



METABOLOMIC CHARACTERISATION OF THE HUMAN
PATHOGEN *STREPTOCOCCUS PNEUMONIAE*

Lavanya Mane

University College London

and

The Francis Crick Institute

PhD Co-Supervisor: Robert S Heyderman

PhD Co-Supervisor: Robert J Wilkinson

A thesis submitted for the degree of

Doctor of Philosophy

University College London

July 2021

Declaration

I, Lavanya Mane, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Streptococcus pneumoniae metabolism is a major determinant of pathogenesis, enabling the pneumococcus to survive in diverse host microenvironments encountered during colonisation of the nasopharynx, transmission, and disease, and pointing to a distinct capacity for metabolic flexibility. However, pneumococcal metabolic pathways have mainly been inferred from genomic and transcriptomic analyses, with experimental elucidation of only the key aspects of energetic and amino acid metabolism.

This thesis describes the characterisation of global carbohydrate and associated metabolism of *S. pneumoniae* by ^{13}C stable isotopologue profiling using the metabolomic techniques of gas chromatography–mass spectrometry (GC-MS) and proton nuclear magnetic resonance (^1H -NMR). It was determined that non-fermentative global metabolism during pneumococcal culture on two niche-specific carbohydrate sources, glucose (abundant in the blood and cerebrospinal fluid) and galactose (abundant in the nasopharynx), proceeds along identical pathways in both high and low carriage prevalence strains (23F and TIGR4, respectively).

Biosynthesis of the pneumococcal polysaccharide capsule, a major virulence factor, is connected to carbohydrate metabolism. Using otherwise isogenic capsule-deficient strains TIGR4 Δ *cps* and 23F Δ *cps*, this thesis also demonstrated differential carbohydrate-derived carbon flow dependent on capsule type – 23F is metabolically more resilient to the loss of capsule compared to TIGR4, which may contribute to their differing carriage prevalence and invasive potential.

In this thesis, novel pathway activities unannotated in the pneumococcal genome corresponding to the presence of a urea cycle, and the potential *de novo* synthesis of glycine and proline, are also described. Finally, unexplained succinate production has led to the generation of a set of hypotheses for future investigation. Overall, the results presented here highlight the power of unbiased metabolomics analysis in uncovering novel and adaptive aspects of microbial metabolism.

Impact Statement

Rapid technological advances in the post-genomic era have catapulted the field of metabolomics to the forefront of biological research. Metabolomics is the study of small molecules involved in metabolism, encompassing a set of essential processes in all living organisms. Improvements in detection methods, analytical techniques, and reference databases have allowed more sophisticated interrogation of the dynamic metabolic pathways in microbes, but the adoption of these technologies to study global pneumococcal metabolism has been relatively limited.

The primary contribution of this thesis has been the establishment of a robust method involving intracellular extraction, sample preparation, and analysis using ^{13}C stable isotopologue profiling to probe metabolic pathways in *Streptococcus pneumoniae*. This will enable the pneumococcal research community to adopt such metabolomics methodologies to study various aspects of metabolism, including adaptation to disease conditions, and will provide a template for similar metabolic studies in other related bacteria.

This is also the first description of a direct comparison between the global metabolic profiles of wild type encapsulated pneumococci with their isogenic capsule-deficient mutants. This work has furthered the link between central carbon metabolism and virulence by showing that carbon flow through pneumococcal metabolic pathways is altered in a serotype-dependent manner in the presence of a capsule. This has implications for understanding and predicting adaptive mechanisms determining, or determined by, the invasive potential of a pneumococcal strain.

This work also described novel metabolic reactions and intermediates, generating a set of hypotheses for further characterisation. In particular, the presence of a urea cycle which may help assimilate nitrogen was shown in this thesis, and helps to illustrate the value of metabolomic analyses in identifying aspects of metabolism which cannot be inferred from genomic studies alone. The implications of the methodology and results discussed in this thesis are wide-ranging, and include aiding in the identification of

novel targets for effective design of vaccines and drugs. Overall, this work will be of interest to biochemists, microbiologists, vaccinologists and drug developers alike.

Acknowledgements

I am immensely grateful to my supervisors, Prof Robert Heyderman and Prof Robert Wilkinson, for taking a chance on me and allowing me the freedom to develop and execute this project. I thank Robert Wilkinson for his continued support and guidance. My deepest gratitude goes to Robert Heyderman, who has been invaluable to the development of this thesis and has always been excited to discuss my work – thank you for your encouragement, mentorship, kindness, and incredible patience.

I would also like to take this opportunity to thank members of my Thesis Committee, Prof Douglas Young, Prof Jerry Brown, Dr Maximiliano Gutierrez, and Dr Moritz Treeck, for their helpful suggestions. Special thanks to Jerry Brown for giving me uninhibited access to the Brown lab's collection of carefully curated pneumococcal strains.

I would also like to thank members of the Wilkinson lab at the Crick and the Heyderman lab at UCL for their camaraderie and friendship – Stacey-Ann Lee, Dr Elizabeth Chan, Dr Asia-Sophia Wolf, Modupeh Betts, Dr Maddalena Cerrone, Dr Alizé Proust, Dr Brenda Kwambana-Adams, Akuzike Kalizang'oma, Dr Rachel Lai, Peter Rossi, Dr Emma Wall, and Dr Katalin Wilkinson. I also extend my gratitude to Dr Andrea Gori, who has always been willing to help, and specifically acknowledge him for performing the Kraken classification analysis; and Modupeh Betts for performing growth curve analyses with Curveball described in this thesis. I would especially like to thank Dr Caroline Weight, who taught me the preliminary techniques for culturing pneumococci, mentored me when I first arrived, and never stopped.

I have been fortunate to have access to a range of specialist facilities at the Crick. I would like to acknowledge the Bioinformatics and Biostatistics facility for performing whole genome sequencing of pneumococcal strains used in this thesis, including Dr Aengus Stewart for help with QC assessment and alignment of sequenced reads.

The Metabolomics facility at the Crick has had an immeasurable impact on this thesis, and I am grateful to the entire team for maintaining the metabolomics equipment used to gather data for this work. I am especially grateful to Dr Nathalie Legrave for training me to eventually be able to independently design, perform, and analyse metabolomic experiments, and for answering every question with great patience. I thank Dr James Ellis for training me to undertake GC-MS data analysis using MANIC. I also thank Dr Paul Driscoll for NMR data acquisition and helping with data interpretation. My most sincere gratitude goes to Dr James MacRae, without whom this project would not exist in the form it does. Dr MacRae has provided invaluable input towards the overall development of this work, has helped to specify the test conditions for intracellular pneumococcal extraction methodology, and has continued to use his specialist experience in providing useful insights.

This time would not have been half as enjoyable or as fruitful without the friends I made along the way. Dr Hua Wang, Dr Daniel Greenwood, Dr Serena Chew, and Dr Mazlina Ismail have helped nurture creativity through discussions about art, philosophy, and everything beyond; Dr Kathleen Dolan, Dr Emma Still, and Claudio Del Fatti have made compassionate housemates; and Dr Emma Milford, Dr Rowan Howell, and Dr Souradeep Basu have been brilliant comrades-in-arms. My deepest gratitude also to Dr Rukshala Illukkumbura, for filling my life with books, and to Dr Vidya Chivukula, whose drive has inspired and friendship has sustained me.

I would like to end by thanking my family. I am immensely grateful to my parents, Vishwajeet and Radha Mane, whose support and encouragement have never faltered; to my brother, Daidipya, for being a source of great joy and comfort; and to my sister, Riddhi, for her unshakeable faith in me and for encouraging me to think deeply. Finally, I would like to thank Ori for being an excellent writing companion, and Parth Ashar – for being unfailingly supportive through these last few months, and for always convincingly pretending to be riveted by pneumococcal metabolism.

Table of Contents

Declaration	2
Abstract	3
Impact Statement	4
Acknowledgements	6
Table of Contents	8
List of Figures	11
List of Tables	14
Abbreviations	15
Chapter 1. Introduction	19
1.1 <i>Streptococcus pneumoniae</i>	19
1.1.1 Microbiology and <i>in vitro</i> cultivation	19
1.1.2 Global burden of pneumococcal disease	20
1.1.3 Pneumococcal colonisation and disease	21
1.2 The outer polysaccharide capsule	25
1.2.1 Pneumococcal capsule metabolism	26
1.2.2 Capsules of <i>S. pneumoniae</i> serotype 4 and 23F	28
1.3 Carbohydrate metabolism in <i>S. pneumoniae</i>	30
1.3.1 Carbohydrate catabolism in the respiratory tract	30
1.3.2 Carbohydrate metabolism and virulence are connected through multifunctional genes	34
1.3.3 Pneumococcal metabolic adaptation to host niches	41
1.3.4 Pneumococcal metabolomics – an emerging field	42
1.4 Metabolomics – an overview	44
1.4.1 Metabolomics techniques	44
1.4.2 Applications for microbial metabolomics	45
1.4.3 Targeting pneumococcal metabolism for therapy	46
1.5 Thesis Aims	48
Chapter 2. Materials & Methods	49
2.1 Bacterial strains, stock preparation and storage	49
2.1.1 Enumerating colony-forming units (CFUs)	50
2.1.2 Microbiological identification of a TIGR4 mucoid colony variant	50
2.1.3 <i>S. pneumoniae</i> TIGR4 whole genome sequencing	51
2.2 <i>S. pneumoniae</i> growth studies in defined culture media	52
2.2.1 Sicard's medium (semi-defined)	52
2.2.1.1 Mucin preparation	54
2.2.2 7H9 ⁺ medium (semi-defined)	54
2.2.3 Chemically Defined Medium (CDM)	56
2.2.4 <i>S. pneumoniae</i> growth curves	58
2.3 Conservation of metabolic genes across <i>S. pneumoniae</i> isolates using nucleotide sequence similarity analysis	59
2.3.1 Selection of <i>S. pneumoniae</i> isolates	59
2.3.2 Selection of metabolic genes to include in analysis	61
2.3.3 Nucleotide sequence similarity analysis	61
2.4 Metabolomics – General Methods	62

2.4.1 Gas chromatography–mass spectrometry (GC-MS)	62
2.4.2 Liquid chromatography–mass spectrometry (LC-MS)	63
2.4.2.1 Data collection and analysis	64
2.5 Semi-targeted extracellular metabolite profiling (footprinting).....	66
2.5.1 Data analysis	66
2.6 Stable isotope labelling metabolomics	67
2.6.1 <i>S. pneumoniae</i> culture in defined media	67
2.6.2 Metabolic quench	67
2.6.3 Intracellular metabolite extraction for semi-targeted GC-MS	68
2.6.4 Extracellular metabolite extraction for semi-targeted GC-MS	69
2.6.5 Data analysis	69
2.6.6 Extracellular targeted proton nuclear magnetic resonance (¹ H-NMR)	70
Chapter 3. Genomic and growth characterisation of <i>S. pneumoniae</i> prior to metabolic assessment	72
3.1 Introduction	72
3.2 Chapter 3 Aims.....	74
3.3 Nucleotide sequence similarity analysis reveals that central metabolism associated genes of <i>S. pneumoniae</i> are highly conserved	75
3.4 7H9⁺ culture medium – a combination of select 7H9 and Sicard’s media components	79
3.4.1 Generating 7H9 ⁺ medium	79
3.4.2 Primary carbon sources in 7H9 ⁺ culture medium	84
3.4.3 7H9 ⁺ medium stability and a standardised inoculation method	88
3.5 Chemically Defined Medium (CDM).....	91
3.5.1 Comparing <i>S. pneumoniae</i> growth in CDM and 7H9 ⁺	92
3.5.2 Growth of <i>S. pneumoniae</i> TIGR4 and 23F wild-type and capsule-deficient strains in glucose and galactose.....	95
3.6 Chapter 3 Discussion.....	97
Chapter 4. Extracellular metabolomic profiling suggests novel metabolic activity in <i>S. pneumoniae</i>	101
4.1 Introduction	101
4.2 Chapter 4 Aims.....	103
4.3 Validation of a metabolic footprinting method	104
4.3.1 Detection of glycolytic metabolites during pneumococcal culture	104
4.3.2 Detection of amino acids and additional metabolites.....	107
4.4 Primary carbon source does not affect the metabolic footprint of <i>S. pneumoniae</i>	110
4.4.1 Extracellular accumulation of potential TCA cycle intermediates	113
4.5 Confirming the production of potential TCA cycle intermediates by <i>S. pneumoniae</i> TIGR4	116
4.5.1 Whole genome sequencing of <i>S. pneumoniae</i> TIGR4	116
4.5.2 Metagenomic classification of <i>S. pneumoniae</i> TIGR4 sequence reads.....	118
4.5.3 Extracellular metabolites of <i>S. pneumoniae</i> ATCC-TIGR4.....	119
4.6 Are there unannotated enzymes capable of TCA cycle-related activity in <i>S. pneumoniae</i> TIGR4?	122
4.7 ¹³C-glucose labelling of the intracellular metabolites of <i>S. pneumoniae</i> TIGR4 and TIGR4Δ<i>cps</i>	125
4.7.1 Intermediates of glycolysis and the pentose phosphate pathway	127

4.7.2 Amino acid labelling patterns.....	129
4.7.3 Potential TCA cycle intermediates in <i>S. pneumoniae</i> TIGR4?	132
4.8 Chapter 4 Discussion.....	134
4.8.1 Glycolysis in THY, 7H9 ⁺ (glucose) and 7H9 ⁺ (galactose).....	135
4.8.2 Amino acid metabolism	135
4.8.3 Additional metabolites	137
4.8.4 Differential carbon cycling in TIGR4Δ <i>cps</i> compared to wild-type TIGR4	139
Chapter 5. ¹³C isotope tracing to determine carbohydrate-derived carbon flow in <i>S. pneumoniae</i>	141
5.1 Introduction	141
5.2 Chapter 5 Aims.....	143
5.3 Standardisation of methods for [U-¹³C₆]-sugar metabolomics	144
5.3.1 Standardised intracellular metabolite extraction method for GC-MS.....	144
5.3.2 Standardised pneumococcal cell numbers for stable isotope labelling	150
5.4 Extracellular fermentation profile of <i>S. pneumoniae</i>.....	157
5.4.1 Carbon source has a similar effect on the fermentation profiles of <i>S. pneumoniae</i> TIGR4 and 23F	157
5.4.2 Capsule-deficient <i>S. pneumoniae</i> show different patterns of response to the carbohydrate supplied.....	159
5.5 GC-MS analysis to determine the complete metabolic profile of <i>S. pneumoniae</i> cultured on ¹³C₆-glucose and ¹³C₆-galactose.....	162
5.5.1 Metabolites of glycolysis and the pentose phosphate pathway	163
5.5.2 Metabolites associated with the urea cycle	167
5.5.3 Succinate	170
5.5.4 Glutamine and associated metabolites	172
5.5.5 ¹³ C-label incorporation into amino acids.....	175
5.5.6 Unlabelled amino acids	179
5.6 Chapter 5 Discussion.....	183
5.6.1 Succinate and fumarate production in <i>S. pneumoniae</i>	183
5.6.2 The effect of carbohydrate source on <i>S. pneumoniae</i> metabolism.....	184
5.6.3 The link between outer polysaccharide capsule and carbohydrate metabolism in <i>S. pneumoniae</i>	185
Chapter 6. Final Discussion.....	188
6.1 Hypothesis-led metabolomics	189
6.1.1 <i>S. pneumoniae</i> displays conserved metabolic adaptation to carbohydrates	189
6.1.2 Capsule type of TIGR4 and 23F dictates differential carbon flow through pneumococcal metabolic pathways	191
6.2 Discoveries from unbiased metabolomics	193
6.2.1 TCA cycle activity in <i>S. pneumoniae</i> was hypothesised and disproved, but discovery of succinate production raises the possibility of alternative pathways.....	193
6.2.2 Evidence for a pneumococcal urea cycle	196
6.2.3 Potential biosynthesis of proline and glycine in <i>S. pneumoniae</i>	197
6.3 Summary, limitations, and outlook	199
Chapter 7. Appendix.....	201
Reference List.....	216
Conference Abstracts	252

List of Figures

Figure 1.1 Polysaccharide repeat units of serotype 4 and 23F capsules.....	28
Figure 1.2 Synthesis of TIGR4 and 23F capsule sugars from glycolytic intermediates..	29
Figure 1.3 Central carbon metabolism of <i>S. pneumoniae</i>	33
Figure 2.1 <i>S. pneumoniae</i> TIGR4 colony variants.....	51
Figure 2.2 Characteristics of <i>S. pneumoniae</i> isolates included in similarity analysis	60
Figure 3.1 Nucleotide sequence similarity analysis of metabolism-associated <i>S. pneumoniae</i> genes.....	77
Figure 3.2 <i>S. pneumoniae</i> TIGR4 growth in Sicard's medium with glucose or mucin..	81
Figure 3.3 <i>S. pneumoniae</i> TIGR4 cultured in combinations of Sicard's and 7H9 media.	83
Figure 3.4 <i>S. pneumoniae</i> TIGR4 growth in 7H9 ⁺ medium with varying amounts of sugar.....	85
Figure 3.5 Optimising mucin preparation in 7H9 ⁺ medium.	87
Figure 3.6 Determining an inoculation method for <i>S. pneumoniae</i> in 7H9 ⁺ medium	89
Figure 3.7 Storage of Solution A delays <i>S. pneumoniae</i> TIGR4 growth in 7H9 ⁺ over time.	90
Figure 3.8 TIGR4 growth in media containing 11 mM glucose.....	91
Figure 3.9 <i>S. pneumoniae</i> TIGR4 and 23F growth in a microplate reader.....	94
Figure 3.10 Growth curve best-fit models using the Curveball analysis script.....	94
Figure 3.11 <i>S. pneumoniae</i> strains cultured on glucose and galactose	96
Figure 3.12 Colony morphology of <i>S. pneumoniae</i> strains	96
Figure 3.13 <i>S. pneumoniae</i> core and accessory genome sizes.....	97
Figure 4.1 Metabolic footprint of <i>S. pneumoniae</i> TIGR4 cultured in THY	105
Figure 4.2 Pyruvate metabolism pathways in <i>S. pneumoniae</i>	106
Figure 4.3 <i>S. pneumoniae</i> metabolic footprint in glucose and galactose (7H9 ⁺ medium)	111
Figure 4.4 Schematic showing <i>S. pneumoniae</i> lacks a classical TCA cycle	115
Figure 4.5 <i>S. pneumoniae</i> growth in 7H9 ⁺ (glucose).....	120
Figure 4.6 Extracellular metabolites of two <i>S. pneumoniae</i> TIGR4 strains grown in 7H9 ⁺ (glucose) medium.	121

Figure 4.7 An example BLAST and conserved domain search analysis.....	124
Figure 4.8 CFU of <i>S. pneumoniae</i> TIGR4± <i>cps</i> per extraction in 7H9 ⁺ (glc) medium ..	126
Figure 4.9 Intracellular glycolytic intermediates of <i>S. pneumoniae</i> TIGR4± <i>cps</i> in 7H9 ⁺ (glc).	128
Figure 4.10 Intracellular labelled amino acids of <i>S. pneumoniae</i> TIGR4± <i>cps</i> in 7H9 ⁺ (glc)	130
Figure 4.11 Intracellular unlabelled amino acids of <i>S. pneumoniae</i> TIGR4± <i>cps</i> in 7H9 ⁺ (glc)	131
Figure 4.12 Intracellular “TCA” intermediates of <i>S. pneumoniae</i> TIGR4± <i>cps</i> in 7H9 ⁺ (glc)	133
Figure 5.1 Post-quench CFU of <i>S. pneumoniae</i> TIGR4 per metabolite extraction method	145
Figure 5.2 Principal Components Analysis (PCA) of samples analysed by LC-MS ...	147
Figure 5.3 Variation and metabolite coverage for each intracellular extraction method	148
Figure 5.4 Extraction efficiency of different extraction methods for annotated metabolites	149
Figure 5.5 <i>S. pneumoniae</i> growth (CFU ml ⁻¹) to determine media swap time points. .	150
Figure 5.6 Testing media swap time points for <i>S. pneumoniae</i> TIGR4Δ <i>cps</i>	152
Figure 5.7 <i>S. pneumoniae</i> CFU at different stages of culturing for stable isotope labelling	154
Figure 5.8 Example chromatograms of different intracellular extract injection amounts.	155
Figure 5.9 Extracellular fermentation profile of <i>S. pneumoniae</i> TIGR4 and 23F	158
Figure 5.10 Comparison of the fermentation profile between WT and Δ <i>cps</i> <i>S. pneumoniae</i> strains cultured in ¹³ C-glucose or ¹³ C-galactose.	161
Figure 5.11 Intermediates of glycolysis and the pentose phosphate pathway – GC-MS.	166
Figure 5.12 Urea cycle in <i>S. pneumoniae</i>	168
Figure 5.13 Metabolites involved in the urea cycle – GC-MS.	170
Figure 5.14 Succinate production and labelling in <i>S. pneumoniae</i> – GC-MS.	171

Figure 5.15 Isotope ratios of glutamine and glutamine-associated metabolites – GC-MS	172
Figure 5.16 Amounts of glutamine, glutamate and 5-oxoproline – GC-MS.	173
Figure 5.17 α -ketoglutarate production and labelling in <i>S. pneumoniae</i> – GC-MS.	174
Figure 5.18 Isotope ratios of amino acids in <i>S. pneumoniae</i> – GC-MS.	178
Figure 5.19 Amino acids depleted from the culture medium during <i>S. pneumoniae</i> growth.	179
Figure 5.20 Amino acids unlabelled during <i>S. pneumoniae</i> growth – GC-MS.	181
Figure 6.1 Summary schematic – Fermentation profiles of <i>S. pneumoniae</i> TIGR4 and 23F on glucose and galactose	190
Figure 6.2 Intracellular carbon labelling patterns in <i>S. pneumoniae</i> metabolites do not correspond to TCA cycle activity	194
Figure 6.3 Stable isotope labelling of urea cycle intermediates in <i>S. pneumoniae</i>	197
Figure 7.1 Extracellular metabolites of <i>S. pneumoniae</i> TIGR4 cultured in 7H9 ⁺ (glucose) and 7H9 ⁺ (galactose).....	211

List of Tables

Table 1.1 Genes and regulators involved in metabolism and virulence of <i>S. pneumoniae</i>	39
Table 2.1 Composition of Sicard's medium.....	53
Table 2.2 Composition of 7H9 ⁺ medium.....	55
Table 2.3 Composition of CDM	57
Table 2.4 Summary of <i>S. pneumoniae</i> isolates included in similarity analysis.....	60
Table 2.5 Intracellular metabolite extraction methods	65
Table 3.1 Composition of chemically defined media.	80
Table 4.1 Characteristics of <i>S. pneumoniae</i> TIGR4 sequenced reads	117
Table 4.2 Metagenomic characterisation of <i>S. pneumoniae</i> TIGR4 sequence reads....	119
Table 4.3 BLAST analysis of <i>S. pneumoniae</i> TIGR4 against TCA cycle enzyme sequences	123
Table 5.1 Summary of intracellular metabolite extraction methods.....	145
Table 7.1 List of <i>S. pneumoniae</i> isolates included in nucleotide sequence similarity analysis of metabolic genes	202
Table 7.2 List of metabolic genes included in nucleotide sequence similarity analysis across 120 <i>S. pneumoniae</i> isolates.....	206
Table 7.3 Extracellular metabolites of <i>S. pneumoniae</i> TIGR4 cultured in THY.....	210
Table 7.4 List of nucleotide variants in two TIGR4 strains compared to reference.....	213

Abbreviations

[U- ¹³ C ₆]-sugar	Uniformly labelled ¹³ C ₆ -sugar; ¹³ C-sugar
1,3-BPG	1,3-biphosphoglycerate
¹ H-NMR	Proton nuclear magnetic resonance
2-PGA	2-phosphoglycerate
23F	<i>S. pneumoniae</i> 23F ST36 clinical isolate OXC-1417-23F
3-PGA	3-phosphoglycerate; glycerate 3-phosphate
5,10-MTHF	5,10-methylenetetrahydrofolate
Ac-CoA	Acetyl-Coenzyme A
Acetyl-P	Acetyl-phosphate
AckA	Acetate kinase
ACN	Acetonitrile
Adh	Alcohol dehydrogenase
AlaT	Alanine transaminase
AspC	Aspartate transaminase
ATCC	American Type Culture Collection
ATCC-TIGR4	<i>S. pneumoniae</i> TIGR4 type strain (BAA-334)
BCAA	Branched chain amino acid
BgaA/C	β-galactosidase A/C
BHI	Brain heart infusion
BLAST	Basic Local Alignment Search Tool
blastn	Nucleotide BLAST
blastp	Protein BLAST
BSA	Bovine serum albumin
BSTFA	N,O-bis(trimethylsilyl)trifluoroacetamide
CBP	Choline-binding proteins
CCR	Carbon catabolite repression
CDM	Chemically defined medium
CFU	Colony-forming units
ChoP	Phosphorylcholine
<i>cps</i>	Capsule biosynthesis locus

CSF	Cerebrospinal fluid
CV	Coefficient of variation
D ₂ O	Deuterium oxide
DHAP	Dihydroxyacetone phosphate; glycerone phosphate
DSS	3-(trimethylsilyl)propane-1-sulfonate, sodium salt
dTDP	Deoxythymidine diphosphate
EC	Enzyme Commission number
EDTA	Ethylenediaminetetraacetic acid
EI	Electron Impact ionisation
Eno	Enolase
F-1,6-bP/ FBP	Fructose-1,6,-bisphosphate
F6P	Fructose-6-phosphate
FucNAc	<i>N</i> -acetylfucosamine
G6P	Glucose-6-phosphate
Gal	Galactose
Gal6P	Galactose-6-phosphate
GalNAc	<i>N</i> -acetylgalactosamine
GAP	Glyceraldehyde 3-phosphate
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
Glc	Glucose
Glc-1-P/ G1P	Glucose-1-phosphate
GlcNAc	<i>N</i> -acetylglucosamine
Glp	α -glycerophosphate; <i>sn</i> -Glycerol-3-phosphate
GlpO	α -glycerophosphate oxidase
GPSC	Global Pneumococcal Sequence Clusters
<i>H. influenzae</i>	<i>Haemophilus influenzae</i>
IC	Intracellular
ISTD	Internal standard
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Liquid chromatography

LC-MS	Liquid chromatography-mass spectrometry
LctO	Lactate oxidase
Ldh	Lactate dehydrogenase
Log ₂ FC	Log ₂ fold change
LRTI	Lower respiratory tract infection
ManNAc	<i>N</i> -acetylmannosamine
Mdh	Malate dehydrogenase
MLST	Multilocus sequence type
MM	Metabolite standard mix
MS/MS	Tandem mass spectrometry
MWCO	Molecular weight cut-off
<i>N. meningitides</i>	<i>Neisseria meningitides</i>
NanA	Neuraminidase A
NCBI	The National Center for Biotechnology Information
NOESY	Nuclear Overhauser Effect Spectroscopy
OD ₆₀₀	Optical density at 600 nm
PBQC	Pooled biological quality control
PBS	Phospho-Buffered Saline
PCA	Principal component analysis
PCV	Pneumococcal conjugate vaccine
PEP	Phosphoenolpyruvate
Pfl	Pyruvate formate lyase
Pgk	Phosphoglycerate kinase
PGM	Porcine gastric mucin
Pgm	Phosphoglucomutase
Ply	Pneumolysin
Ppc	Phosphoenolpyruvate carboxylase
PPP	Pentose phosphate pathway
PPSV	Pneumococcal polysaccharide vaccine
Psp	Pneumococcal surface protein
Pta	Phosphotransacetylase
PTS	Phosphotransferase system

QC	Quality control
S-7-P	Sedoheptulose-7-phosphate
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
<i>S. pneumoniae</i>	<i>Streptococcus pneumoniae</i>
SHMT	Serine hydroxymethyltransferase
SpxB	Pyruvate oxidase
ST	Sequence type (of <i>Streptococcus pneumoniae</i>)
StrH	β -N-acetylglucosaminidase
t0	Time 0
T4	<i>S. pneumoniae</i> TIGR4 strain P1542; TIGR4
T6P	Tagatose-6-phosphate
tblastx	Translated nucleotide BLAST
TBP	Tagatose 1,6-diphosphate
THY	Todd-Hewitt Broth + 0.5% yeast extract
TMCS	Trimethylchlorosilane
UDP	Uridine diphosphate
URT	Upper respiratory tract
v/v	volume/volume
w/v	weight/volume
WGS	Whole genome sequencing
α -GP	α -glycerophosphate
α -KG	α -ketoglutarate; 2-oxoglutarate
α G1P	α -Glucose-1-phosphate
α Gal1P	α -Galactose-1-phosphate

Chapter 1. Introduction

1.1 *Streptococcus pneumoniae*

1.1.1 Microbiology and *in vitro* cultivation

Streptococcus pneumoniae (*S. pneumoniae*; the pneumococcus) are gram-positive, lancet-shaped, non-motile bacteria belonging to the Lactobacillales order of lactic acid-producing bacteria. The diameter of a single cell is 0.5 – 1.25 µm, and pneumococci are commonly observed in pairs as diplococci but may also occur individually or form chains. Classified as facultative anaerobes, *S. pneumoniae* *in vitro* grow best in microaerophilic conditions with 5% CO₂ on a stationary platform. They produce large amounts of endogenous hydrogen peroxide (up to ~ 0.4 mM (Lisher et al., 2017)) but lack the H₂O₂ -neutralising enzyme catalase. Therefore, an external source of catalase must be provided during *in vitro* culture, which is commonly derived from horse or sheep blood during solid culture and animal extracts or bovine liver catalase in liquid culture (McDevitt et al., 2020).

Routine pneumococcal cultivation is performed in complex media, such as THY (Todd Hewitt broth supplemented with 0.5% yeast extract) and BHI (Brain Heart Infusion broth), but a range of chemically defined culture media have also been described (Sanchez-Rosario & Johnson, 2021) and are used to study specific aspects of metabolism. Microbiological identification of *S. pneumoniae* is aided by bile-solubility, optochin-sensitivity, catalase-negativity, and a distinct colony morphology showing surrounding green zones of α-haemolysis (CDC, NCIRD, & WHO, 2011).

More than 100 immunochemically distinct pneumococcal serotypes have been identified to date, characterised by the structure and organisation of the outer polysaccharide capsule (The Global Pneumococcal Sequencing Project). Multilocus sequence typing (MLST) is an additional method to genotypically group *S. pneumoniae* by determining the allelic profiles of seven housekeeping genes generating a sequence type (ST) (Enright & Spratt, 1998). *S. pneumoniae* display a high degree of genetic diversity and are naturally competent, with the capsule biosynthetic locus *cps* being a

particular hotspot for recombination via horizontal gene transfer. Genetic exchange partners extend beyond the species and include oral streptococci, creating a reservoir for the acquisition of antibiotic-resistance and virulence-enhancing genes (Chaguza et al., 2015; Chi et al., 2007; Kilian et al., 2008; Nahm et al., 2019; Salvadori et al., 2019). A description of the serotype and ST of *S. pneumoniae* strains has aided efforts to understand the genetic factors affecting disease potential and global distribution of *S. pneumoniae* strains. Based on whole genome sequences, Global Pneumococcal Sequence Clusters (GPSCs) have been proposed to categorise pneumococcal lineages (Gladstone et al., 2019).

1.1.2 Global burden of pneumococcal disease

S. pneumoniae is a frequent asymptomatic coloniser of the human nasopharynx and a part of the upper respiratory tract microflora, but may travel to other host sites and cause disease – pneumonia in the lower respiratory tract, meningitis in the brain, and septicaemia can be life-threatening, while bronchitis, sinusitis, and otitis media (infection of the middle ear canal) are milder manifestations of pneumococcal disease (Bogaert et al., 2004; Henriques-Normark & Tuomanen, 2013; Weiser, 2010).

The most vulnerable populations for pneumococcal infection are children < 5 years of age, elderly adults > 65 years of age, immunocompromised individuals, and individuals suffering from co-morbidities such as diabetes and chronic lung conditions (Bennett et al., 2013; CDC, 2013). Risk factors for pneumococcal pneumonia include smoking, alcohol misuse, and viral co-infections (particularly influenza (Klugman et al., 2009)). Geographical location also affects disease outcome, with a significant proportion of mortality occurring in the resource-limited settings of low- and middle-income countries.

S. pneumoniae is the most common causative agent of community-acquired pneumonia, alongside *Haemophilus influenzae* and *Staphylococcus aureus*, and remains one of the leading causes of lower respiratory tract infection (LRTI) morbidity and mortality – an

estimated 2.5 million deaths were attributed to LRIs in 2017 worldwide (Roth et al., 2018).

Global surveillance data from 2015 estimated 3.7 million severe *S. pneumoniae* cases and 300,000 – 500,000 deaths from *S. pneumoniae* in children under five, with half of these deaths occurring in India, Pakistan, Nigeria, and the Democratic Republic of the Congo (Wahl et al., 2018). Total deaths from pneumococcal pneumonia decreased by 67 – 75% between 1990 and 2017, but pneumococcal pneumonia-associated deaths in the elderly during the same time period increased by 40 – 80% (Roth et al., 2018). The overall decrease in death associated with *S. pneumoniae* was driven in part by a reduction in child mortality as a result of vaccine rollout, but *S. pneumoniae* continues to be one of the leading causes of vaccine-preventable deaths in children under five globally (Wahl et al., 2018; WHO, 2020).

Pneumococcal conjugate vaccine (PCV) 10 and 13 are routinely used for immunization of infants, and pneumococcal polysaccharide vaccine (PPSV) 23 has shown some efficacy in the elderly (Bigogo et al., 2019; Kent et al., 2019; Leventer-Roberts et al., 2015; Moberley et al., 2013; Musher, 2021; Tuomanen, 2021). These vaccines are designed to confer protection against 10, 13 and 23 serotypes respectively, out of the ~100 known *S. pneumoniae* serotypes, making them only partially protective. The vaccine type strains are chosen based on their prevalence in disease conditions, but non-vaccine serotype replacement is a growing concern (Levy et al., 2019; Lo et al., 2019; Vadlamudi et al., 2019; Weinberger et al., 2011).

1.1.3 Pneumococcal colonisation and disease

Nasopharyngeal colonisation is a prerequisite for disease and acts as a reservoir for pneumococcal dissemination through the population via airborne droplets (Siegel & Weiser, 2015). *S. pneumoniae* carriage rates differ based on geographical location, but overall are highest in infants (30 – 60%), falling with age to <10 – 25% in adults (Goldblatt et al., 2005; Henriques-Normark & Tuomanen, 2013; Regev-Yochay et al., 2004). Carriage may be symptomatic in children (Fan et al., 2016; Rodrigues et al.,

2013) but was found to be asymptomatic in adults in an experimental human pneumococcal challenge model (Trimble et al., 2020).

Mechanisms of colonisation have been extensively studied (recent reviews: Loughran et al., 2019; Siegel & Weiser, 2015; Subramanian et al., 2019). To persist in the nasopharynx, *S. pneumoniae* encode an arsenal of virulence factors to interact with epithelial cells, evade mucosal clearance, and compete with other members of the respiratory microflora (Mitchell & Mitchell, 2010).

Epithelial interactions are facilitated by surface-expressed adhesion factors, which are revealed by the downregulation of the outer polysaccharide capsule (phase variation). These include, but are not limited to, choline-binding proteins (CBPs, e.g., CbpA), pneumococcal surface proteins (e.g., PspA, PspC, PspK, PsaA), phosphorylcholine (ChoP), enolase (Eno), lipoproteins (e.g., SlrA, PmpA), and pneumococcal adherence and virulence factors (e.g., PavA, PavB). Pneumolysin (ply), bacteriocins (pneumocins), autolysins (e.g., LytA, LytB), serine proteases (e.g., PrtA), exoglycosidases (e.g., NanA, NanB), and hydrogen peroxide also aid both colonisation and invasion.

Pneumococcal proteins aiding pathogenesis are associated with the cell surface via covalent linkage of the protein LPXTG motif to the cell wall peptidoglycan (this attachment motif is present in many Gram-positive bacteria) – e.g., NanA and PrtA – or via noncovalent interaction with the cell wall choline (Mitchell, 2000). Choline-binding proteins include the hydrolytic enzymes LytA, LytB, LytC and CbpE, and the complement inhibitors PspA and PspC, contributing to nasopharyngeal colonisation, adherence to alveolar epithelial and endothelial cells, pneumolysin release, bacteraemia, and IgA inhibition (Berry & Paton, 2000; Crain et al., 1990; Dave et al., 2001; Gosink et al., 2000; Kerr et al., 2006; McDaniel et al., 1987; Ren et al., 2003; Rosenow et al., 1997; Zhang et al., 2000).

Pneumococcal biofilm formation has also been reported, which confers increased resistance to antimicrobial peptides and may aid the sharing of genetic information. Recent evidence additionally implicates biofilm formation in the transition to invasive

disease, as *S. pneumoniae* released from biofilms are more virulent than planktonic pneumococci (Allan et al., 2014; Gilley & Orihuela, 2014; Moscoso et al., 2006).

For the host, mucociliary clearance is the first line of defence against *S. pneumoniae* in the respiratory tract. Besides providing a physical barrier, mucus components (mucins – high-MW glycoproteins) serve as scaffolds for antimicrobial proteins to aid the immune system (Corfield, 2015; Fahy & Dickey, 2010; Kim 2012). To avoid clearance, *S. pneumoniae* employ two main strategies – Increasing production of the anti-phagocytic outer polysaccharide capsule (Kim et al., 1999; Weiser, 2010); and digestion of mucins, which serves the dual purpose of disrupting barrier function and providing free oligosaccharides for nutrition (Yesilkaya et al., 2008).

A key protein in this process is the secreted pneumococcal neuraminidase A (nanA), which is anchored to the cell wall peptidoglycan via its LPXTG motif (Cámara et al., 1994), and plays a role in pathogenesis and nutrient acquisition. It cleaves terminal sialic acids on mucin glycoproteins to damage the host, expose host cell surface receptors for pneumococcal adhesion (Andersson et al., 1983), and provide access to subsequent linked oligosaccharides for pneumococcal glycosidases (Yesilkaya et al., 2008). NanA is found in all clinical isolates of *S. pneumoniae* (Kelly et al., 1967; King et al., 2005; O'Toole et al., 1971) and contributes to the establishment of colonisation, biofilm formation, and the development of otitis media (Brittan et al., 2012; Parker et al., 2009; Stahl et al., 1972).

When *S. pneumoniae* successfully travel to the lower respiratory tract, an inflammatory response is initiated involving resident alveolar macrophages and alveolar epithelial cells followed by an influx of neutrophils (Kadioglu & Andrew, 2004). Pneumococcal clearance occurs predominantly via complement and opsonisation-mediated phagocytosis and is usually effective but comes at the cost of damage to host cells – the inflammatory environment leads to lung consolidation and other symptoms of lobar pneumonia (McCullers & Tuomanen, 2001).

Phagocytosis of *S. pneumoniae* induces caspase- and nitrogen oxide-mediated apoptosis in alveolar macrophages. In murine models of pneumococcal infection, this apoptosis is associated with bacterial clearance and reduced inflammation (Ali et al., 2003; Dockrell et al., 2003; Marriott et al., 2006). However, intracellular survival of *S. pneumoniae* in alveolar macrophages, splenic macrophages, dendritic cells, cardiac cells, and alveolar epithelial cells has also been reported (Brissac et al., 2018; Brown et al., 2014; Ercoli et al., 2018; Marriott et al., 2012; Subramanian et al., 2019).

S. pneumoniae can transmigrate across the nasopharyngeal and alveolar epithelia and invade endothelial cells to access the bloodstream, and from here they can travel to other sites in the body (e.g., brain, kidney, heart) to cause serious disease (Weight et al., 2019). The transition from colonisation to an invasive, disease-causing state is dependent on multiple and incompletely understood factors. A recent mouse model described distinct transcriptomic profiles of *S. pneumoniae* during nasopharyngeal colonisation and disease at particular host sites (lungs, blood, heart, kidney) (D'Mello et al., 2020).

1.2 The outer polysaccharide capsule

The *S. pneumoniae* capsule is a major pneumococcal virulence factor, and is the target for all current licensed pneumococcal vaccines (Wilson et al., 2015). It is made of polysaccharide repeat units partially covalently attached to the cell wall, which often form linkages and contain side chains. As mentioned, polysaccharide composition and capsule structure can be highly variable between strains, with > 100 distinct serotypes described (Bentley et al., 2006). This serotype heterogeneity facilitates non-vaccine serotype replacement within vaccinated populations under immune pressure. The capsule biosynthesis genes in *S. pneumoniae* are encoded within a single transcriptional locus (*cps*), consisting a *cps* promoter, monosaccharide synthesis genes, polysaccharide polymerases, transferases, and flippases (Yother, 2011). Genes within the *cps* locus are type-specific, but the entire *cps* locus is housed between two genes not involved in capsular synthesis but highly homologous between *S. pneumoniae* strains – *dexB* and *aliA*. These homologous flanking genes coupled with the natural competence of *S. pneumoniae* aid capsule switching, which can undermine the efficacy of vaccines, which target only a subset of *S. pneumoniae* serotypes.

Most capsules, including those of *S. pneumoniae* serotypes 4 and 23F, are negatively charged and repel anionic mucus in the respiratory tract. Unencapsulated *S. pneumoniae* are rare, and *cps* deletion mutants exhibit attenuated virulence in animal models of infection (Keller et al., 2016; MacLeod & Krauss, 1950; Morona et al., 2004; Watson & Musher, 1990). Capsule production is dynamic and regulated through the stages of colonisation, disease and transmission (Manso et al., 2014). Broadly, a thicker capsule (opaque phase variant) masks surface antigens, prevents complement-mediated phagocytosis and aids penetration of the mucus layer, while a thinner capsule (transparent phase variant) exposes cell wall-associated adhesion factors and promotes colonisation when in close proximity to epithelial cells (Hammerschmidt et al., 2005; Kim et al., 1999; Magee & Yother, 2001; Middleton et al., 2018; Nelson et al., 2007).

1.2.1 Pneumococcal capsule metabolism

Capsule biosynthesis and carbohydrate metabolism in *S. pneumoniae* are interconnected. Due to the diversity in capsular polysaccharide composition, conclusions from relevant studies are type-specific and must be interpreted as such. The precursor nucleotide di-/ tri-phosphate sugars required for capsule synthesis are generated from the early glycolytic intermediates glucose-6-phosphate (G6P) and fructose-6-phosphate (F6P) (**Figure 1.2**). The conversion of G6P to G1P (glucose-1-phosphate) by phosphoglucomutase (*pgm*), which goes on to generate the capsule sugars UDP-glucose and UDP-galactose, is an essential step to direct carbons away from glycolysis and towards capsule synthesis (Hardy et al., 2000). Capsule production is regulated by the availability and identity of carbohydrate sources – fructose is less efficiently converted to G1P than glucose, and glucose catabolism generates greater amounts of UDP-glucose compared to galactose (Troxler et al., 2019; Werren et al., 2021).

Pyruvate generated during glycolysis can be metabolised to fermentation end-products via multiple genes – *ldh*, *spxB*, *pfl*, *pdhC*, *adhE*, *ackA*. These genes (apart from *ackA*, which has not been studied because generation of a deletion mutant has growth-inhibiting off-target effects) are further involved in capsule regulation and are themselves affected by the carbohydrate source. *SpxB*- and *PdhC*-generated acetyl-CoA directly affects the degree of polysaccharide acetylation, contributing to capsule production in serotypes containing acetylated sugars (e.g., serotype 4, but not 23F). Carbon flux through the acetyl-CoA-generating enzymes *spxB*, *pfl*, and *pdhC* is increased when supplied with galactose vs. glucose, further implicating specific carbohydrate sources in regulating both capsule metabolism and glycolytic flux (Al-Bayati et al., 2017; Carvalho et al., 2013a; Echlin et al., 2016; 2020).

Capsule biosynthesis presents a significant energetic cost for the pneumococcus, the degree of which varies by capsular polysaccharide structure and may be connected to carriage prevalence. Studies using capsule switch mutants (same strain background expressing *cps* loci from different serotypes) have shown that the metabolic burden of capsule production is greater in low carriage prevalence serotypes (e.g., serotype 4)

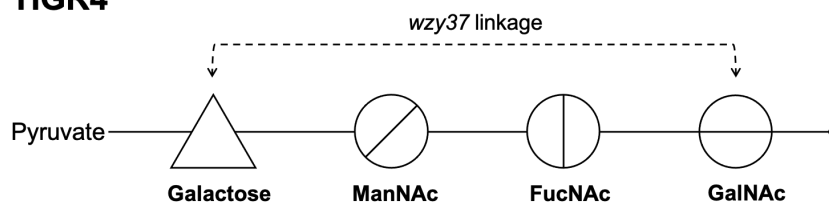
compared to serotypes with high carriage prevalence (e.g., 23F). High carriage prevalence strains with low-burden capsules were also more resistant to neutrophil-mediated death. This evidence suggests that diverting energy away from capsule biosynthesis may help in the establishment of longer-term colonisation and predict serotype-replacement strains which may become dominant after vaccination (Hathaway et al., 2012; Weinberger et al., 2009).

