mycobacteria. *PLoS ONE*, 7(1), pp.3–8.

Serjeant, E.P., Dempsey, B., International Union of Pure and Applied Chemistry; Commission of Equilibrium Data and International Union of Pure and Applied Chemistry; Commission of Electrochemical Data, 1979. *Ionisation constants of organic acids in aqueous solution*. New York: Oxford.

Shah, D., Zhang, Z., Khodursky, A., Kaldalu, N., Kurg, K. and Lewis, K., 2006. Persisters: A distinct physiological state of E. coli. *BMC Microbiology*, 6, pp.1–9.

Shampine, L.F. and Reichelt, M.W., 1997. The MATLAB ode suite. *SIAM Journal of Scientific Computing*, 18(1), pp.1–22.

Shiloh, M.U. and Champion, P.A.D., 2010. To catch a killer. What can mycobacterial models teach us about Mycobacterium tuberculosis pathogenesis? *Current Opinion in Microbiology*, 13(1), pp.1–12.

da Silva, P.E.A. and Palomino, J.C., 2011. Molecular basis and mechanisms of drug

resistance in Mycobacterium tuberculosis: Classical and new drugs. *Journal of Antimicrobial Chemotherapy*, 66(7), pp.1417–1430.

Simner, P.J., Woods, G.L. and Wengenack, N.L., 2016. Mycobacteria. *Microbiology Spectrum*, pp.1–21.

Smeulders, M.J., Keer, J., Speight, R.A. and Williams, H.D., 1999. Adaptation of Mycobacterium smegmatis to stationary phase. *Journal of Bacteriology*, 181(1), pp.270–283.

Spoering, A.L. and Lewis, K., 2001. Biofilms and Planktonic Cells of Pseudomonas aeruginosa Have Similar Resistance to Killing by Antimicrobials. *Society*, 183(23), pp.6746–6751.

Tang, Y.J., Shui, W., Myers, S., Feng, X., Bertozzi, C. and Keasling, J.D., 2009. Central metabolism in mycobacterium smegmatis during the transition from O2-rich to O2-poor conditions as studied by isotopomer-assisted metabolite analysis. *Biotechnology Letters*, 31(8), pp.1233–1240.

Thermo Fisher Scientific, 2010. *ENZYTEC fluid Glycerol*.

Thermo Fisher Scientific, 2016. *ENZYTEC fliud Ammonia*.

Voskuil, M.I., Schnappinger, D., Visconti, K.C., Harrell, M.I., Dolganov, G.M., Sherman, D.R. and Schoolnik, G.K., 2003. Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. *Journal of Experimental Medicine*, 198(5), pp.705–713.

Wakamoto, Y., Dhar, N., Chait, R., Schneider, K., Signorino-Gelo, F., Leibler, S. and McKinney, J.D., 2013. Dynamic persistence of antibiotic-stressed mycobacteria. *Science (New York, N.Y.)*, 339(6115), pp.91–5.

Wang, X., Fu, H., Du, D., Zhou, Z., Zhang, A., Su, C. and Ma, K., 2008. The comparison of pKa determination between carbonic acid and formic acid and its application to prediction of the hydration numbers. *Chemical Physics Letters*, 460(1–3), pp.339–342.

Wayne, L.G. and Hayes, L.G., 1996. An In Vitro Model for Sequential Study of Shiftdown of Mycobacterium tuberculosis through Two Stages of Nonreplicating Persistence. *Infection and Immunity*, 64(6), pp.2062–2069.

Webber, M.A. and Piddock, L.J. V, 2003. The importance of efflux pumps in bacterial antibiotic resistance. *Journal of Antimicrobial Chemotherapy*, 51(1), pp.9–11.

Williams, K.J., Bryant, W.A., Jenkins, V.A., Barton, G.R., Witney, A.A., Pinney, J.W. and Robertson, B.D., 2013. Deciphering the response of Mycobacterium smegmatis to nitrogen stress using bipartite active modules. *BMC Genomics*, 14(436), pp.1–13.

Williams, K.J., Jenkins, V.A., Barton, G.R., Bryant, W.A., Krishnan, N. and Robertson, B.D.,

2015. Deciphering the metabolic response of Mycobacterium tuberculosis to nitrogen stress. *Molecular Microbiology*, 97(6), pp.1142–1157.

Wolschendorf, F., Mahfoud, M. and Niederweis, M., 2007. Porins are required for uptake of phosphates by Mycobacterium smegmatis. *Journal of Bacteriology*, 189(6), pp.2435–2442.

World Health Organization, 2018. *Global tuberculosis report 2018*. France.

Young, D., Stark, J. and Kirschner, D., 2008. Systems biology of persistent infection: Tuberculosis as a case study. *Nature Reviews Microbiology*, 6(7), pp.520–528.

Appendix A: Flow Cytometry (Validation Experiments)

As with the parameter estimation experiments, the replication dynamics of the validation experiments were considered, using the *M. smegmatis*::pTiGc reporter. The OD graphs were contrasted to the number of generations calculated using MFI data (Figure A 1):

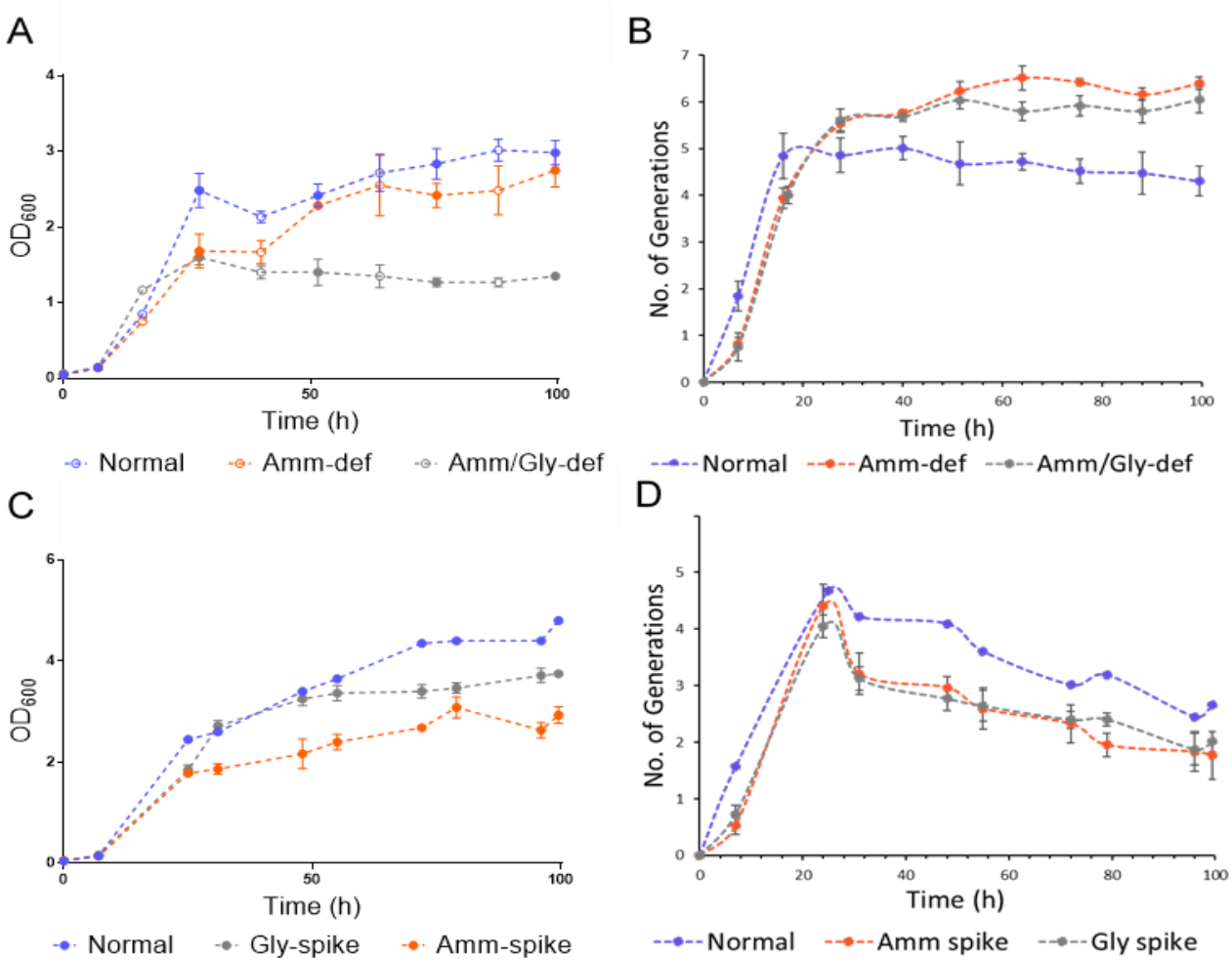

Figure A 1: Comparison of OD (left) and MFI (right) of the validation experiments. A and B: Normal conditions (purple), ammonia deficiency (orange) and ammonia/glycerol deficiency (grey). C and D: Normal conditions (purple), ammonia spiking (orange) and glycerol spiking (grey)

Comparison of bacterial growth using OD values indicate that the highest amount of growth is seen under normal conditions, whilst the least amount of growth would be when both substrates are limiting (Figure A 1 A). But, when it comes to the MFI, the data suggests that the least amount of growth occurred under normal conditions (the figure illustrates that this condition has the lowest number of generations), with bacterial growth under ammonium deficiency showing the highest amount of growth (Figure A 1 B), thereby contradicting the OD data. Additionally, the MFI results of the spiking experiments display a different anomaly; rather than showing the distinct difference in the growth of the bacteria, depending on the environmental condition in question (Figure A 1 C), the MFI shows that all three conditions

follow the same trend, in which the number of generations *decrease* after t = 24 h (Figure A 1 D), rather than stagnating beyond this time point (since the minimum fluorescence should have been reached by t = 16 h). This would suggest that the red fluorescent protein is being synthesized by *M. smegmatis*, which is not the case.