In contrast, evidence is emerging for the role of capsule not just as a biosynthetic cost and virulence factor, but also as a potential carbohydrate store. In nutrient-limiting conditions such as during transmission between hosts, capsular polysaccharides may be catabolised to provide a source of energy. Bacterial transmission potential is, therefore, prolonged for encapsulated bacteria grown in starvation conditions when compared to unencapsulated bacteria (Hamaguchi et al., 2018). Together, the dynamic nature of capsule production, its function as a key virulence factor, and the balancing act between capsular polysaccharide biosynthesis and breakdown demonstrate the intimate link between pneumococcal pathogenesis and central carbohydrate metabolism.

1.2.2 Capsules of *S. pneumoniae* serotype 4 and 23F

In this study, *S. pneumoniae* serotype 4 (TIGR4) and serotype 23F were used (**Figure 1.1**). Both capsules are synthesised via the Wzx/Wzy-dependent pathway. In the cytoplasm, the initial monosaccharide phosphate (UDP-GalNAc in serotype 4 and UDP-glucose in 23F) is linked to the membrane-bound lipid carrier undecaprenyl pyrophosphate, followed by sequential attachment of additional sugars by *cps*-encoded glycosyl-transferases to form the repeat unit. The lipid-linked repeat unit is then transferred from the cytoplasmic side to face outwards by the *wzx* flippase. Here, polysaccharide repeat units are linked by the *wzy* polymerase and attached to the peptidoglycan cell wall to form the mature capsule (Bentley et al., 2006; Yother, 2011).

TIGR4



23F

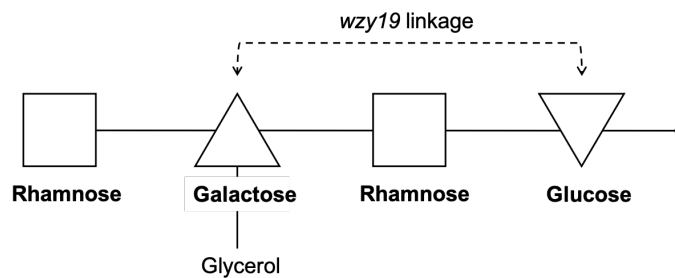


Figure 1.1 Polysaccharide repeat units of serotype 4 and 23F capsules.

Adapted from (Bentley et al., 2006). Dashed lines indicate linkages for polymerisation.

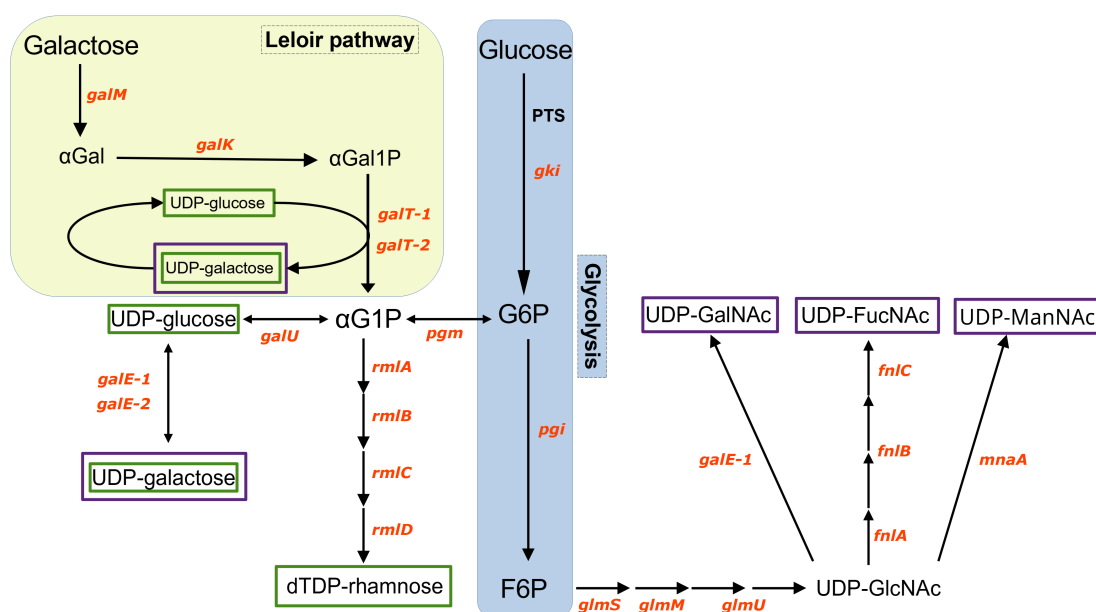


Figure 1.2 Synthesis of TIGR4 and 23F capsule sugars from glycolytic intermediates.

Purple boxes – Monosaccharides of the TIGR4 capsule (UDP-gal, UDP-GalNAc, UDP-FucNAc, UDP-ManNAc). Green boxes – Monosaccharides of the 23F capsule (UDP-glc, UDP-gal, dTDP-rhamnose).

Abbreviations: PTS – phosphotransferase system; G6P – glucose-6-phosphate; F6P – fructose-6-phosphate; αG1P – α-glucose-1-phosphate; αGal1P – α-galactose-1-phosphate; αGal – α-galactose; UDP – uridine diphosphate; GlcNAc – *N*-acetylglucosamine; GalNAc – *N*-acetylgalactosamine; FucNAc – *N*-acetylfucosamine; ManNAc – *N*-acetylmannosamine; dTDP – deoxythymidine diphosphate.

1.3 Carbohydrate metabolism in *S. pneumoniae*

S. pneumoniae is metabolically well-adapted to occupying the nasopharynx (the major site of colonisation), in addition to sites where it causes disease (Leonard & Lalk, 2018). The metabolic demand of maintaining prominent virulence factors such as the polysaccharide capsule, the involvement of metabolic genes in causing disease, and the link to antimicrobial resistance all implicate pneumococcal cellular metabolism as a key player in this adaptation process.

1.3.1 Carbohydrate catabolism in the respiratory tract

The respiratory tract is considered a nutrient-poor environment, as free glucose in the airway surface fluid is actively suppressed by the host as a form of “nutritional immunity” (Rohmer et al., 2011). Normal blood glucose *in vivo* is ~ 5.5 mM, rising to > 10 mM during hyperglycaemia (such as in diabetic patients and in the lung during infection). The glucose concentration in cerebrospinal fluid (CSF) is similar to that in the blood (Leen et al., 2012). Nasal and lung secretions in the uninflamed respiratory tract are estimated to contain ~ 1 mM and < 0.5 mM glucose, respectively. (Pezzulo et al., 2011; Philips et al., 2003; Siegel & Weiser, 2015). However, the respiratory tract mucus lining is rich in membrane-tethered and secreted mucins, which consist of carbohydrate chains linked to a protein backbone, often with a terminal sialic acid (Rose & Voynow, 2006).

The metabolic flexibility of *S. pneumoniae* is exemplified by the production of glycosidases and neuraminidases that degrade host mucin glycoproteins to overcome the lack of free sugars and release energetically utilisable oligosaccharides. Terminal sialic acid cleavage by *S. pneumoniae* neuraminidases is essential for glycosidases to access underlying sugars. (Hoskins et al., 2001; Lanie et al., 2007; Tettelin et al., 2001; Yesilkaya et al., 2008). This breakdown of mucins serves the additional purpose of disrupting barrier function, as mucociliary clearance is the first line of host defence against *S. pneumoniae* in the respiratory tract (Corfield, 2015; Kim 2012). Some of these enzymes also contribute to evading complement-mediated clearance, undermining

other competing upper respiratory tract colonisers such as *H. influenzae* and *N. meningitidis*, and promoting biofilm formation (**Table 1.1**), linking carbohydrate acquisition with pathogenic potential (Robb et al., 2017).

The exoglycosidases neuraminidase A (NanA), β -galactosidase A/C (BgaA/C), β -N-acetylglucosaminidase (StrH) and α -mannosidase (GH92) cleave glycosyl bonds of host mucin glycoproteins to release sialic acid, galactose, N-acetylglucosamine (GlcNAc) and mannose, respectively (Robb et al., 2017). The carbohydrates serve as nutrients for *S. pneumoniae*, with galactose being one of the most abundant components of mucins.

Uptake of the cleaved sugars is facilitated by numerous bacterial transporters, many of which have also been implicated in *S. pneumoniae* pathogenesis. The pneumococcal genome is predicted to encode 21 phosphotransferase (PTS) and 7 ABC transporter systems, allowing it to utilise more than 30 carbohydrates (Bidossi et al., 2012). As such, *S. pneumoniae* has a wider carbohydrate utilisation capacity than *N. meningitidis* and *H. influenzae*, presumably providing a further competitive advantage (Taniai et al., 2008; Tettelin et al., 2001).

S. pneumoniae lacks a respiratory chain and the enzymes required for the tricarboxylic acid (TCA) cycle are not encoded in its genome (Hoskins et al., 2001; Lanie et al., 2007; Tettelin et al., 2001). Imported sugars are therefore metabolised primarily via glycolysis and the pentose phosphate pathway (Leonard et al., 2018; Willenborg & Goethe, 2016). As *S. pneumoniae* is a facultative anaerobe, in the simplest scenario imported carbohydrates undergo homolactic fermentation, in which they are phosphorylated through various steps of the glycolytic pathway to generate pyruvate, which is then converted predominantly to lactic acid (Willenborg & Goethe, 2016). Mixed-acid fermentation, wherein ethanol, formate and acetate can be formed in addition to lactate is also possible under certain conditions, and is dictated by oxygen concentration and carbohydrate availability. This exemplifies the metabolic adaptability of the pneumococcus, allowing it to thrive in distinct host niches (Al-Bayati et al., 2017; Echlin et al., 2020; Paixão et al., 2015b) (**Figure 1.3**).

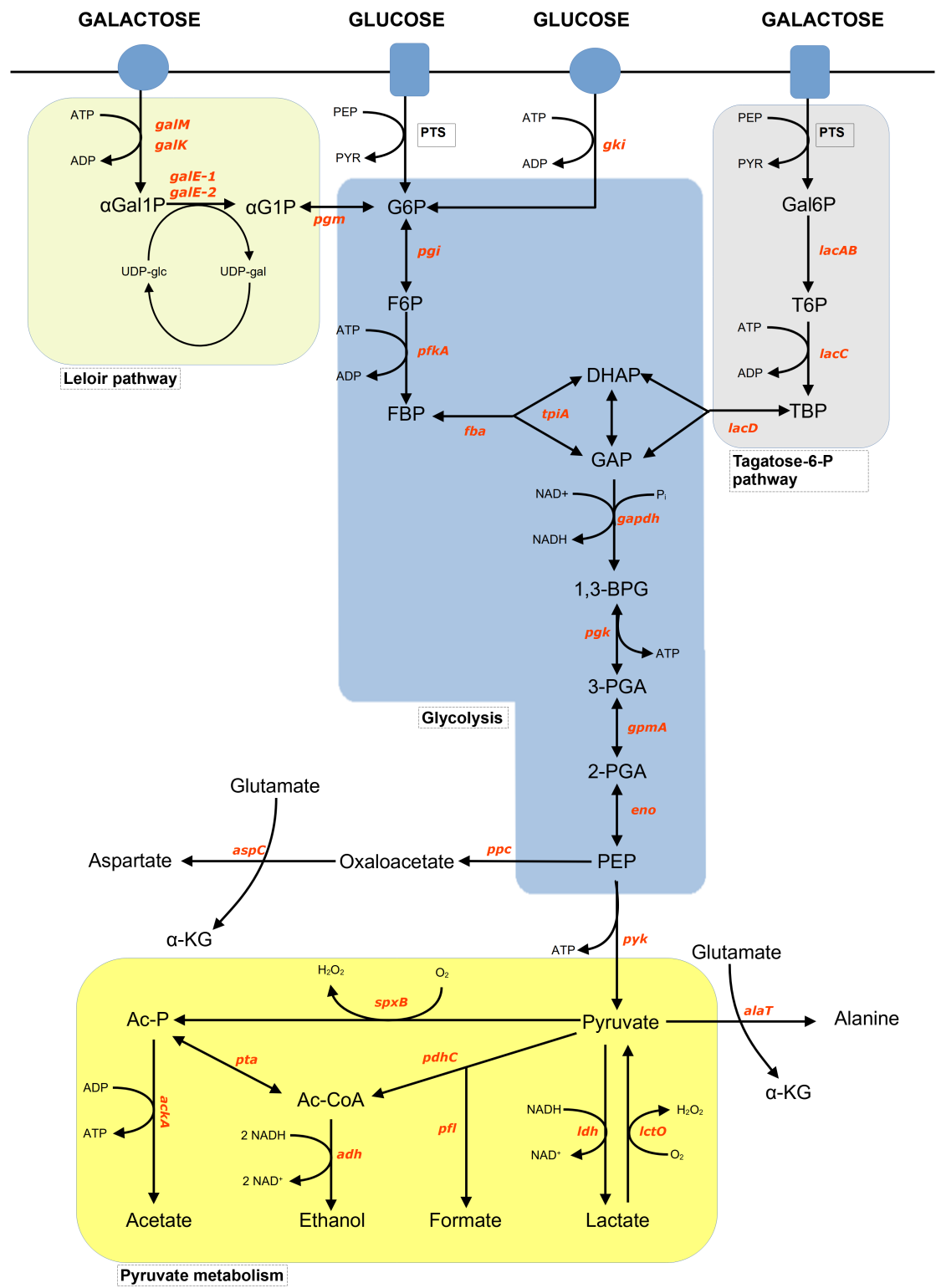


Figure 1.3 Central carbon metabolism of *S. pneumoniae*

Glycolysis is the main energy-generating pathway in *S. pneumoniae*. Blue circles on the cell membrane (solid black line at the top) show ABC transporters and blue rectangles indicate phosphotransferase systems (PTS). Galactose metabolism can be initiated via the Leloir pathway or the tagatose-6-phosphate pathway, before connecting to glycolysis. The pentose phosphate pathway, which generates precursors for purine and pyrimidine biosynthesis, is not shown here. Glycolysis is concluded by pyruvate metabolism. The ratio of fermentation products determines the fermentation profile – homolactic (during growth on glucose) or mixed acid (during growth on galactose). This diagram was generated with reference to genomic annotations and pneumococcal metabolism studies (Carvalho et al., 2011; Härtel et al., 2012; Paixão et al., 2015a; Tettelin et al., 2001).

1.3.2 Carbohydrate metabolism and virulence are connected through multifunctional genes

30% of the ~ 2.2 Mb *S. pneumoniae* genome is devoted to carbohydrate catabolism genes, and genome-wide studies to understand pathogenesis in both mouse models and *in vitro* assays have identified many of these genes to be directly linked to pneumococcal virulence (Aprianto et al., 2018; Hava & Camilli, 2002; Lau et al., 2001; Orihuela et al., 2004; Polissi et al., 1998) (**Table 1.1**). For instance, independent studies have shown that the cytoplasmic glycolytic enzymes GAPDH, phosphoglycerate kinase (*pgk*) and enolase also function as cell surface-associated host adhesion molecules, helping to establish mucosal colonisation and facilitate tissue invasion by degrading the extracellular matrix.

As another example, it has been shown that only mutants of galactose catabolism genes, and not mannose or *N*-acetylglucosamine (GlcNAc), were attenuated in a mouse model of *S. pneumoniae* D39 pneumococcal colonisation and pneumonia (Paixão et al., 2015b). *S. pneumoniae* grown on galactose generates a mixed-acid fermentation profile (i.e., end-products of glycolysis include substantial amounts of acetate and formate, along with lactate), diverging from the mainly homolactic profile when grown on glucose, mannose or GlcNAc. The re-routing of carbon flow towards mixed-acid fermentation is mediated in part by pyruvate formate lyase (*pfl*), which has also been associated with enhanced virulence in a mouse model (Al-Bayati et al., 2017; Yesilkaya et al., 2009). In *Streptococcus mutans*, *pfl* activity was inhibited during aerobic glucose culture, suggesting that the differential fermentation profiles may be observed in other streptococci as well (van Beelen, 1986).

Pyruvate oxidase (*spxB*) is another enzyme at the intersection of metabolic adaptation and multiple aspects of pathogenesis (Spellerberg et al., 1996). SpxB uses oxygen to catalyse the conversion of pyruvate to acetyl-phosphate (acetyl-P), generating H₂O₂ which contributes to lung damage and undermines competing microorganisms colonising the same niche. An acetate kinase can then convert acetyl-P to acetate, generating ATP (**Figure 1.3**). Environmental oxygen concentration decreases during progression of *S. pneumoniae* from nasopharyngeal coloniser to disease-causing

pathogen in the lung, blood and brain, and *spxB* may act as a sensor of environmental oxygenation that aids metabolic adaptation to the host niche (Carvalho et al., 2013a). In addition, *spxB* activity influences capsule production at the transcriptional (expression of the *cps* capsule gene locus) and post-transcriptional (controlling levels of acetyl-CoA for acetylated capsular sugars) levels (Echlin et al., 2016). The release of pneumolysin, a pore-forming toxin and critical virulence factor, is also modulated by *spxB* (Bryant et al., 2016).

The rapid adaptation of its metabolic machinery to diverse host microenvironments renders *S. pneumoniae* robust to the varied environmental and nutritional conditions it encounters (Carvalho et al., 2013b; Price et al., 2012). Master regulators such as CcpA and CodY are at the centre of this regulation, influencing multiple sets of genes and linking pneumococcal metabolism to virulence. Acting either as positive or negative regulators, they control the transcription of a range of virulence factors and sugar-specific operons, including ensuring the hierarchical use of preferential sugars (Carvalho et al., 2011; Iyer et al., 2005; Titgemeyer & Hillen, 2002).

Gene/ regulator	Involvement in pathogenesis	Experimental system	Reference
<i>Host glycoconjugate breakdown for nutrient sugar acquisition</i>			
NanA	Development of OM	Chinchilla; human middle ear effusions	(Diven et al., 1988; Tong et al., 2000; 2002)
	Triggers TGF β signalling to facilitate endothelial invasion	Mouse; hBMEC cell line	(Gratz et al., 2017)
	Undermines competing microbes	<i>In vitro</i> growth and Western blots	(Shakhnovich et al., 2002)
	Biofilm formation in nasal septa	Mouse	(Blanchette et al., 2016)
	Reveals ligands for adherence	Human airway cell lines	(King et al., 2006)
StrH	Limits complement deposition to aid evasion of opsonophagocytic killing	<i>Ex vivo</i> human neutrophil killing assay	(Dalia et al., 2010)
BgaA	Reveals ligands for adherence	Human airway cell lines	(King et al., 2006)
	Biofilm formation in nasal septa	Mouse	(Blanchette et al., 2016)
BgaC	NP survival	Mouse	(Terra et al., 2010)
EndoD	Survival in the blood and host mortality	Mouse	(Robb et al., 2017)
<i>β-glucoside processing for catabolism</i>			
BglA3	Binds host-derived extracellular RNA	A549 human airway cell line	(Zakrzewicz et al., 2016)
	Survival in the lung and host mortality	Mouse	(Terra et al., 2016)
	Attachment	Abiotic surface	

<i>α-galactoside metabolism & carbohydrate transport</i>			
Dldh	Mutant is avirulent in mice, and produces 50% less capsule <i>in vitro</i>	Mouse; ELISA	(Smith et al., 2002)
<i>Galactose metabolism (Leloir pathway)</i>			
GalU	Essential for capsule biosynthesis	<i>In silico</i> genome analysis	(Bonofiglio et al., 2012)
<i>Glycolysis</i>			
Fba	Adhesion to host cadherin receptors	A549 human airway cell line	(Blau et al., 2007)
GAPDH	Binds and sequesters human complement protein C1q	<i>In vitro</i> binding and complement activation	(Terrasse et al., 2012)
	Binds plasminogen to promote host cell adhesion and tissue degradation	A549 human airway cell line; EaHy vascular endothelial cell line	(Attali et al., 2008)
Enolase	Binds human complement inhibitor C4BP to evade complement	<i>In vitro</i> binding and functional assays	(Agarwal et al., 2012)
Pgk	Binds human tissue plasminogen activator to promote tissue invasion	<i>In vitro</i> binding and functional assays	(Fulde et al., 2014)
	Resists MAC-mediating killing to evade complement	<i>Ex vivo</i> and <i>in vitro</i> MAC deposition and complement activation	(Blom et al., 2014)
<i>Regenerates NAD^+ for glycolysis</i>			
Nox	Putative host cell adhesin	A549 human airway cell line	(Muchnik et al., 2013)
<i>Pyruvate metabolism</i>			
Pfl	Survival in NP, lung and blood	Mouse	(Echlin et al., 2020; Yesilkaya et al., 2009)
Pdh	Required for bacteraemia; capsule production	Mouse; Capsule blot assay and ELISA	(Echlin et al., 2020)

	Regulated at the colonisation/invasion interface	Experimental evolution of mouse NP colonisation	(Cooper et al., 2020)
AckA	Mutants are unstable and accumulate mutations in capsule gene loci	Microarray analysis	(Ramos-Montañez et al., 2010)
Ldh	Survival in the blood and lung	Mouse	(Gaspar et al., 2014)
	NP colonisation; reduced bacteraemia and capsule production	Mouse; Capsule blot assay and ELISA	(Echlin et al., 2020)
AdhE	NP survival and host mortality; increased intracellular pneumolysin; promotes inflammation	Mouse; Western blot; ELISA	(Luong et al., 2015)
SpxB	Biofilm formation in nasal septa	Mouse	(Blanchette-Cain et al., 2013)
	Contributes to pneumolysin release	Western blot, haemolysis assay, and gene expression analysis	(Bryant et al., 2016)
	Adheres to host cells	Human endothelial and epithelial cell lines	(Spellerberg et al., 1996)
	Survival in the blood and lung	Rabbit and mouse	(Echlin et al., 2020; Spellerberg et al., 1996)
	Strain-specific effects on capsule formation	Capsule blot and microscopy	(Echlin et al., 2016)
<i>Branched chain amino acid (BCAA) transporter</i>			
LivJHMGF	Survival in the lung and blood	Mouse	(Basavanna et al., 2009)
<i>Branched chain amino acid (BCAA) biosynthesis</i>			
ilvC	Expression of virulence factors <i>ply</i> and <i>lytA</i>	Gene expression	(Kim et al., 2017)
	NP colonisation and host mortality	Mouse	
<i>Regulator of sugar metabolism</i>			

CcpA	Survival in the NP, lung, and blood	Mouse	(Giammarinaro & Paton, 2002; Iyer et al., 2005)
	Transcription of the capsular locus	Long-range PCR	(Giammarinaro & Paton, 2002)
	Represses several metabolism genes also involved in pathogenesis	Transcriptomic and metabolomic analyses	(Carvalho et al., 2011)
<i>Global nutritional regulator</i>			
CodY	Required for colonisation	Mouse, D562 human pharyngeal cell line	(Hendriksen et al., 2008)
	Regulates virulence factors involved in adhesion, H ₂ O ₂ detoxification	DNA microarray, EMSA	(Hajaj et al., 2017; Hendriksen et al., 2008)
<i>Glutamine/ glutamate metabolism regulon</i>			
GlnR	Regulates genes involved in site-specific survival	Mouse	(Al-Bayati et al., 2017; Hendriksen, et al., 2008)
<i>Regulates sugar metabolism genes</i>			
LuxS/AI-2	Promotes hypervirulence	Mouse	(Trappetti et al., 2017)
	Role in competence and biofilm formation	<i>In vitro</i> functional assays	(Trappetti et al., 2011)
	Development of OM	Rat	(Yadav et al., 2018)
<i>Sucrose metabolism system</i>			
Scr	NP survival	Mouse	(Iyer & Camilli, 2007)
Sus	Survival in the lung	Mouse	

Table 1.1 Genes and regulators involved in metabolism and virulence of *S. pneumoniae*

Roles of genes in pathogenesis were determined using deletion mutants, unless stated otherwise.

Abbreviations: OM – otitis media; NP – nasopharynx; hBMEC – human brain microvascular endothelial cells; MAC – membrane attack complex; EMSA – electrophoretic mobility shift assay.

Gene names: NanA – neuraminidase A; StrH – β -*N*-acetylhexosaminidase; BgaA/ C – betagalactosidase A/ C; EndoD – endoglycosidase D; BglA3 – 6-phospho- β -glucosidase; Dldh – dihydrolipoamide dehydrogenase; GalU – UTP-glucose-1-phosphate uridylyltransferase; Fba – fructose-bisphosphate aldolase; GAPDH – glyceraldehyde-3-phosphate dehydrogenase; Pgg – phosphoglycerate kinase; Nox – NADH oxidase; Pfl – pyruvate formate lyase; Pdh – pyruvate dehydrogenase complex; AckA – acetate kinase A; Ldh – lactate dehydrogenase; AdhE – alcohol dehydrogenase E; SpxB – pyruvate oxidase; ilvC – ketol-acid reductoisomerase; Ply – pneumolysin; LytA – autolysin A; CcpA – catabolite control protein A.

1.3.3 Pneumococcal metabolic adaptation to host niches

Within the human host, *S. pneumoniae* encounters diverse microenvironments with changing nutrient availability, pH, osmolarity, and oxygen status (Willenborg & Goethe, 2016), and it is equipped with the sensing and regulatory systems to rapidly adapt its metabolism.

When colonising tissue surfaces in the nasopharynx, lower lung, middle ear and meninges, *S. pneumoniae* exists in a biofilm-like state. The switch to a planktonic physiological state to travel to distant sites after initial nasopharyngeal colonisation, followed again by biofilm formation to establish infection at new sites, are crucially aided by rapid bacterial metabolic adaptation (Allan et al., 2014). *S. pneumoniae* in biofilms are also known to be less sensitive to antimicrobial killing (Chao et al., 2015; Oggioni et al., 2006). Additionally, compared to planktonic pneumococci, biofilm growth is associated with decreased glycolytic metabolism (Allan et al., 2014). It may be hypothesised that decreased glycolysis may in part confer biofilm-associated pneumococci with enhanced antimicrobial resistance, as a study in *E. coli* has shown that the metabolic state of bacteria corresponds to antibiotic efficacy (Lopatkin et al., 2019).

S. pneumoniae can also undermine host immunity while simultaneously benefitting from it once it is physically in the lower respiratory tract, where it causes pneumonia. Type II alveolar cells synthesise surfactant using glycogen, which is an important process for lung health and innate immunity (Ridsdale & Post, 2004; Rooney, 2001). *S. pneumoniae* can migrate across the alveolar epithelium and metabolise glycogen inside type II alveolar epithelial cells for carbon, thus diminishing stores available for surfactant synthesis (Abbott et al., 2010; van Bueren et al., 2007).

A series of studies aiming to dissect niche adaptation of clonal pneumococcal strains isolated from two distinct infection sites – the blood and ear (Trappetti et al., 2013) – discovered that differential tissue tropism could be traced to a set of raffinose metabolism genes (Minhas et al., 2019). Further analysis showed that a SNP in the gene

rafR was responsible for clonally matched *S. pneumoniae* isolates exhibiting a preference for lung infection (local) over that of the ear/brain (systemic). This tropism is likely not directly linked to raffinose metabolism, which in any case is a plant polysaccharide (Hobbs et al., 2019). However, RafR regulates multiple genes involved in *S. pneumoniae* carbohydrate transport and metabolism, which are affected by the SNP in question, highlighting the inter-relatedness of carbohydrate metabolism and transcriptional responses required for niche adaptation. (Minhas et al., 2020).

1.3.4 Pneumococcal metabolomics – an emerging field

The description and regulation of metabolic pathways in *S. pneumoniae* so far has largely been based on genomic annotation of nutrient transporters and essential enzymes involved in metabolism. Large-scale transposon screens and models of infection aimed at understanding pneumococcal pathogenesis by identifying attributes of virulence have often implicated metabolism-associated genes. Some of these genes have been further functionally characterised using deletion mutants, providing invaluable insights into the multifunctionality of metabolic genes and intermediates and providing a novel way of looking at metabolism as a virulence attribute. Characterising the global pneumococcal metabolic network, then, seems to be a necessary requirement to further our understanding of pneumococcal adaptation to distinct host niches during disease progression and the impact of antibiotics and vaccines on this network. Here, recent advances in the techniques of metabolomics, coupled with traditional transcriptomic and biochemical approaches, present a relatively untapped opportunity.

Microbial metabolomics is a fast-rising field, especially in understanding interactions of the gut microbiota and environmental microbes. Metabolomics is also widely used to study prominent human pathogens such as *Mycobacterium*, *Salmonella*, and *Pseudomonas* species (de Carvalho et al., 2010; Eoh, 2014; Kohlstedt & Wittmann, 2019; Tang, 2011; Tounta et al., 2021; Wong et al., 2015; Zhang et al., 2020), but its application to *S. pneumoniae* has been relatively limited.

The framework for employing metabolic analytical techniques such as mass spectrometry (MS), nuclear magnetic resonance (NMR) and isotope tracing in *S. pneumoniae* is still being developed. The primary application of these techniques to *S. pneumoniae* has thus far been to measure concentrations of fermentation end-products during culture on different carbon sources and in different environmental oxygen tensions, and to measure the relative incorporation of carbohydrate-derived phosphates into capsular polysaccharides of different serotypes (Carvalho et al., 2013a; 2013b; Hathaway et al., 2012; Paixão, et al., 2015a; 2015b). Analyses of this kind demonstrate the quantitative potential of metabolomics techniques, but represent only a fraction of its capabilities.

In the only study of its kind, Härtel and colleagues used isotopically labelled glucose to specifically characterise amino acid metabolic pathways in *S. pneumoniae* (Härtel et al., 2012). No studies using stable isotope labelling in the pneumococcus have been published in the decade since. The first studies undertaking global pneumococcal metabolic profiling were published more recently, describing the effects of antibiotic stress on metabolism but looking only at metabolite concentrations (Leonard et al., 2018; 2020). Global metabolic characterisation, stable isotope labelling, and metabolic flux analysis are all techniques that remain underutilised to study pneumococcal adaptation and interactions with the host.

1.4 Metabolomics – an overview

Metabolomics is a branch of systems biology born of the technological advancements made in the post-genomic era. It describes the identification and quantification of small and medium-sized molecules involved in metabolic processes, providing a snapshot of the metabolic state of a closed system under specific conditions (Feist et al., 2008). Providing an unbiased and integrated readout of the genome, transcriptome and the environment, metabolomics can be used to study adaptations to perturbation – in pneumococcal metabolism, this can have implications in niche adaptation, disease-causing potential, antimicrobial resistance, and vaccine development.

1.4.1 Metabolomics techniques

Common metabolomics techniques are mass spectrometry (MS) and nuclear magnetic resonance (NMR). MS is usually coupled with a chromatographic separation technique, such as gas chromatography (GC) or liquid chromatography (LC). Samples generated for metabolomic analyses contain a wide range of compounds, making chromatographic separation critical for improved coverage as metabolites travel across the column at different speeds dictated by their physicochemical properties (e.g., mass, volatility, polarity). A chromatogram with peaks corresponding to specific column retention times (usually in minutes) of individual analytes is then generated. Separation of isomers, such as leucine/ isoleucine and glucose/ galactose, is also achieved by chromatography, which would not be detectable by MS alone (Jang et al., 2018).

Once a compound exits the chromatographic column, it undergoes ionisation to generate a unique fragmentation pattern. The fragments are detected by a mass spectrometer and resolved to generate a mass spectrum, which can then be compared with spectral libraries to aid compound identification. A key application of untargeted metabolomics in drug discovery is the identification of novel compounds, and in such cases, tandem mass spectrometry (MS/MS or MSⁿ) is used to infer chemical structures not determined in spectral libraries (Kind et al., 2018; Smith et al., 2005).

The oldest and most common type of NMR is 1-dimensional (1D) ^1H -NMR, which detects hydrogen atoms and so generates a spectrum of all molecules based on their chemical structure (Beckonert et al., 2007; Markley et al., 2017). Each of these metabolomics techniques identifies different types of metabolites, and an integrated approach is required for maximal metabolite coverage. 1D ^1H -NMR is particularly good at detecting small molecules, but 2D ^1H -NMR can be employed to resolve overlapping signals observed for larger molecules (Caceres-Cortes & Reily, 2010). NMR is also highly quantitative and non-destructive compared to MS techniques. MS techniques, on the other hand, are much more sensitive and can detect a broader range of metabolites. Sample preparation also differs for each technique. Notably, GC analysis requires chemical derivatisation of the sample to improve compound volatility as the mobile phase interacting with the column is an inert gas (nitrogen, hydrogen, or helium) (Lai & Fiehn, 2016).

1.4.2 Applications for microbial metabolomics

While the quantification of metabolites in a microbial system under certain conditions can be useful to develop a picture of the cellular metabolic state, metabolic pathways are in constant flux, and it is rarely especially useful to simply know which metabolites are present and at what concentrations. Therefore, metabolic flux determination using stable isotope labelling provides a more accurate view of physiological metabolism. An isotopically labelled nutrient source (e.g., carbohydrate or nitrogen source) can be traced through microbial metabolic pathways by analysing labelling patterns of metabolic intermediates and products at steady state (Chokkathukalam et al., 2014; Jang et al., 2018).

This can be used to determine how nutrient elemental flow is divided between energy-generation, anabolic metabolism, and other processes (such as towards capsule generation in *S. pneumoniae*), and how this flow is regulated upon perturbation. Studying isotope labelling can also identify novel metabolic pathways, especially where there is a dearth of genomic information, and identify key pathway intermediates for therapeutic targeting of pathogenic bacteria (Peyraud et al., 2009; Tounta et al., 2021).

Studying the intracellular and extracellular microbial metabolome in parallel can additionally help understand microbial interaction with the environment – Studies with respiratory and gut microbes have shown metabolic cooperation between commensal and pathogenic bacteria and described a role for secreted metabolic products in quorum-sensing and the regulation of immunity (Vernocchi et al., 2016; Wong et al., 2015). Metabolomic analyses have also delineated mechanisms of intracellular survival of pathogenic bacteria such as *M. tuberculosis* (Ehrt & Rhee, 2012; Eoh, 2014). Applying these techniques to host cell–*S. pneumoniae* interaction systems will be important for understanding metabolic adaptation in another set of niches, as evidence for intracellular survival of *S. pneumoniae* continues to emerge (Brissac et al., 2018; Brown et al., 2014; Ercoli et al., 2018; Marriott et al., 2012; Subramanian et al., 2019).

1.4.3 Targeting pneumococcal metabolism for therapy

Pneumococcal vaccines have been instrumental in reducing the burden of severe disease worldwide, but are an imperfect solution. As mentioned previously, existing pneumococcal vaccines are targeted to be effective against a subset of the approximately 100 serotypes of *S. pneumoniae* identified, and post-vaccination serotype replacement with non-vaccine serotypes has been extensively reported. This is partly driven by the high genetic recombination frequency of *S. pneumoniae*, and vaccine-induced metabolic shifts where non-vaccine serotypes acquire metabolism- and virulence-associated alleles from vaccine serotypes has been described (Watkins et al., 2015).

Pneumococcal antimicrobial resistance is also highly prevalent, with penicillin-non-susceptible *S. pneumoniae* featuring on the WHO Priority Pathogens List requiring the development of new antibiotics (WHO, 2021). Recent evidence has shown that some mechanisms of antibiotic resistance in *Escherichia coli* are driven by mutations in core metabolic genes (Lopatkin et al., 2021), highlighting an important aspect of bacterial metabolism in adaptive fitness to antibiotic selection pressure.

The essential nature of bacterial metabolism means that targeting metabolic genes or pathways can be an effective therapeutic strategy, and has been emphasised for the pneumococcus as well (Pedram et al., 2020). In *S. pneumoniae*, the glycolytic enzyme GAPDH has been proposed as a broad-spectrum vaccine target independent of pneumococcal serotype, with potential cross-reactive efficacy against *Listeria* and *Mycobacterium* as well (Alvarez-Dominguez et al., 2020; Ling et al., 2004; Teran-Navarro et al., 2021). These insights further underscore the value of detailed metabolomic studies not only to understand basic bacterial physiology and mechanisms of pneumococcal niche adaptation, but as a powerful means to identify targets for effective vaccine and antibiotic design (Schulz & Hammerschmidt, 2014; Tounta et al., 2021).

1.5 Thesis Aims

To successfully colonise the upper respiratory tract and travel to and cause disease in various host sites, *S. pneumoniae* must adapt its metabolism to the diverse range of microenvironments it encounters. Pneumococcal metabolism has largely been inferred from genomic analyses with often incomplete datasets and annotations, with specific metabolic genes highlighted in virulence studies being studied further. The main aim of this thesis was to characterise the central carbon metabolism of *S. pneumoniae* in an unbiased way by using cutting-edge stable isotope labelling metabolomics.

The following chapters assess the effect of niche-specific carbohydrate sources on carbon flow through catabolic and anabolic pathways. I also determined the effect of the outer polysaccharide capsule, a critical virulence factor and the biosynthesis of which is closely connected to carbohydrate metabolism, on the ability of *S. pneumoniae* to assimilate carbohydrate-derived carbons for energetic and biosynthetic metabolism.

The specific aims addressed in this thesis are as follows:

- Identify a chemically defined growth medium to culture *S. pneumoniae* for metabolomic studies.
- Establish robust methods for the extraction, processing, and analysis of *S. pneumoniae* by GC-MS.
- Compare the effects of two carbohydrate sources – glucose, which is abundant in the blood and CSF; and galactose, which is abundant in the upper respiratory tract – on the global intracellular and extracellular pneumococcal metabolome using GC-MS and ^1H -NMR.
- Determine the effect of the capsule on carbon flow through pneumococcal metabolic pathways using ^{13}C stable isotope labelling and capsule-deficient mutant strains.
- Validate results using two different pneumococcal serotypes with distinct capsule compositions and carriage prevalence – *S. pneumoniae* TIGR4 and 23F.

Chapter 2. Materials & Methods

All reagents were purchased at > 98 % purity grade from Sigma-Aldrich (Merck KGaA, Germany) or Alfa Aesar (Fisher Scientific, US), unless stated otherwise. Wherever possible, colour blind-safe hues were used for graphs and figures throughout this thesis.

2.1 Bacterial strains, stock preparation and storage

S. pneumoniae strains used in this study are wild-type TIGR4 (or T4) and 23F, and their otherwise isogenic mutant derivatives lacking the capsule biosynthesis gene locus (Δcps).

All strains were obtained from Jeremy Brown, UCL. Wild-type TIGR4 (serotype 4 clinical isolate P1542 (Hyams et al., 2010; Tettelin et al., 2001)) and TIGR4 Δcps (P1672 (Trzciński et al., 2003)) were kind gifts from Jeffrey Weiser, University of Pennsylvania. Wild-type 23F (ST36 clinical isolate OXC-1417-23F) was a kind gift from Brian Spratt, Imperial College London. A 23F Δcps mutant was constructed in the ST36 background by insertion of the Janus cassette to replace the *cps* locus, as previously described (Trzciński et al., 2003), and was a kind gift from Jeremy Brown, UCL (Hyams et al., 2011).

Working stocks of all strains were prepared in a Biosafety Level 2 laboratory as follows: Cells scraped from a frozen bacterial stock cryovial were streaked on blood agar plates (Columbia agar (Oxoid) supplemented with 5% (v/v) defibrinated horse blood (E&O Laboratories)) and incubated at 37°C/ 5% CO₂ overnight. The lawn of colony-forming units (CFUs) was cultured in 20-25 ml THY (Todd-Howitt broth + 0.5% (v/v) yeast extract; Oxoid) and incubated at 37°C/ 5% CO₂. When the cultures reached exponential growth phase as measured by optical density at 600 nm (OD₆₀₀ 0.4 – 0.6), glycerol was added to a final concentration of 10% (v/v), and single-use aliquots of 1 ml stored in cryovials at –80°C.

Frozen glycerol stocks were tested after 3-6 days of preparation – Bacterial growth was assessed in 25 ml THY cultures by measuring OD₆₀₀ over time, and the colony morphology of scraped stocks on blood agar plates incubated at 37°C/ 5% CO₂ overnight was inspected for microbiological ascertainment of pneumococcal colonies.

2.1.1 Enumerating colony-forming units (CFUs)

Colony-forming units (CFUs) of bacterial stocks were enumerated by making 10-fold serial dilutions of 10 µl culture in 90 µl PBS (in a 96-well plate), and plating 10 µl of each dilution onto blood agar plates. After overnight incubation at 37°C/ 5% CO₂, viable CFUs were counted manually to calculate CFU ml⁻¹.

2.1.2 Microbiological identification of a TIGR4 mucoid colony variant

Stock cultures of *S. pneumoniae* TIGR4 incubated on blood agar plates showed that ~10% of the CFUs were morphologically distinct – larger and with a mucoidal appearance (**Figure 2.1**). To confirm that these were *S. pneumoniae* colonies, a series of established microbiological tests for the identification of *S. pneumoniae* were performed on both the mucoid and normal colony variants.

Gram staining was performed using the Gram Staining Kit (Pro-Lab Diagnostics) and associated standard protocol. Catalase test was carried out by mixing a single bacterial colony with 3% hydrogen peroxide on a microscope slide. Catalase negativity is determined by a lack of oxygen bubbles (product of hydrogen peroxide breakdown by catalase). Optochin sensitivity was determined by placing an optochin disc (5 µg; Oxoid) on a blood agar plate streaked with bacterial colonies and determining the zone of inhibition after overnight incubation at 37°C/ 5% CO₂. Bile solubility was tested by suspending bacterial colonies in a 2% (w/v) sodium deoxycholate solution to obtain a turbid suspension (OD₆₀₀ ~1.0). The suspension was incubated at 37°C/ 5% CO₂ for 10 min, and clearance of turbidity measured by OD₆₀₀.

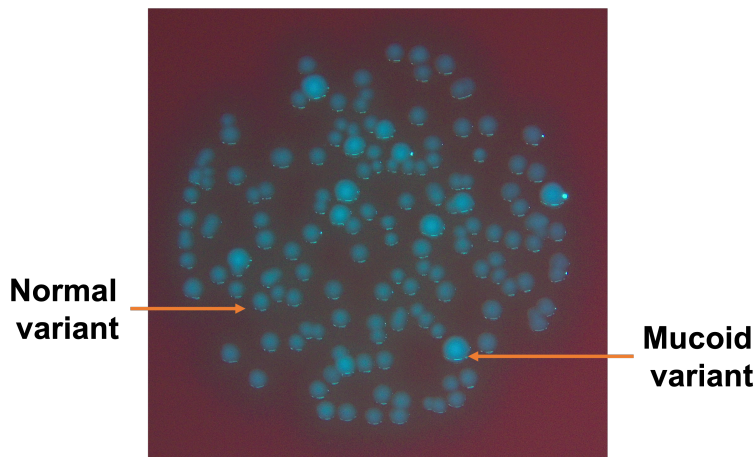


Figure 2.1 *S. pneumoniae* TIGR4 colony variants

S. pneumoniae TIGR4 on a blood agar plate. Mucoid colony variant constituted ~10% of all colony-forming units (CFUs). Photographed using the IC90 E CMOS microscope camera mounted on a M125 C stereo microscope, viewed through a PLANAPO 1.0 x objective lens (Leica).

2.1.3 *S. pneumoniae* TIGR4 whole genome sequencing

An overnight culture (37°C/ 5% CO₂) of TIGR4 in 5 ml THY was centrifuged (3500 rpm, 5 min; Thermo Scientific TX-1000 swing-out rotor). The bacterial pellet was pre-treated with 200 µl EDTA (50 mM) + 0.1% (v/v) sodium deoxycholate and incubated at 37°C for 10 min. DNA extraction was carried out using the Wizard® Genomic DNA Purification Kit (Promega).

Purified genomic DNA was then processed by the Bioinformatics and Biostatistics (BABS) facility at the Francis Crick Institute. Libraries were prepared with the Nextera XT DNA Library Preparation Kit (Illumina). 100 bp paired-end sequencing was then carried out on the HiSeq4000 system (Illumina) at a depth of 2 million reads per sample. Sequencing quality was checked using MULTIQC (Ewels et al., 2016).

Reads were aligned to the TIGR4 reference genome AE005672.3 (Tettelin et al., 2001) using the Burrows-Wheeler Aligner (BWA) (Li & Durbin, 2010). Any low-quality reads were trimmed, and duplicate reads were removed (deduplication). VarScan (Koboldt et al., 2012), BCFtools (Li et al., 2009) and the Integrative Genomics Viewer

(Robinson et al., 2011) were used to detect and visualise genomic variation in the samples compared to the reference. The metagenomic classification platform Kraken (Wood & Salzberg, 2014) was also used to map reads to a database of organisms in an unbiased way, which was done with help from Dr. Andrea Gori, Mucosal Pathogens Research Unit, UCL.