The reason for the discrepancy between the OD and MFI data is yet to be determined. It is probable that there may have been a technical fault with the instrument during analysis, which was missed. Nonetheless, a repetition of this analysis would be required, to see if similar results will be obtained, and if so, determine the cause for this.

Appendix B: Sample Calculations and Additional Data

Calibration Graphs

Estimating the bacterial concentration with flow cytometry data

Flow samples for bacterial counting require that a certain amount of a calibrated suspension of microspheres (Thermo Fischer Scientific) be added to them; this suspension serves as a reference standard for accurate measurement of sample volumes (Molecular Probes Inc, 2004). The standard contains a microsphere concentration of 10^8 beads/mL; 10 µL of the standard was pipetted into each microcentrifuge tube containing 0.5 mL volume of the sample, therefore, 10^6 beads were contained in each tube. As a result, determination of bacterial concentration was based on a ratio of bacterial events to microsphere events, as per the following equation:

$$\frac{bacterial\ events}{bead\ events} * \frac{\#\ of\ beads\ per\ test}{test\ volume} * dilution\ factor\ (bacteria) = bacteria/ml \qquad [B\ 1]$$

Using the data of one of the culture replicates at t = 0 h:

$$\frac{19564}{34966} * \frac{10^6}{0.5\ mL} * 1 = 1119030\ bacteria/mL \qquad [B\ 2]$$

The bacteria are not diluted in any way during sample preparation; thus, its dilution factor is equal to 1.

Estimating the dilution factors for plating (CFU)

At an OD_{600} of 1, the following CFU estimates are expected, depending on the dilution factor (Table B 1):

Table B 1: Expected CFUs for M. smegmatis

Dilution Factor	Expected CFU/mL	Expected CFU/0.1 mL
Undiluted	10^8	10^7
10^{-1}	10^7	10^6
10^{-2}	10^6	10^5
10^{-3}	10^5	10^4
10^{-4}	10^4	10^3
10^{-5}	10^3	10^2
10^{-6}	10^2	10^1
10^{-7}	10^1	10^0

Table B 1 can be used in calculating the expected CFUs for different OD_{600} values that are anticipated as the bacteria grows with time. The approximations are summarized in Table B 2:

Table B 2: Expected OD_{600} and CFU/0.1 mL values at each time point

Time point	Time (Hours)	Expected OD_{600}	Division factor	Expected CFU/0.1 mL
1	0	0.05	20	$5*10^5$
2	3.5	0.1	10	10^6
3	7	0.2	5	$2*10^6$
	10.5	0.4	2.5	$4*10^6$
	14	0.8	1.25	$8*10^6$
4	16	1.26	0.794	$1.26*10^7$
	17.5	1.6	0.625	$1.6*10^7$
5	19.5	2.51	0.398	$2.51*10^7$
	21	3.2	0.3125	$3.2*10^7$
6	23	3.2	0.3125	$3.2*10^7$
7	24	3.2	0.3125	$3.2*10^7$
8	27.5	3.2	0.3125	$3.2*10^7$
9	31	3.2	0.3125	$3.2*10^7$
10	40	3.2	0.3125	$3.2*10^7$

With a starting OD_{600} of 0.05, the OD is expected to double once the generation time of *M. smegmatis* has elapsed, which leads to the values in the third column. The division factor is calculated by dividing the reference OD value (OD_{600} = 1) by the expected OD_{600} values in the third column. Subsequently, the expected CFU/0.1 mL sample is determined by dividing the undiluted CFU value in the shaded cell in Table B 1 by the division factor for each time point. From t = 21 h, the OD is kept constant, as it was anticipated that the cells will have already reached the stationary phase of growth.

To pinpoint the relevant dilution factors, Table B 3 was drawn up:

Table B 3: Expected CFU/0.1mL under various serial dilutions, for each time point

	Timepoint									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Expected OD	0.05	0.1	0.2	1.26	2.51	3.2	3.2	3.2	3.2	3.2
Dilution	**Expected CFU/0.1 mL**									
Undiluted	$5*10^5$	10^6	$2*10^6$	$1.26*10^7$	$2.51*10^7$	$3.2*10^7$	$3.2*10^7$	$3.2*10^7$	$3.2*10^7$	$3.2*10^7$
10^{-1}	$5*10^4$	10^5	$2*10^5$	$1.26*10^6$	$2.51*10^6$	$3.2*10^6$	$3.2*10^6$	$3.2*10^6$	$3.2*10^6$	$3.2*10^6$
10^{-2}	$5*10^3$	10^4	$2*10^4$	$1.26*10^5$	$2.51*10^5$	$3.2*10^5$	$3.2*10^5$	$3.2*10^5$	$3.2*10^5$	$3.2*10^5$
10^{-3}	$5*10^2$	10^3	$2*10^3$	$1.26*10^4$	$2.51*10^4$	$3.2*10^4$	$3.2*10^4$	$3.2*10^4$	$3.2*10^4$	$3.2*10^4$
10^{-4}	50	10^2	$2*10^2$	$1.26*10^3$	$2.51*10^3$	$3.2*10^3$	$3.2*10^3$	$3.2*10^3$	$3.2*10^3$	$3.2*10^3$
10^{-5}	5	10	20	$1.26*10^2$	$2.51*10^2$	$3.2*10^2$	$3.2*10^2$	$3.2*10^2$	$3.2*10^2$	$3.2*10^2$
10^{-6}		1	2	12.6	25.1	32	32	32	32	32
10^{-7}				1.26	2.51	3.2	3.2	3.2	3.2	3.2

The undiluted CFU values derived in Table B 2 are highlighted in the light grey cells. Moving down each column, values on every cell are divided by 10, for further dilution of the samples. The chosen dilution factors are indicated by the darker cells; these were chosen based off colonies that could be counted manually in each petri dish.

Theoretical Phosphorus Uptake

The theoretical amount of phosphorus that would be taken up by the organism within 40 hours of growth is as follows:

$$z = X * \frac{0.4\frac{nmol\ PO_4^{3-}}{min.mg\ cells} * 1000\ mg\ cells * 60\ min * h * 1000\ mg\ P * 30.97\ g\ P}{1*10^9\ nmol\ PO_4^{3-}} \qquad [B\ 3]$$

where *z* is the phosphorus concentration taken up by the cells in mg/L, *X* is the biomass concentration at t = 40h, and *h* is the 40-h mark. Substituting for *X* and *h*, z is found to be **45.45 mg P/L**.

The initial phosphorus concentration, based on the amount of KH_2PO_4 (2.18 g) dissolved in a 1 L solution, was determined to be **496.12 mg P/L**, as follows:

$$2.18\frac{g}{L} * \frac{30.97\ g\ P}{136.086\ g\ KH_2PO_4} * \frac{1000\ mg}{1\ g} = 496.12\frac{mg\ P}{L} \qquad [B\ 4]$$

Molecular Formula of *M. smegmatis*

To illustrate the way the molecular formula of *M. smegmatis* was calculated, the ICP-MS results of the quality control (QC) sample, sulfamethazine, was used:

Table B 4: QC Results of Sulfamethazine

Name	***Weight (mg)***	**N [%]**	**C [%]**	**H [%]**	**S [%]**	**O [%][1]**
Cert Ref Std sulfamethazine		20.13	51.78	5.07	11.52	11.5
QC Analysed	*2.615*	20.3	51.6	5.1	10.6	12.3
% Recovery calculated[2]		***100.7***	***99.7***	***101.1***	***92.1***	**107.4**

[1]The oxygen data was not part of the ICP-MS analysis; rather, the percentages were determined by subtracting the total sum of the percentages of N, C, H and S from 100%

[2] Recovery calculated (%): $100\% + \left(\frac{X\%_{analysis} - X\%_{std}}{X\%_{std}} * 100\%\right)$

NOTE: N = 14.01 g/mol; C = 12.01 g/mol; H = 1.01 g/mol; S = 32.07 g/mol; O = 16 g/mol

Assuming the molecular formula of sulfamethazine is not known, the following procedure was implemented:

1. Calculating the mass of each element, based on the total mass of the sample
2. Converting the mass value to a mole value
3. Dividing the moles of the elements by the mole value of carbon, to get an appropriate molecular equation

Table B 5 below highlights the procedure:

Table B 5: Determination of the Molecular Formula of the QC Sample

	C	**H**	**N**	**O**	**S**
Percentage	51.6	5.1	20.3	12.3	10.6
Mass of Element (mg)	1.350	0.134	0.530	0.323	0.277
Moles of Element	$1.12*10^{-4}$	$1.33*10^{-4}$	$3.786*10^{-5}$	$2.018*10^{-5}$	$8.650*10^{-5}$
Formula	1	1.183	0.337	0.179	0.0769
	12	14.2	4.04	2.15	0.92

As per the last row, the molecular formula calculated was $C_{12}H_{14.2}N_{4.04}O_{2.15}S_{0.92}$, which is very close to the actual molecular formula of sulfamethazine, which is $C_{12}H_{14}N_4O_2S$.