2.2 *S. pneumoniae* growth studies in defined culture media

Commercially available *S. pneumoniae* culture media include THY and brain heart infusion (BHI) broth, which are easy to prepare and able to support excellent pneumococcal growth. However, these media formulations contain compositionally complex substances such as beef heart and calf brain infusions, and therefore lack complete information about individual nutrients (amino acids, vitamins, bases, sugars, minerals).

For accurate interpretation of metabolism studies, it was imperative to culture *S. pneumoniae* in defined growth medium where each component and its concentration were known and could be controlled. I also wanted to compare *S. pneumoniae* metabolism when using either glucose or galactose as the sole carbon source, and formulations of both THY and BHI already contain glucose. Therefore, *S. pneumoniae* growth was tested in three defined culture media formulations – Sicard’s medium (Section 2.2.1), 7H9⁺ medium (Section 2.2.2), and CDM (chemically defined medium) (Section 2.2.3).

2.2.1 Sicard’s medium (semi-defined)

Sicard’s medium (Table 2.1) was originally developed to culture *S. pneumoniae* for transformation studies (Sicard, 1964). More recently, Sicard’s medium supplemented with either glucose (Terra et al., 2016) or mucin (Yesilkaya et al., 2008) as the sole carbon source was shown to support pneumococcal growth *in vitro*. A modified version of Sicard’s medium was also used to mimic the composition of different *S. pneumoniae* niches in the human host (Aprianto et al., 2018).

<i>Minerals and buffer</i>	<i>g L⁻¹</i>	<i>Amino acids</i>	<i>mg L⁻¹</i>
NaCl	5	L-Arginine	200
NH ₄ Cl	2	L-Asparagine	10
KCl	0.4	L-Cysteine HCl	100
Na ₂ HPO ₄	0.12	L-Glutamine	20
MgSO ₄	0.024	Glycine	120
CaCl ₂	0.01	L-Histidine	150
FeSO ₄ ·7H ₂ O	0.0055	L-Isoleucine	6.55
Tris(hydroxymethyl) aminomethane	4.84	L-Leucine	6.55
		L-Lysine	420
		L-Methionine	180
		L-Threonine	175
		L-Valine	5.85
<i>Vitamins</i>	<i>mg L⁻¹</i>	<i>Other</i>	<i>mg L⁻¹</i>
Biotin	0.015	Sodium pyruvate	800
Choline chloride	5	Uracil	1
Nicotinamide	0.6	BSA	160
Calcium pantothenate	2.4	Carbon source*	(varies)
Pyridoxal HCl	0.6		
(-)Riboflavin	0.3		
Thiamine HCl	0.6		

Table 2.1 Composition of Sicard's medium

Adapted from (Sicard, 1964).

* Carbon sources used were D-(+)-glucose, D-(+)-galactose and porcine gastric mucin.

100x stock of *Minerals and buffer* was made and stored at 4°C. 100x stock containing vitamins, cysteine, glutamine, and pyruvate was prepared, filter-sterilised (0.22 µm membrane), and stored in aliquots at -80°C. Fresh Sicard's medium was prepared before each use by combining appropriate volumes of the above solutions and adding the remaining amino acids, along with uracil, bovine serum albumin (BSA) and carbon source. The final Sicard's medium was filter-sterilised through a 0.22 µm membrane.

2.2.1.1 Mucin preparation

Porcine gastric mucin (PGM type III; Sigma-Aldrich) was dissolved in distilled water at 10 mg ml⁻¹ and dialysed against distilled water 300-350x the volume of the solution at 4°C, using SnakeSkin™ dialysis tubing (10kDa MWCO; Thermo Scientific). The PGM solution was distributed into 2 ml aliquots, lyophilised on a FreeZone 4.5 Freeze Dry System (Labconco) for 24 h and dissolved in Middlebrook 7H9 Broth Base (Sigma-Aldrich) at 25 mg ml⁻¹. This was sterilised by autoclaving at 121°C for 15 min and added to the culture medium to a final concentration of 10 mg ml⁻¹ (1% w/v).

2.2.2 7H9⁺ medium (semi-defined)

7H9⁺ medium is an adaptation of Sicard's medium (**Section 2.2.1**). The *Minerals & buffer* component from Sicard's was replaced with the dehydrated, commercially available Middlebrook 7H9 Broth Base (Sigma-Aldrich), and then supplemented with *Vitamins* and *Amino acids* from the Sicard's medium recipe. The complete composition is shown in **Table 2.2**.

7H9⁺ medium was originally created because I found that it supports the growth of both *Streptococcus pneumoniae* and *Mycobacterium bovis*, and we hope to investigate the metabolic interactions between these pathogenic respiratory bacteria in a future project.

7H9 Broth Base[^]	g L⁻¹	Amino acids	mg L⁻¹
(NH ₄) ₂ SO ₄	0.5	L-Arginine	200
Na ₂ HPO ₄	2.5	L-Asparagine	10
KH ₂ PO ₄	1.0	L-Cysteine HCl	100
Sodium citrate	0.1	L-Glutamine	20
MgSO ₄	0.05	Glycine	120
CaCl ₂	0.0005	L-Histidine	150
ZnSO ₄	0.001	L-Isoleucine	6.55
CuSO ₄	0.001	L-Leucine	6.55
Ferric ammonium citrate	0.04	L-Lysine	420
L-Glutamic acid [^]	0.5	L-Methionine	180
Pyridoxine	0.001	L-Threonine	175
Biotin	0.0005	L-Valine	5.85
Vitamins	mg L⁻¹	Other	mg L⁻¹
Choline chloride	5	Sodium pyruvate	800
Nicotinamide	0.6	Uracil	1
Calcium pantothenate	2.4	BSA (fatty acid-free)	5000
Pyridoxal HCl	0.6	Catalase (bovine liver)	3
(-)Riboflavin	0.3	Carbon source [*]	(varies)
Thiamine HCl	0.6		

Table 2.2 Composition of 7H9⁺ medium

[^] 7H9 Broth Base purchased from Sigma-Aldrich.

[^] Note: L-Glutamic acid is an amino acid listed separately.

^{*} Carbon sources used were D-(+)-glucose, D-(+)-galactose and porcine gastric mucin. Concentrations of vitamins, amino acids, uracil and pyruvate were informed by Sicard's medium (Sicard, 1964) (Table 2.1).

100x stock containing vitamins, cysteine, glutamine and pyruvate was prepared, filter-sterilised (0.22 µm membrane), and stored in aliquots at -80°C. The remaining amino acids and uracil were added to a solution of 7H9 Broth Base and stored at 4°C for no longer than 1 week. Stock solutions of fatty acid-free BSA in water (10% w/v) and catalase in 50 mM potassium phosphate buffer (0.1% w/v) were prepared fresh.

Appropriate volumes were combined to prepare 7H9⁺ medium, supplemented with the carbon source, and filter-sterilised (0.22 µm membrane) before use.

2.2.3 Chemically Defined Medium (CDM)

The composition of CDM is shown in **Table 2.3**, modified from a published study (van de Rijn & Kessler, 1980). Although originally developed for culturing Group A streptococci, this CDM formulation has also been adapted to culture *S. pneumoniae* for isotope-labelled metabolomics (Härtel et al., 2012) and metabolism-associated studies (Carvalho et al., 2013a; 2013b; Paixão et al., 2015a; 2015b).

Stocks of CDM components were prepared as described in Protocol S1 of a published study (Chang et al., 2011). The appropriate carbon source was added before use, and CDM filter-sterilised (0.22 µm membrane).

Minerals and buffer	g L⁻¹	Amino acids	mg L⁻¹
K ₂ HPO ₄	0.2	L-Alanine	100
KH ₂ PO ₄	1	L-Arginine	100
NaH ₂ PO ₄ ·2H ₂ O	3.62	L-Aspartic acid	100
Na ₂ HPO ₄	7.3	L-Asparagine	100
MgSO ₄	0.34	L-Cystine	50
MnSO ₄ ·4H ₂ O	0.007	L-Cysteine HCl	650
NaC ₂ H ₃ O ₂	2.7	L-Glutamic acid	100
NaHCO ₃	2.5	L-Glutamine	200
CaCl ₂	0.005	Glycine	100
Fe(NO ₃) ₂ ·9H ₂ O	0.001	L-Histidine	100
FeSO ₄ ·7H ₂ O	0.005	L-Isoleucine	100
Vitamins	mg L⁻¹	L-Leucine	100
<i>p</i> -Aminobenzoic acid	0.2	L-Lysine	100
Biotin	0.2	L-Methionine	100
Choline chloride	1000	L-Phenylalanine	100
Folic acid	0.8	L-Proline	100
Nicotinamide	1	Hydroxy-L-proline	100
β-NADP	2.5	L-Threonine	200
Ca ²⁺ pantothenate	2	L-Tryptophan	100
Pyridoxal HCl	1	L-Tyrosine	100
Pyridoxamine 2HCl	1	L-Valine	100
(-)-Riboflavin	2	Bases	mg L⁻¹
Thiamine HCl	1	Adenine	20
Vitamin B12	0.1	Guanine HCl	20
Other	mg L⁻¹	Uracil	20
Carbon source*	(varies)		

Table 2.3 Composition of CDM

Original formulation (van de Rijn & Kessler, 1980) adapted and prepared as described by Chang *et al.* (Chang *et al.*, 2011).

* Carbon sources used were D-(+)-glucose, D-(+)-galactose, ¹³C-glucose, and ¹³C-galactose.

2.2.4 *S. pneumoniae* growth curves

Multiple conditions for *S. pneumoniae* culture were tested to determine optimal growth in defined culture media, which will be discussed in **Chapter 3**. In general, 25 ml of culture medium was inoculated with 250-500 μ l of frozen bacterial stock and incubated at 37°C/ 5% CO₂ without agitation with slightly loosened tube caps (microaerophilic conditions). Growth was monitored over time by measuring the OD₆₀₀ using a spectrophotometer, and CFUs were enumerated by plating serial dilutions on blood agar (see **Section 2.1.1**).

Growth curves were also generated using the Spark® multimode plate reader (Tecan). 200 μ l of cultures incubated with bacteria as above were added to wells of a 96-well plate (Greiner, U-bottom) and covered with a removable lid. The plate reader was set to 37°C/ 5% CO₂. Every 30 min for 24 h, it was programmed to shake for 3 s and measure OD₆₀₀.

Specific growth rates (μ) and lag phase duration were calculated by linear regression analysis of ln(OD₆₀₀) vs. time plots during exponential phase of growth using GraphPad Prism (Carvalho et al., 2013b).

2.3 Conservation of metabolic genes across *S. pneumoniae* isolates using nucleotide sequence similarity analysis

S. pneumoniae is a highly variable organism. Its natural competence facilitates frequent recombination leading to significant genetic variability among different pneumococcal isolates (Croucher et al., 2011). *S. pneumoniae* also displays diversity in non-antigenic genes, and is classified into multilocus sequence types (MLSTs) by seven metabolic housekeeping genes. Here, I looked at the degree of conservation of metabolism-associated genes across pneumococcal isolates to determine whether using a single strain of *S. pneumoniae* in pilot metabolomics studies would be broadly representative of the species.

2.3.1 Selection of *S. pneumoniae* isolates

A diverse, representative set of *S. pneumoniae* isolates were selected from the web-accessible database PubMLST¹ (Jolley et al., 2018).

With the genomes of < 20 isolates being fully sequenced and assembled, partially sequenced isolates were also included in the analysis. These were sorted according to contig number², N50 contig number³ and N50 contig length⁴, to obtain a list of the top 500 isolates with the highest proportion of their genome sequenced and assembled.

Of these 500, 120 isolates (*Appendix Table 7.1*) were included in the analysis, chosen to represent the greatest diversity – in serotypes, MLSTs (**Table 2.4**), geographical location of isolation, and year of isolation (**Figure 2.2**).

¹ <https://pubmlst.org/>

² One contig ideal, as complete sequence available.

³ N50 contig number = **position** of the shortest contig in a set that together represents 50% of the total assembly length, when all contigs in the assembly are aligned in descending order of length. Smaller the N50 contig number, the more ‘complete’ the assembly.

⁴ N50 contig length = **length** of the shortest contig that encompasses 50% of the total length, when all contigs in the assembly are aligned in descending order of length. Larger the N50 contig length, the more ‘complete’ the assembly.

	<i>Isolates</i>	<i>Serotypes</i>	<i>MLSTs</i>
<i>Vaccine type</i>	66	24	51
<i>Non-vaccine type</i>	54	36	47
<i>Total</i>	120	60	98

Table 2.4 Summary of *S. pneumoniae* isolates included in similarity analysis

Vaccine type isolates include serotypes found in the pneumococcal conjugate vaccines (PCV) 10 and 13, and the pneumococcal polysaccharide vaccine (PPSV) 23.

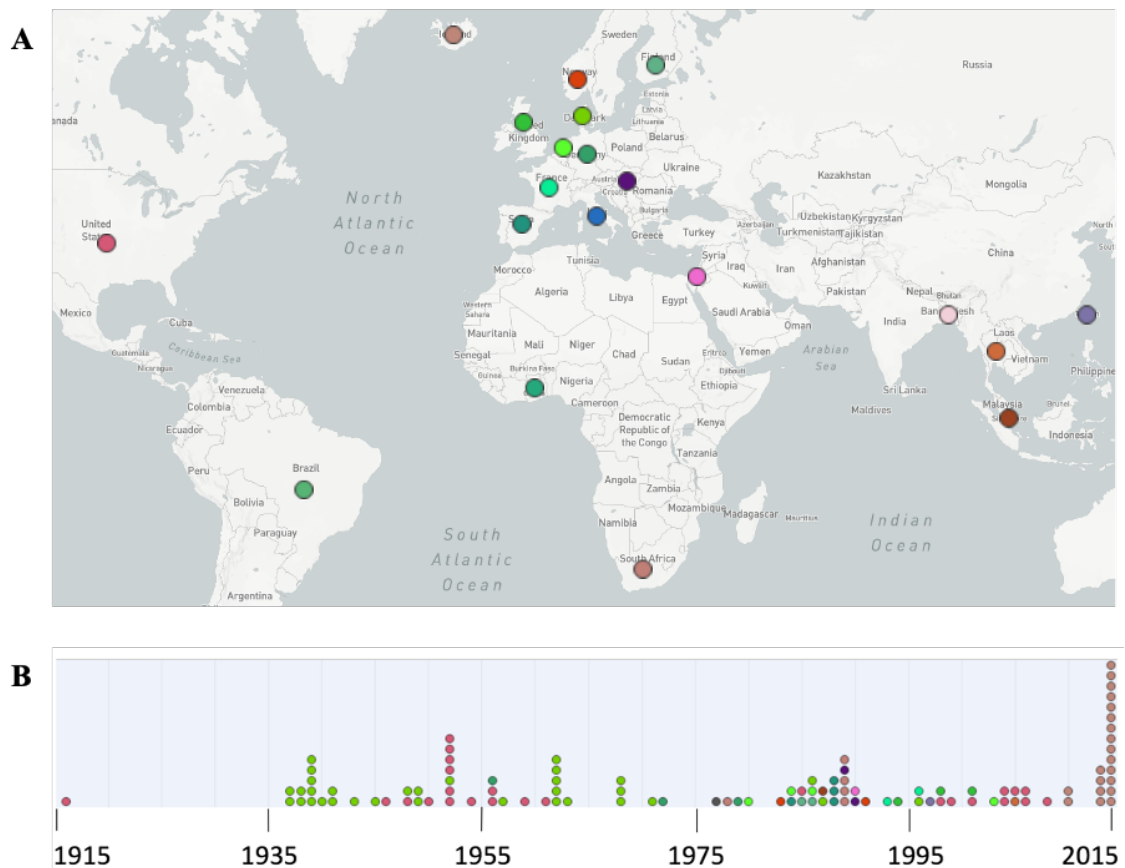


Figure 2.2 Characteristics of *S. pneumoniae* isolates included in similarity analysis

Location (A) and year (B) of isolation. Each dot signifies an individual isolate, and the colour corresponds to country of isolation. Figures generated in Microreact⁵ (Argimón et al., 2016).

⁵ <https://microreact.org/showcase>

2.3.2 Selection of metabolic genes to include in analysis

On the KEGG database⁶ (Kanehisa & Goto, 2000), genes annotated as being involved in ‘carbon metabolism’, ‘glycolysis/ gluconeogenesis’, ‘pyruvate metabolism’, ‘pentose phosphate pathway’ and ‘galactose metabolism’ in *S. pneumoniae* TIGR4 were chosen. There was some overlap in genes between pathways, and the final analysis included 95 genes (n = 109). The complete list of genes can be found in *Appendix Table 7.2*.

2.3.3 Nucleotide sequence similarity analysis

Using the Basic Local Alignment Search Tool (BLAST) function on PubMLST, the presence or absence of each metabolic gene sequence from the TIGR4 genome in the selected isolates was determined. Genes with < 40% similarity in an isolate were considered to be absent.

⁶ <https://www.genome.jp/kegg/pathway.html>

2.4 Metabolomics – General Methods

To avoid chemicals leaching into solvents and contaminating mass spectrometric results, and where glassware could not be used, only Eppendorf® brand plasticware (Safe-Lock tubes and pipette tips) was used for metabolomic sample preparation (Chanotiya et al., n.d.). LC-MS Grade solvents and MilliQ water were used, unless stated otherwise. Samples were kept at 4°C as much as possible during processing, including reagents to be added and any empty tubes. Internal standards and a mix of polar metabolites at known concentrations (metabolite mix, MM), to aid the identification and quantification of metabolites, were prepared by the Metabolomics facility at the Francis Crick Institute.

2.4.1 Gas chromatography–mass spectrometry (GC-MS)

This method has been adapted from that described by MacRae *et al.* (MacRae et al., 2013). Details of bacterial culture and metabolite extraction are described in **Sections 2.5 and 2.6**. Biphasic partitioned samples in a mix of 50 µl chloroform, 150 µl methanol and 150 µl water, containing internal standards – 1 nmol *scyllo*-inositol and/ or 5 nmol norleucine – were vortexed and centrifuged (17,000 x g, 10 min, 4°C) to concentrate protein and debris at the interface between the upper polar (methanol/ water) phase and the lower non-polar (chloroform) phase. The non-polar phase was discarded. The polar phase was transferred to GC vial inserts (250 µl, Agilent), placed in 1.5 ml tubes and dried in a rotary vacuum concentrator (SpeedVac™; Thermo Scientific). The samples were then washed twice with 30 µl methanol and dried after each wash.

Derivatisation of polar metabolites is important to increase compound volatility for GC-MS analysis and improve chromatographic separation. Here, samples were first derivatised with 20 µl of 20 mg ml⁻¹ methoxyamine-HCl in pyridine and left at room temperature overnight (carbonyl functional groups converted to oxime groups). A second silylation derivatisation was then carried out with 20 µl N,O-bis(trimethylsilyl)trifluoroacetamide + 1% trimethylchlorosilane (BSTFA + TMCS), and left for at least 1 hour at room temperature (hydroxyl functional groups converted to

trimethylsilyl). QC samples (metabolite mix; MM) were processed alongside biological samples.

1 µl of each sample was injected (splitless injection, helium gas carrier) in an automated, randomised sequence on the 7890A GC system connected either to a 5975C or 7000C MSD triple-axis detector in electron impact (EI) ionisation mode (Agilent Technologies). Chromatographic separation was performed on a 30 m x 0.25 mm DB-5MS + 10 m DuraGuard capillary column (Agilent).

2.4.2 Liquid chromatography–mass spectrometry (LC-MS)

Liquid chromatography–mass spectrometry (LC-MS) was employed to test various methods to extract intracellular metabolites from *S. pneumoniae*. LC-MS can detect a broader range of metabolic features (100s-1000s) compared to GC-MS (10s-100s), and would provide more information to generate coefficient of variation plots (similar to PCA plots) in order to compare extraction techniques.

25 ml THY was inoculated with 250 µl *S. pneumoniae* TIGR4 stock ($\sim 2 \times 10^7$ CFU) and cultured for 4 – 5 h (OD 0.9 – 1.0) at 37°C/ 5% CO₂. 2 replicate cultures (equivalent to n = 6 units of biomass as described in Leonard *et al.* (Leonard et al., 2018)) were combined in a sterile flask. Bacterial metabolism was quenched by rapid cooling – open culture flasks were placed in a slurry of dry ice and ethanol, and the cultures agitated rapidly with a clean thermometer to $\sim 13^\circ\text{C}$. They were then immediately transferred to an ice/ water bath.

Quenched cultures were divided into 6x 25 ml in 50 ml Falcon tubes, centrifuged (2000 rpm, 25 min, 4°C; Thermo Scientific TX-1000 swing-out rotor), and the supernatant discarded. Bacterial pellets were resuspended in 1 ml ice-cold PBS, transferred to chilled 2 ml tubes, and centrifuged (13,000 x g, 5 min, 4°C). Supernatants were discarded, and the PBS wash was repeated. Samples were centrifuged finally for 1 min (13,000 x g, 4°C) to remove any residual PBS.

Different methods for *S. pneumoniae* intracellular metabolite extraction were tested (discussed in **Chapter 5.3.1** and summarised in **Table 2.5**) and analysed by LC-MS. 50% of the polar phase extracts were transferred to GC vial inserts (250 µl, Agilent) placed inside GC vials (2 ml, Agilent) and dried in a SpeedVac™ rotary evaporator. Dried samples were resuspended in a mix of 25 µl methanol and 25 µl ultrapure water containing 5 µM L-valine-¹³C₅, ¹⁵N as the internal standard. 5 µl from each sample was combined to create a pooled biological QC (PBQC) sample.

2.4.2.1 Data collection and analysis

10 µl of each sample was injected in an automated, randomised sequence on the Vanquish UHPLC system coupled to a Q Exactive Plus Orbitrap mass spectrometer (Thermo Fisher), with a heated electrospray ionisation (HESI-II) probe and positive/negative polarity switching. Chromatographic separation was performed on a SeQuant® Zic-pHILIC (Merck Millipore) column (5 µm particle size, polymeric, 150 x 4.6 mm).

Data was visualised with Xcalibur Qual Browser (Thermo Fisher). Metabolites were identified and quantified by accurate mass and retention time and by comparison to the retention times, mass spectra, and responses of known amounts of authentic standards using TraceFinder software (Thermo Fisher). Progenesis QI (Waters) software was used to compare extraction methods via unsupervised analysis of all metabolite features. Coefficient of variation ($CV = \text{standard deviation} / \text{mean}$ of $n = 6$) plots were generated using GraphPad Prism.









<i>Method</i>	<i>Extraction solvent</i>	<i>Cell lysis</i>	<i>Drying*</i>	<i>Re-extraction (300 µl)</i>	<i>Partitioning (350 µl)</i>
1	400 µl CHCl ₃ + 200 µl MeOH	Pulse sonication 3x 	Spin + dry	2:1 MeOH:H ₂ O	1:3:3 CHCl ₃ :MeOH:H ₂ O
2	450 µl ACN	Pulse sonication 3x 	Spin + dry		1:3:3 CHCl ₃ :MeOH:H ₂ O
3	400 µl CHCl ₃ → 1x sonication → 200 µl MeOH → 1x sonication		Spin + dry	2:1 MeOH:H ₂ O	1:3:3 CHCl ₃ :MeOH:H ₂ O
4	400 µl CHCl ₃ + 200 µl MeOH	Bead homogenisation 	Spin + dry	2:1 MeOH:H ₂ O	1:3:3 CHCl ₃ :MeOH:H ₂ O
5	450 µl ACN	Bead homogenisation 	Spin + dry		1:3:3 CHCl ₃ :MeOH:H ₂ O
6	400 µl CHCl ₃ + 200 µl MeOH	Freeze-thaw 3x 	Spin + dry	2:1 MeOH:H ₂ O	1:3:3 CHCl ₃ :MeOH:H ₂ O
7	450 µl ACN	Freeze thaw 3x 	Spin + dry		1:3:3 CHCl ₃ :MeOH:H ₂ O
8	400 µl MeOH + 400 µl ACN + 200 µl H ₂ O	Bead homogenisation 	Spin + filter* + dry		

Table 2.5 Intracellular metabolite extraction methods

Pulse sonication – 3x 8 min pulses with 12 min rest in between, in water bath sonicator at 4°C.

Bead homogenisation – Extracts transferred to 2 ml screw-top vials (Starstedt) containing 500 µl acid-washed glass beads (150-212 µm, Sigma-Aldrich) and homogenised in a FastPrep-24 homogeniser (MP Biomedicals) using 2x cycles of 40 s sonication at 6 m/s + 300 s rest (4°C).

Freeze-thaw – Tubes dropped into liquid N₂ for 3 min, then thawed at room temperature for 5 min.

* Centrifuged at 13,000 x g for 10 min at 4°C and dried in a SpeedVac™ rotary evaporator.

* Filtered in Spin-X tubes (0.2 µm filter, Costar) by centrifugation at 17,000 x g for 20 min at 4°C.

ACN = acetonitrile; CHCl₃ = chloroform; MeOH = methanol.

Icons were adapted from The Noun Project (<https://thenounproject.com/>).

2.5 Semi-targeted extracellular metabolite profiling (footprinting)

25 ml 7H9⁺ culture medium supplemented with either 11 mM glucose or 11 mM galactose (warmed at 37°C overnight) was inoculated with 250 µl (~ 2-3x10⁷ CFU) *S. pneumoniae* stock and incubated at 37°C/ 5% CO₂. Every 2 hours for 10 to 14 h, 100 µl of the culture was collected and spun (13,000 x g, 10 min, 4°C) to pellet bacteria. The supernatant was collected and stored at –80°C. In parallel for each time point, absorbance (OD₆₀₀) was also measured, and serial dilutions plated on blood agar for CFU determination (see **Section 2.1.1**).

5 µl of the stored supernatant was added to 50 µl chloroform in 1.5 ml tubes. To this, a 295 µl mix of 1:1 (v/v) methanol and water, containing 1 nmol *scyllo*-inositol (internal standard), was added. The now biphasic partitioned samples were vortexed thoroughly and stored at –80°C until further processing for GC-MS, which is detailed in **Section 2.4.1** Gas chromatography–mass spectrometry (GC-MS).

2.5.1 Data analysis

Data was analysed with a combination of programs – MassHunter Qualitative Analysis software and Quantitative Analysis software (Agilent), and the MATLAB (Mathworks) script GAVIN (Behrends et al., 2011). Metabolites were identified by comparing chromatographic retention times and accurate mass against the retention times, mass spectra, and responses of known amounts of authentic standards (a pre-constructed metabolite mix (MM) of ~50 polar metabolites). Metabolites were quantified by integrating peak areas to obtain abundance values (corrected for natural isotope abundance). Absolute abundances, in nmol, were calculated using the molar relative response factor (MRRF) (**Equation 1, page 69**). Plots to visualise data were generated in GraphPad Prism.

2.6 Stable isotope labelling metabolomics

The uniformly stable isotope labelled sugars used in this study were purchased from Goss Scientific – [U- $^{13}\text{C}_6$]-glucose (^{13}C -glc) and [U- $^{13}\text{C}_6$]-galactose (^{13}C -gal). Samples were kept at 4°C as much as possible during processing, and at –80°C for longer term storage. Only Eppendorf brand Safe-Lock tubes and pipette tips were used.

2.6.1 *S. pneumoniae* culture in defined media

25 ml culture medium supplemented with either 11 mM glucose or 11 mM galactose (warmed at 37°C overnight) was inoculated with 250 – 500 μl *S. pneumoniae* stock ($\sim 3\text{--}5 \times 10^7$ CFU) and incubated at 37°C/ 5% CO_2 . At a time point during exponential growth phase (determined by growth curve analyses), which differed for different strains, the cultures were centrifuged (2000 rpm, 25 min, 4°C; Thermo Scientific TX-1000 swing-out rotor) to pellet bacteria. 1 ml of the supernatant was collected for ^1H -NMR and stored at –80°C. The remaining supernatant was discarded. The bacterial pellet was resuspended in 25 ml medium supplemented with ^{13}C -glucose or ^{13}C -galactose, and aliquots were taken for CFU determination. The cultures were incubated for 2 hours at 37°C/ 5% CO_2 with the labelled sugar.

2.6.2 Metabolic quench

Bacterial metabolism was quenched by placing open tubes into a slurry of dry ice and ethanol in an ice bucket and holding them in place with one hand. Cultures were rapidly agitated with a clean thermometer held in the other hand to $\sim 13^\circ\text{C}$, and immediately transferred to an ice/ water bath. The cultures were then centrifuged (2000 rpm, 25 min, 4°C; Thermo Scientific TX-1000 swing-out rotor) and 1 ml of the supernatant was collected for ^1H -NMR and GC-MS analysis of the conditioned media and stored at –80°C. The remaining supernatant was discarded.

2.6.3 Intracellular metabolite extraction for semi-targeted GC-MS

The pellet was resuspended in 1 ml ice-cold PBS, transferred to a chilled 2 ml tube, and centrifuged (13,000 x g, 5 min, 4°C). The supernatant was discarded, and the PBS wash was repeated. Aliquots were taken for CFU determination. Samples were centrifuged finally for 1 min (13,000 x g, 4°C) to remove any residual PBS. Pellets were then resuspended in 1 ml extraction buffer consisting of 400 µl methanol, 400 µl acetonitrile, and 200 µl water, with 1 nmol *scyllo*-inositol as the internal standard.

Samples were vortexed thoroughly and transferred to chilled 2 ml screw-cap vials (Sarstedt) containing 500 µl acid-washed glass beads (150-212 µm, Sigma-Aldrich). Samples were homogenised in a FastPrep-24 homogeniser (MP Biomedicals) using 2x cycles of 40 s sonication at 6 m/s + 300 s rest (4°C), then centrifuged (13,000 x g, 10 min, 4°C). Supernatants were carefully transferred to Spin-X tubes (0.22 µm cellulose acetate filter, Costar) and centrifuged (17,000 x g, 20 min, 4°C) to filter the extracts. The flow-through can be stored at –80°C.

When cultured in 7H9⁺ media, the entirety of the filtered extracts was dried in a SpeedVac™ rotary evaporator. After some testing and optimisation to standardise bacterial CFUs and improve chromatographic peak quality, this step was altered, so that when cultured in CDM, the volume of the filtered extracts corresponding to ~ 2x10⁸ CFU (based on post-quench CFU counts) was dried in a SpeedVac™ rotary evaporator. The dry material was biphasic partitioned in a mix of 50 µl chloroform, 150 µl methanol and 150 µl water, containing 5 nmol norleucine as an additional internal standard.

These biphasic partitioned extracts were vortexed thoroughly and stored at –80°C until further processing for GC-MS analysis of polar metabolites, which is detailed in **Section 2.4.1** Gas chromatography–mass spectrometry (GC-MS).

2.6.4 Extracellular metabolite extraction for semi-targeted GC-MS

5 µl of each extracellular media sample (stored at –80°C from **Section 2.6.2**) was added to 50 µl chloroform in 1.5 ml tubes. To this was added a 295 µl mix of 1:1 (v/v) methanol and water, containing internal standards *scyllo*-inositol (1 nmol) and norleucine (5 nmol). The now biphasic partitioned samples were vortexed thoroughly and stored at –80°C until further processing for GC-MS analysis of polar metabolites, which is detailed in **Section 2.4.1** Gas chromatography–mass spectrometry (GC-MS).

2.6.5 Data analysis

Isotope-labelled metabolomics data was analysed in MATLAB (Mathworks) using a program – MANIC – developed in-house based on a published script (Behrends et al., 2011). Chromatograms and mass spectra were also cross-checked with the Agilent MassHunter Qualitative Analysis software. Metabolites were identified by comparing chromatographic retention times and accurate mass against the retention times, mass spectra, and responses of known amounts of authentic standards (a pre-constructed metabolite mix (MM) of ~50 polar metabolites). Baseline correction was applied to all peaks. Peak areas were integrated to obtain isotope intensities (corrected for natural isotope abundance). Absolute abundances, in nmol, were calculated using the molar relative response factor (MRRF) (**Equation 1**). Isotope ratios for each metabolite peak were also obtained. Plots to visualise data were generated in GraphPad Prism.

$$MRRF = \frac{Area(x) \text{ in MM}}{Area(standard) \text{ in MM}} \times \frac{Quantity(standard) \text{ in MM}}{Quantity(x) \text{ in MM}}$$

Equation 1. Molar relative response factor (MRRF).

MRRF – the relationship between a signal produced by an analyte and the quantity of analyte which produces the signal.

MM – Metabolite Mix; Standard – *scyllo*-inositol; *x* = any analyte.

2.6.6 Extracellular targeted proton nuclear magnetic resonance (^1H -NMR)

^{13}C labelling of fermentation end-products (lactate, acetate, formate, and ethanol) were assessed using ^1H -NMR. Acetate, formate and ethanol are not detectable by GC-MS. NMR spectra were acquired and analysed by Paul Driscoll in the Metabolomics facility at the Francis Crick Institute, and the data was interpreted by me.

NMR samples were prepared as follows: 10 μl 10 mM 3-(trimethylsilyl)propane-1-sulfonate, sodium salt (DSS, internal standard) solution in deuterium oxide (D_2O) was added to 190 μl of undiluted extracellular medium (stored at -80°C from **Section 2.6.2**) in a 1.5 ml microcentrifuge tube, vortexed, and centrifuged (14,000 rpm, 4 min, 4°C). 180 μl of each preparation mixture was loaded into a 3 mm NMR tube. Samples were then loaded into a 96-tube Bruker SampleJet rack.

Metabolite concentrations were assessed using 1D ^1H -NMR spectroscopy on a Bruker Avance III HD spectrometer with a nominal ^1H frequency of 800 MHz equipped with a 5 mm triple resonance CryoProbe and SampleJet platform (in which samples were maintained at 6°C prior to measurement). The Bruker Icon software was employed to provide sample loading, temperature equilibration (at 25°C), radiofrequency tuning, field-frequency locking, homogenisation of the magnetic field, and data acquisition.

Spectra were obtained using the standard 1D nuclear overhauser effect spectroscopy (NOESY) gradient water suppression pulse sequence `noesygpr1d` (Bruker) using a 10 s presaturation/ relaxation delay (50 Hz presaturation field), 10 μs mixing time, and 4 s acquisition time, with a sweep width of 20 ppm. The 14 s repetition rate effectively provided for full inter-scan relaxation of metabolite longitudinal relaxation.

Spectra were processed and integrated in Bruker Topspin 3.6. The free induction decay (FID) was apodized with 1 Hz –line broadening, Fourier transformed. Each spectrum was phase-corrected by hand, and the chemical shift of the DSS methyl group peak set to 0 ppm. Baseline correction was performed by a combination of automatic global polynomial fitting, and local hand-adjusted spline fits close to the residual water resonance, using regions devoid of metabolite peaks.

Metabolite concentrations were obtained by integration of selected resonance intensities from both ^{12}C and ^{13}C species, adopted when overlap with other resonances was absent or minimal. Microsoft Excel calculations were used to scale the peak intensities to the DSS methyl group resonance, taking into account the number of protons that contributed to each peak (e.g., nine for the DSS methyl group peak, three for the ^{12}C -acetate methyl group peak, 1.5 for each ^{13}C -methyl group satellite peak).

The calculations also took account of the α -/ β -anomer population differences for glucose and galactose, obtained by averaging the relative peak intensities for α and β anomer peaks over the whole dataset.

Finally, concentration values were obtained by accounting for the 190-to-200 μl dilution effected on addition of the DSS to the sample, and assuming the DSS concentration in the sample to be 0.5 mM.

Chapter 3. Genomic and growth characterisation of *S. pneumoniae* prior to metabolic assessment

3.1 Introduction

S. pneumoniae can metabolise a wide range of nutritional sugars, contributing to its adaptability in different host microenvironments. One of these sugars is glucose, which is available in the blood at a concentration of ~ 5.5 mM under normal conditions. Glucose is also the main carbon source in complex media (such as THY and BHI) routinely used for laboratory pneumococcal culture. On the other hand, free glucose is actively suppressed in airway surface fluid as a form of nutritional immunity, i.e., to “starve” pathogens, and lung secretions are estimated to contain < 0.5 mM glucose. Therefore, the upper respiratory tract, the location of pneumococcal colonisation, is described as a nutrient-poor environment (Pezzulo et al., 2011; Philips et al., 2003; Siegel & Weiser, 2015).

However, the respiratory tract mucus lining is rich in membrane-tethered and secreted mucins, which consist of carbohydrate chains linked to a protein backbone (Rose & Voynow, 2006). These mucin glycoconjugates lining the respiratory tract are a rich source of carbohydrates for *S. pneumoniae*. The most common components of mucin are *N*-acetylglucosamine, *N*-acetylgalactosamine, fucose, mannose, and galactose (Bansil & Turner, 2006). Galactose released from mucin, therefore, is one of the most abundant sugars available to the pneumococcus in the respiratory tract (Paixão et al., 2015b; Terra et al., 2010).

The two sugars investigated in this study were glucose and galactose, as the most relevant sugars available to *S. pneumoniae* in the blood (invasive disease) and upper respiratory tract (site of colonisation and transmission) respectively. For a metabolomic characterisation study such as this one, it was imperative to first culture bacteria in an appropriate chemically defined medium that would allow for changing components – particularly the main carbohydrate source – while supporting the growth of *S. pneumoniae*.

In this Chapter, combinations of defined culture media and different culture conditions were tested to optimise and standardise pneumococcal growth using *S. pneumoniae* TIGR4. Once culture media and conditions were established, I also tested the growth of *S. pneumoniae* 23F and the capsule-deficient mutants TIGR4 Δ *cps* and 23F Δ *cps* in chemically defined medium (CDM).

S. pneumoniae is naturally competent and has a propensity for horizontal gene transfer, leading to a high recombination rate and allowing the exchange of genetic material between isolates of the same species as well as with other related species, notably *S. mitis* (Hakenbeck et al., 2001; Salvadori et al., 2019).

The *S. pneumoniae* pangenome (~ 5,000 – 7000 genes) is a collection of all orthologous gene clusters in the population. ~ 800 – 1200 (depending on the size of the dataset used) of the genes in the pangenome are present in all pneumococcal isolates, forming the core genome; the remaining genes form the accessory genome, and are distributed unevenly throughout the pneumococcal population. It is therefore estimated that > 75% of the *S. pneumoniae* pangenome consists of accessory genes (Hiller & Sá-Leão, 2018).

An individual pneumococcal isolate consists of ~ 2000 – 2200 genes, with latest data suggesting that approximately 40 – 50% of the genome contains genes from the accessory genome, adding to the genetic diversity of *S. pneumoniae* (Corander et al., 2017). In addition, more than 100 different serotypes of *S. pneumoniae* have been identified to date, with varying colonisation and invasion potential, host niche preferences, and phenotypic behaviour (Dananché et al., 2020; Hausdorff et al., 2005; Minhas et al., 2019; Sjöström et al., 2006; Weight et al., 2019).

Against this backdrop of pneumococcal genetic variability, I wanted to determine the degree of conservation of metabolism-associated genes in the *S. pneumoniae* population in order to decide whether studying the metabolism of two wild-type strains (TIGR4 and 23F) could be generalised to the population. To this end, nucleotide sequence

similarity analysis of metabolism-associated genes across *S. pneumoniae* isolates was undertaken, which is also discussed in this Chapter.

3.2 Chapter 3 Aims

- Determine the level of conservation of metabolic genes across *S. pneumoniae* isolates using nucleotide sequence similarity analysis, to decide whether the metabolomic characterisation of two wild-type strains (TIGR4 and 23F) could provide a reasonable estimation of pneumococcal population metabolism.
- Identify a chemically defined culture medium that supports the growth of *S. pneumoniae* and enables the use of different carbon sources.
 - Test the growth of *S. pneumoniae* TIGR4 in two chemically defined culture media – Sicard’s medium and Chemically Defined Medium (CDM) – and an in-house medium formulation adapted from Sicard’s medium – 7H9⁺ medium.
- Optimise *S. pneumoniae* culture conditions.
- Define the growth parameters of *S. pneumoniae* TIGR4 and 23F, comparing two growth curve analysis techniques – Curveball (Ram et al., 2019) script and regression analysis (GraphPad Prism).
- Compare the growth of all 4 *S. pneumoniae* strains (TIGR4, TIGR4 Δ *cps*, 23F and 23F Δ *cps*) in glucose or galactose as the sole carbon source.

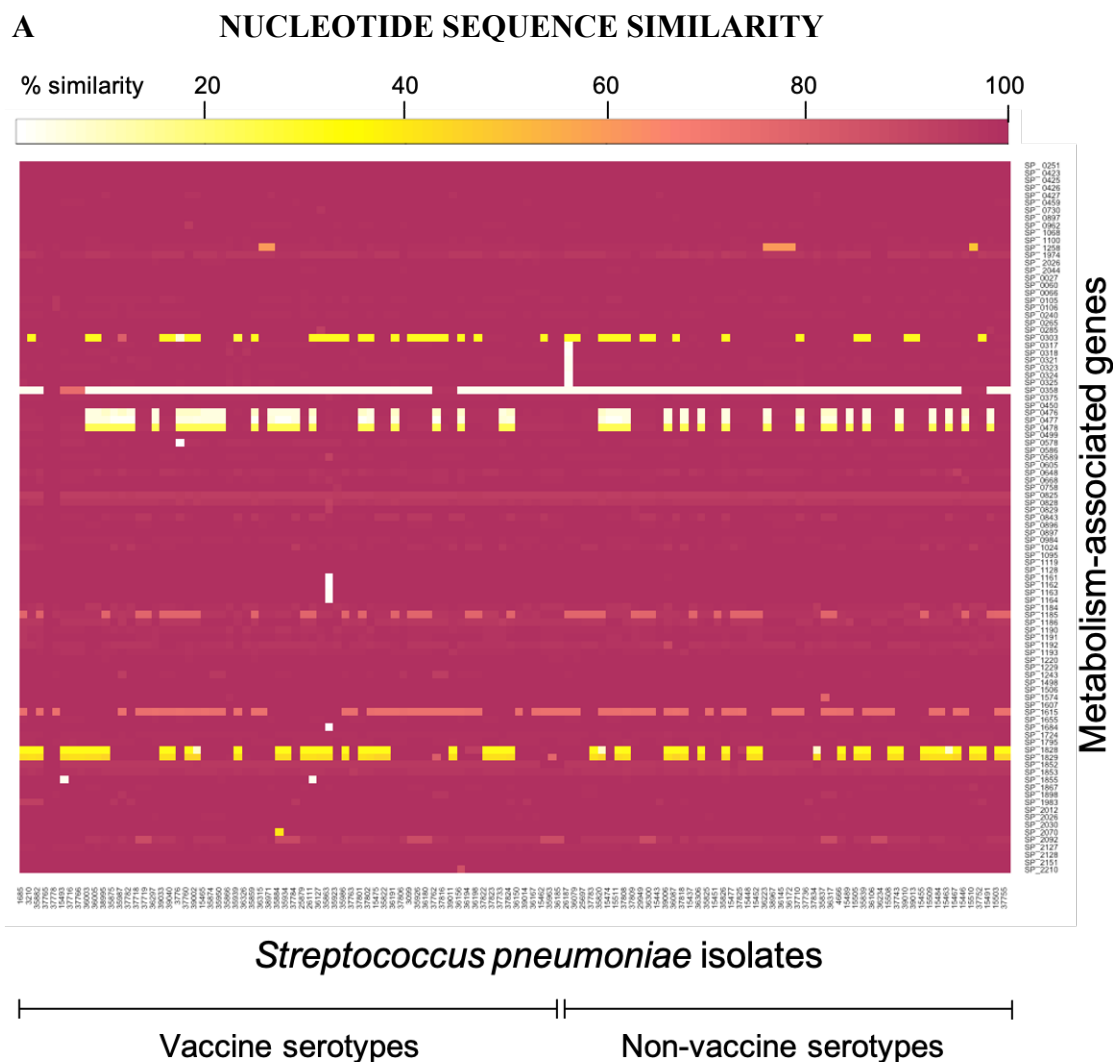
3.3 Nucleotide sequence similarity analysis reveals that central metabolism associated genes of *S. pneumoniae* are highly conserved

As mentioned in the introduction to this Chapter, the high genetic diversity of *S. pneumoniae* prompted us to question whether studying the metabolism of TIGR4 and 23F could be representative of the population. Therefore, preliminary nucleotide sequence similarity analysis was undertaken to determine the degree of conservation of metabolism-associated genes in the pneumococcal population, using *S. pneumoniae* TIGR4 as the reference genome (Tettelin et al., 2001).

A description of the methodology used can be found in **Chapter 2.3**. Briefly, the presence or absence of 95 carbohydrate metabolism genes was determined in 120 sequenced *S. pneumoniae* isolates (*Appendix Table 7.1*) by comparing each genome to the metabolism-associated nucleotide sequences from *S. pneumoniae* TIGR4 and assessing sequence similarity. I also investigated whether there were any patterns in the presence of these genes specific to vaccine serotypes when compared to non-vaccine serotypes.