For *M. smegmatis*, the following results (from ICP-MS and SEM) were obtained (Tables B 6 and B 7):

Table B 6: Results for M. smegmatis (Sample 1 – 6.61mg)

	C	H	N	O	S
Percentage	42.98	6.92	7.86	24.38	0.23
Mass of Element (mg)	2.841	0.457	0.520	1.612	0.0151
Moles of Element	$2.37*10^{-4}$	$4.54*10^{-4}$	$3.709*10^{-5}$	$1.0*10^{-4}$	$4.720*10^{-7}$
Formula	1	1.918	0.157	0.426	0.002

Table B 7: Results for M. smegmatis (Sample 2 – 4.49mg)

	C	H	N	O	S
Percentage	43.29	7.00	7.77	23.75	0.27
Mass of Element (mg)	1.944	0.314	0.349	1.066	0.012
Moles of Element	$1.62*10^{-4}$	$3.12*10^{-4}$	$2.491*10^{-5}$	$6.665*10^{-5}$	$3.809*10^{-7}$
Formula	1	1.926	0.154	0.412	0.00235

An average value was taken across the two biomass samples, resulting in the following formula: $\mathbf{CH_{1.92}N_{0.16}O_{0.42}S_{0.002}}$

Mathematical Model

Calculation of Yield Coefficients

The yield coefficients of glycerol and ammonia were calculated based off the experimental data obtained from the carbon-deficient and nitrogen-deficient environmental conditions (respectively). This value is expressed as a ratio of biomass produced to the substrate consumed:

$$Y_{X/S} = \frac{\left(\frac{dX}{dt}\right)}{\left(\frac{dS}{dt}\right)} \qquad [B\ 5]$$

Therefore, the yield coefficients were determined by dividing the difference in the biomass produced between t = 0 h and t = 40 h by the substrate used within the same time frame:

$$Y_{X/C} = \frac{\left(\frac{dX}{dt}\right)}{\left(\frac{dC}{dt}\right)} = \frac{X_{t=40h} - X_{t=0h}}{C_{t=0h} - C_{t=40h}} = \frac{0.736 - 0.224}{1.059 - 0.003} = 0.484 \qquad [B\ 6]$$

$$Y_{X/N} = \frac{\left(\frac{dX}{dt}\right)}{\left(\frac{dN}{dt}\right)} = \frac{X_{t=40h} - X_{t=0h}}{N_{t=0h} - N_{t=40h}} = \frac{1.006 - 0.224}{0.036 - 0} = 21.575 \qquad [B\ 7]$$

The yield coefficient of CO_2 was determined through regression of the model.

Statistical Analysis

The RMSE and NRMSE of the response variables, under each environmental condition, were calculated (Table B 8):

Table B 8: Fitting statistics, with the SSE representing the sum of squared errors. Left: parameter estimation experiments; right: model validation experiments

		SSE	RMSE	NRMSE (%)		SSE	RMSE	NRMSE (%)
Normal	X	0.590	0.243	19.953	**Normal**	1.226	0.350	35.686
	C	5.352	0.732	31.692		4.005	0.633	23.712
	N	0.009	0.031	26.752		0.021	0.046	27.629
	pH	0.009	0.031	3.531		0.115	0.107	13.399
Gly-def	X	0.011	0.033	6.059	**Amm-def**	0.295	0.172	19.206
	C	0.021	0.045	4.286		2.065	0.454	19.283
	N	0.011	0.034	31.415		0.000	0.001	2.987
	pH	0.069	0.083	16.381		0.050	0.071	12.671
Amm-def	X	0.305	0.175	19.827	**Amm/gly-def**	0.046	0.068	13.324
	C	2.283	0.478	17.242		0.043	0.066	6.008
	N	0.000	0.001	2.118		0.001	0.011	31.369
	pH	0.068	0.083	17.191		0.046	0.067	13.959
pH 4.6	X	0.009	0.029	70.417	**Amm spike**	1.404	0.375	37.526
	C	1.273	0.357	28.500		3.352	0.579	20.061
	N	0.020	0.045	54.351		0.001	0.012	4.393
	pH	0.004	0.021	63.007		0.110	0.105	5.283
pH 5.5	X	0.051	0.072	11.065	**Gly spike**	0.226	0.150	12.309
	C	0.382	0.196	10.129		3.929	0.627	5.795
	N	0.008	0.029	29.316		0.042	0.065	41.770
	pH	0.384	0.196	43.228		0.045	0.067	3.383

The highlighted cells indicate the response variables that have a low NRMSE, signifying a good fit with the proposed mathematical model.

Oxygen Dynamics

Parameter Estimation

Regression of the mathematical model with oxygen as an added response variable yields the following optimized values (Table B 9):

Table B 9: Initial and optimized values of growth parameters

Parameter	Description	Initial Values	Optimized values	Units
μ_{max}	Maximum specific growth rate	0.15	0.2500	h^{-1}
k_C	Saturation constant of glycerol	0.08	0.6187	g/L
k_N	Saturation constant of ammonia	0.01	0.0001	g/L
k_d	Microbial death rate	0.001	0.0010	h^{-1}
pH_{LL}	Lower limit of pH inhibition	4.6	4.2520	-
α	Carrying capacity coefficient	1.5	6.8116	-
I_{val}	Empirical lower inhibition coefficient	3	2.6259	-
$Y_{X/CD}$	Yield coefficient of CO_2	0.35	0.3945	-
k_O	Saturation constant of oxygen	0.01	0.0045	g/L

The shaded cell represents the extra parameter added to the regression analysis.

Estimation of Constant Model Parameters

To estimate the k_La of oxygen under experimental conditions, a dimensionally-consistent empirical correlation was used (Klöckner and Büchs, 2012):

$$k_L a = 0.5 * d^{\frac{73}{36}} * n * d_0^{\frac{1}{4}} * V_L^{-\frac{8}{9}} * \mathbb{D}^{\frac{1}{2}} * v^{-\frac{13}{54}} * g^{-\frac{7}{54}} \qquad [B\ 8]$$

The parameters, and their respective values, are defined in Table B 10:

Table B 10: Parameters of the k_La correlation

Parameter	Definition	Units	Value used in Equation
d	maximum inner flask diameter	m	0.082
n	shaking frequency[1]	s^{-1}	3.333
d_0	shaking diameter	m	0.1
V_L	filling volume	m^3	0.0001
$\mathbb{D}$	diffusion coefficient of oxygen[2]	$m^2.s^{-1}$	$2.52*10^{-9}$
v	kinematic viscosity of the media[3]	$m^2.s^{-1}$	$6.951*10^{-7}$
g	acceleration due to gravity	$m.s^{-2}$	9.81

[1] *n* = rpm/60 sec = 200/60
[2] Estimated from (Han and Bartels, 1996), at a temperature of 35.1ºC
[3] Estimated using parameters of water, interpolated at 37ºC. The dynamic viscosity (μ) of water, as a function of temperature (in Kelvin), was initially determined as follows: $\mu(T) = A * 10^{B/(T-C)}$, where A is $2.414*10^5$Pa.s, B is 247.8K, C is 140K and T is 310.15K (37ºC) (Engineers Edge, 2019). The resulting value was divided by the density of water at 37ºC, to obtain the kinematic viscosity

The resulting k_La, in units of h^{-1}, was calculated to be **86.26**.

An estimate of the yield coefficient of oxygen was calculated based on the metabolic equation of *M. smegmatis* (the elemental balance was done without nitrogen to simplify the calculations):

$$aC_3H_8O_3 + bO_2 \rightarrow cCH_{1.92}O_{0.42} + dCO_2 + eH_2O \qquad [B\ 9]$$

The yield coefficient of glycerol is known ($Y_{X/C}$ = 0.484); this can be used to determine its stoichiometric coefficient. Assuming that the stoichiometric coefficient of the biomass is 1:

$$0.484 = \frac{X_{mass}}{C_{mass}} = \frac{c*MW_X}{a*MW_C} = \frac{1*22.89}{a*92.094} \qquad [B\ 10]$$

Therefore, *a* is determined to be **0.514**.

The elements in Equation [B 9] are balanced as follows:

- Carbon: 3(0.514) = 1 + *d*; *d* = 0.542
- Hydrogen: 8(0.514) = 1.92 + 2*e*; *e* = 1.096
- Oxygen: 3(0.514) + 2*b* = 0.42 + 2(0.542) + 1.096; *b* = 0.529

Therefore, the estimated yield coefficient of oxygen is: $Y_{X/O}$ = 1(22.89)/0.529(32) = **1.352**.

Appendix C: MATLAB Code

Main Script (Parameter Estimation)

```
%% ---------------PARAMETER ESTIMATION SCRIPT
clear;
clc;
[normal,gly,amm,ph46,ph55] = EX_data_1;                              %Reading of Excel data
%% ---------------CONSTANTS, PARAMETERS AND INITIAL CONDITIONS
%Constant model parameters
n.t_lag = 7;                                                         %lag time (hours)
n.Y.X_C = 0.484;                                                     %YC (biomass/glycerol)
n.Y.X_N = 21.575;                                                    %YC (biomass/ammonia)
n.Y.X_O = 1.352;                                                     %YC (biomass/oxygen) (hypothetical situation)
n.kla_O2 = 86.256;                                                   %O2 kla (1/h)
n.Csat_O2 = 7.2673;                                                  %C* (mgO2/L)

%Initial conditions (add initial value of O2 for hypothetical case)
n.init.normal_c = [0.229 5.389 0.951 0.439/1000];                    %Normal initial concentrations [X C N_L CD]
n.init.glydef_c = [0.224 1.059 0.959 0.439/1000];                    %Gly-def initial concentrations
n.init.ammdef_c = [0.224 6.057 0.036 0.439/1000];                    %Amm-def initial concentrations
n.init.ph46_c = [0.222 5.620 1.019 0.439/1000];                      %pH 4.6 initial concentrations
n.init.ph55_c = [0.226 5.225 0.974 0.439/1000];                      %pH 5.5 initial concentrations

%Initial model parameters (guessed)
n.u_max = 0.15;                                                      %[p1]Max. respiratory growth rate
n.k_C = 0.08;                                                        %[p2]Glycerol saturation constant
n.k_N = 0.01;                                                        %[p3]Ammonia saturation constant
n.k_d = 0.0001;                                                      %[p4]Death rate
n.pH.LL = 4.6;                                                       %[p5]Lower limit of pH inhibition
n.alpha = 1.5;                                                       %[p6]Alpha value for f(X)
n.I_val = 3;                                                         %[p7]Value in pH inhibition term
n.Y.X_CD = 0.35;                                                     %[p8]YC (biomass/CO2)
n.k_O = 0.001;                                                       %[p9]Oxygen saturation coefficient (hypothetical situation)

%Upper and lower boundaries for optimization [u_max k_C k_N k_d pH_LL alpha I_val Y.X_CD k_O*]
lb = [0.005 0.01 0.0001 0.00001 2 1 1 0.1];                          %lower boundary (overall)(k_O: 0.0001)
ub = [0.25 1 0.05 0.001 8 13 5 1];                                   %upper boundary (overall)(k_O: 1)
p0 = [n.u_max n.k_C n.k_N n.k_d n.pH.LL n.alpha n.I_val n.Y.X_CD];   %Initial conditions

%Time for integration
n.time = [0 40];                                                     %in hours
n.time_spec = [0 3.5 7 16 19.5 23 24 27.5 31 40];                    %integration at specific time points (for optimization)

%pKa values for pH calculation
```

```
n.pH.pka1 = 6.86;                                                    %pKa of KH2PO4
n.pH.pka2 = 3.13;                                                    %pka of C6H8O7
n.pH.pka3 = 4.76;                                                    %pka of (C6H7O7)-
n.pH.pka4 = 6.40;                                                    %pka of (C6H6O7)2-
n.pH.pka5 = 9.25;                                                    %pka of NH3
n.pH.pka6 = 14.15;                                                   %pka of C3H8O3
n.pH.pka7 = 6.35;                                                    %pka of CO2
n.pH.pka8 = 10.33;                                                   %pka of (HCO3)-
n.pH.pka9 = 14;                                                      %pka of H2O
n.pH.conc.KH2PO4 = 2.18;                                             %Constant KH2PO4 concentration (g/L)
n.pH.conc.C6H8O7 = 2;                                                %Constant C6H8O7 concentration (g/L)
n.pH.UL = 7.4;                                                       %Upper limit of pH inhibition
n.pH_M = [0 14];                                                     %pH limits

n0 = n;                                                              %To save the pre-optimized parameters
%% ---------------DATA FOR OPTIMIZATION
%MAXIMUM AND MINIMUM VALUES
%Normal conditions
normal.max.X = max(normal.X); normal.min.X = min(normal.X);          %biomass
normal.max.C = max(normal.C); normal.min.C = min(normal.C);          %glycerol
normal.max.N = max(normal.N); normal.min.N = min(normal.N);          %ammonia
normal.max.pH = max(normal.pH); normal.min.pH = min(normal.pH);      %pH

%Gly-def conditions                                                  %Amm-def conditions
gly.max.X = max(gly.X); gly.min.X = min(gly.X);                      amm.max.X = max(amm.X); amm.min.X = min(amm.X);
gly.max.C = max(gly.C); gly.min.C = min(gly.C);                      amm.max.C = max(amm.C); amm.min.C = min(amm.C);
gly.max.N = max(gly.N); gly.min.N = min(gly.N);                      amm.max.N = max(amm.N); amm.min.N = min(amm.N);
gly.max.pH = max(gly.pH); gly.min.pH = min(gly.pH);                  amm.max.pH = max(amm.pH); amm.min.pH = min(amm.pH);

%pH 4.6                                                              %pH 5.5
ph46.max.X = max(ph46.X); ph46.min.X = min(ph46.X);                  ph55.max.X = max(ph55.X); ph55.min.X = min(ph55.X);
ph46.max.C = max(ph46.C); ph46.min.C = min(ph46.C);                  ph55.max.C = max(ph55.C); ph55.min.C = min(ph55.C);
ph46.max.N = max(ph46.N); ph46.min.N = min(ph46.N);                  ph55.max.N = max(ph55.N); ph55.min.N = min(ph55.N);
ph46.max.pH = max(ph46.pH); ph46.min.pH = min(ph46.pH); ph55.max.pH = max(ph55.pH); ph55.min.pH = min(ph55.pH);

%NORMALIZATION
%Overall max and min
n.max_ov.X = max([normal.max.X; gly.max.X; amm.max.X; ph46.max.X; ph55.max.X]);
n.max_ov.C = max([normal.max.C; gly.max.C; amm.max.C; ph46.max.C; ph55.max.C]);
n.max_ov.N = max([normal.max.N; gly.max.N; amm.max.N; ph46.max.N; ph55.max.N]);
n.max_ov.pH = max([normal.max.pH; gly.max.pH; amm.max.pH; ph46.max.pH; ph55.max.pH]);
n.min_ov.X = min([normal.min.X; gly.min.X; amm.min.X; ph46.min.X; ph55.min.X]);
n.min_ov.C = min([normal.min.C; gly.min.C; amm.min.C; ph46.min.C; ph55.min.C]);
n.min_ov.N = min([normal.min.N; gly.min.N; amm.min.N; ph46.min.N; ph55.min.N]);
n.min_ov.pH = min([normal.min.pH; gly.min.pH; amm.min.pH; ph46.min.pH; ph55.min.pH]);

%Normalized data (normal conditions)
normal.norma.X = (normal.X-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);      %Norm. biomass
```

```
normal.norma.C = (normal.C-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);          %Norm. glycerol
normal.norma.N = (normal.N-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);          %Norm. ammonia
normal.norma.pH = (normal.pH-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);     %pH

%Normalized data (glycerol-deficiency)                                    %Normalized data (ammonia-deficiency)
gly.norma.X = (gly.X-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);                amm.norma.X = (amm.X-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);
gly.norma.C = (gly.C-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);                amm.norma.C = (amm.C-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);
gly.norma.N = (gly.N-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);                amm.norma.N = (amm.N-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);
gly.norma.pH = (gly.pH-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);           amm.norma.pH = (amm.pH-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);

%Normalized data (pH 4.6)                                                 %Normalized data (pH 5.5)
ph46.norma.X = (ph46.X-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);              ph55.norma.X = (ph55.X-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);
ph46.norma.C = (ph46.C-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);              ph55.norma.C = (ph55.C-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);
ph46.norma.N = (ph46.N-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);              ph55.norma.N = (ph55.N-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);
ph46.norma.pH = (ph46.pH-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);         ph55.norma.pH = (ph55.pH-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);

n0 = n;
%% ---------------OPTIMIZATION: lsqnonlin (Local Minimum)
options = optimoptions('lsqnonlin','Display','iter');
[p,resnorm,residual,exitflag,output] = lsqnonlin(...
    @(p)OPT_overall_4var(p,normal,gly,amm,ph46,ph55,n),p0,lb,ub,options);       %Check OPT_overall_4var for n.k_O
n.u_max = p(1); n.k_C = p(2); n.k_N = p(3); n.k_d = p(4); n.pH.LL = p(5);
n.alpha = p(6); n.I_val = p(7); n.Y.X_CD = p(8); %n.k_O = p(9);
%% ---------------OPTIMIZATION: MultiStart (Global Minimum)
prob_ms = createOptimProblem('lsqnonlin','x0',p0,'objective',@(p)OPT_overall_4var(p,normal,gly,amm,ph46,ph55,n),'lb',lb,'ub',ub);
[glo_min_ms,SSE_glo_ms] = run(MultiStart,prob_ms,50);
n.u_max = glo_min_ms(1); n.k_C = glo_min_ms(2); n.k_N = glo_min_ms(3);
n.k_d = glo_min_ms(4); n.pH.LL = glo_min_ms(5); n.alpha = glo_min_ms(6);
n.I_val = glo_min_ms(7); n.Y.X_CD = glo_min_ms(8); n.k_O = glo_min_ms(9);
%% ---------------PLOTS: DATA AND PRE-OPTIMIZATION
%Normal
figure(1); clear n0.pH.alk; n0.pH.alk = 7.15;                             %Alkalinity value
[T.a, R.a] = ode15s(@(t,D)Dynamics_pH(t,D,n0),n.time,n.init.normal_c);
r = 3; c = 5; z = 1; [~,~,~,~] = exp_data_plots(normal,r,c,z);            %r: row; c: column; z: number of plot (left to right)
pH.a = pH_fig2(T.a,R.a,n0);                                               %Determination of media pH
[~,~,~,~,~] = pre_opt_plots(r,c,z,T.a,R.a,pH.a);                          %Function: plots with unoptimized parameters

%Glycerol-deficient
clear n0.pH.alk; n0.pH.alk = 6.74;
[T.b, R.b] = ode15s(@(t,D)Dynamics_pH(t,D,n0),n.time,n.init.glydef_c);
z = z + 5; [~,~,~,~] = exp_data_plots(gly,r,c,z);
pH.b = pH_fig2(T.b,R.b,n0);
[~,~,~,~,~] = pre_opt_plots(r,c,z,T.b,R.b,pH.b);

%Ammonia-deficient
clear n0.pH.alk; n0.pH.alk = 6.65;
[T.c, R.c] = ode15s(@(t,D)Dynamics_pH(t,D,n0),n.time,n.init.ammdef_c);
z = z + 5; [~,~,~,~] = exp_data_plots(amm,r,c,z);
```

```
pH.c = pH_fig2(T.c,R.c,n0);
[~,~,~,~,~] = pre_opt_plots(r,c,z,T.c,R.c,pH.c);

%pH 4.6
figure(2); clear pH.d n0.pH.alk; n0.pH.alk = 2.574;
[T.d, R.d] = ode15s(@(t,D)Dynamics_pH(t,D,n0),n.time,n.init.ph46_c);
r = 2; z = 1; [~,~,~,~] = exp_data_plots(ph46,r,c,z);
pH.d = pH_fig2(T.d,R.d,n0);
[~,~,~,~,~] = pre_opt_plots(r,c,z,T.d,R.d,pH.d);

%pH 5.5
clear pH.e n0.pH.alk; n0.pH.alk = 3.81;
[T.e, R.e] = ode15s(@(t,D)Dynamics_pH(t,D,n0),n.time,n.init.ph55_c);
z = z + 5; [~,~,~,~] = exp_data_plots(ph55,r,c,z);
pH.e = pH_fig2(T.e,R.e,n0);
[~,~,~,~,~] = pre_opt_plots(r,c,z,T.e,R.e,pH.e);
%% ---------------PLOTS: POST-OPTIMIZATION
%Normal
figure(1); clear pH.a n.pH.alk; n.pH.alk = 7.15; r = 3; z = 1;
[T.a, R.a] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.normal_c);
pH.a = pH_fig2(T.a,R.a,n);
[~,~,~,~,~] = post_opt_plots(r,c,z,T.a,R.a,pH.a);

%Glycerol-deficient
clear pH.b n.pH.alk; n.pH.alk = 6.74; z = z + 5;
[T.b, R.b] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.glydef_c);
pH.b = pH_fig2(T.b,R.b,n);
[~,~,~,~,~] = post_opt_plots(r,c,z,T.b,R.b,pH.b);

%Ammonia-deficient
clear pH.c n.pH.alk; n.pH.alk = 6.65; z = z + 5;
[T.c, R.c] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.ammdef_c);
pH.c = pH_fig2(T.c,R.c,n);
[~,~,~,~,~] = post_opt_plots(r,c,z,T.c,R.c,pH.c);

%pH 4.6
figure(2); clear pH.d n.pH.alk; n.pH.alk = 2.574; r = 2; z = 1;
[T.d, R.d] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.ph46_c);
pH.d = pH_fig2(T.d,R.d,n);
[~,~,~,~,~] = post_opt_plots(r,c,z,T.d,R.d,pH.d);

%pH 5.5
clear pH.e n.pH.alk; n.pH.alk = 3.81; z = z + 5;
[T.e, R.e] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.ph55_c);
pH.e = pH_fig2(T.e,R.e,n);
[~,~,~,~,~] = post_opt_plots(r,c,z,T.e,R.e,pH.e);
%legend('Data A', 'Data B', 'Post-opt Sim'); %Add 'Pre-opt Sim', if plotting with unoptimized parameters
```