Pneumococcal conjugate vaccine (PCV) 10 and 13 are routinely used for immunization of infants, and pneumococcal polysaccharide vaccine (PPSV) 23 has shown some efficacy in the elderly. These vaccines are designed to confer protection against 10, 13 and 23 serotypes respectively, out of the ~100 known *S. pneumoniae* serotypes, making them only partially protective. The vaccine type strains are chosen based on their prevalence in disease conditions.

This nucleotide sequence similarity analysis showed that the central carbon metabolism-associated genes (as annotated in the KEGG database (Kanehisa & Goto, 2000) for TIGR4; *Appendix Table 7.2*) are highly conserved at the genome level across a diverse set of *S. pneumoniae* isolates, and there was no discernible pattern of gene presence/absence in vaccine type compared to non-vaccine type groups (**Figure 3.1A**).



B **GENES WITH < 80% SIMILARITY IN > 40% OF ISOLATES**

<i>Gene</i>	<i>Annotation</i>	<i>Function</i>	<i>Pathway</i>	<i>Sequence similarity</i>	
bglA-1	6-phospho-β-glucosidase	Arbutin-6-P hydrolysis	Glycolysis	< 40%	
cap4J	UDP-glucose 4-epimerase	UDP-gal ↔ UDP-glc;	Galactose metabolism	< 10%	
galE-2		Involved in capsule synthesis		< 50%	
galT-1	Gal-1-P uridylyltransferase	UDP-glc + gal-1P ↔ glc-1P + UDP-gal		< 50%	
lacE-1	Lactose-specific PTS	Lactose → Lactose-6-P		< 40%	
lacE-2				< 80%	
lacF-1				< 20%	
lacG-1	6-phospho-β-galactosidase	Lactose-6-P → D-gal-6-P	PPP	< 10%	
recP-1	Transketolase			< 80%	

Figure 3.1 Nucleotide sequence similarity analysis of metabolism-associated *S. pneumoniae* genes

(A) False colour representation of percentage sequence similarity, generated in RStudio. 95 gene sequences involved in carbon metabolism pathways in *S. pneumoniae* TIGR4 were compared to the genomes of 120 pneumococcal isolates, using NCBI BLAST. (B) Genes that showed < 80% sequence similarity in > 40% of isolates. All genes have one or more functional homologs in each *S. pneumoniae* isolate. Abbreviations: UDP – uridine diphosphate; PTS – phosphotransferase system; gal – galactose; glc – glucose; PPP – pentose phosphate pathway; P – phosphate.

Of the 95 genes analysed, 9 showed a sequence similarity of < 80% in > 40% of pneumococcal isolates (cut-offs defined for convenience, not empirically) and are summarised in **Figure 3.1B**. 7 genes showed a sequence similarity of < 50% in many isolates and were considered absent. However, further analysis revealed that their absence would be unlikely to have a functional consequence on the associated metabolic pathways either because an alternative or orthologous gene was present, or because the gene was absent in only a proportion of isolates with the same serotype, so it could be an artefact of sequencing or assembly. These genes are listed below:

***bglA-1* and *bglA-2* – 6-phospho- β -glucosidases**

(*bglA-2* only absent in one isolate – ST14 JJA)

These hydrolyse arbutin-6-P and salicin-6-P for entry into the glycolytic pathway.

Arbutin is an extracellular glycoside (glucose bound to another functional group), found in wheat, pear skins and some bacteria. Salicin is an extracellular alcoholic glucoside (glucose bound to an alcohol) and has anti-inflammatory properties. As these glucosidases are obscure and non-essential, they were not investigated further.

***cap4J* – UDP-glucose 4-epimerase**

Interconverts UDP-galactose and UDP-glucose, either of which can be incorporated into specific capsule types. *cap4J* is probably serotype-specific or sequence type-specific, as it is present only in serotypes 4, 12F, 44 and 45, with 75% similarity with serotype 5. At least one of the other annotated UDP-glucose 4-epimerases (e.g. *galE-1*, *galE-2*, SP_1867) is present in all isolates.

lacF-1, lacG-1, lacE-1 – Whenever absent, absent together.

lacF-1 and *lacF-2* – lactose-specific PTS IIA components (lactose → lactose-6-P)

lacE-1 and ***lacE-2*** – lactose-specific PTS IIBC components (lactose → lactose-6-P)

lacG-1 and *lacG-2* – 6-phospho-β-galactosidase (lactose-6-P → D-gal-6-P)

Presumed non-consequential because homologous genes were present.

galE-2* and *galT-1 – Whenever absent, absent together.

galE-2 – UDP-glucose 4-epimerase (UDP-galactose ↔ UDP-glucose)

galT-1 – galactose-1-phosphate uridylyltransferase (UDP-glc + gal-1-P ↔ glc-1-P + UDP-gal), i.e., converts ingested galactose to glucose, as part of the Leloir pathway.

Non-consequential, because alternatives were always present.

As no striking differences in the presence/ absence of metabolism-associated genes were observed, we decided to forego planned downstream analyses to include allelic profile determination, sequence alignment, and phylogenetic tree construction. I recognise that nucleotide sequence similarity analysis by BLAST, as performed here, merely confirms the presence of a gene without providing any information about its activity, interactions, or transcriptional regulation under different conditions.

Nevertheless, the high degree of conservation of metabolism-associated genes in a geographically, temporally and serotypically diverse set of *S. pneumoniae* isolates led us to conclude that studying pneumococcal metabolism using *S. pneumoniae* TIGR4 could be generalised to the pneumococcal population, and that if necessary, any key findings would only have to be experimentally confirmed in a limited number of strains.

3.4 7H9⁺ culture medium – a combination of select 7H9 and Sicard's media components

3.4.1 Generating 7H9⁺ medium

With the aim of identifying a defined medium formulation to culture *S. pneumoniae* for subsequent metabolism studies, I reviewed the literature and decided to test the semi-defined Sicard's medium, which was designed for pneumococcal culture (Sicard, 1964). The preparation of Sicard's medium is detailed in **Chapter 2.2.1**, and its composition is summarised in **Table 3.1A**.

The original formulation contains 4 g L⁻¹ (22 mM; 0.4% w/v) glucose, and has been used to culture *S. pneumoniae* in a recent study looking at nutrient metabolism (Terra et al., 2016). The carbon source can also be easily replaced, and Sicard's medium supplemented with 10 g L⁻¹ (1% w/v) porcine gastric mucin as the sole carbon source (no glucose) has also been shown to support pneumococcal growth (Terra et al., 2010; Yesilkaya et al., 2008). A modified version of the Sicard's recipe was used in another study aiming to loosely mimic different host niches of *S. pneumoniae in vitro* (Aprianto et al., 2018).

Its adaptability, defined components, demonstrated compatibility with mucin, and recent use by other members of the pneumococcal research community made Sicard's medium a good initial candidate to culture *S. pneumoniae*.

A SICARD'S MEDIUM				B 7H9 MEDIUM	
<i>Minerals and buffer</i>	<i>g L⁻¹</i>	<i>Solution A</i>	<i>mg L⁻¹</i>	<i>7H9 Broth Base (Sigma-Aldrich)</i>	<i>g L⁻¹</i>
NaCl	5	L-Arginine	200	(NH ₄) ₂ SO ₄	0.5
NH ₄ Cl	2	L-Asparagine	10	Na ₂ HPO ₄	2.5
KCl	0.4	Glycine	120	KH ₂ PO ₄	1.0
Na ₂ HPO ₄	0.12	L-Histidine	150	Sodium citrate	0.1
MgSO ₄	0.024	L-Isoleucine	6.55	MgSO ₄	0.05
CaCl ₂	0.01	L-Leucine	6.55	CaCl ₂	0.0005
FeSO ₄ ·7H ₂ O	0.0055	L-Lysine	420	ZnSO ₄	0.001
Tris(hydroxymethyl)	4.84	L-Methionine	180	CuSO ₄	0.001
aminomethane		L-Threonine	175	Ferric ammonium citrate	0.04
<i>Other</i>	<i>g L⁻¹</i>	L-Valine	5.85	L-Glutamic acid	0.5
Linoleic acid*	10 ⁻⁷	Uracil	1	Pyridoxine	0.001
Spermidine phosphate*	10 ⁻⁵	<i>Solution B</i>	<i>mg L⁻¹</i>	Biotin	0.0005
<i>Supplements</i>	<i>g L⁻¹</i>	Biotin	0.015	<i>ADC Growth Supplement (Sigma-Aldrich)</i>	<i>g L⁻¹</i>
BSA	0.16	Choline chloride	5	BSA	5
D-(+)-glucose ¹	4	Nicotinamide	0.6	Dextrose (glucose)	2
Porcine gastric mucin ²	10	Calcium pantothenate	2.4	Catalase	0.003
		Pyridoxal HCl	0.6	<i>Other</i>	<i>% (v/v)</i>
		(-)-Riboflavin	0.3	Glycerol	0.5
		Thiamine HCl	0.6	Tween-80/ Tyloxapol	0.25
* Replaced by BSA		L-Cysteine HCl	100		
¹ Carbon source in (Sicard, 1964).		L-Glutamine	20		
² Carbon source in (Yesilkaya et al., 2008)		Sodium pyruvate	800		

Table 3.1 Composition of chemically defined media.

Sicard's medium (A) (Sicard, 1964) was completed with either glucose or mucin. 7H9 medium (B) was tested on its own; supplemented with Sicard's Solutions A and B; and with BSA, catalase, glycerol, or Tween-80 (a detergent) omitted.

A preliminary experiment to replicate growth in Sicard's supplemented with glucose or mucin using *S. pneumoniae* TIGR4 was successful (**Figure 3.2**). An increase in CFU ml⁻¹ of 2 log₁₀ units is characteristic of *S. pneumoniae* cultured in Sicard's medium, as shown by Yesilkaya and colleagues, where OD₆₀₀ was not reported (Yesilkaya et al., 2008).

However, the maximal OD₆₀₀ (OD-max) of 0.3-0.4 observed in **Figure 3.2** is lower than that reported in another study (Terra et al., 2016), which may be due to differences in culture technique. Terra and colleagues (2016) seeded Sicard's medium with a starter culture of *S. pneumoniae* D39 in chemically defined medium (CDM), while here I inoculated Sicard's medium with viable TIGR4 colonies plated on blood agar. The effect of different inoculation techniques will be discussed later in this Chapter. BSA

was not required for growth of TIGR4 in Sicard's medium containing either glucose or mucin (**Figure 3.2**).

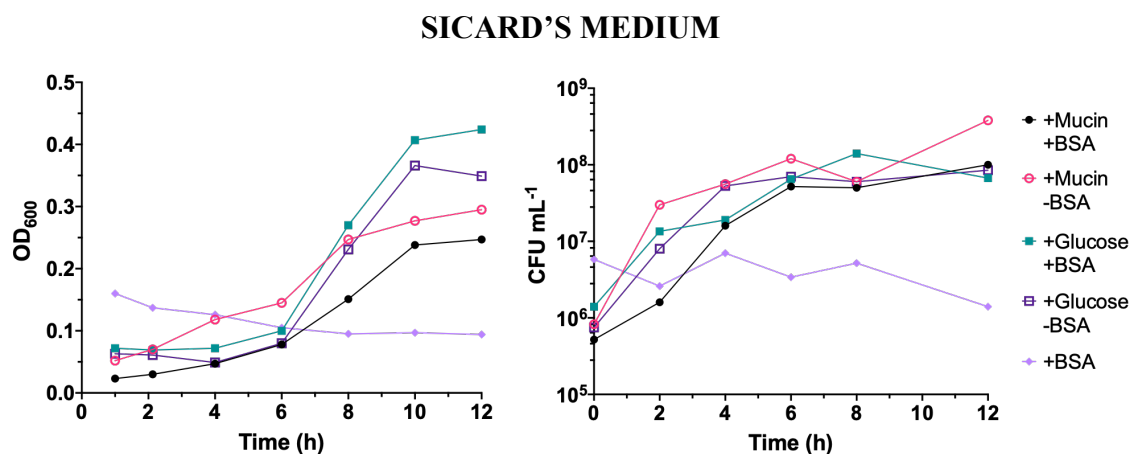


Figure 3.2 *S. pneumoniae* TIGR4 growth in Sicard's medium with glucose or mucin.

OD₆₀₀ (left) and CFU mL⁻¹ (right) show that TIGR4 growth is supported in Sicard's medium with glucose or mucin, regardless of the presence of BSA. 0.4% (22 mM) glucose and 1% porcine gastric mucin were used. 15 ml medium was inoculated with bacterial colonies streaked on blood agar plates overnight. n = 1.

In addition to Sicard's medium, growth of *S. pneumoniae* was also tested in 7H9 medium, which is routinely used to culture mycobacteria⁷. The preparation of 7H9 medium is more efficient than that of Sicard's medium, and it also contains defined components – it is assembled by combining the dehydrated 7H9 Broth Base, ADC Growth Supplement (consisting of BSA, glucose and catalase), glycerol and Tween-80/Tyloxapol (detergents used to prevent mycobacterial clumping in liquid culture) (**Table 3.1B**).

However, 7H9 medium lacks essential amino acids and vitamins and is unable to support pneumococcal growth on its own (**Figure 3.3**). So, I supplemented 7H9 medium with Solution A and Solution B from Sicard's medium (**Table 3.1A**), and

⁷ 7H9 medium was used because an initial study objective was to look at metabolic interactions between *S. pneumoniae* and *Mycobacterium tuberculosis*, which was not ultimately realised for this thesis.

found that the growth rate and OD-max of *S. pneumoniae* TIGR4 in this supplemented 7H9 formulation were higher than in Sicard's medium (**Figure 3.3A**). This was despite Sicard's medium containing twice the amount of glucose (4 g L^{-1} compared to 2 g L^{-1}), indicating that other component(s) of 7H9 media must play a key role in combination with Solutions A and B.

To determine what these components may be, TIGR4 was cultured in variations of supplemented 7H9 medium with individual media components omitted (**Figure 3.3B**). Glycerol, Tween-80, and catalase are dispensable for pneumococcal growth in supplemented 7H9, but BSA is essential in this context. In contrast, TIGR4 growth was unaffected in Sicard's medium lacking BSA, as shown previously in **Figure 3.2**. This is likely because 7H9 medium contains more than 30x the amount of BSA than Sicard's medium (0.5% and 0.016% respectively), and so the growth-promoting effect of BSA in 7H9 medium was more pronounced.

A new medium formulation – called 7H9⁺ – was thus developed by combining components of 7H9 and Sicard's medium. 7H9⁺ contains 7H9 Broth Base, Solutions A and B (from Sicard's medium), BSA, and catalase, supplemented with a carbon source (glucose, galactose, or mucin); the complete composition and preparation method are detailed in **Chapter 2.2.2**. Of all the formulations tested, 7H9⁺ medium best supported the overall growth of *S. pneumoniae* and was chosen for subsequent optimisation experiments.

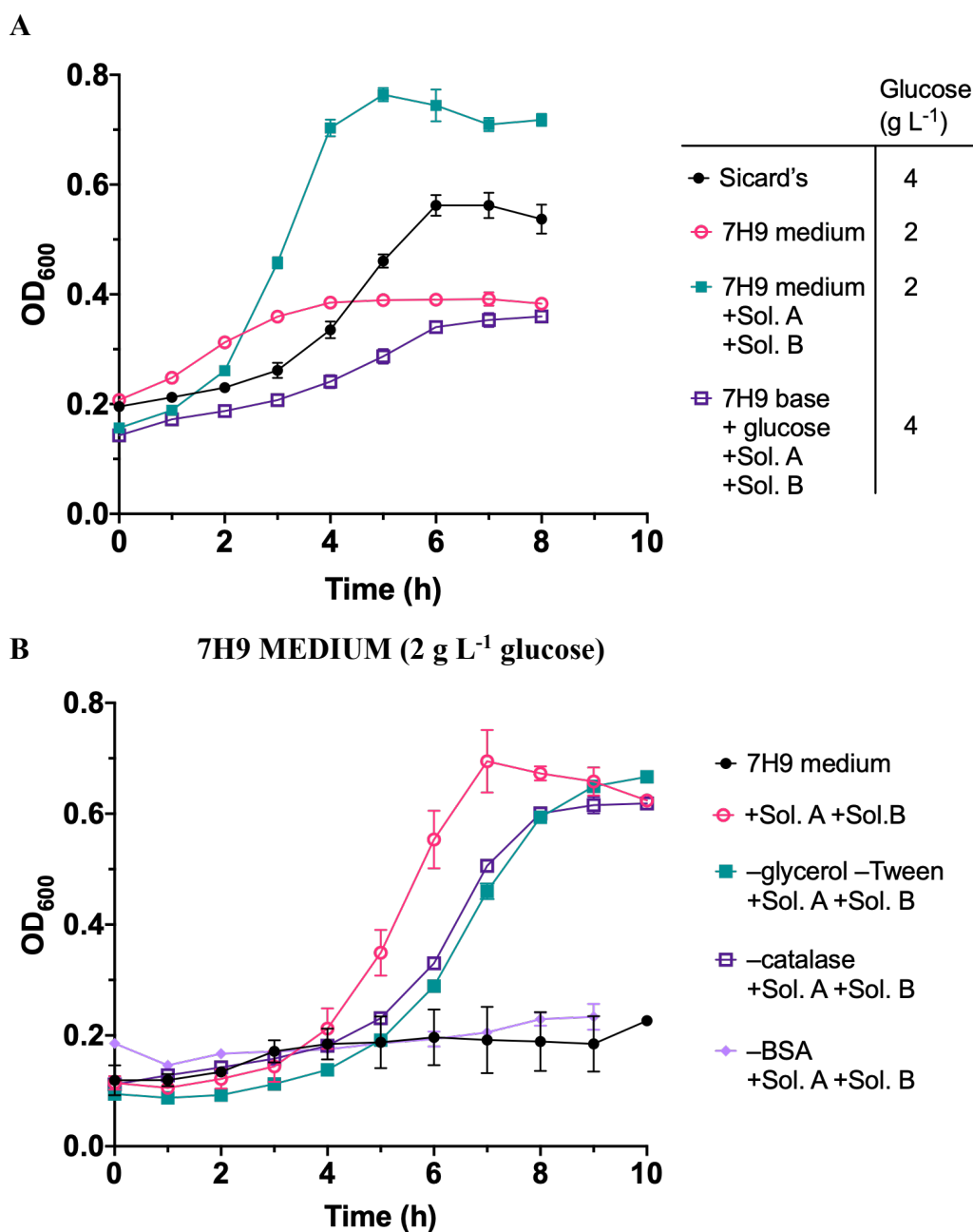


Figure 3.3 *S. pneumoniae* TIGR4 cultured in combinations of Sicard's and 7H9 media.

7H9 medium supplemented with Solutions A and B from Sicard's medium supported the best overall growth of TIGR4. In this context, BSA was essential for growth, but glycerol, Tween-80 and catalase were not. Plots depict mean and standard deviation (error bars). $n = 2$.

(A) 15 ml medium was inoculated with bacterial colonies streaked on blood agar plates overnight. (B) 0.2% (11 mM) glucose used. 5 ml starter culture of the appropriate media formulation was inoculated with bacterial colonies streaked on blood agar plates overnight, incubated for 30 min, then diluted to OD 0.1 in 15 ml fresh culture medium. The “-” notation in the legend denotes components that were omitted from the media formulation.

3.4.2 Primary carbon sources in 7H9⁺ culture medium

Having generated the base formulation of 7H9⁺ medium, the growth of *S. pneumoniae* TIGR4 on glucose and galactose as primary carbon sources was tested. The concentration of sugars in published pneumococcal metabolism studies are often in excess, ~ 30 mM to 60 mM (Carvalho, et al., 2013a; 2013b; Leonard et al., 2018; Paixão, et al., 2015a; 2015b). In the GC-MS metabolomics central to this project, excess sugars may lead to saturated chromatographic peaks and risk inaccurate quantification. These concentrations also exceed *in vivo* levels – normal blood glucose is ~ 5.5 mM and lung secretions are estimated to contain < 0.5 mM glucose. The glycan content of the respiratory epithelium and levels of galactose and mucin, however, are difficult to estimate.

To mimic physiological glucose concentration as closely as possible, a range of concentrations were tested to determine the minimum amount of sugar that would support pneumococcal growth in 7H9⁺ medium (**Figure 3.4**). Growth curves in 30 mM and 11 mM glucose were very similar (**Figure 3.4A**), with the growth rate decreasing rapidly with further decreasing glucose concentration (**Figure 3.4B**). Growth was slower in galactose, but the OD-max (and by approximation the final biomass) at 11 mM was similar for both carbon sources (**Figure 3.4A**). 11 mM, therefore, was deemed the optimal concentration of glucose and galactose for subsequent metabolism experiments.

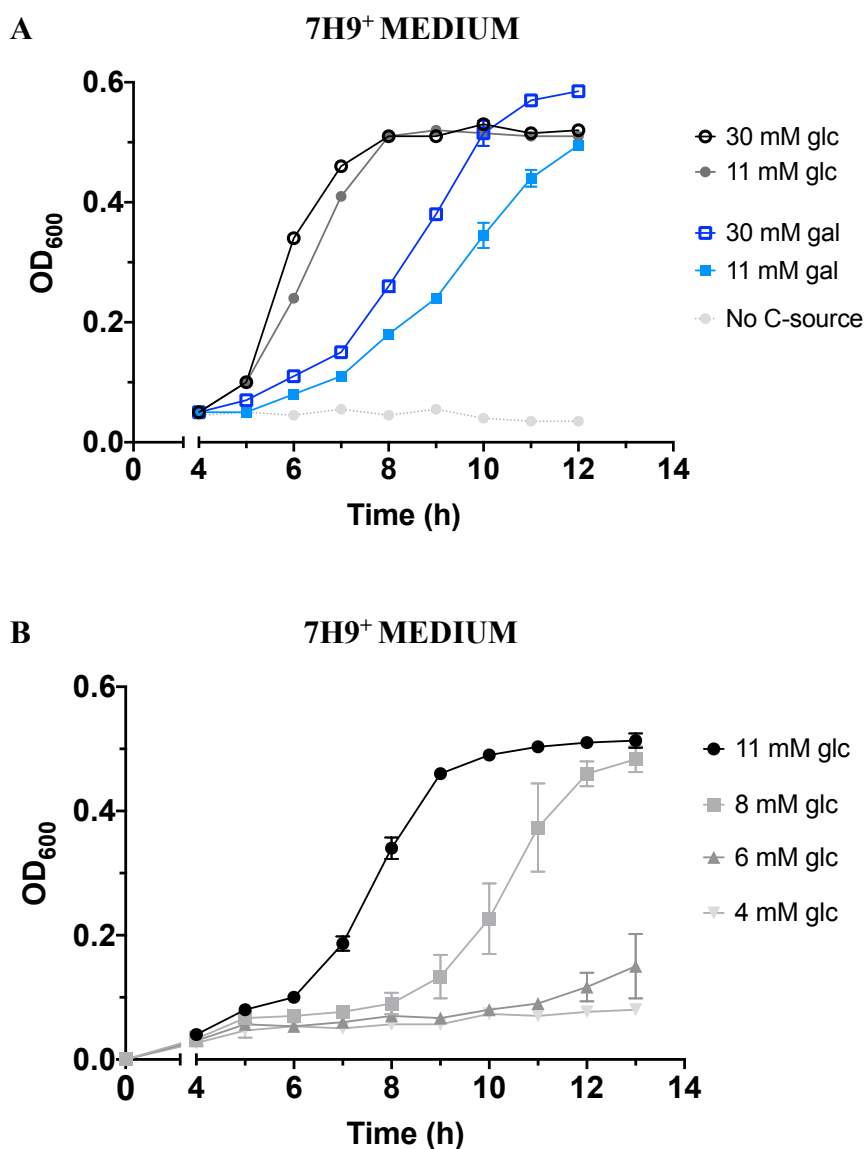


Figure 3.4 *S. pneumoniae* TIGR4 growth in 7H9⁺ medium with varying amounts of sugar.

5 ml medium was inoculated with 50 μ l centrifuged bacterial stock. Plots depict mean and standard deviation (error bars). n = 2-3.

Abbreviations: C-source = carbon source; glc = glucose; gal = galactose.

1% (w/v) porcine gastric mucin has been shown to support pneumococcal growth in Sicard's medium at a rate comparable to glucose (**Figure 3.2**) (Kahya et al., 2017; Yesilkaya et al., 2008), but I was unable to replicate this in 7H9⁺ medium using published protocols (**Figure 3.5A**). Therefore, modifications to the preparation and sterilisation of mucin were tested.

Dialysing a solution of mucin against water to remove low molecular weight sugars is an important preparation step, but the low solubility of mucin in water prevented high concentration stocks to be generated, posing the additional challenge of heavily diluted 7H9⁺ media components in the final growth medium. After dialysis, therefore, the mucin solution (in water) was either buffer exchanged into 7H9⁺ using a spin protein concentrator column or lyophilised to remove the water and then dissolved in 7H9⁺ medium (**Figure 3.5B**). The final mucin solution was sterilised by autoclaving as filter-sterilisation through a 0.22 µm filter caused significant loss of mucin.

The buffer exchange method was unsuccessful as mucin stuck to the concentrator column. Lyophilisation, on the other hand, allowed for a more concentrated mucin solution in the balanced salt solution of 7H9⁺ medium to be obtained without significant loss of material. Therefore, a reliable method to study *S. pneumoniae* growth on mucin as the primary carbon source in 7H9⁺ medium has been developed (**Figure 3.5B**).

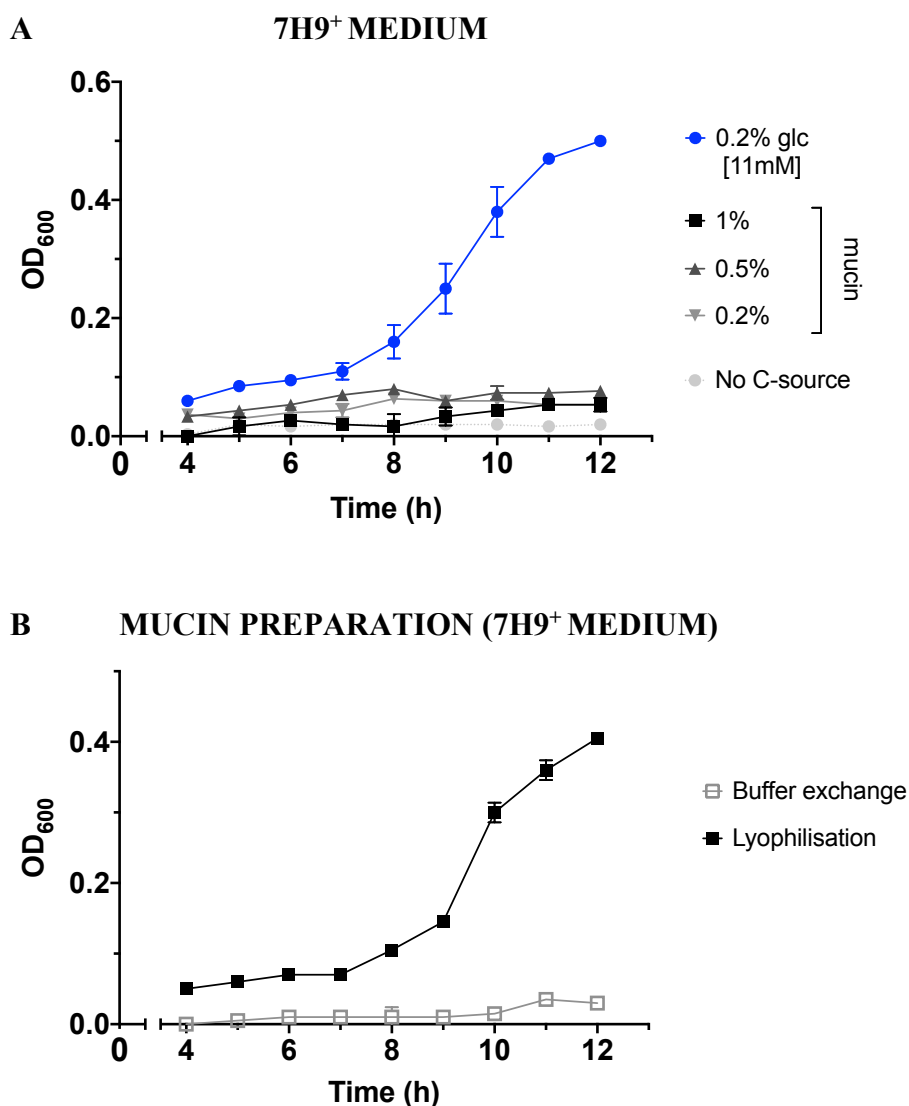


Figure 3.5 Optimising mucin preparation in 7H9⁺ medium.

S. pneumoniae TIGR4 fails to grow when cultured in different concentrations of (porcine gastric) mucin after dialysis against water (A), likely as a result of dilution of media components. Of two methods to concentrate dialysed mucin solution (buffer-exchange using a protein concentrator column and lyophilisation), lyophilisation rescued TIGR4 growth in 1% (w/v) mucin in 7H9⁺ medium (B). 5 ml medium was inoculated with 50 μ l washed bacterial stock. Plots depict mean and standard deviation (error bars). $n = 3$.

3.4.3 7H9⁺ medium stability and a standardised inoculation method

Various bacterial inoculation methods have been reported in the literature, which can affect the growth profile of *S. pneumoniae* in culture. Bacteria can be adapted before inoculating experimental culture medium by pre-culturing on solid medium (such as blood agar plates) or establishing a starter culture in complex (such as THY) or defined (such as 7H9⁺) media. Because this study aims to investigate pneumococcal metabolism, pre-adapting the bacteria to complex undefined media was not considered to be ideal.

Growing pneumococcal stocks in a starter culture of the appropriate defined medium and then diluting the culture to OD ~ 0.1 in fresh experimental medium allowed for standardisation of starting OD (**Figure 3.3B**), but proved to be an inefficient method in the context of this study due to the slower growth rate in Sicard's and 7H9⁺ media; obtaining sufficient bacterial biomass (mid-late log phase) requires 6-8 hours of pre-culture, which would be followed by sample collection for metabolic footprinting (discussed in Chapter 4) every one or two hours for 10-12 hours.

Therefore, I opted for direct inoculation with working pneumococcal stocks stored at -80°C. This allowed for standardisation of the bacterial inoculum and was also more time efficient. **Figure 3.6A** shows that centrifuging stocks or washing once with PBS had an increasingly detrimental effect on bacterial growth in 7H9⁺ – this may be a result of bacterial loss after centrifugation, as viable bacterial colonies were observed in the supernatant (**Figure 3.6B**). An attempt to mitigate this loss by doubling the initial inoculum from 5x10⁷ to 1x10⁸, however, was unable to rescue growth (**Figure 3.6A**), suggesting that potential bacterial loss due to centrifugation or washing was not the reason for poor growth in 7H9⁺ medium.

Stocks of *S. pneumoniae* were generated and frozen in THY. Good growth was observed in 25 ml 7H9⁺ when inoculated directly with ~ 500 µl stock (i.e., 2% v/v), (**Figure 3.6A**). To determine whether centrifuging stocks stressed bacteria and thus led to poor growth in 7H9⁺, or whether the residual THY in the stock was responsible for driving pneumococcal growth in 7H9⁺, I tested the growth of centrifuged bacterial

stocks supplemented with 2% (v/v) THY (i.e., the amount of THY present when directly inoculating 7H9⁺ with frozen stock). **Figure 3.6A** shows that supplementation with 2% (v/v) THY rescued the growth of centrifuged pneumococcal stocks, suggesting that residual THY in the stock is likely kick-starting growth of *S. pneumoniae* TIGR4 in 7H9⁺ media.

Because THY is a rich, complex medium with some undefined components, it is difficult to determine the exact components supplementing pneumococcal growth in 7H9⁺. However, the amount of glucose in both THY and 7H9⁺ is equivalent (11 mM), so it can be concluded that the differences observed in growth were not related to the main carbon source under investigation.

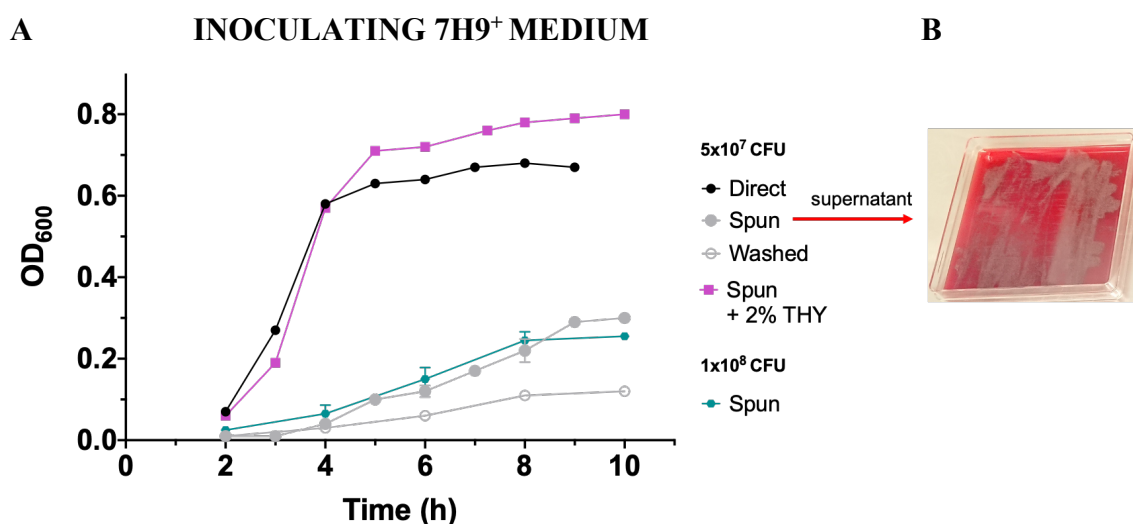


Figure 3.6 Determining an inoculation method for *S. pneumoniae* in 7H9⁺ medium

25 ml 7H9⁺ medium (+ 11mM glucose) was inoculated with 500 μ l (5x10⁷ CFU) or 1 ml (1x10⁸ CFU) *S. pneumoniae* TIGR4 stock. Plots depict mean and standard deviation (error bars). n = 1-2

(A) Medium was inoculated directly, or with stock treated in the following ways: spun – centrifuged at 13,000 x g, 10 min, and pellet resuspended in 7H9⁺; washed – centrifuged (13,000 x g, 10 min) and washed in PBS, then resuspended in 7H9⁺; spun + 2% THY – centrifuged (13,000 x g, 10 min), pellet resuspended in 7H9⁺, and medium supplemented with 2% (v/v) (i.e., 500 μ l) THY.

(B) The supernatant of a centrifuged stock vial was plated on blood agar and incubated (37°C/ 5% CO₂) overnight. Visible viable colonies suggest bacterial loss on centrifugation.

A drawback of using 7H9⁺ medium was that one of its components, Solution A (**Table 3.1A**), is unstable over time, even when stored at 4°C. **Figure 3.7** shows that only 3 days after preparing Solution A, the growth of *S. pneumoniae* TIGR4 in 7H9⁺ was negatively affected – with time, the lag phase of growth increased. Therefore, fresh 7H9⁺ was prepared before each use by adding the individual ingredients of Solution A directly. This made 7H9⁺ media preparation time-consuming. Together with the fact that growth (as measured by change in CFU) in 7H9⁺ was modest, I decided to compare pneumococcal growth in another defined culture medium – Chemically Defined Medium (CDM) (**Table 2.3**) (van de Rijn & Kessler, 1980).

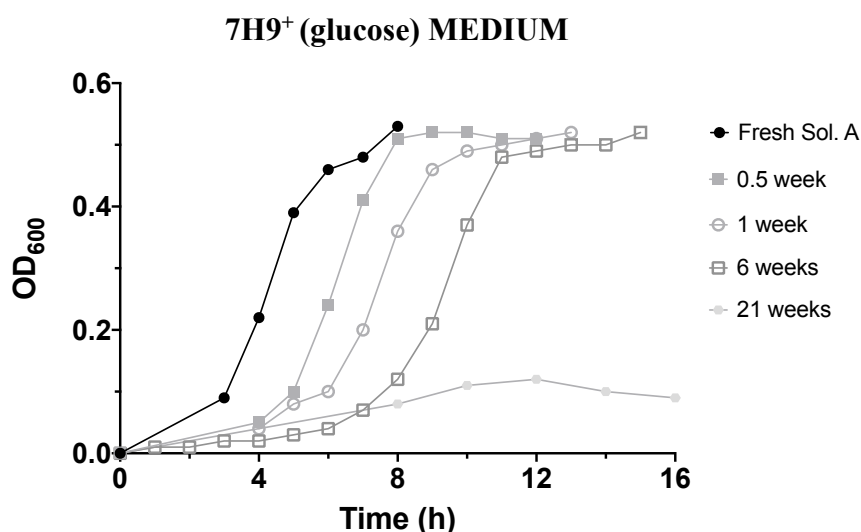


Figure 3.7 Storage of Solution A delays *S. pneumoniae* TIGR4 growth in 7H9⁺ over time.

7H9⁺ (+ 11 mM glucose) was inoculated with TIGR4 stock. The medium was prepared with fresh Solution A, or Solution A stored at 4°C for the indicated duration. n = 1.

3.5 Chemically Defined Medium (CDM)

CDM has been used to culture *S. pneumoniae* to study aspects of its metabolism (Carvalho et al., 2013a; 2013b; Paixão et al., 2015a; 2015b; Härtel 2012). The composition of CDM is shown in **Table 2.3**, modified from the original publication where it was developed for Group A streptococcal culture (van de Rijn & Kessler, 1980), and prepared according to Protocol S1 from a study by Chang et al. (Chang et al., 2011).

With stock solutions prepared, CDM was relatively more time-efficient to assemble compared to 7H9⁺ medium. It also does not contain BSA, allowing more control over amino acid concentrations. Finally, the OD-max, and therefore bacterial yield, was higher when *S. pneumoniae* TIGR4 was cultured in CDM and mimicked growth in THY (**Figure 3.8**). This was of particular importance because the greater bacterial biomass would aid in the eventual detection of intracellular metabolites present at a low concentration or undergoing rapid metabolic turnover.

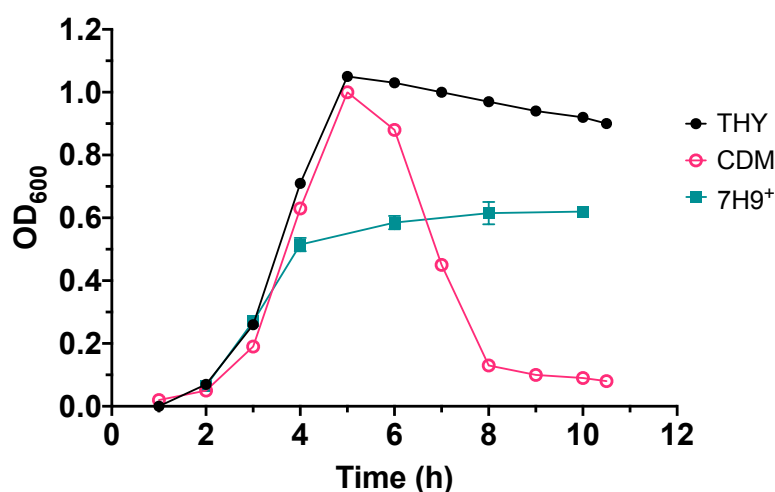


Figure 3.8 TIGR4 growth in media containing 11 mM glucose.

25 ml culture medium was inoculated with $\sim 3\text{--}5 \times 10^7$ CFU *S. pneumoniae* TIGR4 stock. THY is a complex medium; CDM and 7H9⁺ are chemically defined media formulations (Chapter 2.2). Plots depict mean and standard deviation (error bars). n = 1-2.

3.5.1 Comparing *S. pneumoniae* growth in CDM and 7H9⁺

The growth of *S. pneumoniae* TIGR4 and 23F in CDM and 7H9⁺ containing glucose (glc) or galactose (gal) was compared at high throughput using a microplate reader to generate and analyse growth curves (**Figure 3.9**). CDM containing either sugar supported the growth of both strains similarly, with OD-max being comparable between TIGR4 and 23F (**Figure 3.9A, B, G**). In 7H9⁺ medium, however, 23F has a longer lag phase and reaches approximately half the OD-max of TIGR4, suggesting poorer growth (**Figure 3.9C, D, G**).

While both CDM(glc) and 7H9⁺(glc) supported the growth of TIGR4 and 23F to some extent, the growth rate in CDM(glc) was higher (**Figure 3.9E**). In galactose, TIGR4 grew at a similar rate in both culture media, and 23F had a slightly higher growth rate in CDM(gal) (**Figure 3.9F**). Growth rates were calculated by applying linear regression to plots of ln(OD₆₀₀) vs. time, as described by Carvalho and colleagues (Carvalho et al., 2013b). Lag phase duration and OD-max were determined from growth curve plots (GraphPad Prism).

The growth curve analysis program Curveball (Ram et al., 2019) was also employed (with help from Modupeh Betts, Mucosal Pathogens Research Unit, UCL), designed to work with the large amounts of growth data obtained from microplate readers and make predictions about microbial competition experiments. The script chooses one out of six best-fit growth models to provide various growth parameters, including growth rate, lag duration and OD-max (**Figure 3.10**). However, none of the models allow for a “death” phase as seen in CDM(glc) cultures (**Figure 3.10A, B**), therefore under-reporting the OD-max. The best-fit model for 23F cultured in 7H9⁺(gal) was also inappropriate, as it smoothed over the unusual growth curve depicting two exponential growth phases (**Figure 3.10H**).

Therefore, the growth parameters reported in **Figure 3.9E – G** were those calculated by linear regression of the plots of ln(OD₆₀₀) vs. time and the growth curves plotted in GraphPad Prism.