Main Script (Validation of Proposed Model)

```
%% ---------------VALIDATION SCRIPT
clear;
clc;
[normal_r,amm_r,ammgly,ammsp,glysp] = EX_data_2;        %Reading of Excel data
%% ---------------CONSTANTS, PARAMETERS AND INITIAL CONDITIONS
%Constant model parameters
n.t_lag = 7;                                            %lag time (hours)
n.Y.X_C = 0.484;                                        %YC (biomass/glycerol)
n.Y.X_N = 21.575;                                       %YC (biomass/ammonia)
n.max_ov.X = 1.4462;                                    %Maximum biomass (based on data used for optimization)

%Initial conditions
n.init.normal_rc = [0.223 5.922 1.027 0.439/1000];      %Normal (repetition) initial concentrations [X C N_L CD]
n.init.amm_rc = [0.222 5.511 0.032 0.439/1000];         %Amm-def (repetition) initial concentrations
n.init.ammgly_c = [0.229 1.132 0.037 0.439/1000];       %Amm & gly-def initial concentrations
n.init.ammsp_c = [0.227 5.352 0.028 0.439/1000];        %Amm-def spiking initial concentrations
n.init.glysp_c = [0.224 1.078 0.972 0.439/1000];        %Gly-def spiking initial concentrations

%Optiized model parameters
n.u_max = 0.1455;                                       %[p1]Max. respiratory growth rate
n.k_C = 0.3828;                                         %[p2]Glycerol saturation constant
n.k_N = 1.0027e-4;                                      %[p3]Ammonia saturation constant
n.k_d = 0.001;                                          %[p4]Death rate
n.pH.LL = 3.4687;                                       %[p5]Lower limit of pH inhibition
n.alpha = 12.7492;                                      %[p6]Alpha value for f(X)
n.I_val = 4.0413;                                       %[p7]Value in pH inhibition term
n.Y.X_CD = 0.3687;                                      %[p8]YC (biomass/CO2)

%Time for integration
n.time = [0 99.5];                                      %in hours
n.time_spec = [0 7 16 27.5 40 51.5 64 75.5 88 99.5];    %Normal conditions and nutrient deficiency

%pKa values for pH calculation
n.pH.pka1 = 6.86;                                       %pKa of KH2PO4
n.pH.pka2 = 3.13;                                       %pka of C6H8O7
n.pH.pka3 = 4.76;                                       %pka of (C6H7O7)-
n.pH.pka4 = 6.40;                                       %pka of (C6H6O7)2-
n.pH.pka5 = 9.25;                                       %pka of NH3
n.pH.pka6 = 14.15;                                      %pka of C3H8O3
n.pH.pka7 = 6.35;                                       %pka of CO2
n.pH.pka8 = 10.33;                                      %pka of (HCO3)-
n.pH.pka9 = 14;                                         %pka of H2O
n.pH.conc.KH2PO4 = 2.18;                                %Constant KH2PO4 concentration (g/L)
n.pH.conc.C6H8O7 = 2;                                   %Constant C6H8O7 concentration (g/L)
```

```
n.pH.UL = 7.4;                                          %Upper limit of pH inhibition
n.pH_M = [0 14];                                        %pH range
%% ---------------NUTRIENT DEFICIENCY
%Normal (repetition)
figure(3); clear n.pH.alk; n.pH.alk = 7.2; r = 3; c = 5; z = 1;
[T.a, R.a] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.normal_rc);
[~,~,~,~] = exp_data_plots_val(normal_r,r,c,z);
pH.a = pH_fig2(T.a,R.a,n);
[~,~,~,~,~] = val_plots(r,c,z,T.a,R.a,pH.a);

%Ammonia-deficiency (repetition)
clear n.pH.alk; n.pH.alk = 7.05; z = z + 5;
[T.b, R.b] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.amm_rc);
[~,~,~,~] = exp_data_plots_val(amm_r,r,c,z);
pH.b = pH_fig2(T.b,R.b,n);
[~,~,~,~,~] = val_plots(r,c,z,T.b,R.b,pH.b);

%Ammonia and glycerol deficiency
clear n.pH.alk; n.pH.alk = 7.05; z = z + 5;
[T.c, R.c] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time,n.init.ammgly_c);
[~,~,~,~] = exp_data_plots_val(ammgly,r,c,z);
pH.c = pH_fig2(T.c,R.c,n);
[~,~,~,~,~] = val_plots(r,c,z,T.c,R.c,pH.c);
%% ---------------SPIKING EXPERIMENTS
%AMM-DEF SPIKING (VALIDATION)
figure('Name','Ammonia Deficiency: Spiking'); clear pH.d n.pH.alk; n.pH.alk = 7.05;
n.init.ammsp_c = [0.227 5.352 0.028 0.439/1000];
%Integration before nutrient spiking
[T.d, R.d] = ode15s(@(t,D)Dynamics_pH(t,D,n),[0 24],n.init.ammsp_c);
r = 3; c = 2; z = 1;
subplot(r,c,z); plot(T.d, R.d(:,1), '-r','LineWidth',2); hold on;
errorbar(ammsp.t, ammsp.X, ammsp.e.X,'ob'); hold on; xlabel('Time (Hours)'); ylabel('Biomass (g/L)');
subplot(r,c,z+1); plot(T.d, R.d(:,2), '-r','LineWidth',2); hold on;
errorbar(ammsp.t, ammsp.C, ammsp.e.C,'ob'); hold on; xlabel('Time (Hours)'); ylabel('Glycerol Conc. (g/L)');
subplot(r,c,z+2); plot(T.d, R.d(:,3), '-r','LineWidth',2); hold on;
errorbar(ammsp.t, ammsp.N, ammsp.e.N,'ob'); hold on;  xlabel('Time (Hours)'); ylabel('NH_3 Conc. (EX) (g/L)');
subplot(r,c,z+3); plot(T.d, R.d(:,4), '-r','LineWidth',2); hold on; xlabel('Time (Hours)'); ylabel('CO_2 Conc. (g/L)');
pH.d = pH_fig2(T.d,R.d,n); subplot(r,c,z+4); plot(T.d, pH.d', '-r','LineWidth',2); hold on;
errorbar(ammsp.t, ammsp.pH, ammsp.e.pH,'ob'); hold on; xlabel('Time (Hours)'); ylabel('pH');

%Integration after nutrient spiking
subplot(r,c,z+2); plot([24 24],[R.d(length(R.d),3) 0.2775],'--r','LineWidth',2); hold on;     %NH3 spike
n.init.ammsp_c = [R.d(length(R.d),1) R.d(length(R.d),2) 0.2775 R.d(length(R.d),4)];
clear T.d R.d pH.d; [T.d, R.d] = ode15s(@(t,D)Dynamics_pH(t,D,n),[24 55],n.init.ammsp_c);
subplot(r,c,z); plot(T.d, R.d(:,1), '--r','LineWidth',2); hold on;                            %Biomass
subplot(r,c,z+1); plot(T.d, R.d(:,2), '--r','LineWidth',2); hold on;                          %Glycerol
subplot(r,c,z+2); plot(T.d, R.d(:,3), '--r','LineWidth',2); hold on;                          %Exogenous ammonia
subplot(r,c,z+3); plot(T.d, R.d(:,4), '--r','LineWidth',2); hold on;                          %CO2
```

```
pH.d = pH_fig2(T.d,R.d,n); subplot(r,c,z+4); plot(T.d, pH.d', '--r','LineWidth',2); hold on;  %pH

%integration after HCl addition
subplot(r,c,z+4); plot([55 55],[pH.d(length(pH.d)) 5.2],':r','LineWidth',2); hold on; %HCl spike
n.init.ammsp_c = [R.d(length(R.d),1) R.d(length(R.d),2) R.d(length(R.d),3) R.d(length(R.d),4)];
clear T.d R.d pH.d n.pH.alk; n.pH.alk = 3.8;
[T.d, R.d] = ode15s(@(t,D)Dynamics_pH(t,D,n),[55 99.5],n.init.ammsp_c);
subplot(r,c,z); plot(T.d, R.d(:,1), ':r','LineWidth',2); hold off;
subplot(r,c,z+1); plot(T.d, R.d(:,2), ':r','LineWidth',2); hold off;
subplot(r,c,z+2); plot(T.d, R.d(:,3), ':r','LineWidth',2); hold off;
subplot(r,c,z+3); plot(T.d, R.d(:,4), ':r','LineWidth',2); hold off;
pH.d = pH_fig2(T.d,R.d,n); subplot(r,c,z+4); plot(T.d, pH.d', ':r','LineWidth',2); hold off;

%GLY-DEF SPIKING (VALIDATION)
figure('Name','Glycerol Deficiency: Spiking'); clear pH.e n.pH.alk; n.pH.alk = 7.05;
n.init.glysp_c = [0.224 1.078 0.972 0.439/1000];
%Integration before nutrient spiking
[T.e, R.e] = ode15s(@(t,D)Dynamics_pH(t,D,n),[0 24],n.init.glysp_c);
subplot(r,c,z); plot(T.e, R.e(:,1), '-r','LineWidth',2); hold on;
errorbar(glysp.t, glysp.X, glysp.e.X,'ob'); hold on; xlabel('Time (Hours)'); ylabel('Biomass (g/L)');
subplot(r,c,z+1); plot(T.e, R.e(:,2), '-r','LineWidth',2); hold on;
errorbar(glysp.t, glysp.C, glysp.e.C,'ob'); hold on; xlabel('Time (Hours)'); ylabel('Glycerol Conc. (g/L)');
subplot(r,c,z+2); plot(T.e, R.e(:,3), '-r','LineWidth',2); hold on;
errorbar(glysp.t, glysp.N, glysp.e.N,'ob'); hold on; xlabel('Time (Hours)'); ylabel('NH_3 Conc. (EX) (g/L)');
subplot(r,c,z+3); plot(T.e, R.e(:,4), '-r','LineWidth',2); hold on; xlabel('Time (Hours)'); ylabel('CO_2 Conc. (g/L)');
pH.e = pH_fig2(T.e,R.e,n); subplot(r,c,z+4); plot(T.e, pH.e', '-r','LineWidth',2); hold on;
errorbar(glysp.t, glysp.pH, glysp.e.pH,'ob'); hold on; xlabel('Time (Hours)'); ylabel('pH');

%Integration after nutrient spiking
subplot(r,c,z+1); plot([24 24],[R.e(length(R.e),2) 11.18],'--r','LineWidth',2); hold on;      %Glycerol spike
n.init.glysp_c = [R.e(length(R.e),1) 11.18 R.e(length(R.e),3) R.e(length(R.e),4)];
clear T.e R.e pH.e; [T.e, R.e] = ode15s(@(t,D)Dynamics_pH(t,D,n),[24 55],n.init.glysp_c);
subplot(r,c,z); plot(T.e, R.e(:,1), '--r','LineWidth',2); hold on;
subplot(r,c,z+1); plot(T.e, R.e(:,2), '--r','LineWidth',2); hold on;
subplot(r,c,z+2); plot(T.e, R.e(:,3), '--r','LineWidth',2); hold on;
subplot(r,c,z+3); plot(T.e, R.e(:,4), '--r','LineWidth',2); hold on;
pH.e = pH_fig2(T.e,R.e,n); subplot(r,c,z+4); plot(T.e, pH.e', '--r','LineWidth',2); hold on;

%integration after HCl addition
subplot(r,c,z+4); plot([55 55],[pH.e(length(pH.e)) 5.2],':r','LineWidth',2); hold on; %HCl spike
n.init.glysp_c = [R.e(length(R.e),1) R.e(length(R.e),2) R.e(length(R.e),3) R.e(length(R.e),4)];
clear T.e R.e pH.e n.pH.alk; n.pH.alk = 3.8;
[T.e, R.e] = ode15s(@(t,D)Dynamics_pH(t,D,n),[55 99.5],n.init.glysp_c);
subplot(r,c,z); plot(T.e, R.e(:,1), ':r','LineWidth',2); hold off;
subplot(r,c,z+1); plot(T.e, R.e(:,2), ':r','LineWidth',2); hold off;
subplot(r,c,z+2); plot(T.e, R.e(:,3), ':r','LineWidth',2); hold off;
subplot(r,c,z+3); plot(T.e, R.e(:,4), ':r','LineWidth',2); hold off;
pH.e = pH_fig2(T.e,R.e,n); subplot(r,c,z+4); plot(T.e, pH.e', ':r','LineWidth',2); hold off;
```