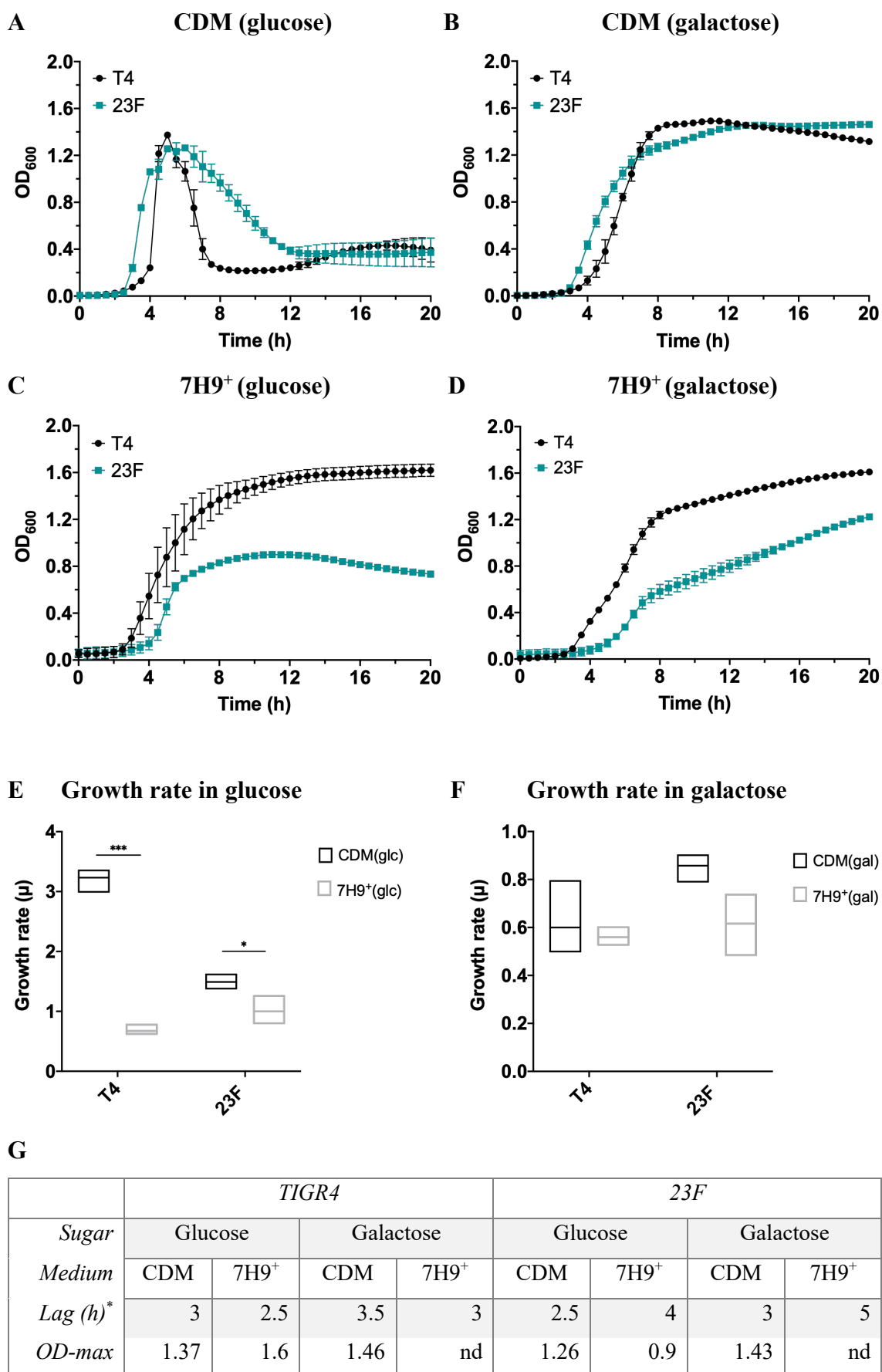
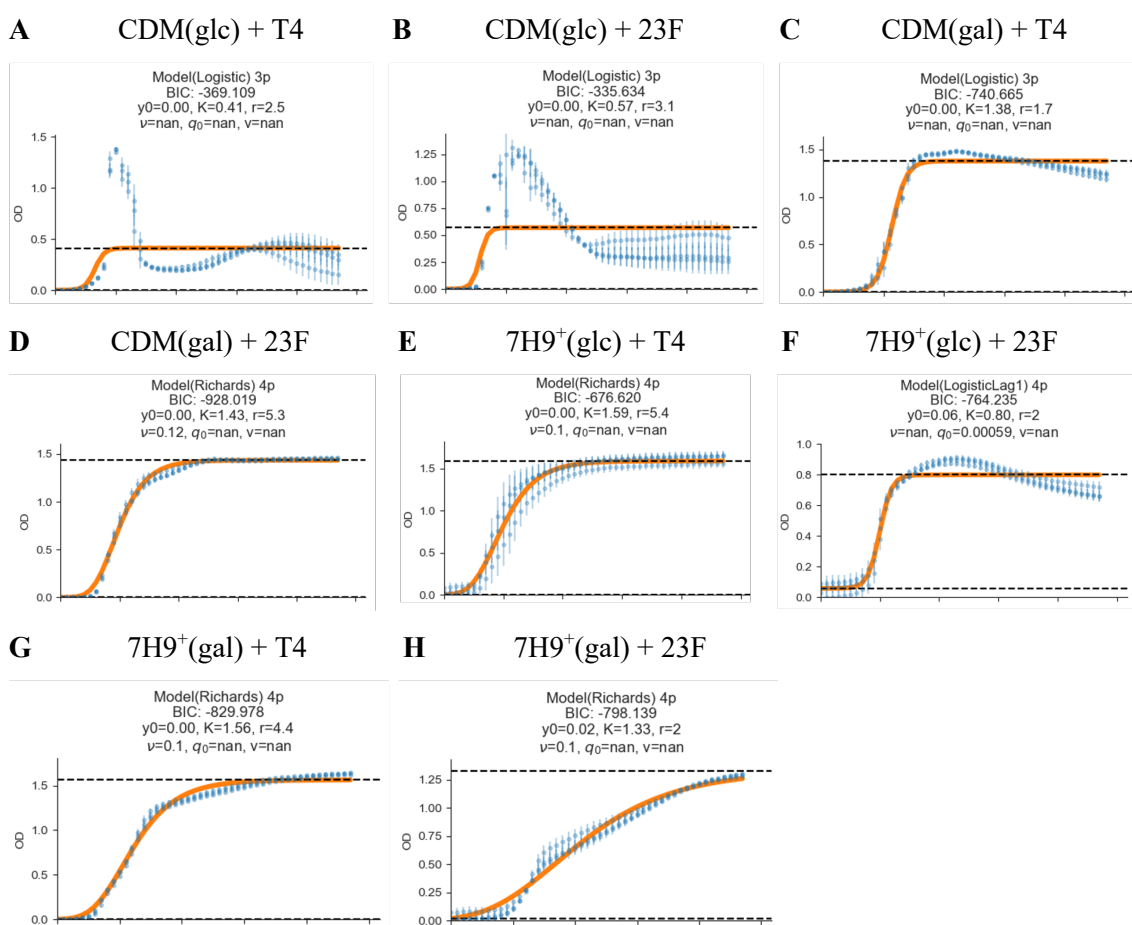


Figure 3.9 *S. pneumoniae* TIGR4 and 23F growth in a microplate reader

Each well of the 96-well plate contained 200 μ l culture medium with $\sim 2\text{-}4 \times 10^5$ CFU *S. pneumoniae* TIGR4 or 23F stock (equivalent to 25 ml culture medium inoculated with $\sim 3\text{-}5 \times 10^7$ CFU stock). $n = 3$.

(A – D) Growth curves were generated using a Tecan multimode plate reader set to 37°C/ 5% CO₂. Plots depict mean and standard deviation (error bars). (E – F) Growth rates were calculated by linear regression analysis of $\ln(\text{OD}_{600})$ vs time during exponential phase of growth. Statistical significance was determined using multiple t-tests and the Holm-Šidák test to correct for multiple comparisons. (G) Lag phase duration (* to the nearest 0.5 h) and OD-max are tabulated; nd – not defined.

**Figure 3.10 Growth curve best-fit models using the Curveball analysis script**

Growth curves generated using a microplate reader (Tecan Spark® multimode plate reader) are depicted as blue data points. The best-fit model for each growth curve as determined by the Curveball analysis script is denoted by an orange line.

Each well of the 96-well plate contained 200 μ l culture medium with $\sim 2\text{-}4 \times 10^5$ CFU *S. pneumoniae* TIGR4 or 23F stock (equivalent to 25 ml culture medium inoculated with $\sim 3\text{-}5 \times 10^7$ CFU stock). $n = 3$.

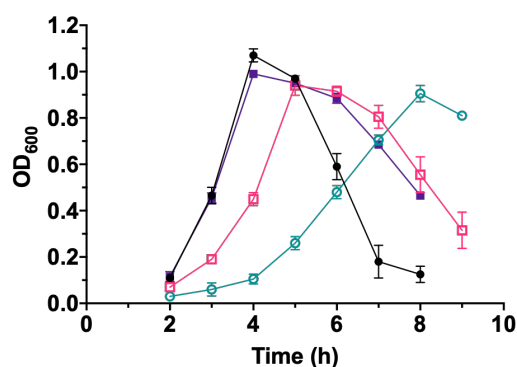
3.5.2 Growth of *S. pneumoniae* TIGR4 and 23F wild-type and capsule-deficient strains in glucose and galactose

To confirm CDM as the culture medium of choice for ^{13}C -labelling metabolomics, I tested its ability to support the growth of *S. pneumoniae* on both glucose and galactose. The 4 strains of interest – TIGR4, TIGR4 Δ *cps*, 23F and 23F Δ *cps* – were able to grow comparably in CDM containing glucose or galactose (**Figure 3.11**), with growth in galactose being delayed by 1.5-2 hours for all strains.

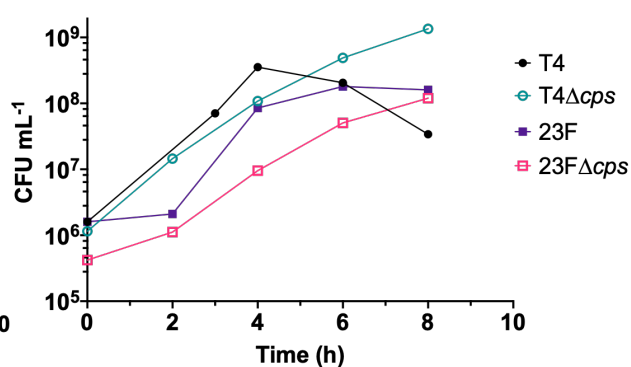
An important discrepancy between OD and CFU ml $^{-1}$ patterns, particularly for TIGR4 Δ *cps*, in **Figure 3.11** can be explained by the differences in colony size between the strains – given the same number of viable colonies (CFU), the OD (optical density, a measure of bacterial density rather than bacterial count) of TIGR4 Δ *cps* is expected to be lower than that of the wild-type due to their smaller size (**Figure 3.12**).

CDM (glucose)

A

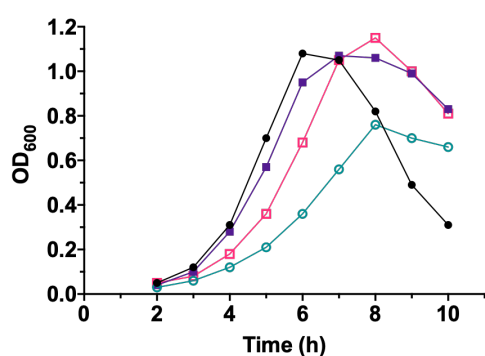


B



CDM (galactose)

C



D

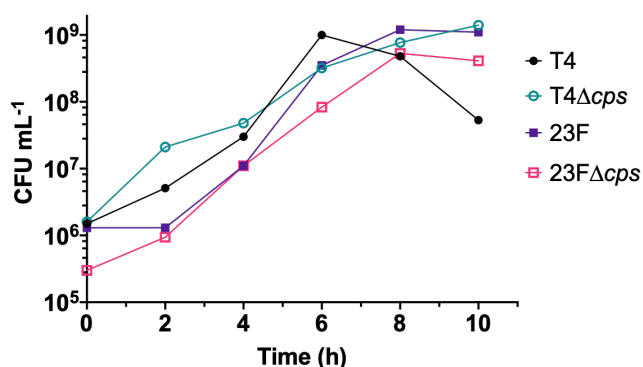


Figure 3.11 *S. pneumoniae* strains cultured on glucose and galactose

25 ml culture medium was inoculated with $\sim 3\text{-}5 \times 10^7$ CFU *S. pneumoniae* stock. 11 mM glucose and galactose were used. The decrease in OD₆₀₀ compared to CFU ml⁻¹ is likely a result of plotting on a linear vs. a log₁₀ scale, respectively. Plots depict mean and standard deviation (error bars). n = 1-3.

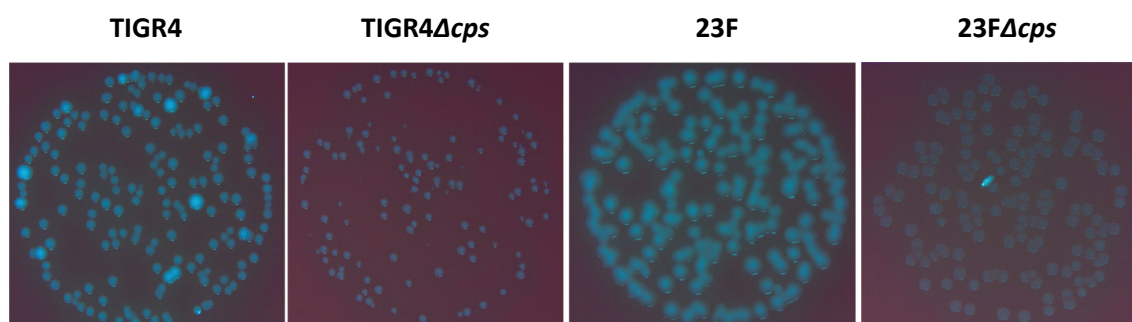


Figure 3.12 Colony morphology of *S. pneumoniae* strains

S. pneumoniae on a blood agar plate. Photographed using the IC90 E CMOS microscope camera mounted on a M125 C stereo microscope, viewed through a PLANAPO 1.0 x objective lens (Leica).

3.6 Chapter 3 Discussion

This Chapter detailed the optimisation of *S. pneumoniae* culture conditions, building a foundation for subsequent metabolomic characterisation.

A high degree of genetic variation is observed in the pneumococcal population as a result of natural competence and horizontal gene transfer, reflected in the large size of its accessory (or dispensable) genome (**Figure 3.13**).

***S. PNEUMONIAE* PANGENOME (POPULATION)**

CORE GENOME 800 – 1200	ACCESSORY GENOME 5000 – 7000
---------------------------	---------------------------------

***S. PNEUMONIAE* ISOLATE (SINGLE GENOME)**

CORE GENES 800 – 1200	ACCESSORY GENES 800 – 1200
--------------------------	-------------------------------

~ 2000 – 2200 GENES

Figure 3.13 *S. pneumoniae* core and accessory genome sizes

An individual isolate of *S. pneumoniae* has ~ 2000 – 2200 genes, with a 50/50 or 60/40 distribution of core genes and accessory genes. The core genome is conserved in all isolates, but the accessory genes can vary between isolates. The totality of accessory genes from all isolates in a population forms the accessory genome, which together with the core genome is termed the pangenome. The number of genes depicted have been estimated based on data from various studies (Corander et al., 2017; Donati et al., 2010; Hiller & Sá-Leão, 2018); the exact numbers vary with the size of the dataset analysed.

Therefore, I first compared nucleotide sequences of central carbon metabolism-associated genes of TIGR4 from the KEGG database (Kanehisa & Goto, 2000) to the genomes of 120 diverse *S. pneumoniae* isolates (complete and partial assemblies included). This analysis revealed a high level of conservation and sequence similarity for all pneumococcal genes annotated as being involved in carbon metabolism, glycolysis/ gluconeogenesis, pyruvate metabolism, galactose metabolism, and the pentose phosphate pathway.

This was perhaps not an unexpected result, as known central metabolism genes have been shown to be essential in transposon mutagenesis and CRISPRi studies of *S. pneumoniae* D39, R6 and TIGR4 (Camilli & van Opijnen, 2012; Liu et al., 2017; van Opijnen et al., 2009), and so they likely belong to the core genome. However, the preliminary nucleotide sequence similarity analysis of central carbon metabolic genes carried out in this Chapter provides confidence that metabolic characterisation of *S. pneumoniae* using TIGR4 and 23F as model strains could be generalised to the population as a whole.

Having justified the use of TIGR4 and 23F, I then set out to generate a culture medium with defined components which would allow accurate interpretation of metabolomics data and would allow for the switching of the primary carbon source between glucose and galactose.

7H9⁺ medium was created by combining the vitamins and amino acid solutions from Sicard's medium (semi-defined culture medium generated for *S. pneumoniae* culture) (Sicard, 1964) with the mineral and buffer solutions of 7H9 medium (proprietary culture medium for mycobacterial culture). Although Sicard's medium did support pneumococcal growth, the yield was higher with 7H9⁺. Therefore, subsequent optimisation of culture conditions was performed in 7H9⁺ inoculated with *S. pneumoniae* TIGR4.

The optimal concentration of glucose and galactose was determined to be 11 mM. This was the lowest concentration that supported pneumococcal growth in these *in vitro* culture conditions – an important consideration in order to avoid saturating the chromatographic column with excess sugars during metabolomic analysis. 11 mM is also closer to the glucose concentration reported in the blood (~ 5.5 mM) (Perez Diaz et al., 1977).

This Chapter also discussed the comparison of different bacterial inoculation methods, eventually concluding that direct inoculation from frozen stocks into pre-warmed

culture medium was the most efficient method that also allowed for standardisation of the inoculum.

However, 7H9⁺ medium had two main drawbacks – BSA was found to be an essential component without which TIGR4 would not grow in 7H9⁺, and the instability of the amino acid solution of 7H9⁺ prevented the long-term storage of the medium and required a time-intensive media preparation step every 48 hours. The practical implications of these drawbacks were that the exact concentration of amino acids in 7H9⁺ would not be quantifiable because BSA consists of polypeptide chains; and fresh medium would have to be prepared every 48 hours, making the workflow inefficient.

These limitations of 7H9⁺ medium were mitigated by using a chemically defined medium (CDM) adapted for use by the pneumococcal community. CDM does not contain BSA; its stock solutions are stable for > 6 months, making media preparation much more efficient than 7H9⁺; and the biomass yield and growth rate of *S. pneumoniae* were higher when cultured in CDM compared to in 7H9⁺.

While much of the pneumococcal cultures in this Chapter as well as in subsequent metabolomics chapters were generated in 20 – 25 ml volumes, the growth rate comparisons were performed by generating growth curves of 200 µl cultures using a microplate reader. This not only facilitated multiple conditions and biological replicates to be tested in parallel in a high throughput manner, but also generated growth curve data in a format compatible for input into the growth curve analysis program Curveball (Ram et al., 2019).

Curveball attempts to fit the growth curve data to one of six pre-defined growth models, generating information about various growth parameters (including growth rate) in the process. We found that the program worked well for growth curves with well-defined areas of lag, exponential, and stationary phases, providing useful insights about growth curve characteristics relatively quickly. However, these best fit models were not appropriate for some types of growth curves observed in our dataset, such as those with

a death phase (due to bacterial lysis upon nutrient depletion, for example) or those with a less well-defined and gradual transition between the phases of growth.

Therefore, I pneumococcal growth rates were finally calculated using regression analysis – the slope of a linear regression trendline fitted to plots of $\ln(\text{OD}_{600})$ versus time – as mentioned in an another study (Carvalho et al., 2013b).

After determining the superiority of CDM over 7H9⁺ medium in supporting TIGR4 growth, I then confirmed that 23F, TIGR4 Δcps and 23F Δcps were also able to grow comparably in CDM supplemented with either 11 mM glucose or galactose in 25 ml cultures, thereby achieving the final aim set out for this Chapter.

Chapter 4. Extracellular metabolomic profiling suggests novel metabolic activity in *S. pneumoniae*

4.1 Introduction

As *S. pneumoniae* grows in the liquid culture medium, it alters the extracellular medium composition by taking up essential nutrients and releasing certain products of metabolism, generating a ‘footprint’ of its metabolic activity. Metabolic footprinting is a relatively straightforward, non-destructive technique that allows continuous sampling from the same culture, helping to generate a dynamic portrait of the extracellular metabolome under a specific set of conditions over time.

Studying the metabolite composition of the conditioned culture medium can also provide information about how pneumococci might interact with their environment – for example, by releasing metabolic end-products potentially involved in quorum sensing, or altering the physicochemical properties and nutrient availability of the extracellular space to generate an atmosphere more or less conducive for the survival of other species. Metabolic footprinting, therefore, has far-reaching implications in helping to build a picture of the space in which microbial communities reside (Douglas, 2020).

In the nasopharynx, *S. pneumoniae* must establish colonisation with a range of other microorganisms also occupying the same niche, whereas deeper sites of disease such as the blood and cerebrospinal fluid (CSF) are sterile. In this Chapter, gas chromatography–mass spectrometry (GC-MS) was used to analyse the metabolic footprint of *S. pneumoniae* TIGR4 cultured in either glucose (abundant in the blood and CSF) or galactose (abundant in the nasopharynx). The primary aim was to determine whether carbohydrate-specific, and therefore perhaps niche-specific, differences were evident in the extracellular metabolome.

During the course of these footprinting studies, extracellular accumulation of the tricarboxylic acid cycle (TCA cycle; a.k.a. Krebs cycle) intermediates succinate, fumarate and α -ketoglutarate (a.k.a. 2-oxoglutarate) was detected proportional to

pneumococcal growth and metabolic activity. The *S. pneumoniae* genome is not annotated to encode the enzymes required to operate a functional TCA cycle (Hoskins et al., 2001; Lanie et al., 2007; Tettelin et al., 2001), rendering these findings inexplicable in the context of existing knowledge of pneumococcal metabolism. This led us to hypothesise that *S. pneumoniae* TIGR4 operates a partial or non-canonical TCA cycle.

Lastly, the intracellular pneumococcal metabolome was also investigated by GC-MS to complement and validate the extracellular metabolic footprinting analysis. I performed isotope tracing analysis with *S. pneumoniae* TIGR4 and TIGR4 Δ *cps* using uniformly labelled ^{13}C -glucose ([U- $^{13}\text{C}_6$]-glc) to determine whether other TCA cycle intermediates not detected in the conditioned culture medium could be detected as part of the intracellular metabolome, and if the intracellular labelling patterns supported the presence of novel TCA-like activity.

4.2 Chapter 4 Aims

- Validate the metabolic footprinting method and determine the extracellular polar metabolome by GC-MS of *S. pneumoniae* TIGR4 when cultured in the complex medium THY.
- Investigate whether the primary carbon source – glucose or galactose – modifies the metabolic footprint of *S. pneumoniae* TIGR4 cultured in the semi-defined 7H9⁺ medium.
- Validate the extracellular accumulation of potential TCA cycle intermediates observed in *S. pneumoniae* TIGR4 cultured in 7H9⁺ medium, by studying an additional TIGR4 type strain (ATCC BAA-334):
 - Whole genome sequencing of TIGR4 to investigate any significant nucleotide variants compared to the parent reference genome.
 - Bioinformatic taxonomic classification of TIGR4 sequencing fragments to verify the purity of bacterial stocks.
 - Metabolic footprinting analysis of ATCC BAA-334 TIGR4 type strain.
- Perform a small-scale preliminary analysis to determine whether enzymatic domains involved in a classical TCA cycle might be present but unannotated in the *S. pneumoniae* TIGR4 genome, by investigating nucleotide and protein sequence homology (NCBI BLAST) and any conserved domains (NCBI Conserved Domain Database).
- Employ stable isotopologue profiling using uniformly labelled glucose ([U-¹³C₆]-glucose) to analyse the intracellular metabolite labelling patterns of *S. pneumoniae* TIGR4 cultured in 7H9⁺ medium, with a particular focus on uncovering whether succinate and fumarate identified in the extracellular footprinting analysis were produced via a variation of the TCA cycle.
 - Compare to the intracellular labelling patterns of a TIGR4Δ*cps* strain to determine the effect of the capsule on central carbon and amino acid metabolism.

4.3 Validation of a metabolic footprinting method

S. pneumoniae stocks in this study were cultured and stored in the complex THY medium, widely used to culture *S. pneumoniae* in the research community (Todd & Hewitt, 1932). Therefore, the metabolic footprint of *S. pneumoniae* TIGR4 cultured in THY over a 14-hour time-course experiment was analysed. This also provided an opportunity to validate the standard metabolic footprinting method used by the Metabolomics facility at the Crick (MacRae et al., 2013), which is described in **Chapter 2 Materials & Methods**.

4.3.1 Detection of glycolytic metabolites during pneumococcal culture

The primary carbohydrate source listed in the proprietary composition of Todd-Hewitt (TH) broth is dextrose (Oxoid). However, footprinting analysis found that THY contains at least two sugars, glucose and fructose, and that these sugars are taken up simultaneously by the pneumococcus (**Figure 4.1A**).

It is known that *S. pneumoniae* utilises carbon catabolite repression (CCR) to hierarchically metabolise preferential sugars (e.g., glucose) before non-preferred ones (e.g., galactose and fructose) (Fleming, 2016). The simultaneous depletion of fructose and glucose from the culture medium in **Figure 4.1A** suggests that transport of the non-preferred sugar fructose into pneumococcal cells is not inhibited under CCR, but likely exerts its action through regulation of genes and regulators involved in the metabolism of non-preferred sugars (Fleming & Camilli, 2016; Fleming et al., 2015).

Pyruvate is a key intermediate product of carbohydrate catabolism via glycolysis. The depletion of pyruvate from the culture medium observed in **Figure 4.1A** suggests that it can also be taken up by *S. pneumoniae* from the environment, entering the glycolytic pathway downstream of glucose/ fructose in THY. Pyruvate can then be further metabolised to fermentation products (lactate, acetate, formate, ethanol) and contribute to the biosynthesis of alanine by a transamination reaction (Echlin et al., 2020; Härtel et al., 2012) (**Figure 4.2**).

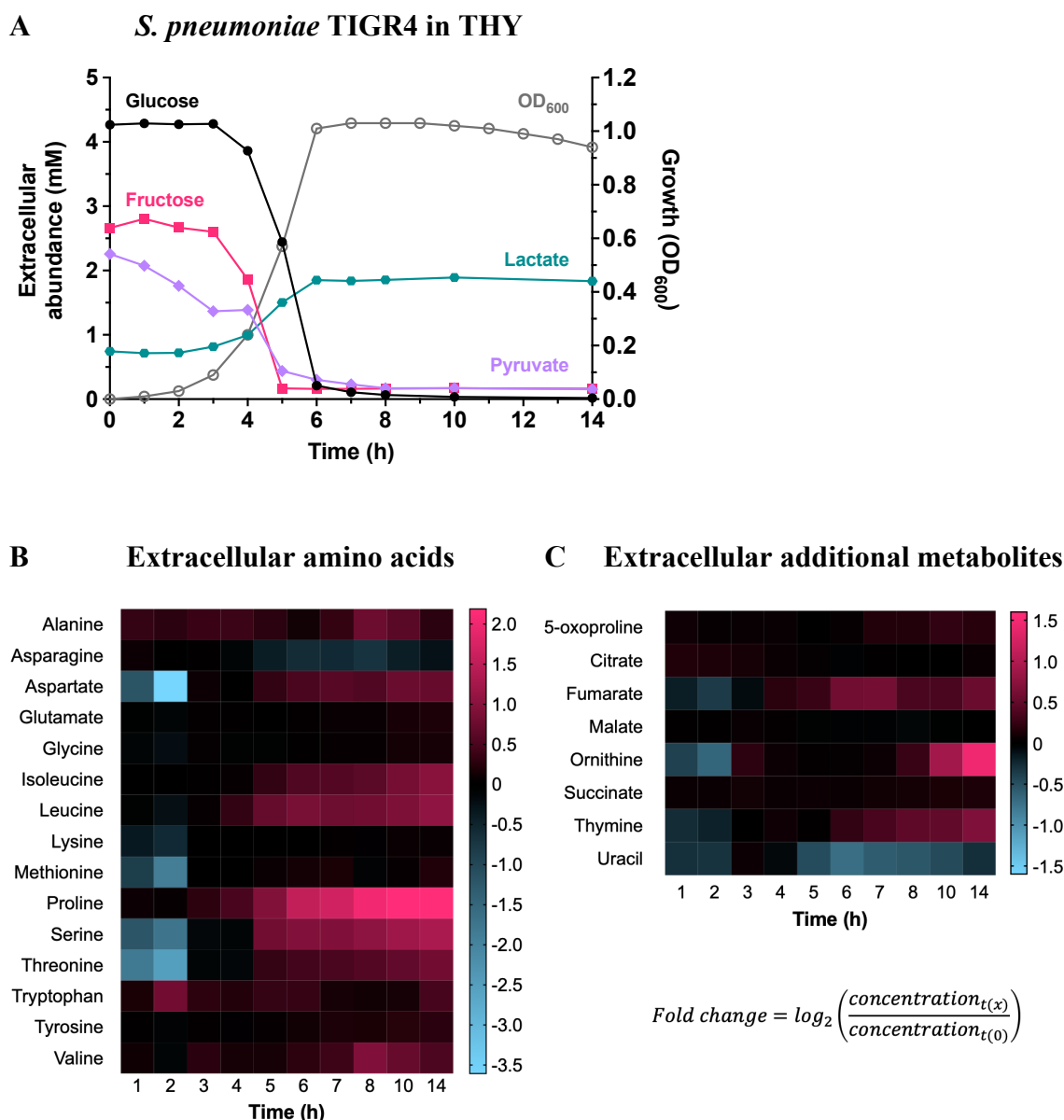


Figure 4.1 Metabolic footprint of *S. pneumoniae* TIGR4 cultured in THY

25 ml THY was inoculated with $\sim 3\text{-}5 \times 10^7$ CFU *S. pneumoniae* TIGR4 stock, and growth measured by OD_{600} . The metabolic composition of THY over time was analysed by GC-MS. $n = 1$.

The extracellular concentration was calculated by comparing the absolute abundance of each metabolite to the internal standard (A). Colour maps represent the \log_2 fold change in extracellular concentration for each metabolite (B, C), calculated using the embedded equation. Absolute concentration values are reported in *Appendix Table 7.3*. Amino acids not detected – arginine, cysteine, glutamine, histidine, lysine, and phenylalanine.

This footprinting analysis also detected the production and extracellular accumulation of lactate (**Figure 4.1A**) with similar dynamics to the consumption of glucose, fructose and pyruvate, and the rate of growth (as measured by OD₆₀₀). Together, these data suggest that pneumococcal uptake of glucose, fructose and pyruvate leads largely to lactate production. It is worth noting that metabolic pathways are in constant flux and the concentration of extracellular lactate does not reflect the total amount of lactate produced by pneumococcal cells – for e.g., lactate may be converted back to pyruvate through the action of lactate oxidase (*lctO*) (**Figure 4.2**).

Lactate is a major end-product of anaerobic glycolysis (fermentation), the primary energy-generating pathway in *S. pneumoniae*. The generation of lactate from pyruvate by lactate dehydrogenase (*ldh*) is a crucial step in carbohydrate catabolism as the reaction also regenerates the reducing agent NAD⁺ to power subsequent rounds of glycolysis (**Figure 4.2**). Others have shown that *S. pneumoniae* grown on glucose under anaerobic conditions undergoes homolactic fermentation, i.e., pyruvate is predominantly converted to lactate to complete glucose catabolism via glycolysis (Paixão et al., 2015a), which is supported by this data.

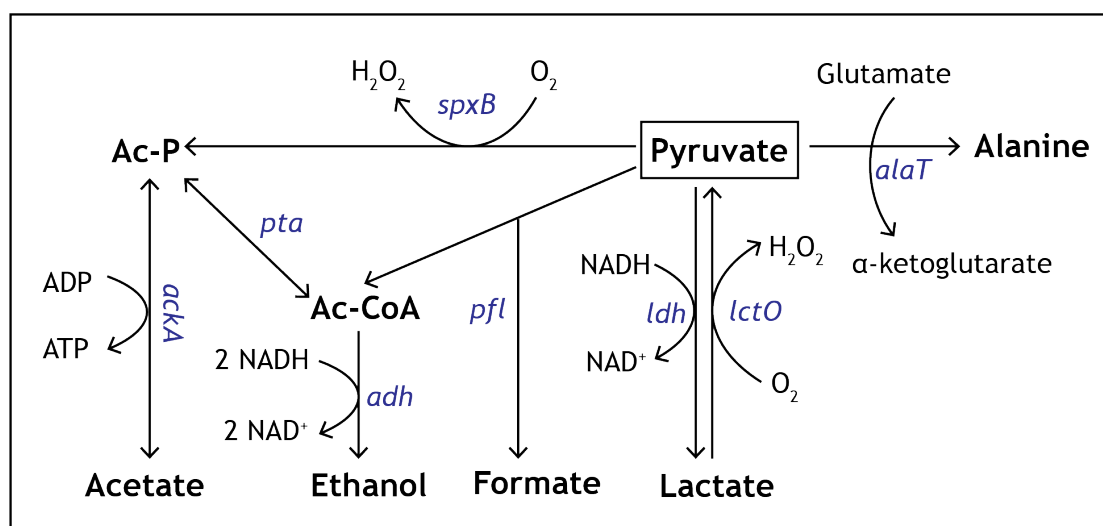


Figure 4.2 Pyruvate metabolism pathways in *S. pneumoniae*

These pathways were recreated from genome annotations and various studies in the literature. Abbreviations: Ac-P = acetyl-phosphate; Ac-CoA = acetyl-CoA.

However, pyruvate can also be metabolised by different enzymes to generate other products of pneumococcal fermentation – formate, acetate and ethanol; these metabolites are not detectable by GC-MS, making it difficult to independently determine the final fermentation profile of *S. pneumoniae* carbohydrate metabolism using GC-MS alone. The next Chapter will describe the use of ^1H -NMR to quantify levels of extracellular formate, acetate and ethanol.

4.3.2 Detection of amino acids and additional metabolites

15 amino acids were detected in the spent culture medium THY during the growth of TIGR4 (**Figure 4.1B**). Although the TIGR4 genome shows an incomplete pathway for proline biosynthesis derived from glutamate, extracellular proline accumulation was observed over time, suggesting *de novo* synthesis (**Figure 4.1B**). Experiments with isotopically labelled glutamate will be required to confirm whether proline biosynthesis proceeds from environmentally acquired glutamate through a series of reactions that are currently unannotated. A group of *S. pneumoniae* virulence factors, pneumococcal surface proteins (Psp), contain a cross-reactive proline-rich domain, which may point to the importance of proline biosynthesis. The proline-rich region has also been proposed as a broad-spectrum vaccine target as it elicits a protective antibody response in mice (Daniels et al., 2010; Mukerji et al., 2018).

Previously, an isotope labelling study with *S. pneumoniae* D39 Δ *cps* in a chemically defined medium containing 55.5 mM [U- $^{13}\text{C}_6$]-glucose showed that the bacterium preferentially acquires branched chain amino acids (BCAA) isoleucine, leucine and valine from its environment and cannot grow in their absence, despite containing BCAA biosynthetic enzymes (Härtel et al., 2012). However, in the footprinting analysis reported here, *S. pneumoniae* TIGR4 cultured in the complex medium THY showed an increase in the extracellular levels of isoleucine and leucine (**Figure 4.1B**). This indicates that the regulation of BCAA biosynthesis may vary with pneumococcal strains or the type of culture medium used, which requires further confirmation.

Härtel and colleagues (Härtel et al., 2012) found that *S. pneumoniae* was also auxotrophic for glycine in their study conditions. In agreement, another study profiling the metabolome of *S. pneumoniae* TIGR4 Δ *cps* in a modified RPMI medium (Leonard et al., 2018), showed that extracellular levels of isoleucine, leucine, valine and glycine decreased over time. These studies contrast with the observation here that when TIGR4 is grown in THY, there is no noticeable depletion of extracellular glycine (**Figure 4.1B**).

The discrepancies emphasise that insights gleaned about pneumococcal amino acid metabolism must be interpreted within the context of the specific media and culture conditions, for the results of one *in vitro* study may not necessarily be applicable to a different pneumococcal strain growing in different culture conditions (Sanchez-Rosario & Johnson, 2021). It must also be noted that this pilot experiment in THY was only performed once, as the primary aim was to validate the metabolic footprinting method for *S. pneumoniae*, and should be repeated to ensure that the differential observations about amino acid uptake and biosynthesis are upheld.

Aspartate, serine, and threonine accumulated extracellularly, as would be expected (**Figure 4.1B**). All three amino acids incorporate carbons from glycolytic carbohydrate catabolism, with serine biosynthesis occurring through a unique hydroxymethylation of glycine involving formate. The glycolytic intermediate phosphoenolpyruvate (PEP) is converted to oxaloacetate via PEP carboxylase (*ppc*), which undergoes a transamination reaction via *aspC* to generate aspartate, and aspartate can in turn be converted to threonine (Härtel et al., 2012).

The same temporal extracellular pattern seen for aspartate is also observed in fumarate and ornithine (**Figure 4.1C**), all three of which are intermediates of an intact urea cycle (a.k.a. ornithine cycle). Although TIGR4 genome analysis shows an incomplete urea cycle (Chapter 5 **Figure 5.12A**), another study has provided supporting evidence for the presence of a pneumococcal urea cycle (Leonard et al., 2018).

Other metabolites detected in this footprinting analysis include potential tricarboxylic acid (TCA) cycle intermediates citrate, succinate, and malate. Genomic inference suggests the absence of the TCA cycle and electron transport chain in the pneumococcus (Hoskins et al., 2001; Lanie et al., 2007; Tettelin et al., 2001; Willenborg & Goethe, 2016), which is supported by this data showing that extracellular concentrations of citrate, succinate and malate in *S. pneumoniae*-conditioned THY remained unchanged (**Figure 4.1C**).

Uracil depletion was also observed over time – uracil is an important component of uridine diphosphate (UDP-) sugars, which are precursors of outer polysaccharide capsule biosynthesis (Carvalho et al., 2013b) (Chapter 1 **Figure 1.2**).

4.4 Primary carbon source does not affect the metabolic footprint of *S. pneumoniae*

It has previously been reported that *S. pneumoniae* grown on the non-preferred sugar galactose undergoes mixed acid fermentation through the increased action of pyruvate formate lyase (*pfl*) generating formate and acetate, competing with lactate dehydrogenase (*ldh*) to generate lactate (**Figure 4.2**). On the fast-metabolising and preferred sugar glucose, on the other hand, predominantly homolactic fermentation is observed (Al-Bayati et al., 2017; Carvalho et al., 2013b; Paixão et al., 2015a; 2015b). Additionally, glucose and galactose are differentially abundant in the blood and the respiratory tract, respectively. Therefore, I wanted to compare whether the metabolic footprint of *S. pneumoniae* (beyond fermentation) was altered when supplied with either glucose or galactose as the primary carbon source.

S. pneumoniae TIGR4 was cultured in the semi-defined 7H9⁺ medium containing either 11 mM glucose or 11 mM galactose. Time-course metabolic footprinting analyses by GC-MS were carried out to compare the extracellular metabolite composition of *S. pneumoniae*-conditioned culture medium over time (**Figure 4.3**).

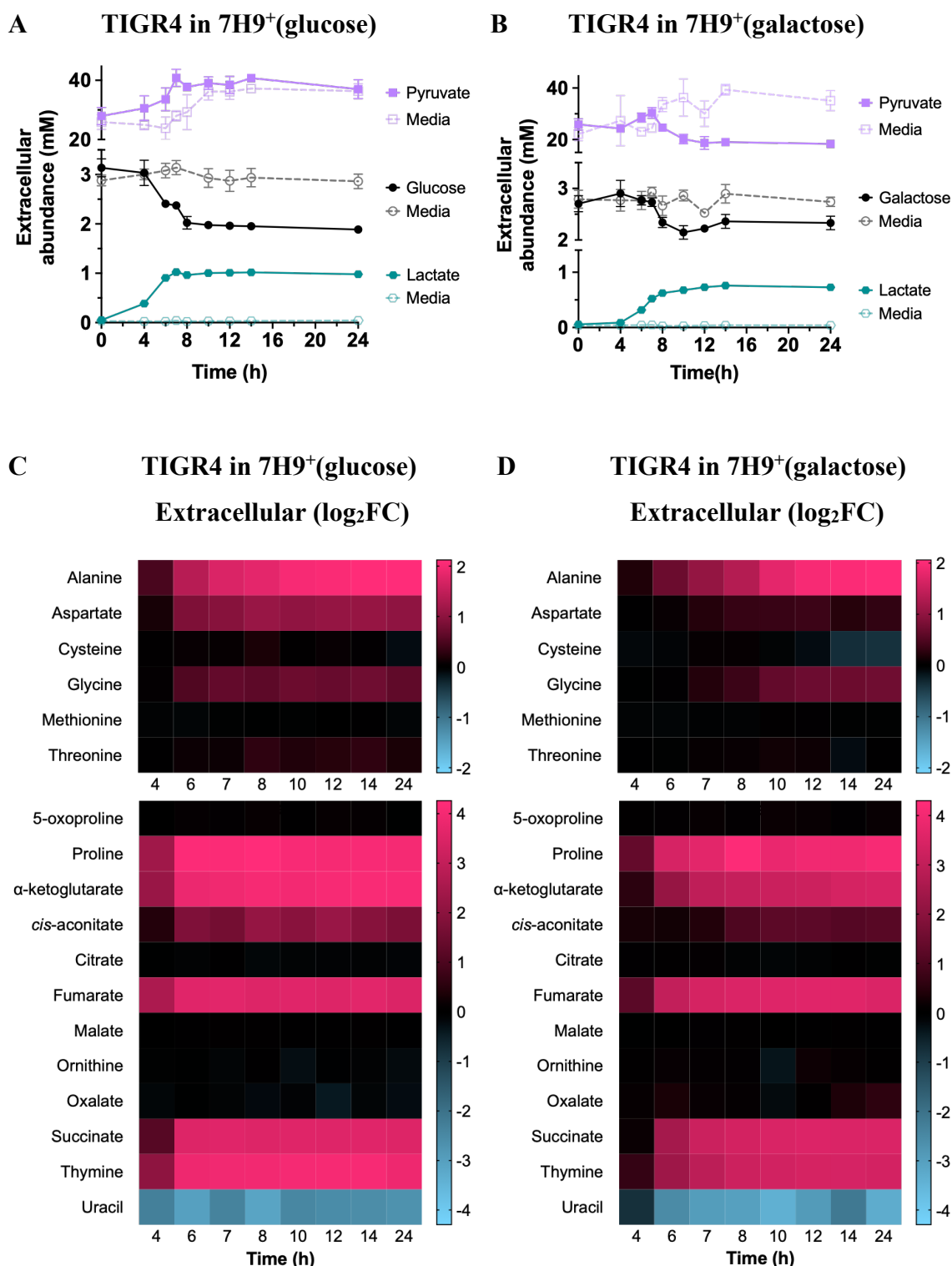


Figure 4.3 *S. pneumoniae* metabolic footprint in glucose and galactose (7H9⁺ medium)

Extracellular metabolite concentrations were calculated by comparing the absolute abundance of each metabolite to the internal standard. (A, B) Glycolytic metabolites over time compared to their concentration in the uninoculated medium (“media”). Plots depict mean and standard deviation. (C, D) Colour maps represent the mean log₂ fold change (log₂FC) in extracellular metabolite concentration over time, relative to t₀. n = 3.

I saw that although 11 mM of sugar was the minimal concentration required for the pneumococcal growth in 7H9⁺ (described in **Chapter 3**), the carbohydrate chromatographic peaks at this concentration were saturated, and so the extracellular abundances of glucose and galactose were under-reported as measured by GC-MS (**Figure 4.3A, B**). Nevertheless, it was apparent that neither glucose nor galactose were completely depleted in 7H9⁺ medium, suggesting a different limiting factor leading to premature growth arrest. The emergence of extracellular lactate during growth on galactose was delayed when compared to growth in glucose, which is in line with the longer lag phase observed for *S. pneumoniae* cultured in galactose discussed in the previous Chapter.

Colour maps depicting the mean log₂ fold change (log₂FC) of three biological replicates in extracellular metabolite concentrations relative to time 0 (t₀) revealed that the metabolic footprints of *S. pneumoniae* TIGR4 were comparable when grown in glucose or galactose (**Figure 4.3C, D**). Plots showing the actual extracellular amount of each detectable metabolite over time can be found in *Appendix Figure 7.1*.

Of the amino acids, a relatively high accumulation of alanine and proline was observed in the extracellular space with pneumococcal growth in both glucose and galactose. Alanine is generated from pyruvate through a transamination reaction (Härtel et al., 2012). Proline is likely generated from exogenous glutamate, but the proline biosynthesis pathway in *S. pneumoniae* genome is incomplete (Tettelin et al., 2001). The extracellular accumulation of proline was also observed for TIGR4 cultured in THY (**Figure 4.1B**).

Amino acids not shown on the colour maps in **Figure 4.3** either because their abundance was too low to be accurately quantified or because no change in concentration was observed over time compared to the negative control (medium only) were asparagine, glutamine, glutamate, isoleucine, leucine, lysine, phenylalanine, serine and valine. The absolute concentrations of these amino acids are plotted in *Appendix Figure 7.1*. Histidine, tryptophan, and tyrosine were not quantified due to low quality

peaks, and arginine is not detectable using the GC-MS derivatisation method used in this study.

4.4.1 Extracellular accumulation of potential TCA cycle intermediates

The most surprising finding in this footprinting analysis was the extracellular accumulation of succinate and *cis*-aconitate in both glucose- and galactose-supplemented *S. pneumoniae* (**Figure 4.3C, D**). Succinate and *cis*-aconitate are intermediates of the TCA cycle, a pathway absent in *S. pneumoniae* according to genomic analyses. Other intermediates of the TCA cycle which also showed a corresponding increase in log₂FC were α -ketoglutarate (α -KG) and fumarate (illustrated in **Figure 4.4**).

It is known that α -KG can be generated from exogenous glutamate through the action of an aspartate transaminase (*aspC*), which simultaneously converts oxaloacetate to aspartate (Härtel et al., 2012). Oxaloacetate is also a TCA cycle intermediate, but in *S. pneumoniae* it can be generated through carboxylation of the glycolytic intermediate phosphoenolpyruvate (PEP) (Tettelin et al., 2001) (**Figure 4.4**).

Intracellular fumarate production in TIGR4 Δ *cps* was also observed by Leonard and colleagues, who hypothesised that fumarate may be an intermediate of the urea cycle generated during ornithine production (Leonard et al., 2018).

The production of succinate by TIGR4, however, cannot be explained by current genomic knowledge of pneumococcal metabolic pathways. The intracellular isotope labelling study with *S. pneumoniae* D39 by Härtel and colleagues did not mention succinate or fumarate, but this is likely because they were specifically looking for amino acids in Selected Ion Monitoring (SIM) mode of GC-MS and did not look for TCA intermediates (Härtel et al., 2012). The global profiling study by Leonard *et al* did detect intracellular succinate in TIGR4 Δ *cps*, but this was not specifically commented on or investigated further (Leonard et al., 2018).

Although extracellular accumulation of citrate and malate, the remaining TCA cycle intermediates, was not observed, we cannot rule out the possibility of their production without looking at the intracellular metabolome (discussed in **Section 4.7**).

Variant TCA cycles have been reported in other bacteria adapting to particular niches. For example, *Helicobacter pylori*, which lives in anaerobic/ microaerophilic environments, performs a noncyclic branched pathway culminating in succinate production by α -ketoglutarate oxidase (α -KG \rightarrow succinate) and fumarate reductase (fumarate \rightarrow succinate) (Pitson et al., 1999). *Escherichia coli* is capable of doing the same under anaerobic conditions, while operating a full TCA cycle under aerobic conditions (Guest, 1995). Other members of the *Streptococcus* spp, *S. gordonii* and *S. mutans*, encode a fragmentary TCA cycle (**Figure 4.4**), although its function is unknown (Willenborg & Goethe, 2016).

In this context and on the basis of the footprinting analysis described in this section, I therefore hypothesised that a version of the TCA cycle is present in *S. pneumoniae* TIGR4. The remainder of this Chapter will describe how this hypothesis was tested.

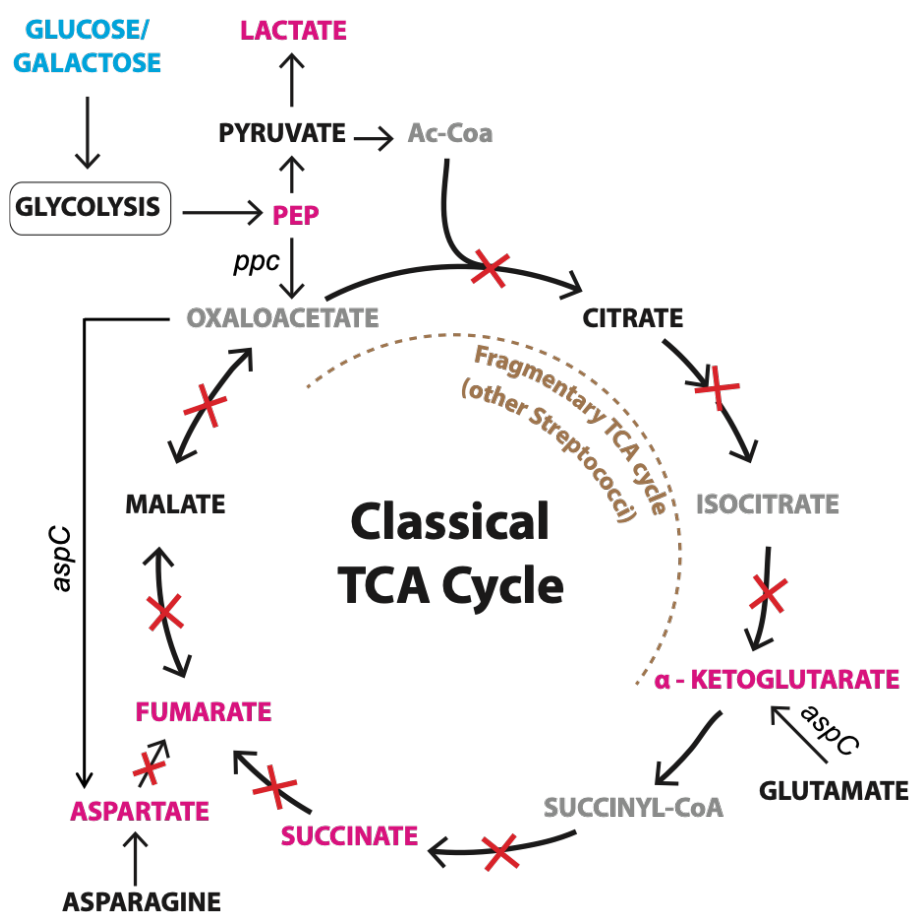


Figure 4.4 Schematic showing *S. pneumoniae* lacks a classical TCA cycle

This schematic was constructed using the genomic information available for *S. pneumoniae* TIGR4 from the KEGG metabolic database (Kanehisa & Goto, 2000). Enzymes not annotated in the TIGR4 genome are indicated by a red cross. Over 24h of TIGR4 growth on glucose or galactose as primary carbon sources, extracellular metabolite levels that increased are shown in magenta; that decreased are shown in blue; that remained unchanged are shown in black; and undetected metabolites are shown in grey.

4.5 Confirming the production of potential TCA cycle intermediates by *S. pneumoniae* TIGR4

Due to the unexpected nature of potential TCA cycle intermediates accumulating in the extracellular medium during TIGR4 growth, I wanted to confirm the purity of the TIGR4 stock used (i.e., whether a contaminant was producing these unexpected intermediates) and validate these results using the TIGR4 type strain ATCC BAA-334 purchased from the American Type Culture Collection (ATCC).

First, whole genome sequencing (WGS) of the laboratory TIGR4 strain P1542 (“TIGR4”) and the TIGR4 type strain (“ATCC-TIGR4”) was carried out to determine whether the laboratory strain P1542 had developed significant genomic variation. We then also used the metagenomic classification system Kraken (Wood & Salzberg, 2014) to rule out the possibility of any contaminating bacterial species in our stocks.

4.5.1 Whole genome sequencing of *S. pneumoniae* TIGR4

The protocols for *S. pneumoniae* DNA extraction, library preparation and paired-end sequencing are described in **Chapter 2.1.3**. Paired-end (100 bp) genomic sequencing at a depth of 2 million reads was carried out on the Illumina HiSeq4000 system for the laboratory TIGR4 and ATCC-TIGR4 strains. Sequenced reads were aligned to the TIGR4 reference genome AE005672.3 (Tettelin et al., 2001). Characteristics of read alignment, reference genome coverage and nucleotide variation are highlighted in **Table 4.1**.

We obtained excellent alignment and coverage of the reference genome AE005672.3 – 98.9% of the laboratory TIGR4 reads and 98.7% of the ATCC-TIGR4 reads were aligned to the reference genome, with 0.01% of the reads for both reads being unmapped. The reference genome coverage with both strains was 99%. The number of nucleotide variants for both strains when aligned to the reference were similar – 65 in laboratory P1542 TIGR4 and 63 in ATCC-TIGR4. A complete list of the variants and their chromosomal positions can be found in *Appendix Table 7.4*.

These results confirm that the laboratory TIGR4 strain has not developed any significant genomic variation compared to the originally sequenced TIGR4 genome, and that information about metabolic pathways inferred from the reference TIGR4 genome (Tettelin et al., 2001) can be applied to the laboratory TIGR4 strain used in this thesis. This also confirms that neither mutations nor gene acquisition by horizontal gene transfer explain the appearance of the unexpected metabolic intermediates identified in this study.

	TIGR4-lab (P1542)	ATCC-TIGR4 (BAA-334)
<i>Read alignment characteristics</i>		
Raw R1 & R2	13491872	22321162
Trimmed R1 & R2	13360874	22070676
Aligned R1 & R2	13339352	22028391
% Aligned	98.9	98.7
Deduplicated R1 & R2	11168874	17276849
% Deduplicated	82.8	77.4
Unmapped R1 & R2	1490	3208
% Unmapped	0.01	0.01
<i>Reference genome characteristics</i>		
Total Genome Size	2160842	2160842
Zero Coverage	990	1
Non-Zero Coverage	2159852	2160841
100x Coverage¹	99%	99%
<i>Nucleotide variation compared to reference</i>		
Number of variants	65	63

Table 4.1 Characteristics of *S. pneumoniae* TIGR4 sequenced reads

100 bp paired-end sequencing was performed. R1 and R2 refer to complementary reads. Reads were aligned to the TIGR4 reference genome AE005672.3 (Tettelin et al., 2001).

¹100x coverage indicates the percentage of the reference genome covered by the sequenced reads at 100x sequencing depth.

4.5.2 Metagenomic classification of *S. pneumoniae* TIGR4 sequence reads

The raw sequenced reads from *S. pneumoniae* TIGR4 and ATCC-TIGR4 were also analysed with the metagenomic classification platform Kraken (Wood & Salzberg, 2014), typically used for microbiome analysis. Kraken analysis was performed with guidance from Dr. Andrea Gori, Mucosal Pathogens Research Unit, UCL. Kraken assigns taxonomic labels to short DNA sequences, essentially identifying the bacterial species to which each read belongs. It is faster than machine learning techniques, and its accuracy is comparable to NCBI's MegaBLAST.

TIGR4 BAA-334 was included as a control, as it was purchased from ATCC as a pure stock of the TIGR4 type strain. Kraken analysis showed that ~ 99.3% of the classified reads of both strains were assigned to the *Streptococcus* spp. genus (**Table 4.2**).

The only species besides *Streptococcus pneumoniae* to have a notable number of sequenced reads assigned was *Streptococcus mitis*, and even then, at a negligible amount of 0.01% for both strains tested (**Table 4.2**). *S. mitis* is an oral commensal and a frequent genetic exchange partner for *S. pneumoniae* through horizontal gene transfer (mainly virulence and antibiotic resistance genes) (Salvadori et al., 2019). Therefore, this is likely a misassignment.

66 - 67% of all reads were assigned to *S. pneumoniae* in this analysis (**Table 4.2**).

Although this may not intuitively seem like an indicator of a single-species sample, it is of greater importance that this percentage was consistent in the control ATCC-TIGR4 sample, indicating that the reads mapped to taxa likely arise from unavoidable sources – such as the DNA purification kit (Salter et al., 2014) – and not the bacterial stock itself.

Therefore, this analysis confirms the purity of the laboratory TIGR4 stock, providing confidence that the unusual accumulation of TCA cycle intermediates in conditioned culture medium observed was indeed a consequence of the metabolic activity of *S. pneumoniae* rather than any other bacterial species.

Taxonomy	Number of reads assigned		% reads covered by clade rooted at this taxon	
	<i>TIGR4-lab</i> (P1542)	<i>ATCC-TIGR4</i> (BAA-334)	<i>TIGR4-lab</i> (P1542)	<i>ATCC-TIGR4</i> (BAA-334)
unclassified	439579	748662	6.58	6.78
root	1596	2560	93.42	93.22
cellular organisms	21	28	93.4	93.19
Bacteria	8072	13050	93.4	93.19
Terrabacteria group	5277	8608	93.27	93.07
Firmicutes	6121	10374	93.19	92.99
Bacilli	6730	10685	93.1	92.9
Lactobacillales	10282	16752	93	92.8
Streptococcaceae	8897	14662	92.84	92.65
Streptococcus	1733110	2873243	92.71	92.52
<i>S. pneumoniae</i>	4145026	6816316	66.74	66.45
<i>S. mitis</i>	424	754	0.01	0.01

Table 4.2 Metagenomic characterisation of *S. pneumoniae* TIGR4 sequence reads

The most relevant taxa accounting for > 99% of reads assigned by the taxonomic classification program Kraken (Wood & Salzberg, 2014) are shown.

4.5.3 Extracellular metabolites of *S. pneumoniae* ATCC-TIGR4

After genomically validating the accumulation of potential TCA cycle intermediates in the extracellular medium as a consequence of *S. pneumoniae* TIGR4 growth in 7H9⁺ medium, I hypothesised that these same metabolites would also accumulate during the growth of other TIGR4 strains under the same conditions.

To test this, the metabolic footprint of the TIGR4 type strain BAA-334 cultured in 7H9⁺ medium (+11 mM glucose) was compared to that of the laboratory TIGR4 strain P1542 used thus far. Both strains exhibit similar growth curves (**Figure 4.5**), which correspond to glucose depletion and lactate accumulation in the culture medium (**Figure 4.6A, B**).

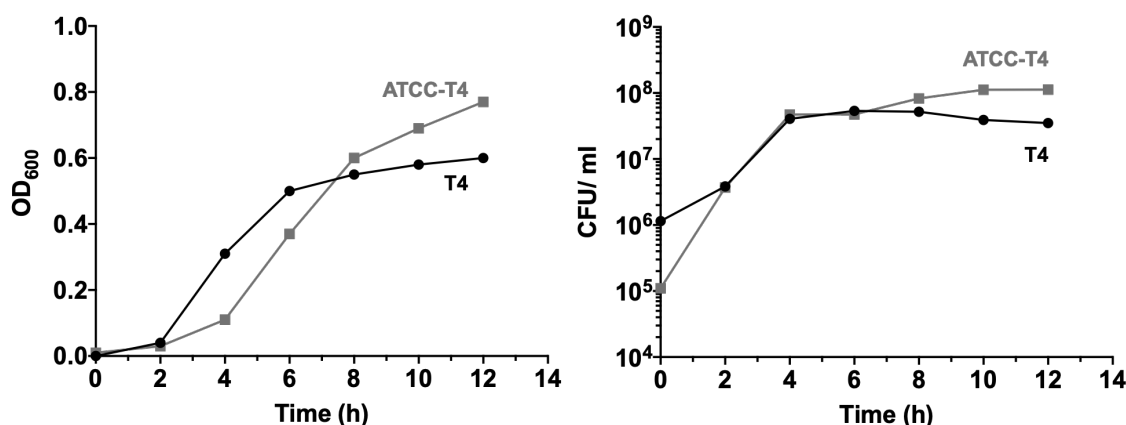


Figure 4.5 *S. pneumoniae* growth in 7H9⁺(glucose)

25 ml 7H9⁺ medium with 11 mM glucose was inoculated with $\sim 3\text{--}5 \times 10^7$ CFU *S. pneumoniae* stock. Growth was monitored by OD₆₀₀ (A) and viable CFU ml⁻¹ (B). n = 1.

The extracellular accumulation of potential TCA cycle metabolites and related amino acids aspartate and glutamate were comparable in both strains within this experiment, providing further validation of the findings described earlier in this Chapter (**Figure 4.6C, D**).

However, some differences were observed in this experiment in comparison with the metabolic footprint of *S. pneumoniae* TIGR4 in **Section 4.4** (**Figure 4.3**). The log₂FC in extracellular succinate was lower than that of fumarate, and α -ketoglutarate was not shown here due to low quality chromatographic peaks. It is worth noting that α -ketoglutarate is known to be notoriously difficult to detect by GC-MS because it rapidly interconverts with glutamate and 5-oxoproline.

These discrepancies are likely methodological rather than biological in nature, either as a result of lower pneumococcal biomass leading to an overall lower abundance in extracellular metabolites, or dynamic metabolite processing that cannot be captured by looking at the metabolic footprint alone. Therefore, definitive conclusions could not be made without analysing intracellular metabolite levels and performing isotope tracing studies, which are described later in **Section 4.7** of this Chapter.

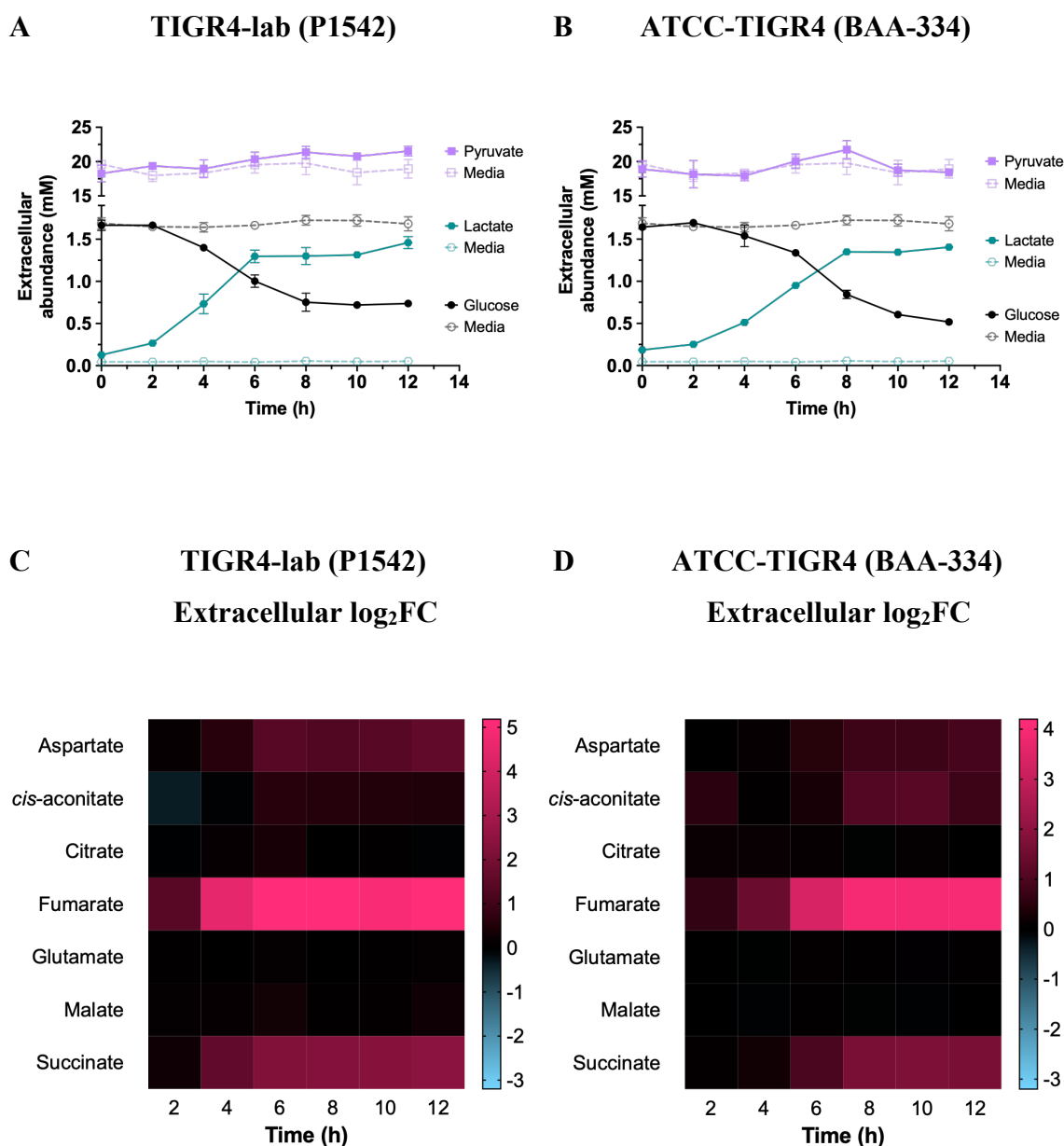


Figure 4.6 Extracellular metabolites of two *S. pneumoniae* TIGR4 strains grown in 7H9⁺(glucose) medium.

Metabolite concentrations were calculated by comparing the absolute abundance of each metabolite to the internal standard. (A, B) Glycolytic metabolites over time compared to their concentration in the uninoculated medium (“media”). Plots depict mean and standard deviation. (C, D) Colour maps represent the mean log₂ fold change (log₂FC) in extracellular “TCA” metabolite concentration over time, relative to t₀. n = 4.

4.6 Are there unannotated enzymes capable of TCA cycle-related activity in *S. pneumoniae* TIGR4?

Genomic analyses can provide a wealth of information about an organism's physiology, but inferences about phenotypic consequences are heavily reliant on the degree and accuracy of genome annotation. While the *S. pneumoniae* TIGR4 genome is reasonably well-annotated, gene annotations are largely made based on sequence homology of known proteins. Of the enzymes involved in operating a classical TCA cycle, TIGR4 is reported to encode only α -ketoglutarate dehydrogenase (α -KG \rightarrow succinyl-CoA). I wanted to determine whether there were any hints for TCA-like enzyme activity encoded within the pneumococcal genome that may not have been detected by automated computational methods used to annotate genomes.

As a preliminary investigation, I compared nucleotide and protein sequences of the *S. pneumoniae* TIGR4 genome (AE005672.3) with TCA cycle enzyme sequences from *Escherichia coli*⁸ and *Bacillus subtilis*⁹ using NCBI Basic Local Alignment Search Tool (BLAST) to determine sequence homology. This was followed by a conserved domain search on the NCBI Conserved Domain Database (CDD) of the top 5 protein sequence alignments for each gene regardless of the degree of homology. *E. coli* was chosen because it is exceptionally well-characterised and not closely related to *S. pneumoniae* phylogenetically, while *B. subtilis* was selected as a closer relative of *S. pneumoniae* operating a TCA cycle.

Nucleotide BLAST (blastn)¹⁰, protein BLAST (blastp)¹¹ and translated nucleotide BLAST (tblastx)¹² did not yield significant sequence homology for any of the TCA cycle nucleotide or protein sequences from *E. coli* and *B. subtilis* tested, with the exception of the malate dehydrogenase (*mdh*) of *B. subtilis* (malate \leftrightarrow oxaloacetate)

⁸ Query gene list for *E. coli* = https://www.genome.jp/kegg-bin/show_pathway?eco00020

⁹ Query gene list for *B. subtilis* = https://www.genome.jp/kegg-bin/show_pathway?bsu00020

¹⁰ blastn = searches nucleotide database using a nucleotide query

¹¹ blastp = searches protein database using a protein query

¹² tblastx = searches translated nucleotide databases using the translated nucleotide query

(Table 4.3). This was the only condition with an alignment score greater than that for α -ketoglutarate dehydrogenase, which is known to be present in TIGR4.

A conserved domain search of the TIGR4 tblastx alignment with *B. subtilis mdh* revealed the protein motif ‘lactate/ malate dehydrogenase’ (Figure 4.7), indicating potential malate dehydrogenase activity which has not been described for *S. pneumoniae* before. However, this activity remains to be experimentally tested. In conclusion, I found that there are likely no conserved domains with canonical enzymatic activity that may be involved in TCA cycle transformations in *S. pneumoniae*.

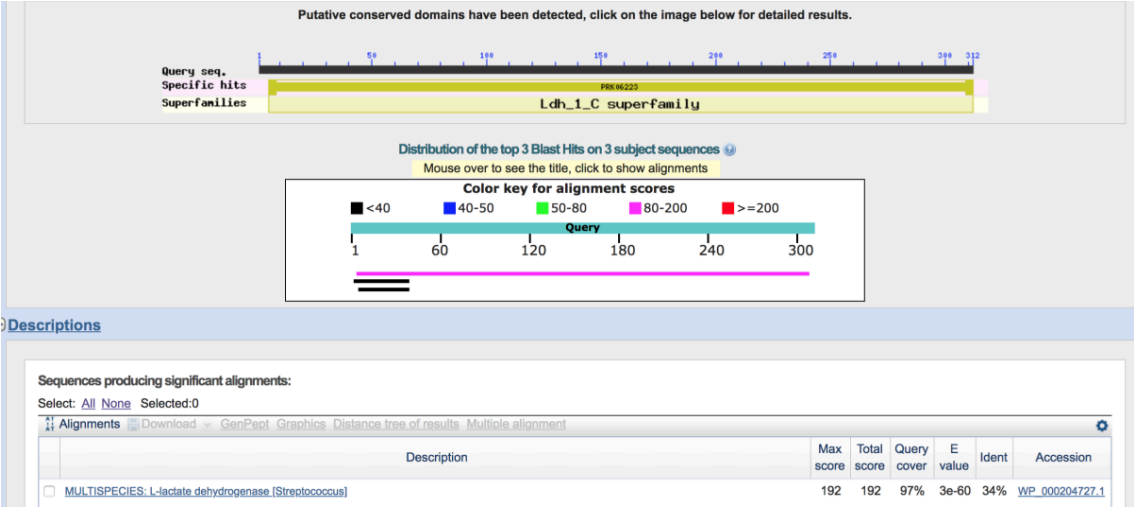
ENZYME	Alignment to <i>S. pneumoniae</i> TIGR4 genome (score) ¹							
	GENE	<i>E. coli</i> K-12 MG1655			<i>B. subtilis</i> subsp. subtilis 168			GENE
		blastn	tblastx	blastp	blastn	tblastx	blastp	
Citrate synthase	<i>gltA</i>	0.014	0.031	0.028	0.017	0.035	0.000	<i>citA</i>
					0.017	0.037	0.000	<i>mmgD</i>
					0.018	0.045	0.000	<i>citZ</i>
Aconitase	<i>acnB</i>	0.006	0.021	0.000	0.007	0.018	0.000	<i>citB</i>
	<i>acnA</i>	0.007	0.022	0.000				
	<i>ybhJ</i>	0.009	0.017	0.017				
Isocitrate dehydrogenase	<i>icd</i>	0.015	0.044	0.034	0.016	0.041	0.000	<i>icd</i>
α -ketoglutarate dehydrogenase	<i>sucA</i>	0.007	0.013	0.014	0.007	0.013	0.014	<i>sucA</i>
	<i>sucB</i>	0.014	0.167	0.221	0.015	0.156	0.179	<i>odhB</i>
Succinyl-CoA synthetase	<i>sucC</i>	0.016	0.036	0.000	0.017	0.034	0.034	<i>sucC</i>
	<i>sucD</i>	0.020	0.042	0.042	0.019	0.038	0.000	<i>sucD</i>
Succinate dehydrogenase	<i>sdhA</i>	0.011	0.026	0.026	0.011	0.028	0.023	<i>sdhA</i>
	<i>sdhB</i>	0.023	0.049	0.000	0.023	0.053	0.000	<i>sdhB</i>
	<i>sdhC</i>	0.043	0.085	0.000	0.032	0.057	0.000	<i>sdhC</i>
	<i>sdhD</i>	0.046	0.110	0.095				
Fumarate reductase	<i>frdA</i>	0.010	0.023	0.026				
	<i>frdB</i>	0.021	0.050	0.000				
	<i>frdC</i>	0.045	0.080	0.000				
	<i>frdD</i>	0.044	0.103	0.000				
Fumarase	<i>fumA</i>	0.012	0.022	0.026	0.015	0.040	0.028	<i>fumC</i>
	<i>fumB</i>	0.011	0.026	0.025				
	<i>fumC</i>	0.012	0.028	0.027				
Malate dehydrogenase	<i>mdh</i>	0.019	0.062	0.000	0.020	0.198	0.308	<i>mdh</i>
	<i>mgo</i>	0.011	0.023	0.000				

Table 4.3 BLAST analysis of *S. pneumoniae* TIGR4 against TCA cycle enzyme sequences

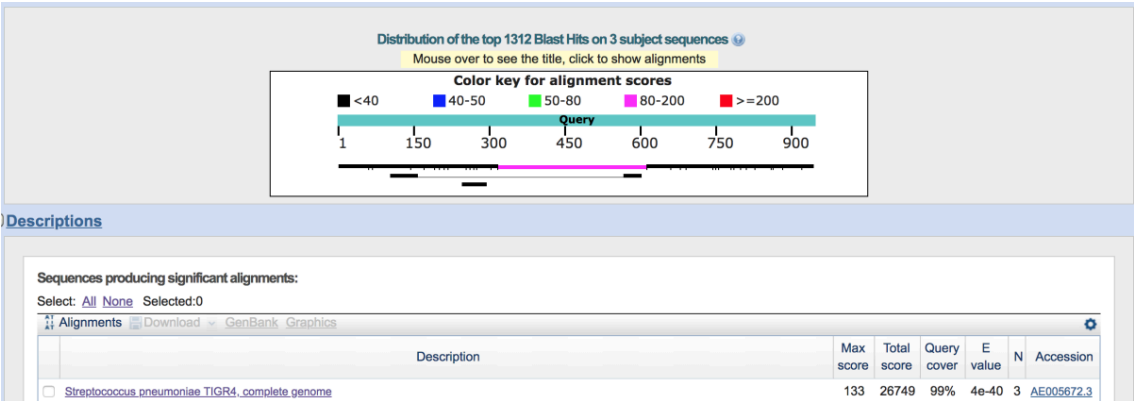
For each sequence aligned to the TIGR4 genome, a ‘max score’ is generated by the BLAST scoring algorithm, depicting the best single alignment.

¹Alignment scores in this table are calculated as $\frac{\text{Max score (query-subject alignment)}}{\text{Max score (query-query alignment)}}$. The maximum possible score, therefore, is 1 and corresponds to 100% sequence homology.

A blastp – Query = *B. subtilis mdh* amino acid sequence



B tblastx – Query = *B. subtilis mdh* nucleotide sequence



C Conserved domain search of the magenta aligned region from (B) above

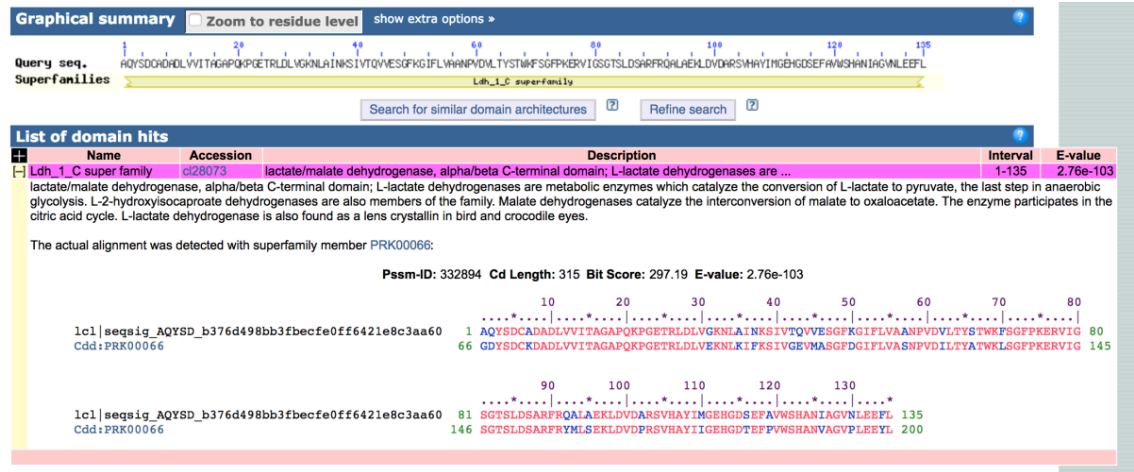


Figure 4.7 An example BLAST and conserved domain search analysis.

The analysis workflow is illustrated here using malate dehydrogenase (*mdh*) from *Bacillus subtilis* subsp. *subtilis* 168.

4.7 ¹³C-glucose labelling of the intracellular metabolites of *S. pneumoniae* TIGR4 and TIGR4Δ*cps*

To complement extracellular metabolite profiling of *S. pneumoniae* TIGR4, the intracellular metabolome was analysed. This was done by using uniformly labelled ¹³C-glucose ([U-¹³C₆]-glucose), which helps illustrate carbohydrate metabolism by tracing the route of labelled carbon atoms through metabolic pathways.

Aim 1: Determine whether the *S. pneumoniae* capsule affects carbon flow through associated non-glycolytic intermediates and amino acids by analysing all the intracellular polar metabolites that can reliably be quantified by GC-MS, using a TIGR4 capsule-deficient mutant (TIGR4Δ*cps*). The overlap between glycolysis and capsule biosynthesis has been well-documented (Ayoola et al., 2019; Carvalho et al., 2013a; Echlin et al., 2016; 2020; Hamaguchi et al., 2018; Hardy et al., 2000; Hathaway et al., 2012; Troxler et al., 2019).

Hypothesis: The lack of a capsule will direct carbon flow towards anabolic pathways as the energetic requirement for capsule production is mitigated.

Aim 2: Determine whether succinate and fumarate, which accumulated extracellularly in earlier metabolic footprinting experiments, are also detectable inside metabolising pneumococci, and whether they incorporate labelled carbon atoms derived from ¹³C-glucose in a manner indicative of TCA cycle-like activity.

S. pneumoniae was grown to early-mid log phase under standard conditions in 7H9⁺ medium containing 11 mM glucose (unlabelled), following which bacteria were pelleted by centrifugation and an equal volume of 7H9⁺ medium containing 11 mM ¹³C-glucose was swapped in. Bacteria were cultured in the labelled medium for 2 hours to allow achievement of metabolic steady state, at which point metabolism was rapidly quenched.

The amount of viable colony-forming units obtained for metabolite extraction after quenching was comparable between TIGR4 and TIGR4Δ*cps* (**Figure 4.8**). Bacterial cells were then lysed, and further sample preparation steps preceded the analysis of intracellular metabolites by GC-MS (detailed in **Chapter 2.5**).

The stable isotope labelling study in *S. pneumoniae* by Härtel and colleagues (Härtel et al., 2012) provides a reference against which to assess my results. That study used the lab-adapted *S. pneumoniae* D39Δ*cps*, culturing it in excess [U-¹³C₆]-glucose at 55.5 mM (compared to 11 mM glucose/ galactose used in my experiments), and placed a specific focus on amino acid metabolism.

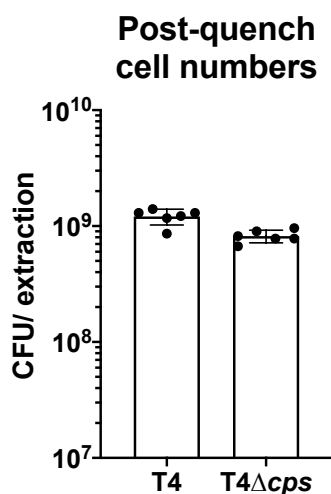


Figure 4.8 CFU of *S. pneumoniae* TIGR4±*cps* per extraction in 7H9⁺(glc) medium

25 ml 7H9⁺(glc) was inoculated with ~ 5-7x10⁷ CFU of *S. pneumoniae* stock and cultured to early-mid log phase. Bacteria were exposed to ¹³C-glucose medium for 2 h. CFU were determined after metabolic quenching. Column heights depict mean and error bars depict standard deviation. n = 6.

4.7.1 Intermediates of glycolysis and the pentose phosphate pathway

The intracellular analysis detected multiple glycolytic intermediates and one intermediate of pentose phosphate pathway – ribulose-5-phosphate (**Figure 4.9**). Most of these metabolites were fully labelled, i.e., all their carbon atoms were derived from the supplied ^{13}C -glucose, which supports glycolysis being the main carbohydrate catabolism pathway.

While nearly 100% of glucose and lactate in TIGR4 were labelled, the label percentage in TIGR4 Δ *cps* was 61.8 ± 6.9 %. This may indicate greater gluconeogenic flux or repeated carbon cycling through glycolysis/ gluconeogenesis in the capsule-deficient mutant, as some unlabelled glucose from the pre-culture seems to have persisted.

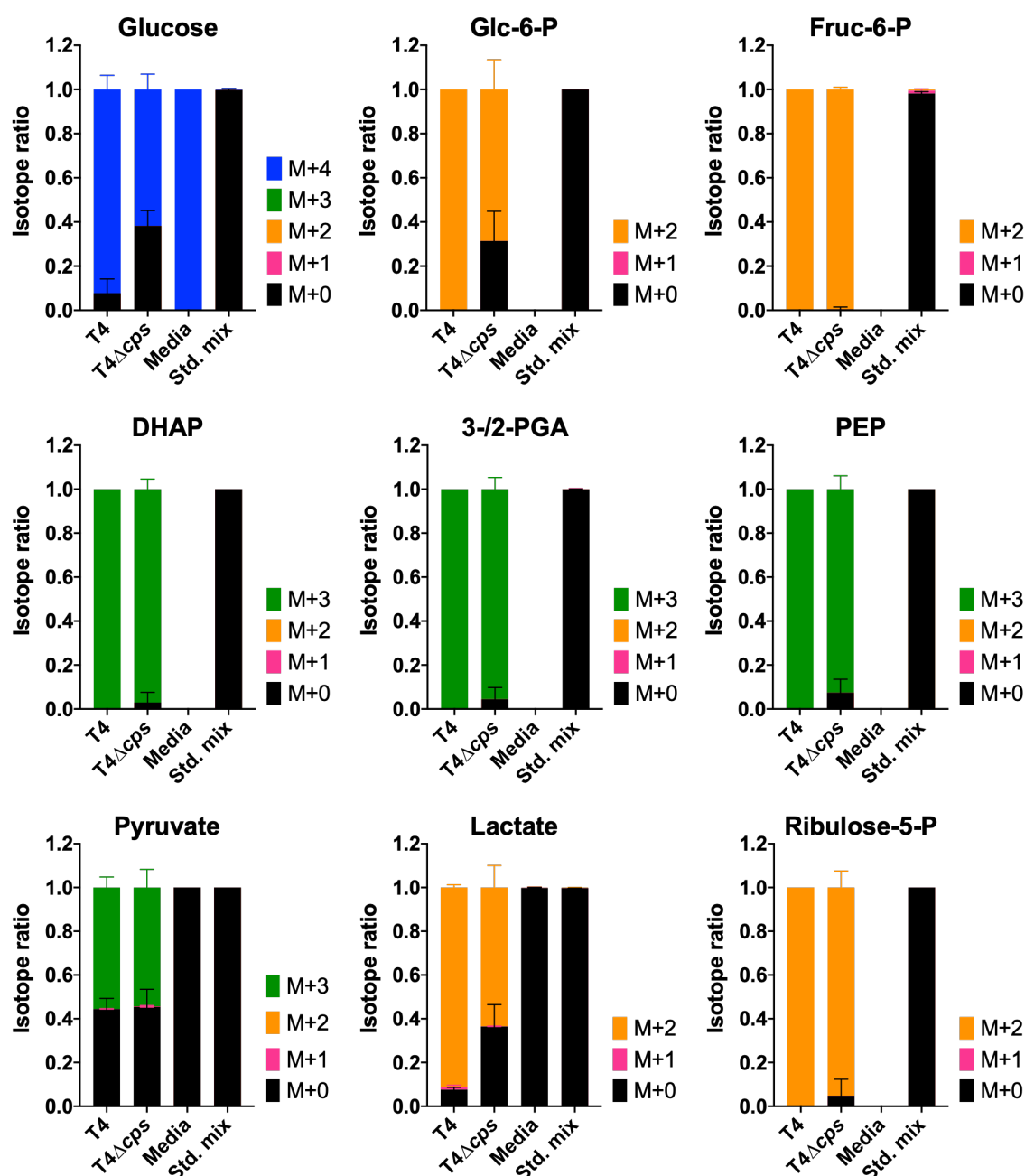


Figure 4.9 Intracellular glycolytic intermediates of *S. pneumoniae* TIGR4±cps in 7H9⁺(glc).

Isotope ratios denote ¹³C-label incorporation. M+[x] indicates [x] number of carbon atoms labelled with ¹³C carbons from ¹³C-sugar ('M' indicates the monoisotopic mass containing no ¹³C atoms). T4 = TIGR4; Media = uninoculated culture medium; Std. mix = standard mix containing known amount of each metabolite, to visualise natural isotopic abundance;

Glc-/Fruc-6-P = glucose/ fructose 6-phosphate; DHAP = dihydroxyacetone phosphate; 3-PGA = 3-phosphoglycerate/ glycerate 3-phosphate; PEP = phosphoenolpyruvate.

4.7.2 Amino acid labelling patterns

Alanine is generated from pyruvate by an alanine transaminase, and this is evident in the pyruvate (**Figure 4.9**) and alanine (**Figure 4.10**) labelling patterns of both TIGR4 and TIGR4 Δ *cps*. Biosynthetic pathways for the aromatic amino acids tryptophan, tyrosine and phenylalanine branch out from the pentose phosphate pathway (Härtel et al., 2012). Tryptophan chromatographic peaks in this experiment were of low quality so could not be quantified, but tyrosine and phenylalanine exhibited complete labelling in both strains (**Figure 4.10**).

Interestingly, glycine, asparagine, aspartate, and threonine were labelled in the wild-type TIGR4 but not the capsule-deficient mutant (**Figure 4.10**), suggesting that the TIGR4 Δ *cps* mutant does not prioritise the synthesis of these amino acids and likely preferentially obtains them from the environment.

An incomplete glycine biosynthesis pathway is encoded in the *S. pneumoniae* genome, and no significant label in glycine was observed by Härtel *et al.* (Härtel et al., 2012). This contrasts with the observation that ~ 40% of total glycine in this TIGR4 dataset was labelled in the present experiments.

Aspartate and asparagine can interconvert, and threonine is synthesised from aspartate, so the labelling patterns for these three amino acids remained consistent within each strain. The glycolytic intermediate PEP (M+3) is converted to oxaloacetate by *ppc*, which is then converted to aspartate (M+3) by *aspC* (**Figure 4.4**). The lack of ^{13}C label in aspartate for TIGR4 Δ *cps*, therefore, further supports the theory of carbon cycling through glycolysis/ gluconeogenesis/ pentose phosphate pathway in the mutant, as carbon flow is directed away from oxaloacetate.

S. pneumoniae is auxotrophic for leucine, isoleucine, and valine despite encoding hypothetical pathways for their *de novo* synthesis (Härtel et al., 2012). *S. pneumoniae* also encodes incomplete pathways for the biosynthesis of cysteine, glutamine, and glutamate, and likely for lysine and proline. In agreement, no significant label

incorporation was observed into these amino acids in both TIGR4 and TIGR4 Δ *cps* (Figure 4.11).

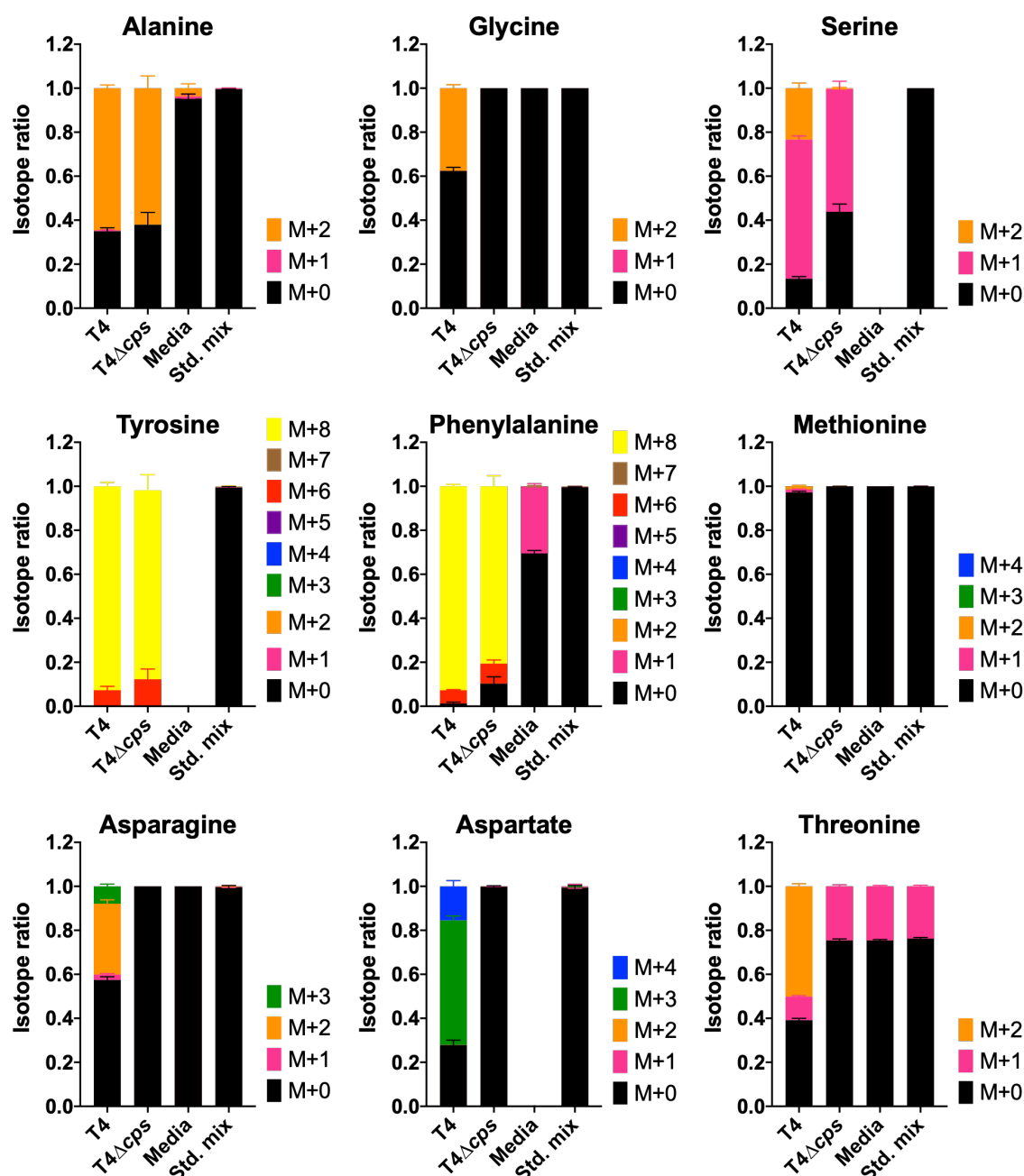


Figure 4.10 Intracellular labelled amino acids of *S. pneumoniae* TIGR4 \pm *cps* in 7H9⁺(glc)

Isotope ratios denote ^{13}C -label incorporation. M+[x] indicates [x] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). T4 = TIGR4; Media = uninoculated culture medium; Std. mix = standard mix containing known amount of each metabolite, to visualise natural isotopic abundance.