Function File for Growth Dynamics

```
function dDdt = Dynamics_pH(t,D,n)
%GROWTH KINETICS
%Summary of model parameters
%D(1) = X (Bacterial biomass)
%D(2) = C (Glycerol concentration)
%D(3) = N_L (Ammonia concentration - exogenous)
%D(4) = CD (Carbon dioxide concentration)
%D(5) = O (Oxygen concentration) [hypothetical]

X = D(1); C = D(2); N_L = D(3); CD = D(4); %O = D(5);

%Calculation of actual pH term
pH = fzero(@(pH_val) pH_trend(pH_val,n,CD),n.pH_M);

%Inhibition term, due to acidic pH
I = exp(-n.I_val*((pH-n.pH.UL)/(n.pH.UL-n.pH.LL))^2);

%Growth rate
u = n.u_max*(1-exp(-t/n.t_lag))*(C/(C+n.k_C))*(N_L/(N_L+n.k_N))*I;
%u = n.u_max*(1-exp(-t/n.t_lag))*(C/(C+n.k_C))*(N_L/(N_L+n.k_N))*(O/(O+n.k_O))*I;  %Growth rate with O2

f_X = X-((exp(n.alpha*X)-1)/(exp(n.alpha*n.max_ov.X)-1));   %f(X)

dX = (u-n.k_d)*f_X;                     %Biomass
dC = -((u/n.Y.X_C)*f_X);                %Glycerol
dN_L = -((u/n.Y.X_N)*f_X);              %Ammonia
dCD = (u/n.Y.X_CD)*f_X;                 %Carbon dioxide
%dO = (n.kla_O2*((n.Csat_O2/1000)-O))-((u/n.Y.X_O)*X);     %Oxygen consumption

dDdt = [dX; dC; dN_L; dCD];
end
```

Function Files for pH Modelling

- **pH_trend**

```
function Z = pH_trend(pH_val,n,CD)
%Function to minimize the implicit function (ionic balance)
%CD = Carbon dioxide concentration

H = 10^(-pH_val);           %Changing the pH value to H+ concentration

%ka values (10^(-pka))
ka1 = 10^(-n.pH.pka1);      %ka of KH2PO4
ka2 = 10^(-n.pH.pka2);      %ka of C6H8O7
ka3 = 10^(-n.pH.pka3);      %ka of (C6H7O7)-
ka4 = 10^(-n.pH.pka4);      %ka of (C6H6O7)2-
ka7 = 10^(-n.pH.pka7);      %ka of CO2
ka9 = 10^(-n.pH.pka9);      %ka of H2O

%Concentration of charges based on H+ ions
KHPO4 = n.pH.conc.KH2PO4/((H/ka1)+1);                                   %Equation 1
C6H5O7 = n.pH.conc.C6H8O7/((H^3/(ka2*ka3*ka4))+(H^2/(ka3*ka4))+(H/ka4)+1);  %Equation 2
C6H6O7 = (H/ka4)*C6H5O7;                                                %Equation 3
C6H7O7 = (H/ka3)*C6H6O7;                                                %Equation 4
HCO3 = CD/((H/ka7)+1);                                                  %Equation 8
OH_w = (ka9/H);                                                         %Equation 9

%Charge balance
Z = (OH_w) + HCO3 + KHPO4 + (3*C6H5O7) + (2*C6H6O7) + C6H7O7 - n.pH.alk - H;
end
```

- **pH_fig2**

```
function Z = pH_fig2(T,R,n)
%Function for establishing pH values with time

for i = 1:length(T)
     CD = R(i,4);
     pH(i) = fzero(@(pH_val) pH_trend(pH_val,n,CD), n.pH_M);
end

Z = pH;
end
```