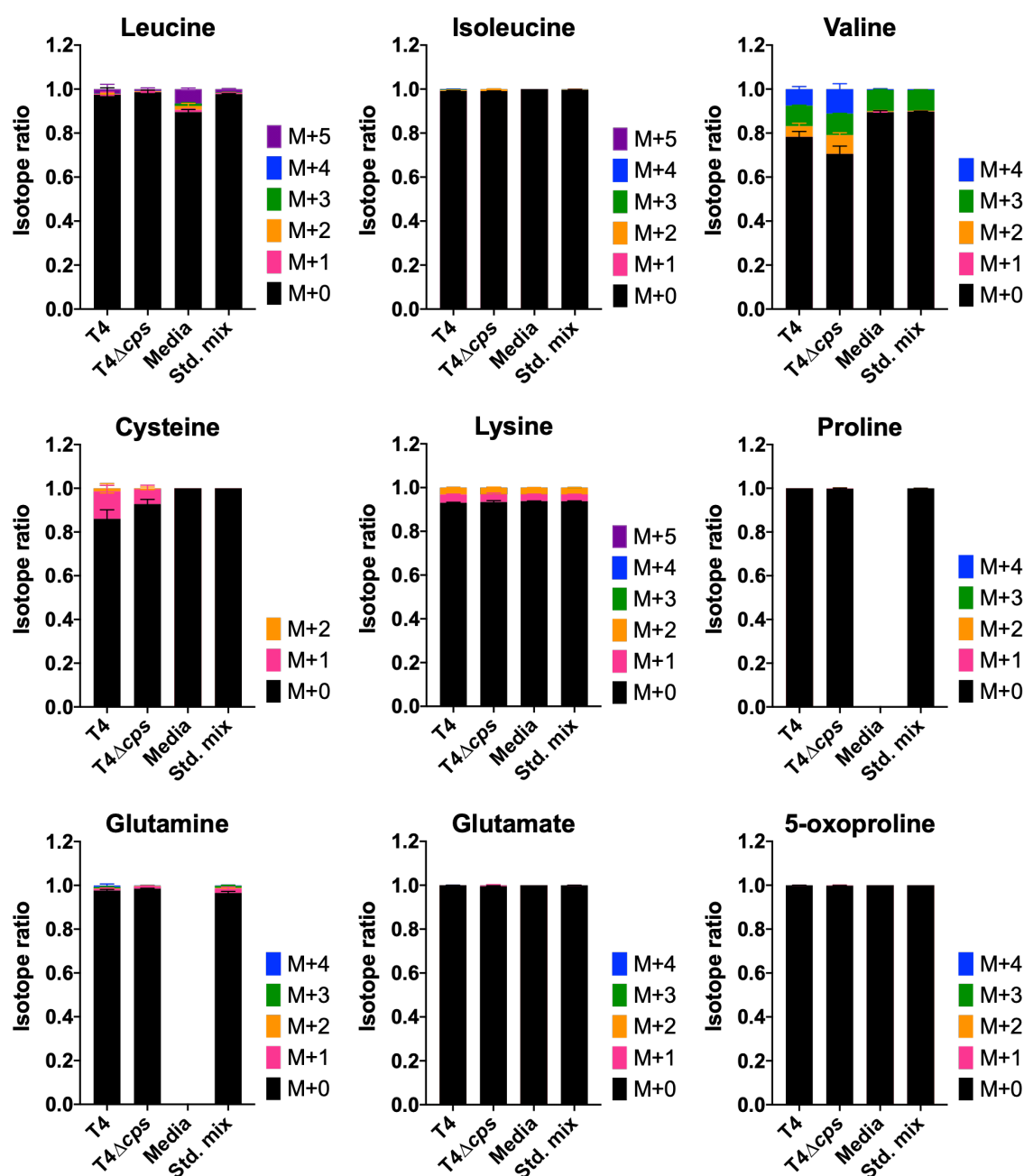


Figure 4.11 Intracellular unlabelled amino acids of *S. pneumoniae* TIGR4±cps in 7H9⁺(glc)

Isotope ratios denote ^{13}C -label incorporation. M+[x] indicates [x] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). T4 = TIGR4; Media = uninoculated culture medium; Std. mix = standard mix containing known amount of each metabolite, to visualise natural isotopic abundance.

4.7.3 Potential TCA cycle intermediates in *S. pneumoniae* TIGR4?

Isotope ratios of citrate, succinate and fumarate are shown in **Figure 4.12**. α -ketoglutarate was detected at a very low chromatographic intensity and could not be accurately quantified. Malate was also detected at low concentration, but it was unlabelled. Citrate was also unlabelled.

Succinate and fumarate, the two most interesting candidates due to their inexplicable extracellular accumulation in a bacterial species that does not operate a TCA cycle, were detectable intracellularly and were only labelled in TIGR4, with no labelling in TIGR4 Δ *cps*.

An important indication of potential TCA cycle activity would be [M+2] and [M+4] labelling in succinate and fumarate, hinting at carbons obtained from [U- $^{13}\text{C}_2$]-acetyl CoA labelled from [U- $^{13}\text{C}_6$]-glucose via glycolysis (Chapter 6 **Figure 6.2**). However, no [M+2] labelling was observed in TIGR4 succinate and fumarate. 3.5 ± 0.97 % of total succinate and 10.2 ± 1.2 % of total fumarate were [M+4] labelled. Comparatively, a larger proportion of [M+3] labelling was observed – 17.2 ± 3.6 % of succinate and 39.3 ± 4.3 % of fumarate (**Figure 4.12**).

Fumarate labelling corresponded to aspartate labelling – 56.8 ± 2 % of aspartate was [M+3] labelled, and 15.5 ± 2.6 % of aspartate was [M+4] labelled (**Figure 4.10**). It is possible that aspartate and fumarate are connected through the urea cycle, but other intermediates of the urea cycle – ornithine, citrulline and arginine – were not detected, so definitive conclusions could not be made at this stage.

The source(s) of succinate in *S. pneumoniae* therefore remain elusive. 79.2 ± 3.8 % of intracellular succinate was unlabelled, but without information about isotope label incorporation in α -ketoglutarate, it is not possible to determine whether the production of unlabelled succinate was a consequence of exogenous glutamate being converted to α -ketoglutarate and then to succinate. The sources of [M+3] and [M+4] succinate and fumarate isotopes also remain to be determined and require further work.

In conclusion, the operation of a classical pneumococcal TCA cycle has been disproved by using isotopic carbon tracing. However, the production of fumarate and succinate point to novel metabolic reaction(s) requiring further investigation. Additionally, the pathways involved in generating isotopically labelled succinate and fumarate may be potentially adaptive, as this labelling was not observed in TIGR4 Δ *cps*.

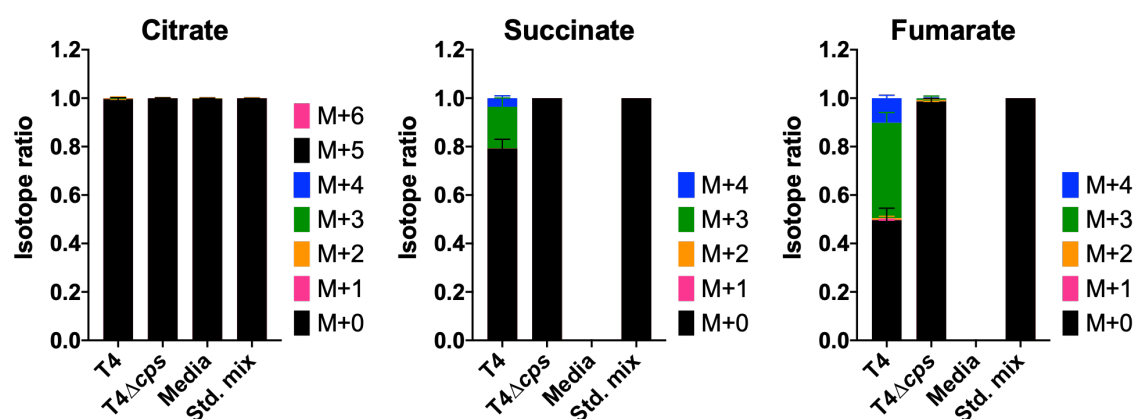


Figure 4.12 Intracellular “TCA” intermediates of *S. pneumoniae* TIGR4 \pm *cps* in 7H9⁺(glc)

Isotope ratios denote ^{13}C -label incorporation. M+[*x*] indicates [*x*] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar (‘M’ indicates the monoisotopic mass containing no ^{13}C atoms). T4 = TIGR4; Media = uninoculated culture medium; Std. mix = standard mix containing known amount of each metabolite, to visualise natural isotopic abundance.

4.8 Chapter 4 Discussion

This chapter described a metabolomics study of *S. pneumoniae* TIGR4 cultured in the semi-defined 7H9⁺ medium.

The extracellular metabolite composition (metabolic footprint) of TIGR4 over time was analysed by GC-MS. I studied the metabolic footprint when cultured in the complex medium THY, and 7H9⁺ medium containing 11 mM of either glucose or galactose as the primary carbon source. Stable isotope labelling by GC-MS of intracellular metabolites using [U-¹³C₆]-glucose was also performed for TIGR4 and TIGR4Δ*cps* cultured in 7H9⁺ medium.

When present above the detection threshold, GC-MS analysis was able to reliably quantify 17 amino acids (arginine, histidine and tryptophan were not detected); several intermediates of glycolysis; the nucleotide bases thymine and uracil; and some other metabolites potentially involved in novel undiscovered metabolic reactions (succinate, fumarate, α-ketoglutarate). Of the fermentation end-products, lactate was detectable but formate, acetate, and ethanol were not detectable by GC-MS.

The results from this Chapter were compared to two published *S. pneumoniae* metabolism studies. One is an isotope labelling study by GC-MS using [U-¹³C₆]-glucose and [1,2-¹³C₂]-glucose in *S. pneumoniae* D39Δ*cps* by Härtel *et al.* (Härtel *et al.*, 2012), focusing exclusively on amino acid metabolism. They used 55.5 mM [U-¹³C₆]-glucose and 27.8 mM [1,2-¹³C₂]-glucose in a chemically defined culture medium (CDM). The second study describes the intracellular and extracellular metabolome of *S. pneumoniae* TIGR4Δ*cps* using GC-MS, HPLC-MS and ¹H-NMR (Leonard *et al.*, 2018). They used 30.5 mM glucose in a modified RPMI medium (RPMI_{modi}) to culture *S. pneumoniae*.

4.8.1 Glycolysis in THY, 7H9⁺(glucose) and 7H9⁺(galactose)

In THY culture medium, fructose was detected along with glucose. Todd-Hewitt (TH) broth contains 2 g L⁻¹ glucose, so fructose is probably a component of yeast extract (THY). To our knowledge, the exact composition of yeast extract has not been defined, and the presence of fructose in THY is an important observation showing that the carbohydrate content of THY is not limited to 2 g L⁻¹ glucose.

S. pneumoniae TIGR4 encodes several phosphotransferase and ABC transporters for glucose, fructose and galactose uptake (Bidossi et al., 2012; Fleming & Camilli, 2016). Glucose and fructose were entirely depleted from THY. Conversely, in 7H9⁺ medium, glucose and galactose were not depleted completely, leaving ~ 2 mM in the medium by the time the bacteria reached stationary growth phase.

Intermediates of glycolysis were detected intracellularly and exhibited ¹³C labelling corresponding to glucose catabolism, and extracellular lactate accumulation was observed. This confirmed the well-established role of glycolysis as the major energy-generating and carbohydrate catabolism pathway in *S. pneumoniae*.

The overall metabolic footprint of TIGR4 when cultured in galactose was very similar to that in glucose (excluding fermentation end-product ratios, which are known to be different), suggesting that metabolic processing of the two carbohydrates abundant in different host sites proceeds along the same secondary and anabolic pathways and highlights pneumococcal metabolic flexibility.

4.8.2 Amino acid metabolism

100% labelling was observed in phenylalanine and tyrosine, as they are derived from the pentose phosphate pathway (Härtel et al., 2012). Alanine labelling was also as expected – it is generated from labelled pyruvate by a transaminase (*alaT*; SP_1994).

Intact biosynthetic pathways for the branched chain amino acids (BCAA) isoleucine, leucine and valine are present in the TIGR4 genome. However, isotope labelling experiments showed no label incorporation in isoleucine and leucine when cultured in 7H9⁺ with [U-¹³C₆]-glucose, suggesting these were obtained from the culture medium via the dedicated ABC transporter LivJHMGF (Basavanna et al., 2009) rather than being synthesised *de novo*. This corroborates the observation by Härtel *et al.* showing that D39 preferentially obtains BCAA from the environment and cannot grow in CDM lacking BCAA (Härtel et al., 2012). Leonard *et al* also saw uptake of BCAA by TIGR4Δ*cps* in RPMI_{modi} medium.

Relatively large amounts of extracellular proline corresponding to TIGR4 growth were observed when cultured in THY, 7H9⁺(glucose) and 7H9⁺(galactose). This was also seen in TIGR4Δ*cps* (Leonard et al., 2018). Lack of label incorporation in proline suggests that its synthesis does not involve carbohydrate metabolism intermediates (**Figure 4.11**) (Härtel et al., 2012). Proline can theoretically be synthesised from exogenously supplied glutamate, which would support the results here, but *S. pneumoniae* encodes an incomplete proline biosynthetic pathway (Hoskins et al., 2001; Lanie et al., 2007; Tettelin et al., 2001). These results provide solid precedence for further studies with isotopically labelled glutamate, which will be necessary to determine the existence of a glutamate-derived proline biosynthesis pathway.

Glycine, an essential amino acid, also has an incomplete biosynthetic pathway in *S. pneumoniae* genomes and is taken up from the environment. In D39Δ*cps*, glycine (unlabelled) is then converted to serine by hydroxymethylation (*glyA*), incorporating a carbon atom from a formate-derived precursor, thus generating [M+1] labelled serine (Härtel et al., 2012). However, ~ 40% of intracellular glycine and ~ 25% of intracellular serine in TIGR4 were [M+2] labelled, suggesting potential reversibility of *glyA* (SP_1024) to generate glycine from serine also by hydroxymethylation. This activity has been reported in *Streptococcus thermophilus* (Vidal et al., 2005) and *Listeria monocytogenes* (Eisenreich et al., 2006).

4.8.3 Additional metabolites

The *S. pneumoniae* genome does not encode a TCA cycle. To our surprise, extracellular succinate, *cis*-aconitate, α -ketoglutarate and fumarate accumulation corresponding to TIGR4 growth was observed in glucose- and galactose-containing 7H9⁺ medium – these intermediates may participate in the TCA cycle. In TIGR4-conditioned THY, however, only extracellular fumarate accumulation was seen, while succinate levels remained unchanged, and *cis*-aconitate and α -ketoglutarate were not detected. Malate and citrate levels remained unchanged in all culture media tested.

α -ketoglutarate can be generated from glutamate in *S. pneumoniae* D39, and fumarate may be part of the urea cycle, but succinate and *cis*-aconitate production by TIGR4 cannot be explained by genome analyses. Bacteria adapting to an anaerobic environment, such as *H. pylori* and *E. coli*, operate a variant noncyclic TCA cycle (Guest, 1995; Pitson et al., 1999). A fragmented TCA cycle with an unknown function is encoded in the streptococci *S. mutans* and *S. gordonii* (Willenborg & Goethe, 2016).

Therefore, we hypothesised that *S. pneumoniae* could plausibly perform some variation of a TCA cycle. Because there is no precedent for this in pneumococcal genomes or the published literature, I first decided to carry out a series of experiments to determine the validity of these results.

First, whole genome sequencing of the TIGR4 stock (strain P1542) and a type strain (ATCC TIGR4 BAA-334) was performed, which confirmed the laboratory TIGR4 stock had not developed any significant genome variation when compared to the ATCC-TIGR4 strain and the publicly available reference TIGR4 genome AE005672.3 (Tettelin et al., 2001). Taxonomic classification of the sequenced reads (Wood & Salzberg, 2014) confirmed that no contaminating bacterial species were present in the pneumococcal stocks used in these experiments.

Lastly, the metabolic footprint of the ATCC-TIGR4 strain showed similar extracellular accumulation of succinate and fumarate as for the laboratory TIGR4 strain. Together, these validation experiments proved that the unexpected observations from metabolic

footprinting experiments were a consequence of pneumococcal metabolic activity in the $7\text{H}9^+$ culture medium.

To conclusively determine potential TCA cycle activity, stable isotope labelling of intracellular TIGR4 metabolites using $[\text{U}-^{13}\text{C}_6]$ -glucose was performed. No label incorporation was seen in citrate and malate, *cis*-aconitate was undetectable, and α -ketoglutarate was present at low abundance so its labelling pattern could not be accurately determined. $[\text{M}+2]$ isotopes of fumarate and succinate were not present, and only small percentages of $[\text{M}+4]$ isotopes were present. $[\text{M}+2]$ and $[\text{M}+4]$ isotopes of TCA cycle intermediates are essential markers of TCA cycle activity (Jang et al., 2018).

Thus, the existence of a classical TCA cycle in *S. pneumoniae* TIGR4 has been definitively disproved. However, sources of unlabelled and $[\text{M}+3]$ labelled succinate remain unresolved. One potential source of unlabelled succinate could be α -ketoglutarate generated by transamination of exogenous (unlabelled) glutamate.

Another possibility is a reaction involving formate and fumarate producing $[\text{M}+3]$ labelled succinate and CO_2 , as seen in *E. coli* undergoing anaerobic respiration but not described for *S. pneumoniae* thus far.

A third possibility might be $[\text{M}+3]$ labelled succinate arising from fumarate via fumarate reductase activity. However, nucleotide and protein sequence homology and a conserved domain search of fumarate reductase enzymes encoded in *E. coli* and *B. subtilis* did not yield any hits in the TIGR4 genome. In fact, this (admittedly limited) analysis did not show any homology or conserved domains in TIGR4 which might point to succinate or fumarate production via the TCA cycle.

It is worth noting that Leonard *et al.* detected intracellular succinate production in TIGR4 Δ *cps*, but did not look into this further (Leonard et al., 2018). Succinate and fumarate were not discussed in the stable isotope labelling study by Härtel *et al.*, but this is likely because they were looking specifically for amino acids by Selected Ion Monitoring (SIM) in GC-MS (Härtel et al., 2012).

With regard to fumarate, the proportion of [M+3] isotopes corresponded to [M+3] labelling in aspartate, suggesting a link between these metabolites. This is most likely through the urea cycle, as aspartate and fumarate are both intermediates of the urea cycle. In support of this hypothesis, Leonard *et al.* observed decreasing extracellular arginine concentration correlating with increasing extracellular ornithine and fumarate concentrations in TIGR4 Δ *cps*, which would be expected of urea cycle activity (Leonard et al., 2018). In TIGR4-conditioned THY medium, the extracellular accumulation patterns of ornithine, aspartate and fumarate were nearly identical. Together, these data suggest an intact urea cycle, which has not yet been described for *S. pneumoniae* TIGR4.

In conclusion, I believe that sources of unlabelled and [M+3] labelled succinate and fumarate warrant further investigation, and will be described in the next Chapter.

4.8.4 Differential carbon cycling in TIGR4 Δ *cps* compared to wild-type TIGR4

Carbon flow through metabolic pathways between wild-type (TIGR4) and unencapsulated (TIGR4 Δ *cps*) *S. pneumoniae* has not been compared before. In this Chapter, stable isotope labelling using [U-¹³C₆]-glucose revealed several differences between the wild-type and capsule-deficient mutant strains of TIGR4.

Complete labelling of the glycolytic intermediates glucose 6-phosphate, fructose 6-phosphate, dihydroxyacetone phosphate, 3-phosphoglycerate, phosphoenolpyruvate and pyruvate, and the pentose phosphate pathway (PPP) intermediate ribulose 5-phosphate, was observed for both TIGR4 and TIGR4 Δ *cps*. Alanine, tyrosine and phenylalanine, amino acids derived from these intermediates, also exhibited identical labelling patterns in the two strains.

Glucose and lactate were entirely labelled in TIGR4, but showed ~ 60% labelling in TIGR4 Δ *cps* – 38.2 ± 6.9 % of glucose and 36.5 ± 10 % of lactate in TIGR4 Δ *cps* were unlabelled, indicating potential persistence of unlabelled glucose from the pre-culture

before [U- $^{13}\text{C}_6$]-glucose was supplied. Similarly, glycine, serine, asparagine, aspartate, succinate, and fumarate were shown to incorporate ^{13}C label in TIGR4 but not in TIGR4 Δ *cps*, suggesting that the production of these metabolites or their precursors are not prioritised or not essential in TIGR4 Δ *cps*.

Together, these observations suggest cyclical carbon flow within glycolysis/gluconeogenesis and associated amino acids, directing carbons away from oxaloacetate production from PEP. This disproves the hypothesis set out in this Chapter that the reduced metabolic burden of capsule production in TIGR4 Δ *cps* would lead to increased carbon flow to anabolic pathways of amino acids and other associated metabolites.

Whether these carbon atoms are re-directed towards fermentative energetic metabolism will be studied using ^1H -NMR, which can detect all fermentation end-products, in the following Chapter. Capsule biosynthesis genes have been shown to exhibit wide-ranging effects on virulence and metabolism (Ogunniyi et al., 2002), and it remains to be determined whether the deletion of the capsular locus in TIGR4 Δ *cps* has an effect on other metabolic enzymes in this context.

Chapter 5. ^{13}C isotope tracing to determine carbohydrate-derived carbon flow in *S. pneumoniae*

5.1 Introduction

In the preceding Chapter, I found differential labelling of secondary (non-glycolytic) metabolites in *S. pneumoniae* TIGR4 Δcps compared to its WT isogenic parent strain TIGR4, cultured in 7H9 $^{+}$ (^{13}C -glucose) and 7H9 $^{+}$ (^{13}C -galactose). This suggests that capsulation may determine the pattern of carbon flow in *S. pneumoniae*. I therefore hypothesised that the reduced labelling of secondary metabolites in TIGR4 Δcps is a result of carbon diversion towards end-products of fermentation via glycolytic breakdown of carbohydrates. This hypothesis is tested in this Chapter using ^1H -NMR to detect ^{13}C -label incorporation from ^{13}C -glucose or ^{13}C -galactose into lactate, formate, acetate and ethanol, i.e., end-products of fermentation.

However, TIGR4 Δcps grows at a sub-optimal level in 7H9 $^{+}$ broth compared to its WT isogenic parent TIGR4, which may be a potential confounder in interpreting ^{13}C label incorporation. Therefore, in this Chapter I used a different, more growth-friendly chemically defined medium CDM (derived from van de Rijn & Kessler, 1980) to confirm the differential carbon flow in WT vs. Δcps TIGR4 observed in 7H9 $^{+}$ – the need for this validation is further supported by a recent review summarising the sometimes conflicting results obtained in pneumococcal studies using different culture media (Sanchez-Rosario & Johnson, 2021). An additional benefit of using CDM is its widespread prevalence as a culture medium in the existing pneumococcal literature, particular in another ^{13}C -glucose stable isotope labelling study carried out in *S. pneumoniae* D39 (Härtel et al., 2012).

In this chapter, I also sought to determine whether carbon flow through pneumococcal metabolic pathways observed with TIGR4 would be generalisable to other strains, particularly in relation to capsule deletion in a different strain and serotype background. Therefore, I performed stable isotope labelling analysis of *S. pneumoniae* 23F and 23F Δcps in parallel with TIGR4 and TIGR4 Δcps , as described. 23F is much more

prevalent in nasopharyngeal carriage in the population (Brueggemann et al., 2003; Usuf et al., 2014), and its growth characteristics in CDM culture are comparable with those of TIGR4 (**Chapter 3.5.2**). Compared to the serotype 4 capsule, the 23F capsule is biochemically distinct in its structure and composition (Chapter 1 **Figure 1.1**), may have a similar energetic cost for production (determined by the number of high-energy bonds per polysaccharide repeat unit), but may be nutritionally more costly to produce (determined by the number of carbon atoms per polysaccharide repeat unit) (Hathaway et al., 2012; Weinberger et al., 2009).

The main aim of the analyses described in this Chapter was to trace carbon flow through pneumococcal carbohydrate metabolism pathways using [U- $^{13}\text{C}_6$]-glucose and [U- $^{13}\text{C}_6$]-galactose to obtain a complete picture of intracellular (with GC-MS) and extracellular (with GC-MS and ^1H -NMR) labelling of pneumococcal metabolites associated with carbohydrate utilisation.

5.2 Chapter 5 Aims

- Establish a method for *S. pneumoniae* intracellular metabolite extraction for GC-MS analysis.
- Determine a method to standardise the amount of intracellular extract for GC-MS analysis from 4 *S. pneumoniae* strains (TIGR4, TIGR4 Δ *cps*, 23F and 23F Δ *cps*) so that absolute intracellular amounts and isotope ratios of metabolites can be compared.
- Compare the extracellular fermentation profiles of *S. pneumoniae* WT strains (TIGR4 and 23F) cultured on CDM(¹³C-glucose) vs. CDM(¹³C-galactose), by employing ¹H-NMR to determine ¹³C-label incorporation into fermentation end-products.
- Describe the effect of the absence of capsule (using strains TIGR4 Δ *cps* and 23F Δ *cps*) on the extracellular fermentation profiles established during culture on CDM(¹³C-glucose) and CDM(¹³C-galactose), using ¹H-NMR.
- Use GC-MS analysis to trace the metabolic route(s) of isotopically labelled carbon atoms derived from ¹³C-labelled glucose and galactose in the intracellular and extracellular space of all 4 *S. pneumoniae* strains.
 - Determine the relationship between central carbon and capsule metabolism.
 - Determine whether the carbohydrate source affects the balance of amino acid biosynthesis vs. uptake.
 - Determine the mode of production of succinate and fumarate by analysing labelling patterns of related metabolites.

5.3 Standardisation of methods for [U-¹³C₆]-sugar metabolomics

5.3.1 Standardised intracellular metabolite extraction method for GC-MS

In comparison with extracellular metabolite analysis of an organism growing in culture *in vitro*, studying the intracellular metabolome requires additional processing – this includes rapid quenching of metabolism at the time at which the metabolome is to be investigated, and bacterial cell lysis to release intracellular contents into an appropriate solvent/ solvent mix to extract metabolites of interest.

Robust methods that had been tested on *S. pneumoniae* were not available in the literature at the time of experimental analysis. Therefore, as a prelude to stable isotope labelling of intracellular *S. pneumoniae* metabolites, a standardised protocol for the extraction of intracellular metabolites was established using *S. pneumoniae* TIGR4 cultured in THY medium.

In order to determine the combination of extraction solvent and cellular lysis method that would provide the greatest extraction efficiency across all detectable metabolite features, a total of 8 combinations of extraction techniques were tested using liquid chromatography-mass spectrometry (LC-MS). These included three extraction solvents – chloroform + methanol (2:1 v/v) ('CM') or acetonitrile ('ACN') or methanol + acetonitrile + water (2:2:1 v/v) ('2:2:1') – and three methods for pneumococcal cell lysis – pulse sonication in a chilled water bath ('SON'), freeze-thaw cycles in liquid nitrogen ('FT') or homogenisation with glass beads using a FastPrep-24 homogeniser ('beads') (Table 5.1). The complete protocol is described in Chapter 2.4.2 Materials & Methods.

Similar CFU counts after metabolic quenching (Figure 5.1) and tight clustering of the pooled biological quality control (PBQC) samples on principal component analysis (PCA) plots (Figure 5.2A, B) indicate technical comparability across the intracellular metabolite extraction conditions tested. I also observed clustering of replicate injections according to the extraction and lysis technique employed (Figure 5.2C, D). With the same extraction solvent, cellular lysis by sonication or freeze-thawing showed

overlapping clusters, while lysis by bead homogenisation formed separate clusters. Extraction method 8 (MeOH + ACN + H₂O (2:2:1) with bead homogenisation) showed the tightest clustering, i.e., least variation in abundance between replicates, in both positive- and negative-ion mode of LC-MS (**Figure 5.2C, D**).

<i>Method</i>	<i>Extraction solvent</i>	<i>Cell lysis</i>	<i>Abbreviation*</i>
1	CHCl ₃ + MeOH (2:1)	Pulse sonication 3x	1/ CM-SON
2	ACN	Pulse sonication 3x	2/ ACN-SON
3	CHCl ₃ → 1x sonication → + MeOH → 1x sonication		3/ C-SON-M-SON
4	CHCl ₃ + MeOH (2:1)	Bead homogenisation	4/ CM-Beads
5	ACN	Bead homogenisation	5/ ACN-Beads
6	CHCl ₃ + MeOH (2:1)	Freeze thaw 3x	6/ CM-FT
7	ACN	Freeze thaw 3x	7/ ACN-FT
8	MeOH + ACN + H ₂ O (2:2:1)	Bead homogenisation	8/ 2:2:1-Beads

Table 5.1 Summary of intracellular metabolite extraction methods

Details can be found in *Chapter 2.4.2 Materials & Methods*.

CHCl₃ = chloroform; ACN = acetonitrile; MeOH = methanol.

*These abbreviations will be used to indicate the extraction method in subsequent figures.

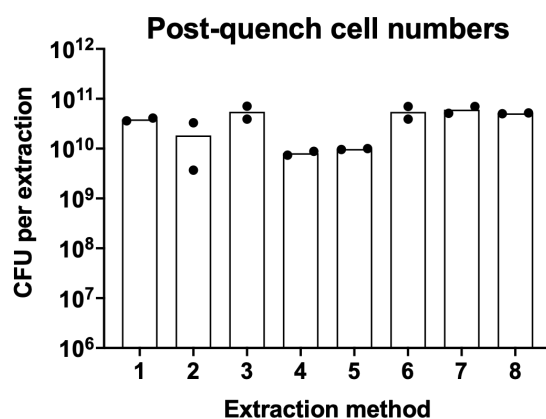


Figure 5.1 Post-quench CFU of *S. pneumoniae* TIGR4 per metabolite extraction method

25 ml THY was inoculated with ~2-3 × 10⁷ CFU of *S. pneumoniae* stock and cultured to OD₆₀₀ 0.9 – 1.0 (4 h). For each extraction replicate (1/6), 8 ml of the culture was quenched (therefore, 2 × 25 ml cultures for each extraction method, i.e., 2 biological replicates). Individual dots represent CFU determinations of 2 biological replicates and column heights depict the mean.

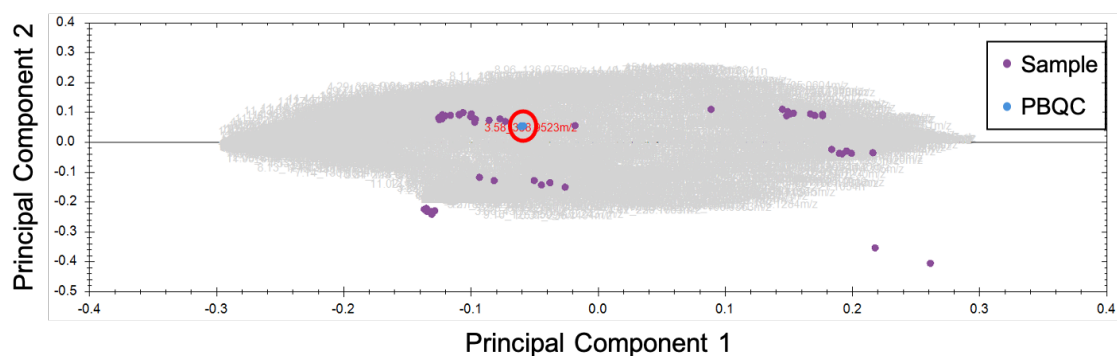
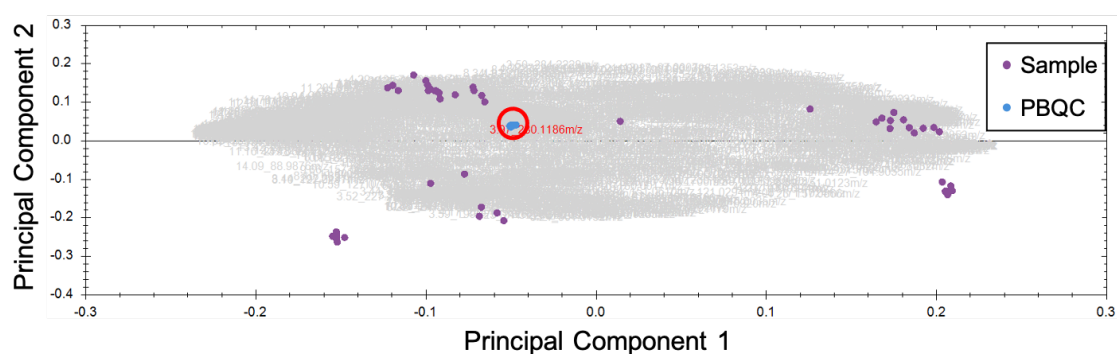
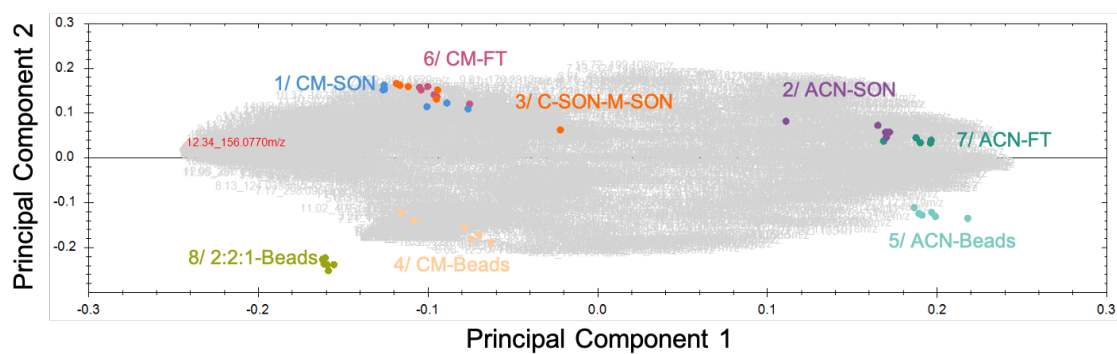
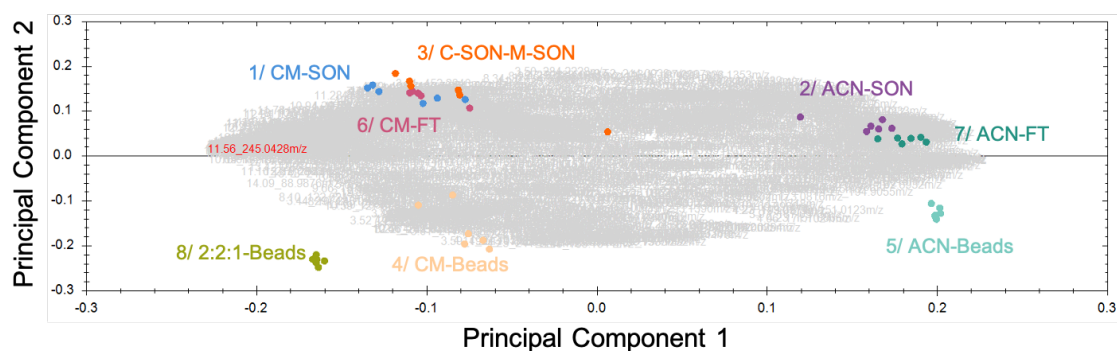
A Principal Components Analysis grouped by sample vs. control (+ve ion mode)**B Principal Components Analysis grouped by samples vs. control (-ve ion mode)****C Principal Components Analysis grouped by extraction method (+ve ion mode)****D Principal Components Analysis grouped by extraction method (-ve ion mode)**

Figure 5.2 Principal Components Analysis (PCA) of samples analysed by LC-MS

PCA plots generated in Progenesis QI (Waters) using normalised abundance of all detected metabolite features. Normalisation was performed using total ion abundance (*area of metabolite/sum of all areas*). Tight clustering of replicate injections indicates low variation in abundance. PBQC control samples cluster together (A, B). Samples obtained using the 8 different extraction methods are colour-coded and labelled (C, D). $n = 6$. For details of abbreviated extraction methods, refer to Table 5.1. CM-SON (blue); ACN-SON (purple); C-SON-M-SON (orange); CM-Beads (peach); ACN-Beads (turquoise); CM-FT (magenta); ACN-FT (dark green); 2:2:1-Beads (light green).

~3000 metabolite features (fragments) were detected here. For every metabolite feature detected, variation of total ion abundance within each extraction method was determined by calculating a coefficient of variation (CV; *standard deviation/mean* of $n = 6$). The metabolite features (unannotated) were then ordered by CV (smallest to largest) to assess the metabolite coverage (ideally, high) and variation (ideally, low) for each extraction method. I found that the extraction solvent had the greatest effect on extraction efficiency, regardless of the pneumococcal cell lysis method (**Figure 5.3**). Extractions with acetonitrile alone were the least efficient, with low overall metabolite coverage and high CV. Chloroform + methanol extractions fared better, and the most efficient intracellular extraction overall was observed with method 8 (MeOH + ACN + H₂O (2:2:1) coupled with bead homogenisation) (**Figure 5.3**).

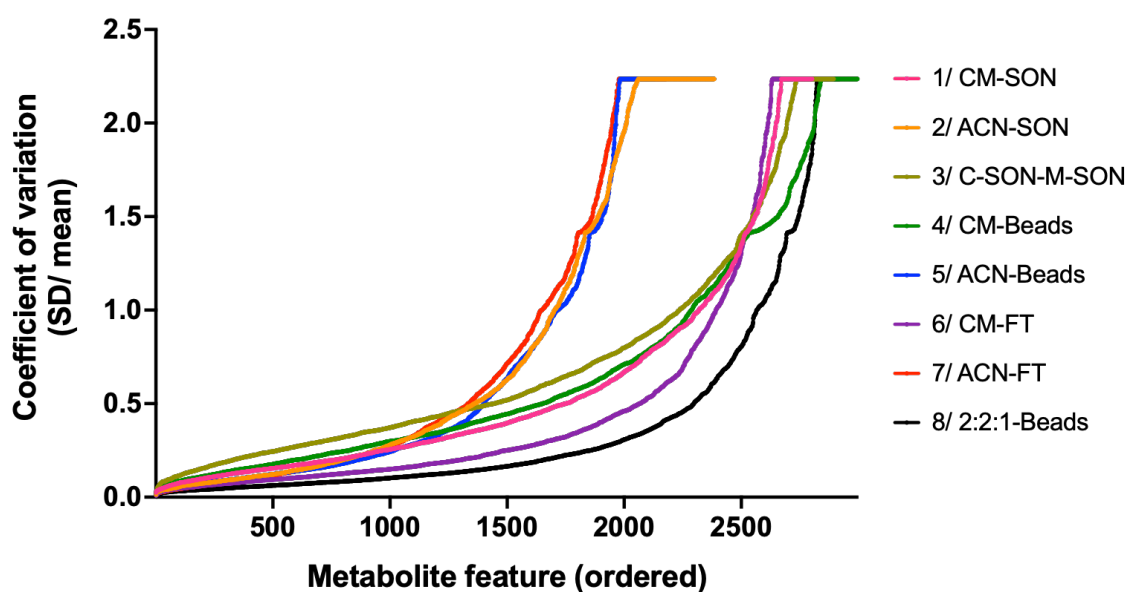


Figure 5.3 Variation and metabolite coverage for each intracellular extraction method

This plot was generated using total ion abundance of metabolite features (ion fragments) detected by LC-MS for 6 replicate cultures of *S. pneumoniae* TIGR4 per extraction method. For details of extraction methods abbreviated in the legend, refer to Table 5.1. Coefficient of variation denotes technical variation. Metabolite features are ordered for each extraction method, meaning that feature numbers do not correspond to the same feature for each extraction method. SD = standard deviation.

From the 2000-3000 detected metabolite features, 34 metabolites of interest (which were to be analysed in subsequent ^{13}C -labelling experiments) were then manually annotated, and their CV and relative abundance compared for each extraction method (**Figure 5.4**). Extraction methods 6 (chloroform + MeOH (2:1) with lysis by freeze-thawing) and 8 (MeOH + ACN + H₂O (2:2:1) with lysis by bead homogenisation) showed the highest intracellular extraction efficiencies for the metabolites of interest. Method 8 was comparatively less efficient at detecting intermediates of glycolysis and the pentose phosphate pathway (DHAP, F-1,6-bP, Glc-1-P, PEP, 2-PGA, and S-7-P), but carbon flow from the supplied carbohydrate through glycolysis is undebated so the detection of these metabolites is not critical for this study.

Extraction method 8 had the shortest sample processing time, which is an additional benefit for metabolomics. Therefore, extraction in MeOH + ACN + H₂O with cell lysis by bead homogenisation (Method 8) was chosen for pneumococcal intracellular metabolite extraction – it was highly reproducible (**Figure 5.2C, D**), showed the greatest extraction efficiency across all detected metabolite features (**Figure 5.3**), and excellent extraction efficiency for specific annotated metabolites which will be studied using ^{13}C -labelling experiments (**Figure 5.4**).

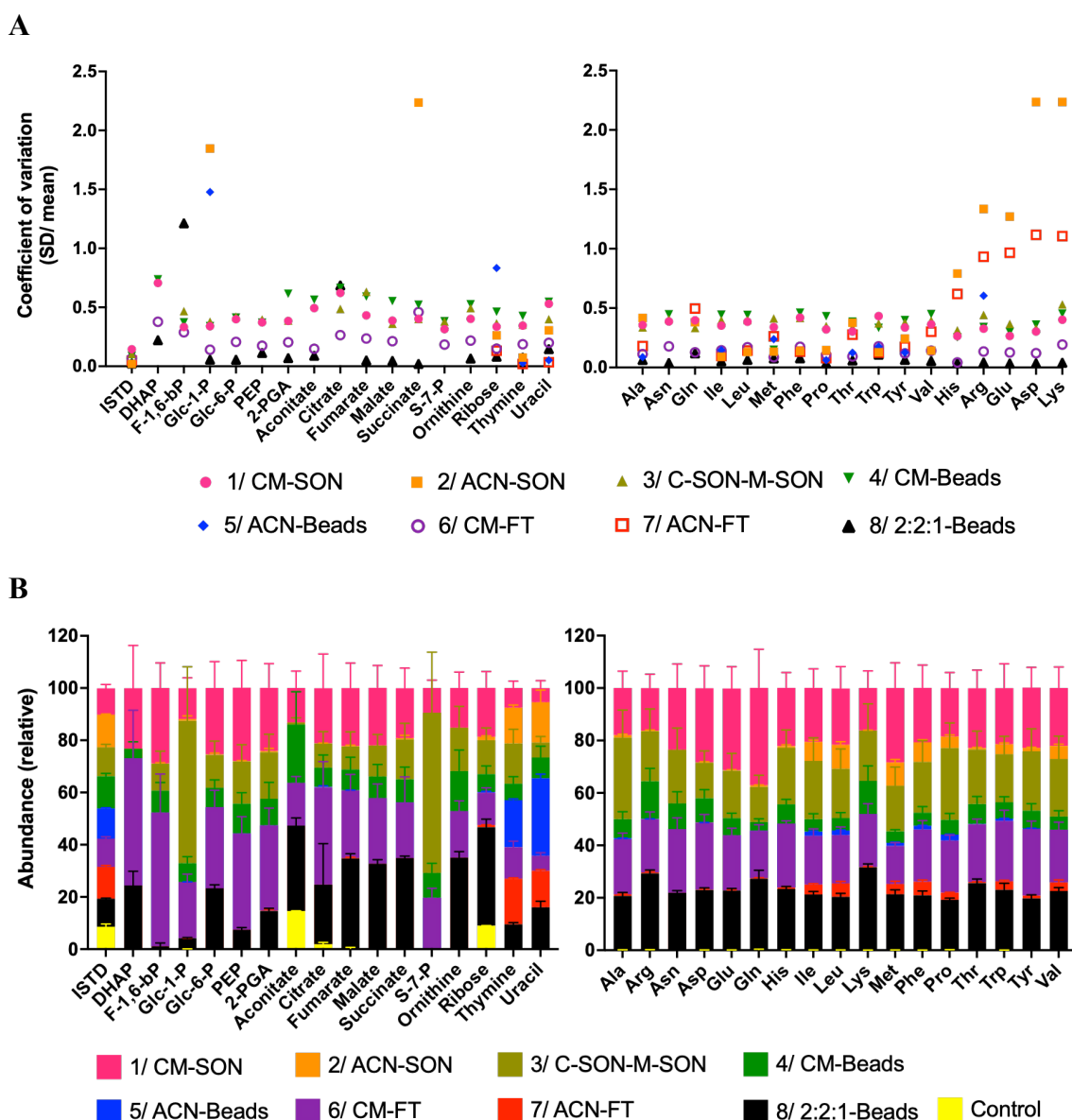


Figure 5.4 Extraction efficiency of different extraction methods for annotated metabolites

Metabolite annotation performed in Progenesis QI after the LC-MS analysis workflow. A – Technical variation in metabolite abundance is denoted by coefficient of variation, calculated using 6 replicate cultures of *S. pneumoniae* TIGR4 for each extraction method. SD = standard deviation. B – Relative abundance of each metabolite is depicted according to the extraction method used. Control = no bacteria. n = 6. For details of extraction methods abbreviated in the legends, refer to Table 5.1. Abbreviations: ISTD = internal standard (*scyllo*-inositol); DHAP = dihydroxyacetone phosphate; F-1,6-bP = fructose-1,6-bisphosphate; glc-1/6-P = glucose-1/6-phosphate; PEP = phosphoenolpyruvate; 2-PGA = 2-phosphoglycerate; S-7-P = sedoheptulose-7-phosphate.