Function File for Calculation of Residuals (Optimization)

```
function Error = OPT_overall_4var(p,normal,gly,amm,ph46,ph55,n)

n.u_max = p(1); n.k_C = p(2); n.k_N = p(3); n.k_d = p(4); n.pH.LL = p(5);
n.alpha = p(6); n.I_val = p(7); n.Y.X_CD = p(8); %n.k_O = p(9);

clear n.pH.alk; n.pH.alk = 7.15;
[T1, D1] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time_spec,n.init.normal_c);
pH.norm = pH_fig2(T1,D1,n);                                          %normal
clear n.pH.alk; n.pH.alk = 6.74;
[T2, D2] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time_spec,n.init.glydef_c);
pH.gly = pH_fig2(T2,D2,n);                                           %gly-def
clear n.pH.alk; n.pH.alk = 6.65;
[T3, D3] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time_spec,n.init.ammdef_c);
pH.amm = pH_fig2(T3,D3,n);                                           %amm-def
clear n.pH.alk; n.pH.alk = 2.574;
[T4, D4] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time_spec,n.init.ph46_c);
pH.ac_46 = pH_fig2(T4,D4,n);                                         %pH 4.6
clear n.pH.alk; n.pH.alk = 3.81;
[T5, D5] = ode15s(@(t,D)Dynamics_pH(t,D,n),n.time_spec,n.init.ph55_c);
pH.ac_55 = pH_fig2(T5,D5,n);                                         %pH 5.5

%Normalize simulation results (normal conditions)
norma.normal.X = (D1(:,1)-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);     %Biomass
norma.normal.C = (D1(:,2)-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);     %Glycerol
norma.normal.N = (D1(:,3)-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);     %Ammonia
norma.normal.pH = (pH.norm'-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);%pH

%Normalize simulation results (glycerol deficiency)
norma.gly.X = (D2(:,1)-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);
norma.gly.C = (D2(:,2)-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);
norma.gly.N = (D2(:,3)-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);
norma.gly.pH = (pH.gly'-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);

%Normalize simulation results (ammonia deficiency)
norma.amm.X = (D3(:,1)-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);
norma.amm.C = (D3(:,2)-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);
norma.amm.N = (D3(:,3)-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);
norma.amm.pH = (pH.amm'-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);

%Normalize simulation results (pH 4.6)
norma.ph46.X = (D4(:,1)-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);
norma.ph46.C = (D4(:,2)-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);
norma.ph46.N = (D4(:,3)-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);
norma.ph46.pH = (pH.ac_46'-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);
```

```
%Normalize simulation results (pH 5.5)
norma.ph55.X = (D5(:,1)-n.min_ov.X)./(n.max_ov.X-n.min_ov.X);
norma.ph55.C = (D5(:,2)-n.min_ov.C)./(n.max_ov.C-n.min_ov.C);
norma.ph55.N = (D5(:,3)-n.min_ov.N)./(n.max_ov.N-n.min_ov.N);
norma.ph55.pH = (pH.ac_55'-n.min_ov.pH)./(n.max_ov.pH-n.min_ov.pH);

%Unweighted error
E.normal.X = norma.normal.X - normal.norma.X;                    %Biomass error
E.normal.C = norma.normal.C - normal.norma.C;                    %Glycerol error
E.normal.N = norma.normal.N - normal.norma.N;                    %Ammonia error
E.normal.pH = norma.normal.pH - normal.norma.pH;                 %pH error
E.gly.X = norma.gly.X - gly.norma.X;
E.gly.C = norma.gly.C - gly.norma.C;
E.gly.N = norma.gly.N - gly.norma.N;
E.gly.pH = norma.gly.pH - gly.norma.pH;
E.amm.X = norma.amm.X - amm.norma.X;
E.amm.C = norma.amm.C - amm.norma.C;
E.amm.N = norma.amm.N - amm.norma.N;
E.amm.pH = norma.amm.pH - amm.norma.pH;
E.ph46.X = norma.ph46.X - ph46.norma.X;
E.ph46.C = norma.ph46.C - ph46.norma.C;
E.ph46.N = norma.ph46.N - ph46.norma.N;
E.ph46.pH = norma.ph46.pH - ph46.norma.pH;
E.ph55.X = norma.ph55.X - ph55.norma.X;
E.ph55.C = norma.ph55.C - ph55.norma.C;
E.ph55.N = norma.ph55.N - ph55.norma.N;
E.ph55.pH = norma.ph55.pH - ph55.norma.pH;

err = [E.normal.X; E.normal.C; E.normal.N; E.normal.pH; E.gly.X; E.gly.C; E.gly.N; E.gly.pH;...
    E.amm.X; E.amm.C; E.amm.N; E.amm.pH; E.ph46.X; E.ph46.C; E.ph46.N; E.ph46.pH;...
    E.ph55.X; E.ph55.C; E.ph55.N; E.ph55.pH];

%Matrix containing weights (1/var)
w = sqrt([normal.v.X; normal.v.C; normal.v.N; normal.v.pH; gly.v.X; gly.v.C; gly.v.N; gly.v.pH;...
    amm.v.X; amm.v.C; amm.v.N; amm.v.pH; ph46.v.X; ph46.v.C; ph46.v.pH; ph46.v.N;...
    ph55.v.X; ph55.v.C; ph55.v.N; ph55.v.pH]);

%Weighted error
Error = w.*err;
end
```

Function Files for Reading Data and Plotting

- **exp_data_plots (plotting of data from parameter estimation experiments)**

```
function [bm,gly,amm,pH] = exp_data_plots(data,r,c,z)
%differentiation of cultures
culA.time = [data.t(1:3); data.t(7:9)];                    culB.time = [data.t(4:6); data.t(10)];
culA.X = [data.X(1:3); data.X(7:9)];                       culB.X = [data.X(4:6); data.X(10)];
culA.e.X = [data.e.X(1:3); data.e.X(7:9)];                 culB.e.X = [data.e.X(4:6); data.e.X(10)];
culA.C = [data.C(1:3); data.C(7:9)];                       culB.C = [data.C(4:6); data.C(10)];
culA.e.C = [data.e.C(1:3); data.e.C(7:9)];                 culB.e.C = [data.e.C(4:6); data.e.C(10)];
culA.N = [data.N(1:3); data.N(7:9)];                       culB.N = [data.N(4:6); data.N(10)];
culA.e.N = [data.e.N(1:3); data.e.N(7:9)];                 culB.e.N = [data.e.N(4:6); data.e.N(10)];
culA.pH = [data.pH(1:3); data.pH(7:9)];                    culB.pH = [data.pH(4:6); data.pH(10)];
culA.e.pH = [data.e.pH(1:3); data.e.pH(7:9)];              culB.e.pH = [data.e.pH(4:6); data.e.pH(10)];

bm = subplot(r,c,z); errorbar(culA.time, culA.X, culA.e.X,'or'); hold on; errorbar(culB.time, culB.X, culB.e.X,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('Biomass (g/L)');
gly = subplot(r,c,z+1); errorbar(culA.time, culA.C, culA.e.C,'or'); hold on; errorbar(culB.time, culB.C, culB.e.C,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('Glycerol Conc. (g/L)');
amm = subplot(r,c,z+2); errorbar(culA.time, culA.N, culA.e.N,'or'); hold on; errorbar(culB.time, culB.N, culB.e.N,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('NH_3 Conc. (EX) (g/L)');
pH = subplot(r,c,z+4); errorbar(culA.time, culA.pH, culA.e.pH,'or'); hold on; errorbar(culB.time, culB.pH, culB.e.pH,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('pH');
end
```

- **pre_opt_plots (plotting of pre-optimized simulations)**

```
function [bm,gly,amm,co2,ph] = pre_opt_plots(r,c,z,T,R,pH)
bm = subplot(r,c,z); plot(T, R(:,1), '-'); hold on;     %Biomass
gly = subplot(r,c,z+1); plot(T, R(:,2), '-'); hold on;  %Glycerol
amm = subplot(r,c,z+2); plot(T, R(:,3), '-'); hold on;  %Exogenous ammonia
co2 = subplot(r,c,z+3); plot(T, R(:,4), '-'); hold on; xlabel('Time (Hours)'); ylabel('CO_2 Conc. (g/L)');      %CO2
ph = subplot(r,c,z+4); plot(T, pH', '-'); hold on;      %pH
%ox = subplot(r,c,z+5); plot(T, R(:,5), '-'); hold on; xlabel('Time (Hours)'); ylabel('O_2 (g/L)');              %O2
end
```

- **post_opt_plots (plotting of post-optimized simulations)**

```
function [bm,gly,amm,co2,ph] = post_opt_plots(r,c,z,T,R,pH)
bm = subplot(r,c,z); plot(T, R(:,1), '-k'); hold off;    %Biomass plot
gly = subplot(r,c,z+1); plot(T, R(:,2), '-k'); hold off; %Glycerol plot
amm = subplot(r,c,z+2); plot(T, R(:,3), '-k'); hold off; %Exogenous ammonia plot
co2 = subplot(r,c,z+3); plot(T, R(:,4), '-k'); hold off; xlabel('Time (Hours)'); ylabel('CO_2 Conc. (g/L)'); %CO2
ph = subplot(r,c,z+4); plot(T, pH', '-k'); hold off;     %pH
%ox = subplot(r,c,z+5); plot(T, R(:,5), '-k'); hold off; xlabel('Time (Hours)'); ylabel('O_2 (g/L)');         %O2
end
```

- **exp_data_plots_val (plotting of data from validation experiments)**

```
function [bm,gly,amm_ex,pH] = exp_data_plots_val(data,r,c,z)
%Culture A
culA.time = [data.t(1:2); data.t(4); data.t(6); data.t(8); data.t(10)];
culA.X = [data.X(1:2); data.X(4); data.X(6); data.X(8); data.X(10)];
culA.e.X = [data.e.X(1:2); data.e.X(4); data.e.X(6); data.e.X(8); data.e.X(10)];
culA.C = [data.C(1:2); data.C(4); data.C(6); data.C(8); data.C(10)];
culA.e.C = [data.e.C(1:2); data.e.C(4); data.e.C(6); data.e.C(8); data.e.C(10)];
culA.N = [data.N(1:2); data.N(4); data.N(6); data.N(8); data.N(10)];
culA.e.N = [data.e.N(1:2); data.e.N(4); data.e.N(6); data.e.N(8); data.e.N(10)];
culA.pH = [data.pH(1:2); data.pH(4); data.pH(6); data.pH(8); data.pH(10)];
culA.e.pH = [data.e.pH(1:2); data.e.pH(4); data.e.pH(6); data.e.pH(8); data.e.pH(10)];

%Culture B
culB.time = [data.t(3); data.t(5); data.t(7); data.t(9)];
culB.X = [data.X(3); data.X(5); data.X(7); data.X(9)];
culB.e.X = [data.e.X(3); data.e.X(5); data.e.X(7); data.e.X(9)];
culB.C = [data.C(3); data.C(5); data.C(7); data.C(9)];
culB.e.C = [data.e.C(3); data.e.C(5); data.e.C(7); data.e.C(9)];
culB.N = [data.N(3); data.N(5); data.N(7); data.N(9)];
culB.e.N = [data.e.N(3); data.e.N(5); data.e.N(7); data.e.N(9)];
culB.pH = [data.pH(3); data.pH(5); data.pH(7); data.pH(9)];
culB.e.pH = [data.e.pH(3); data.e.pH(5); data.e.pH(7); data.e.pH(9)];

bm = subplot(r,c,z); errorbar(culA.time, culA.X, culA.e.X,'or'); hold on; errorbar(culB.time, culB.X, culB.e.X,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('Biomass (g/L)');
gly = subplot(r,c,z+1); errorbar(culA.time, culA.C, culA.e.C,'or'); hold on; errorbar(culB.time, culB.C, culB.e.C,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('Glycerol Conc. (g/L)');
amm_ex = subplot(r,c,z+2); errorbar(culA.time, culA.N, culA.e.N,'or'); hold on; errorbar(culB.time, culB.N, culB.e.N,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('NH_3 Conc. (EX) (g/L)');
pH = subplot(r,c,z+4); errorbar(culA.time, culA.pH, culA.e.pH,'or'); hold on; errorbar(culB.time, culB.pH, culB.e.pH,'vb'); hold on;
xlabel('Time (Hours)'); ylabel('pH');
end
```

- **val_plots (plotting of validation simulations)**

```
function [bm,gly,amm,co2,ph] = val_plots(r,c,z,T,R,pH)
bm = subplot(r,c,z); plot(T, R(:,1), '-'); hold off;
gly = subplot(r,c,z+1); plot(T, R(:,2), '-'); hold off;
amm = subplot(r,c,z+2); plot(T, R(:,3), '-'); hold off;
co2 = subplot(r,c,z+3); plot(T, R(:,4), '-'); xlabel('Time (Hours)'); ylabel('CO_2 Conc. (g/L)');
ph = subplot(r,c,z+4); plot(T, pH', '-'); hold off;
end
```

- **EX_data_1 (reading of parameter estimation data from Microsoft Excel)**

```
function [a,b,c,d,e] = EX_data_1

%Normal conditions
normal.t = xlsread('DATA.xlsx','PARAMETER ESTIMATION','A3:A12');    %Time
normal.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','B3:B12');    %Biomass
normal.v.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','C3:C12');  %Weight_biomass
normal.e.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','D3:D12');  %SE_biomass
normal.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','E3:E12');    %Glycerol
normal.v.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','F3:F12');  %Weight_glycerol
normal.e.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','G3:G12');  %SE_glycerol
normal.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','H3:H12');    %Ammonia
normal.v.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','I3:I12');  %Weight_ammonia
normal.e.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','J3:J12');  %SE_ammonia
normal.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','K3:K12');   %pH
normal.v.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','L3:L12'); %Weight_pH
normal.e.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','M3:M12'); %SE_pH

%Glycerol deficiency
gly.t = xlsread('DATA.xlsx','PARAMETER ESTIMATION','A3:A12');
gly.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','N3:N12');
gly.v.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','O3:O12');
gly.e.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','P3:P12');
gly.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','Q3:Q12');
gly.v.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','R3:R12');
gly.e.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','S3:S12');
gly.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','T3:T12');
gly.v.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','U3:U12');
gly.e.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','V3:V12');
gly.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','W3:W12');
gly.v.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','X3:X12');
gly.e.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','Y3:Y12');

%Ammonia deficiency
amm.t = xlsread('DATA.xlsx','PARAMETER ESTIMATION','A3:A12');
amm.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','B15:B24');
amm.v.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','C15:C24');
amm.e.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','D15:D24');
amm.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','E15:E24');
amm.v.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','F15:F24');
amm.e.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','G15:G24');
amm.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','H15:H24');
amm.v.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','I15:I24');
amm.e.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','J15:J24');
amm.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','K15:K24');
amm.v.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','L15:L24');
amm.e.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','M15:M24');

%pH 4.6
ph46.t = xlsread('DATA.xlsx','PARAMETER ESTIMATION','A3:A12');
ph46.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','N15:N24');
ph46.v.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','O15:O24');
ph46.e.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','P15:P24');
ph46.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','Q15:Q24');
ph46.v.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','R15:R24');
ph46.e.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','S15:S24');
ph46.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','T15:T24');
ph46.v.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','U15:U24');
ph46.e.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','V15:V24');
ph46.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','W15:W24');
ph46.v.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','X15:X24');
ph46.e.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','Y15:Y24');

%pH 5.5
ph55.t = xlsread('DATA.xlsx','PARAMETER ESTIMATION','A3:A12');
ph55.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','B27:B36');
ph55.v.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','C27:C36');
ph55.e.X = xlsread('DATA.xlsx','PARAMETER ESTIMATION','D27:D36');
ph55.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','E27:E36');
ph55.v.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','F27:F36');
ph55.e.C = xlsread('DATA.xlsx','PARAMETER ESTIMATION','G27:G36');
ph55.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','H27:H36');
ph55.v.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','I27:I36');
```

```
ph55.e.N = xlsread('DATA.xlsx','PARAMETER ESTIMATION','J27:J36');
ph55.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','K27:K36');
ph55.v.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','L27:L36');
ph55.e.pH = xlsread('DATA.xlsx','PARAMETER ESTIMATION','M27:M36');

a = normal; b = gly; c = amm; d = ph46; e = ph55;
end
```

- **EX_data_2 (reading of validation data from Microsoft Excel)**

```
function [a,b,c,d,e] = EX_data_2

%Normal conditions (repetition)
normal_r.t = xlsread('DATA.xlsx','VALIDATION','A3:A12');     %Time
normal_r.X = xlsread('DATA.xlsx','VALIDATION','B3:B12');     %Biomass
normal_r.e.X = xlsread('DATA.xlsx','VALIDATION','C3:C12');   %SE_biomass
normal_r.C = xlsread('DATA.xlsx','VALIDATION','D3:D12');     %Glycerol
normal_r.e.C = xlsread('DATA.xlsx','VALIDATION','E3:E12');   %SE_glycerol
normal_r.N = xlsread('DATA.xlsx','VALIDATION','F3:F12');     %Ammonia
normal_r.e.N = xlsread('DATA.xlsx','VALIDATION','G3:G12');   %SE_ammonia
normal_r.pH = xlsread('DATA.xlsx','VALIDATION','H3:H12');    %pH
normal_r.e.pH = xlsread('DATA.xlsx','VALIDATION','I3:I12'); %SE_pH

%Ammonia deficiency (repetition)
amm_r.t = xlsread('DATA.xlsx','VALIDATION','J3:J12');
amm_r.X = xlsread('DATA.xlsx','VALIDATION','K3:K12');
amm_r.e.X = xlsread('DATA.xlsx','VALIDATION','L3:L12');
amm_r.C = xlsread('DATA.xlsx','VALIDATION','M3:M12');
amm_r.e.C = xlsread('DATA.xlsx','VALIDATION','N3:N12');
amm_r.N = xlsread('DATA.xlsx','VALIDATION','O3:O12');
amm_r.e.N = xlsread('DATA.xlsx','VALIDATION','P3:P12');
amm_r.pH = xlsread('DATA.xlsx','VALIDATION','Q3:Q12');
amm_r.e.pH = xlsread('DATA.xlsx','VALIDATION','R3:R12');

%Ammonia and glycerol deficiency (validation)
ammgly.t = xlsread('DATA.xlsx','VALIDATION','A15:A24');
ammgly.X = xlsread('DATA.xlsx','VALIDATION','B15:B24');
ammgly.e.X = xlsread('DATA.xlsx','VALIDATION','C15:C24');
ammgly.C = xlsread('DATA.xlsx','VALIDATION','D15:D24');
ammgly.e.C = xlsread('DATA.xlsx','VALIDATION','E15:E24');
ammgly.N = xlsread('DATA.xlsx','VALIDATION','F15:F24');
ammgly.e.N = xlsread('DATA.xlsx','VALIDATION','G15:G24');
ammgly.pH = xlsread('DATA.xlsx','VALIDATION','H15:H24');
ammgly.e.pH = xlsread('DATA.xlsx','VALIDATION','I15:I24');

%Spiking - Ammonia deficiency (validation)
ammsp.t = xlsread('DATA.xlsx','VALIDATION','J15:J24');
ammsp.X = xlsread('DATA.xlsx','VALIDATION','K15:K24');
ammsp.e.X = xlsread('DATA.xlsx','VALIDATION','L15:L24');
ammsp.C = xlsread('DATA.xlsx','VALIDATION','M15:M24');
ammsp.e.C = xlsread('DATA.xlsx','VALIDATION','N15:N24');
ammsp.N = xlsread('DATA.xlsx','VALIDATION','O15:O24');
ammsp.e.N = xlsread('DATA.xlsx','VALIDATION','P15:P24');
ammsp.pH = xlsread('DATA.xlsx','VALIDATION','Q15:Q24');
ammsp.e.pH = xlsread('DATA.xlsx','VALIDATION','R15:R24');

%Spiking - Glycerol deficiency (validation)
glysp.t = xlsread('DATA.xlsx','VALIDATION','A27:A36');
glysp.X = xlsread('DATA.xlsx','VALIDATION','B27:B36');
glysp.e.X = xlsread('DATA.xlsx','VALIDATION','C27:C36');
glysp.C = xlsread('DATA.xlsx','VALIDATION','D27:D36');
glysp.e.C = xlsread('DATA.xlsx','VALIDATION','E27:E36');
glysp.N = xlsread('DATA.xlsx','VALIDATION','F27:F36');
glysp.e.N = xlsread('DATA.xlsx','VALIDATION','G27:G36');
glysp.pH = xlsread('DATA.xlsx','VALIDATION','H27:H36');
glysp.e.pH = xlsread('DATA.xlsx','VALIDATION','I27:I36');

a = normal_r; b = amm_r; c = ammgly; d = ammsp; e = glysp;
end
```