5.3.2 Standardised pneumococcal cell numbers for stable isotope labelling

To perform stable isotope labelling of pneumococcal metabolic pathways, *S. pneumoniae* was first cultured in CDM (Chapter 2 **Table 2.3**) containing unlabelled carbohydrate (glc or gal), then pelleted by centrifugation, and the unlabelled CDM replaced with CDM containing ^{13}C -labelled glucose or galactose. Pneumococci were cultured in the ^{13}C -sugar-containing CDM for 2 hours before quenching metabolism.

As discussed in **Chapter 3.5**, the 4 strains used in this study – TIGR4, TIGR4 Δ *cps*, 23F and 23F Δ *cps* – have slightly different patterns of growth, and the relationship between OD₆₀₀ (measuring culture density) and CFU (number of colony-forming units, a measure of biomass) is not equivalent among these strains, likely due to differing cell sizes. In order to minimise the variability of bacterial cell numbers, I first set out to standardise CFU across strains at two points in the labelling process – by pre-culturing each strain (in unlabelled CDM) to a similar CFU before swapping in CDM containing ^{13}C -labelled sugar (2×10^7 CFU ml⁻¹, corresponding approximately to mid-log phase of growth; **Figure 5.5**), and by injecting onto the GC-MS instrument an amount of intracellular extract corresponding to the same CFU for each strain.

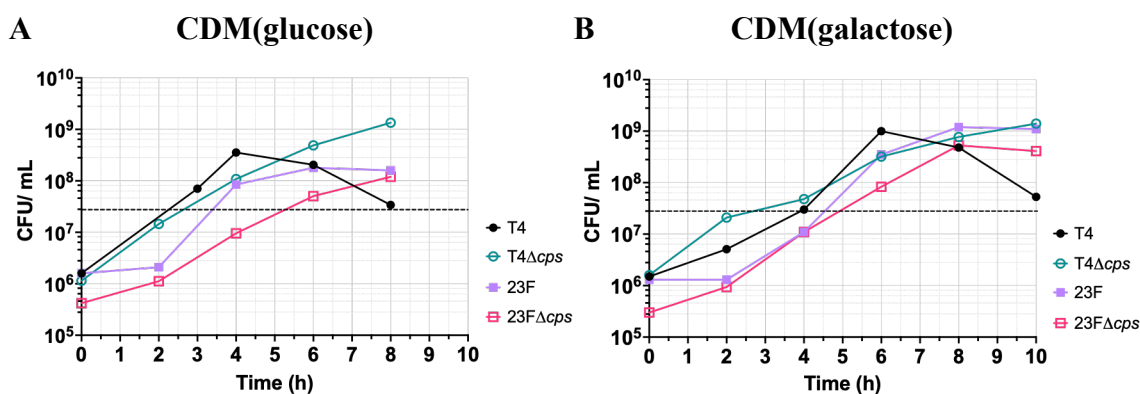


Figure 5.5 *S. pneumoniae* growth (CFU ml⁻¹) to determine media swap time points.

25 ml CDM containing 11 mM glucose (A) or galactose (B) was inoculated with ~ 3 - 5×10^7 CFU *S. pneumoniae* stock. Figure reproduced as in *Figure 3.11B*, D. n = 1-3. Dashed lines at 2×10^7 CFU ml⁻¹ represent the corresponding x-axis time points for each strain. These time points were chosen as the duration for pre-culture in unlabelled CDM before swapping in CDM containing ^{13}C -glucose/ -galactose for subsequent stable isotope labelling experiments.

Growth curves depicting CFU ml⁻¹ for all strains cultured in CDM containing glucose or galactose were used to determine the time points corresponding to 2x10⁷ CFU ml⁻¹, i.e., the time points at which CDM containing ¹³C-sugar would be swapped in (**Figure 5.5**). While *S. pneumoniae* TIGR4, 23F and 23FΔ*cps* continued to grow successfully after media swapping at the predicted time points, TIGR4Δ*cps* did not – swapping media at 2x10⁷ CFU ml⁻¹ for TIGR4Δ*cps* led to no further growth in CDM containing ¹³C-labelled sugar, which hampered our ability to trace the ¹³C₆ isotope marker through metabolic pathways.

Therefore, media swapping at later time points was tested in TIGR4Δ*cps* (**Figure 5.6**). It was found that, compared to the other strains, TIGR4Δ*cps* required the establishment of a greater biomass before media swapping to be able to recover from the shock of media swapping and continue to grow in CDM containing ¹³C-labelled sugar. Swapping culture medium at 4 h in CDM(glucose) and at 6 h in CDM(galactose) allowed for subsequent growth of TIGR4Δ*cps* during the 2 h culture in CDM containing ¹³C-sugar (**Figure 5.6**), and these time points were thus used for the stable isotope labelling studies of TIGR4Δ*cps* discussed in this Chapter.

Although this meant that TIGR4Δ*cps* cell numbers at the point of media swapping were higher than that of the other three strains (TIGR4, 23F and 23FΔ*cps* CFU were comparable; **Figure 5.7A, B**), the post-quench cell numbers were similar for all strains cultured in either CDM(glucose) or CDM(galactose) (**Figure 5.7E, F**). Additionally, the growth of all strains during 2 h culture in CDM containing ¹³C-labelled sugars was also comparable, as measured by the difference in CFU between media swapping and metabolic quenching (**Figure 5.7C, D**).

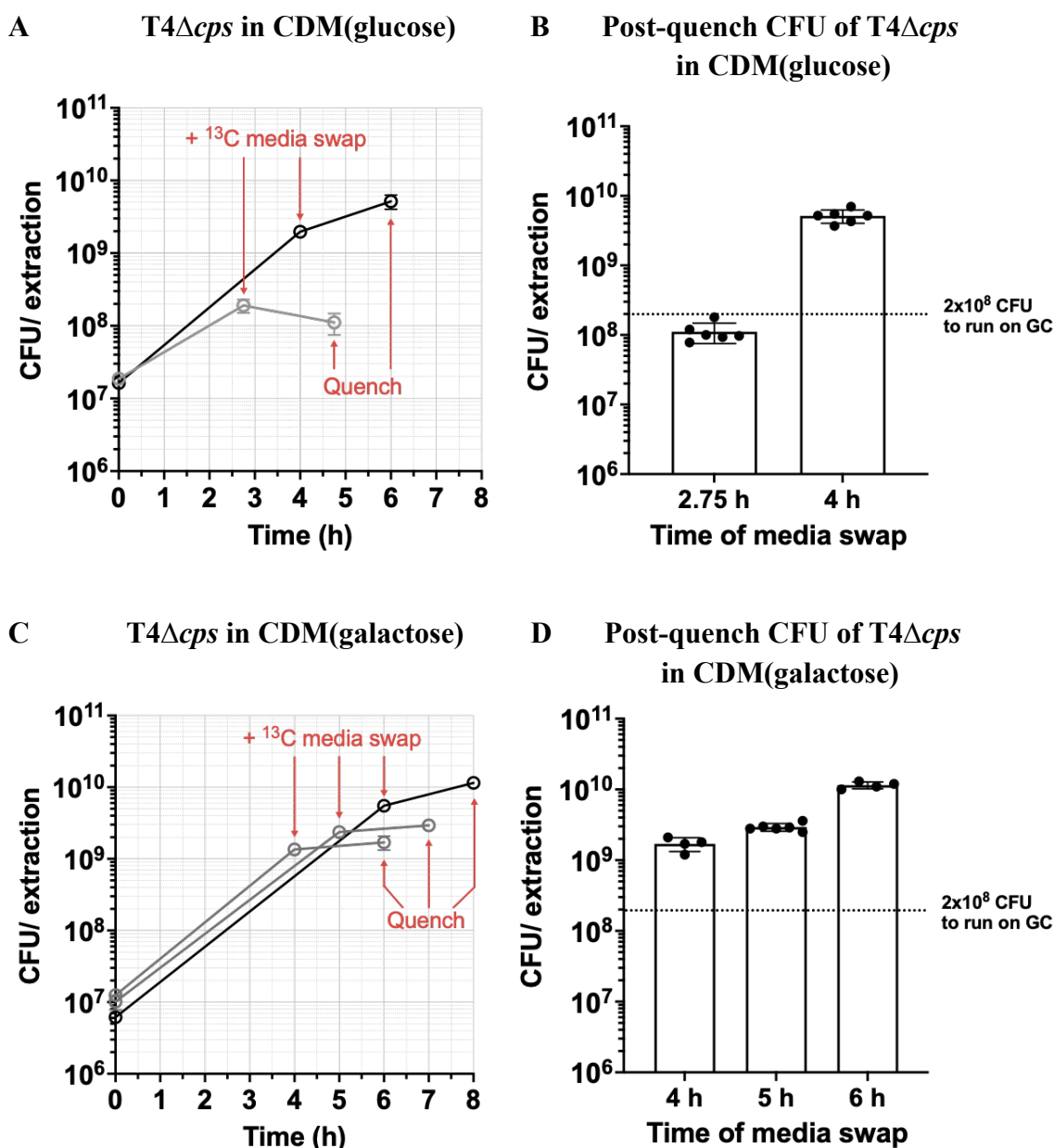


Figure 5.6 Testing media swap time points for *S. pneumoniae* TIGR4Δcps.

~ 5×10^7 initial CFU of TIGR4Δcps stock was cultured in 25 ml CDM(glucose) (A) or CDM(galactose) (C). CDM containing ^{13}C -labelled sugar was swapped in at the points indicated by “+ ^{13}C -media swap” and CFU after metabolic quenching is annotated as “Quench” (A, C). After media swapping, *S. pneumoniae* was always cultured for 2 h in the CDM containing ^{13}C -labelled sugar, then the metabolism quenched. (B, D) Post-quench CFU at the different swap time points tested. Individual dots represent technical replicates. Column heights depict the mean, and error bars depict standard deviation. $n = 2-3$. Based on this data it was decided to swap culture medium after 4 h of TIGR4Δcps pre-culture in CDM(glucose) and after 6 h in CDM(galactose).

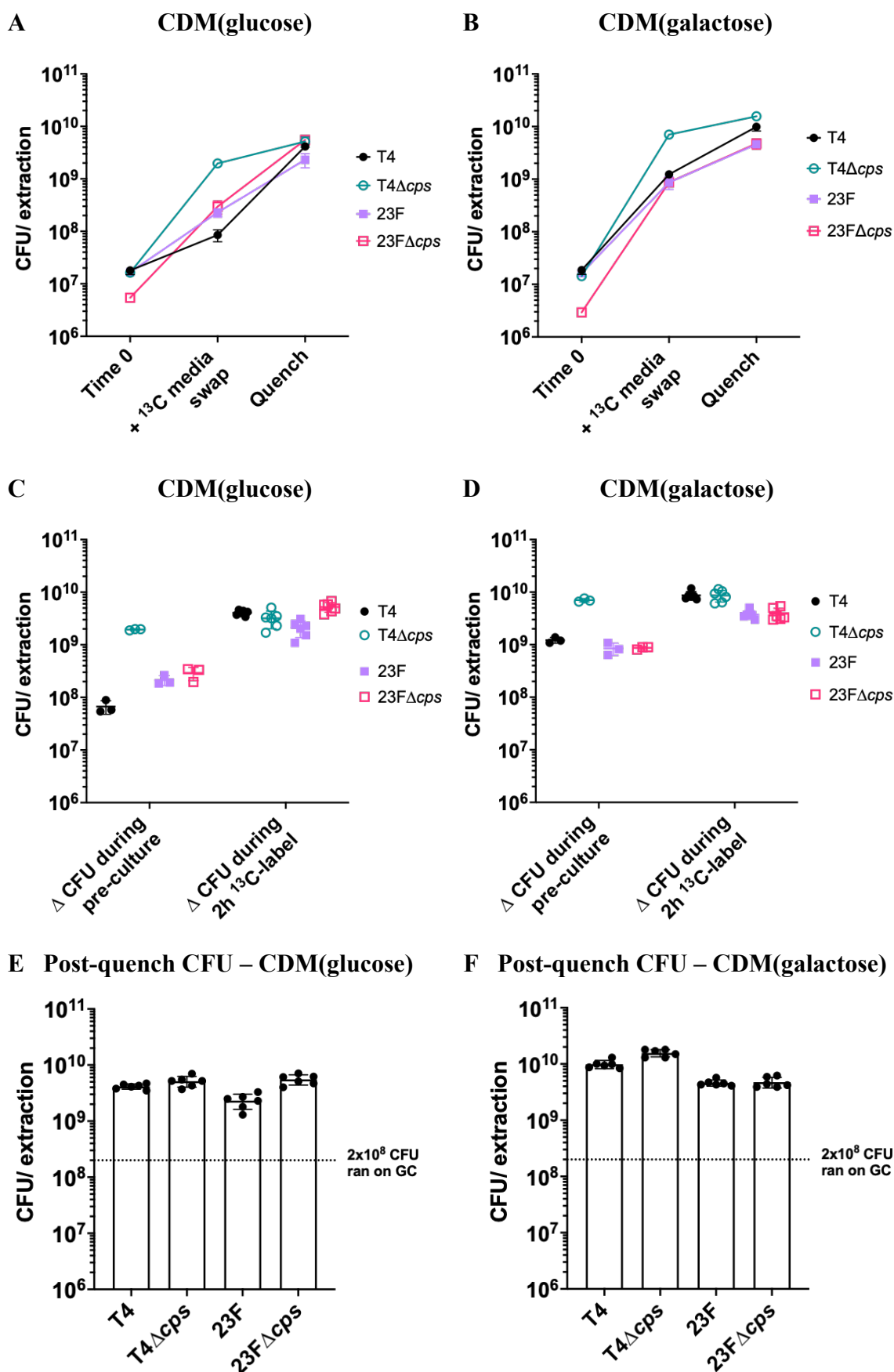


Figure 5.7 *S. pneumoniae* CFU at different stages of culturing for stable isotope labelling

(A, B) *S. pneumoniae* CFU at different stages of culturing for stable isotope labelling. n = 1-3. (C, D) Change in CFU (Δ CFU) during pre-culture and 2 h culture in ^{13}C -sugar determined by calculating the difference in CFU between “Time 0” and “+ ^{13}C media swap”, and between “+ ^{13}C media swap” and “Quench”, respectively. n = 3. (E, F) Post-quench CFU were comparable across *S. pneumoniae* strains cultured in either CDM(^{13}C -glucose) or CDM(^{13}C -galactose). n = 3. Mean and standard deviation are depicted for all plots.

Standardisation of pneumococcal cell numbers was achieved at a further stage by injecting cellular extracts corresponding to the same CFU for all strains, which would allow for the reliable comparison of metabolite labelling patterns across all strains.

In deciding the amount of intracellular extract to be injected onto the GC column, I found that injecting the entirety of the intracellular extract obtained from a 25 ml culture caused saturation of the chromatographic column (**Figure 5.8A**). Concentrated samples also lead to ion suppression during electron impact ionisation in mass spectrometry, interfering with metabolite detection and accurate quantification. I also avoided split injection – where only a fraction of the derivatised sample can be injected onto the GC column – as this does not necessarily correlate with actually loading that fraction of the sample, and opted instead to standardise pneumococcal samples before injection.

Conversely, injecting extract from 1×10^8 CFU compromised metabolite detection due to low concentration of metabolites (**Figure 5.8B**). I finally found that injecting intracellular extract from 2×10^8 CFU provided the best results across all strains (**Figure 5.8C**) – a clean, unsaturated chromatogram; an unclogged injector and column; improved resolution and detection of chromatographic peaks; and reliable quantification of metabolite isotopes without the requirement for post-acquisition data manipulation. Therefore, intracellular extract from 2×10^8 CFU *S. pneumoniae* was consistently injected for all metabolic labelling experiments.

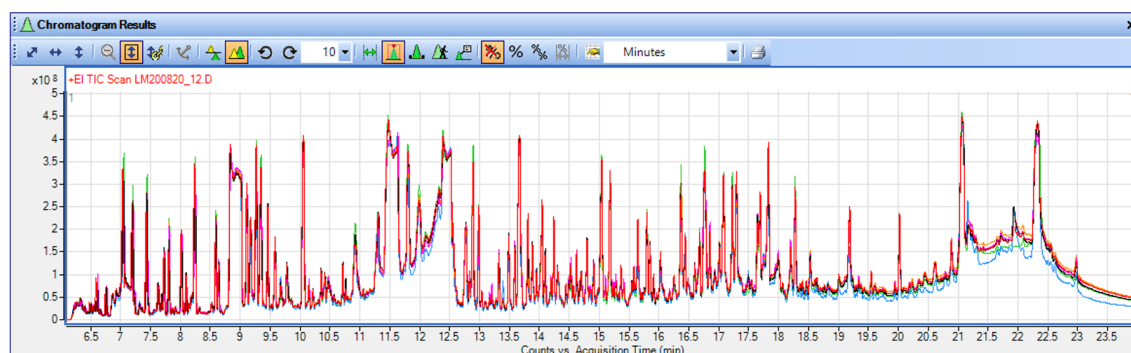
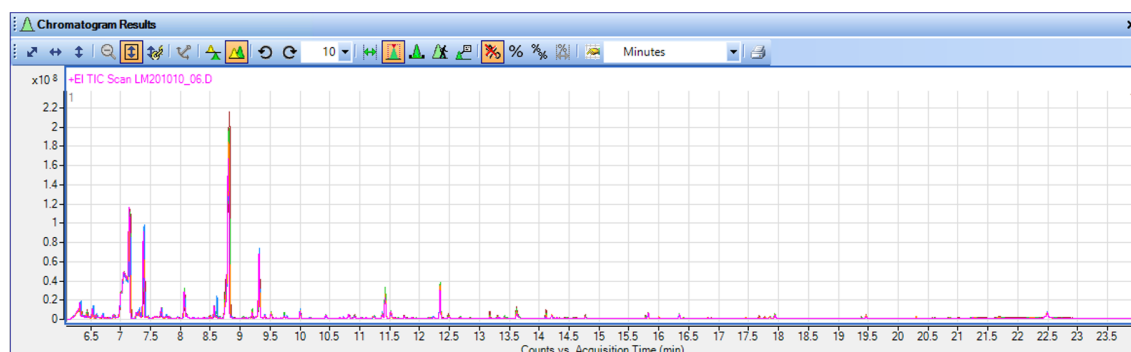
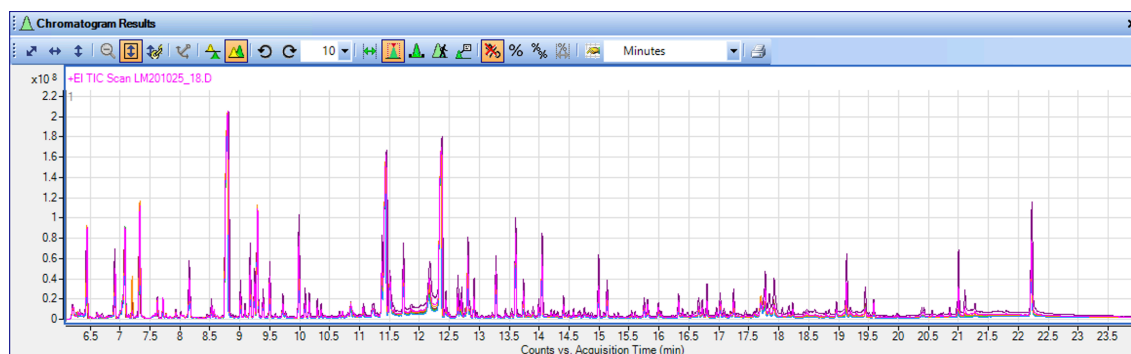
A Overloaded chromatographic column**B Underloaded chromatographic column****C “Goldilocks” chromatographic column loading**

Figure 5.8 Example chromatograms of different intracellular extract injection amounts.

Chromatograms visualised in Agilent MassHunter Qualitative Analysis software. These chromatogram screenshots are representative examples of a concentrated sample leading to column overloading and saturated peaks (A), a sample with insufficient intracellular extract (B), and a balanced sample concentration (C).

In conclusion, the challenge posed by differential growth patterns of the four strains used in this study was mitigated by standardising cell numbers post-quench and at the point of injection onto the GC-MS instrument, thereby allowing direct comparisons between metabolite abundance and labelling patterns to be made between each WT/ Δcps pair of *S. pneumoniae* TIGR4 and 23F cultured in glucose or galactose (**Figure 5.7**).

5.4 Extracellular fermentation profile of *S. pneumoniae*

The end-products of *S. pneumoniae* anaerobic fermentation are lactate, formate, acetate and ethanol (Neijssel et al., 1997; Yesilkaya et al., 2009). These are generated in varying ratios, depending on the activity of the specific enzymes of pyruvate catabolism (Chapter 4 **Figure 4.2**). Of these, only lactate is detectable by GC-MS, but ^1H -NMR is additionally able to detect formate, acetate and ethanol. Here, ^1H -NMR was employed to study the complete extracellular fermentation profile of *S. pneumoniae* TIGR4 and 23F, and their respective capsule-deficient mutants (TIGR4 Δ *cps* and 23F Δ *cps*).

Hypothesis 1: *S. pneumoniae* TIGR4 and 23F will display homolactic fermentation on ^{13}C -glucose and mixed acid fermentation on ^{13}C -galactose, as has been previously observed in *S. pneumoniae* D39, showing that carbohydrate-dependent differential fermentation is conserved in *S. pneumoniae* strains.

Hypothesis 2: Biosynthesis of the TIGR4 and 23F capsules come at a similar energetic cost (Weinberger et al., 2009). I hypothesise that that the Δ *cps* strains will divert more carbons towards fermentation end-products as compared to their WT counterparts, owing to a lack of the burden of capsule biosynthesis.

5.4.1 Carbon source has a similar effect on the fermentation profiles of *S. pneumoniae* TIGR4 and 23F

When glucose was supplied as the main carbohydrate source, it was catabolised predominantly to lactate in both TIGR4 and 23F (**Figure 5.9A, B**). This changed during growth of the WT strains on galactose, where formate, acetate and ethanol were also produced in an approximate ratio of 4:4:1 (**Figure 5.9C, D**).

Previous studies in *S. pneumoniae* D39 have shown that the fermentation profile differs when cultured on galactose compared to glucose – lactate production is predominant on glucose (homolactic fermentation), while a mixture of lactate, acetate, formate and ethanol are observed during growth on galactose (mixed acid fermentation) (Carvalho et

al., 2013b; Paixão et al., 2015a; 2015b). Here, we show that this pattern of fermentation end-products is also true for *S. pneumoniae* TIGR4 and 23F, which suggests that differential fermentation is a conserved adaptation to growth on the two carbon sources abundant in different human microenvironments (glucose in the blood/ CSF and galactose in the respiratory tract).

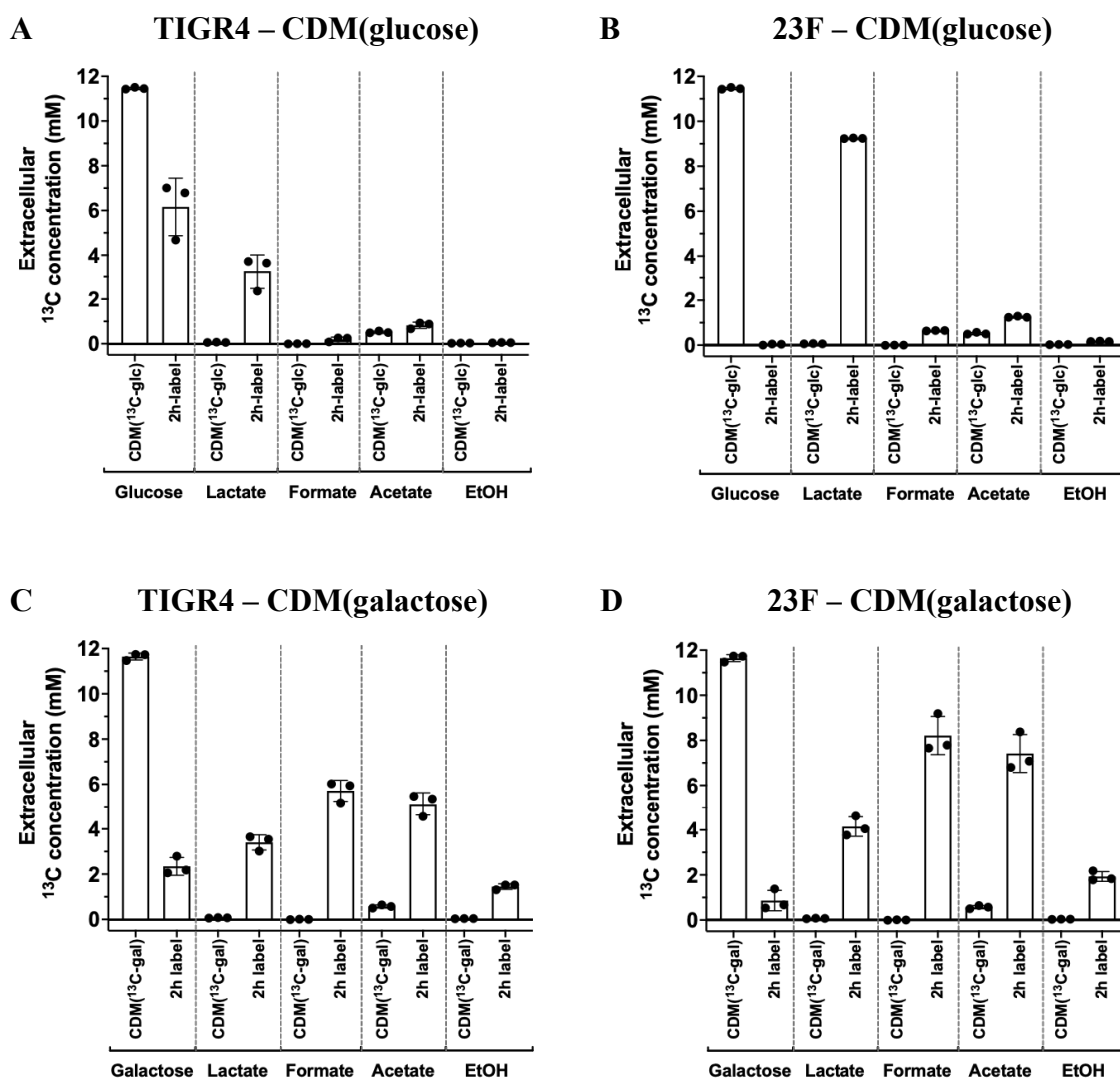


Figure 5.9 Extracellular fermentation profile of *S. pneumoniae* TIGR4 and 23F

S. pneumoniae was cultured in CDM containing ¹³C-glucose (A, B) or ¹³C-galactose (C, D) for 2 hours. The extracellular concentrations of ¹³C-labelled glucose/ galactose, lactate, formate, acetate and ethanol were measured by ¹H-NMR. Mean and standard deviation of each metabolite in the uninoculated medium (“CDM(¹³C-glc/gal)”) and after 2 h culture in ¹³C-labelled CDM (“2h label”) are depicted here. n = 3.

5.4.2 Capsule-deficient *S. pneumoniae* show different patterns of response to the carbohydrate supplied

During culture on glucose, while TIGR4 displayed strictly homolactic fermentation (3.2 ± 0.8 mM lactate), TIGR4 Δcps produced small amounts of formate (1.5 ± 0.3 mM), acetate (1.5 ± 0.4 mM) and ethanol (0.3 ± 0.06 mM) along with lactate (2.2 ± 0.5 mM). Similar amounts of formate (1.3 ± 0.3 mM), acetate (1.4 ± 0.6 mM) and ethanol (0.3 ± 0.06 mM) were generated by TIGR4 Δcps cultured on galactose (**Figure 5.10A**), suggesting that a certain level of non-lactate fermentation products are generated in TIGR4 Δcps regardless of the carbohydrate supplied.

The fermentation profile of 23F Δcps , on the other hand, was similar to that of 23F in that it was dependent on the carbohydrate substrate – homolactic fermentation was observed on glucose, and mixed acid fermentation on galactose (**Figure 5.10B**). On galactose, 23F Δcps generated significantly higher amounts of formate, acetate and ethanol compared to 23F (**Figure 5.10B**).

In contrast, TIGR4 Δcps grown on galactose produced significantly lower amounts of formate, acetate and ethanol compared to TIGR4, and significantly more lactate (5.5 ± 1.6 mM in TIGR4 Δcps , compared to 3.4 ± 0.3 mM in TIGR4) (**Figure 5.10A**). This data suggests that TIGR4 Δcps exhibits a semi-mixed acid fermentation profile on both glucose and galactose, with a shift towards increased lactate production on galactose.

Overall, all strains displayed a level of mixed acid fermentation on galactose, which is affected by the lack of a capsule. Moreover, the lack of a capsule has a differential effect based on the genetic background of its isogenic counterpart, suggesting that differential fermentation patterns are not only linked to the presence/ absence or biosynthetic burden of the capsule, but potentially also the genetic context and polysaccharide makeup of the capsule – this is expanded on in **Section 5.6** (Chapter 5 Discussion).

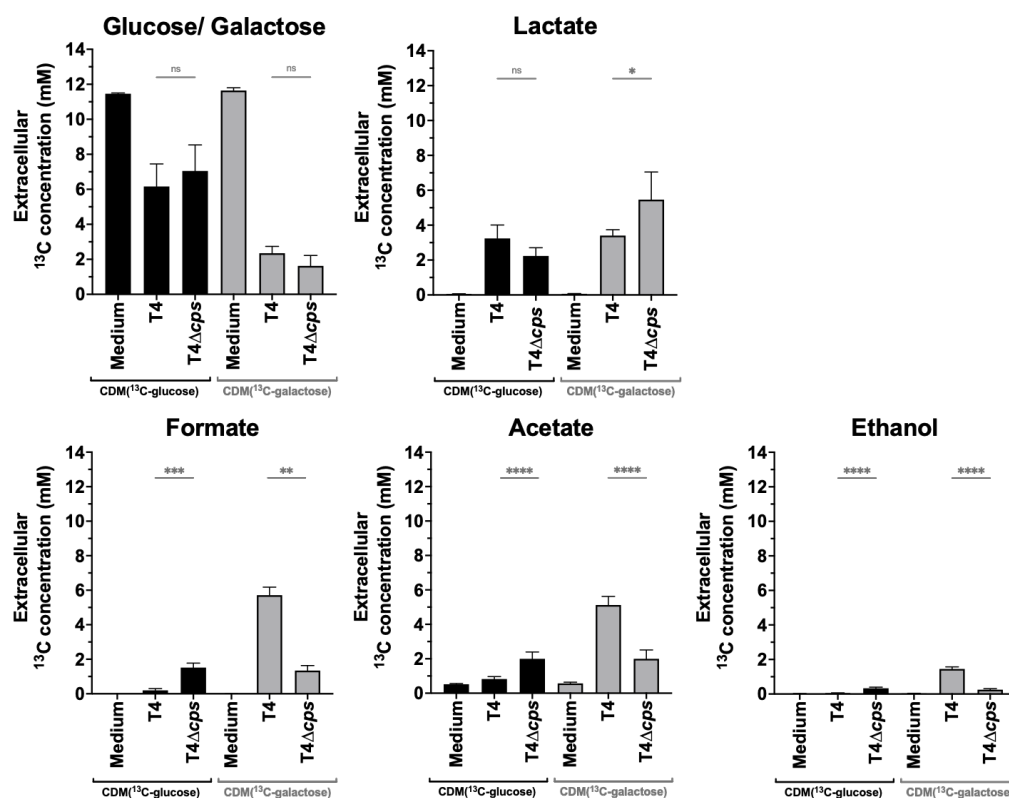
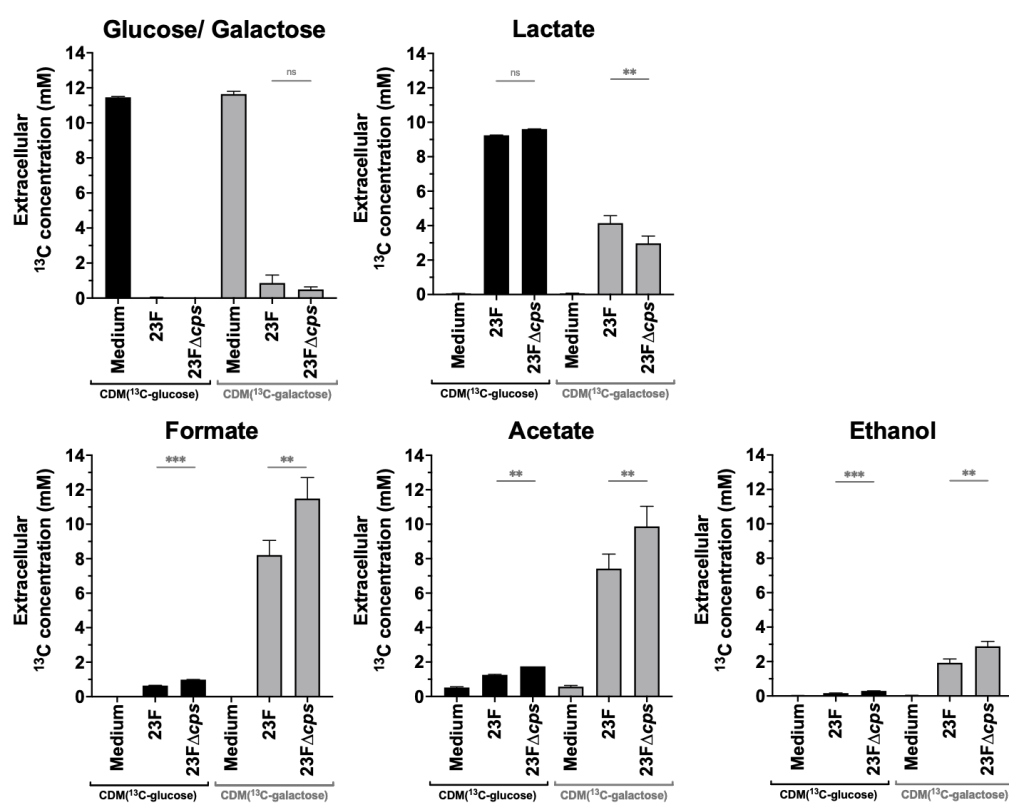
A Fermentation products of T4 and T4 Δ cps cultured in glucose and galactose**B Fermentation products of 23F and 23F Δ cps cultured in glucose and galactose**

Figure 5.10 Comparison of the fermentation profile between WT and Δcps *S. pneumoniae* strains cultured in ^{13}C -glucose or ^{13}C -galactose.

S. pneumoniae TIGR4, TIGR4 Δcps (A) and 23F, 23F Δcps (B) were cultured in CDM containing ^{13}C -glucose (black bars) or ^{13}C -galactose (grey bars) for 2 hours. The extracellular concentrations of ^{13}C -labelled glucose/ galactose, lactate, formate, acetate and ethanol were measured by ^1H -NMR. Column heights depict the mean concentration and error bars denote standard deviation. Statistical analysis – one-way ANOVA with Šidák correction for multiple comparisons – was performed on log-transformed data, and significant differences between WT and Δcps pairs are indicated by asterisks. $n = 3$. ns = not significant ($p > 0.05$); * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$.

5.5 GC-MS analysis to determine the complete metabolic profile of *S. pneumoniae* cultured on $^{13}\text{C}_6$ -glucose and $^{13}\text{C}_6$ -galactose

In conjunction with analysing the fermentation profiles of *S. pneumoniae* TIGR4 and 23F and their capsule-deficient derivatives using [U- $^{13}\text{C}_6$]-glucose and [U- $^{13}\text{C}_6$]-galactose, I also studied stable isotope labelling of metabolites associated with carbohydrate metabolism to obtain a complete picture of carbon flow through pneumococcal metabolic pathways, using GC-MS. In this section, I traced the route of labelled carbon atoms from ^{13}C -glucose and ^{13}C -galactose to describe the labelling patterns of glycolytic intermediates, potential metabolites of the urea cycle, potential “TCA” intermediates, and amino acids, with particular focus on determining whether deletion of the capsular biosynthetic locus (using TIGR4 Δcps and 23F Δcps) leads to divergent carbon flow.

Hypothesis 1: Owing to pneumococcal promiscuity in carbohydrate utilisation, carbon flow will be unaffected when cultured in ^{13}C -glucose or ^{13}C -galactose.

Hypothesis 2: While capsule biosynthesis comes at an energetic cost to the pneumococcus, it may also act as a carbon store (Hamaguchi et al., 2018). In this system with a limited amount of carbohydrate (11 mM) and no external stress (such as from immune cells) requiring increased capsule production, I hypothesise that the capsule of WT encapsulated strains may act as an additional carbon source for the synthesis of secondary metabolites (such as amino acids), diluting the carbohydrate-derived ^{13}C label. The unencapsulated Δcps mutants will only incorporate carbons from the ^{13}C -sugar supplied, exhibiting complete labelling of essential metabolites.

Hypothesis 3: Succinate is a by-product of anaerobic fermentation derived from formate, rather than an intermediate of a derivative TCA cycle, and so its labelling pattern will correspond to that of formate.

Hypothesis 4: Fumarate is a by-product of the urea cycle, rather than an intermediate of a derivative TCA cycle, and so its labelling pattern will correspond to that of aspartate.

5.5.1 Metabolites of glycolysis and the pentose phosphate pathway

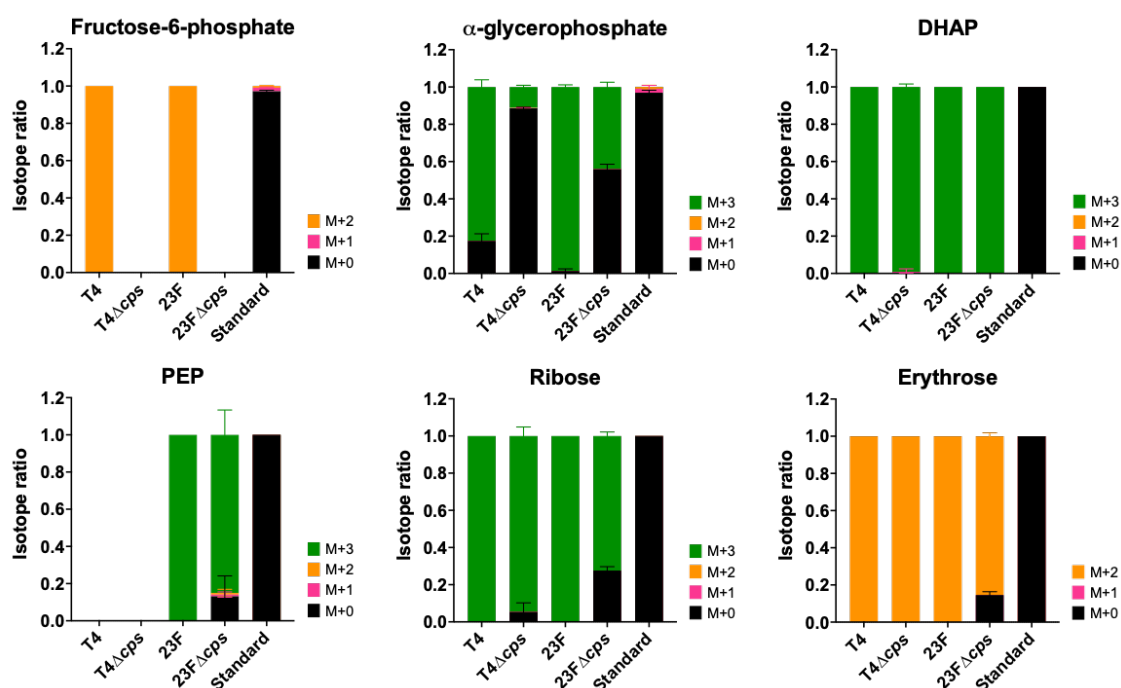
When the following intermediates of glycolysis and the pentose phosphate pathway (PPP) were detected intracellularly, they showed complete labelling – galactose-6-phosphate (Gal6P), fructose-6-phosphate (F6P), dihydroxyacetone phosphate (DHAP), phosphoenolpyruvate (PEP), ribose and erythrose (**Figure 5.11A, B**). Gal6P was only detected during ^{13}C -galactose culture in all strains (**Figure 5.11B**), which was expected as Gal6P is an intermediate of pneumococcal galactose metabolism via the tagatose-6-phosphate pathway. F6P was only detectable in TIGR4 and 23F during ^{13}C -glucose culture (**Figure 5.11A**); the low abundance of F6P in TIGR4 Δcps and 23F Δcps made quantification unreliable. PEP was undetectable in the TIGR4 strain pairs during ^{13}C -glucose culture (**Figure 5.11A**) but showed complete labelling in ^{13}C -galactose culture (**Figure 5.11B**). In the 23F strain pairs, PEP exhibited labelling in both ^{13}C -glucose and ^{13}C -galactose cultures.

α -glycerophosphate (Glp) is a product of glycerol degradation (EC 2.7.1.30), and can also be generated from DHAP (EC 1.1.1.94, EC 1.1.1.8). Similar amounts of intracellular Glp were detected in the WT 23F strain when cultured in either ^{13}C -glucose or ^{13}C -galactose (**Figure 5.11C**). However, the percentage labelled differed based on the carbohydrate – ~ 80-100% of intracellular Glp in the WT strains cultured in glucose was labelled, falling to ~ 20% in galactose (**Figure 5.11A, B**). 23F Δcps showed similarly carbohydrate-dependent labelling, with unlabelled Glp in galactose but ~ 45% labelled in glucose, even though a higher amount of Glp was detected in galactose (**Figure 5.11C**). TIGR4 Δcps -generated Glp showed minimal labelling (< 10%) in both carbohydrate sources. Glp is not an intermediate of glycolysis, only derived from the glycolytic intermediate DHAP, but why the overall labelling of Glp was reduced in galactose culture is uncertain. It is clear, however, that Glp is generated during glucose and galactose culture (**Figure 5.11C**). This highlights its importance as a substrate for the enzyme α -glycerophosphate oxidase (*glpO*), conserved in all *S. pneumoniae* strains and implicated in nasopharyngeal colonisation and the development of pneumococcal meningitis (Mahdi et al., 2017; Mahdi et al., 2012).

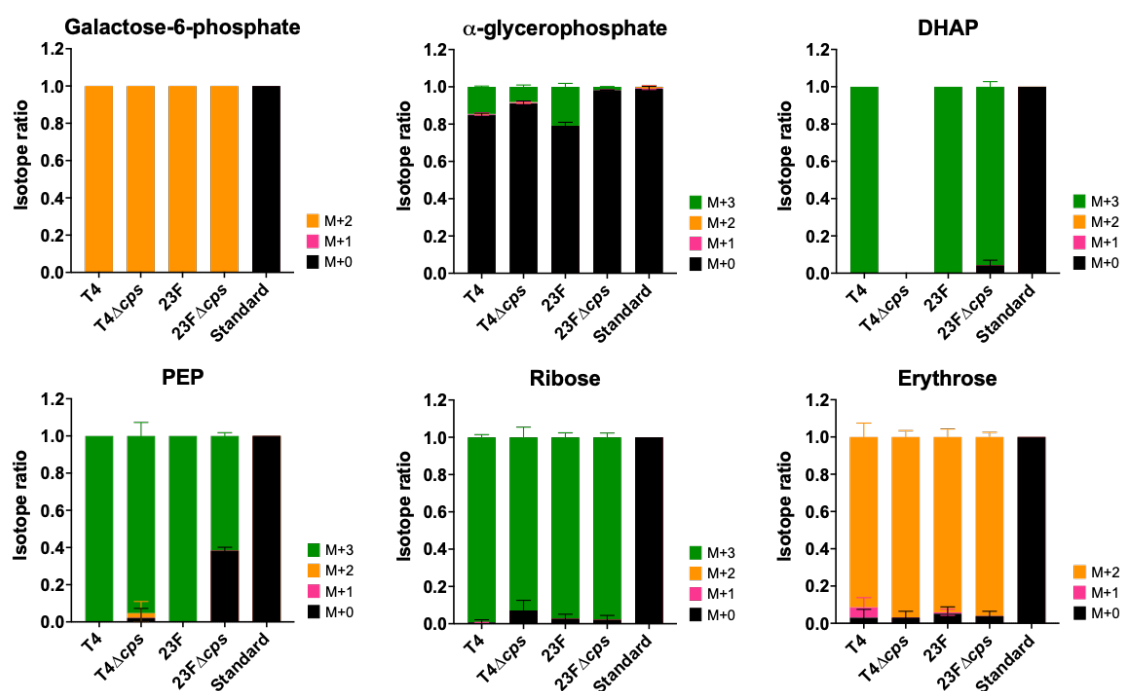
Complete labelling of extracellularly detected pyruvate, lactate, erythrose and ribose was also observed in all conditions (**Figure 5.11D, E**). The glycolytic metabolites Gal6P, F6P, DHAP, PEP and α -GP were not detected in the extracellular medium or in the uninoculated culture medium.

Carbohydrate breakdown in *S. pneumoniae* is known to proceed via glycolysis and the pentose phosphate pathway (PPP), and this is confirmed here by the complete labelling of intermediates of these pathways detectable by GC-MS.

A CDM(¹³C-glucose) – Intracellular



B CDM(¹³C-galactose) – Intracellular



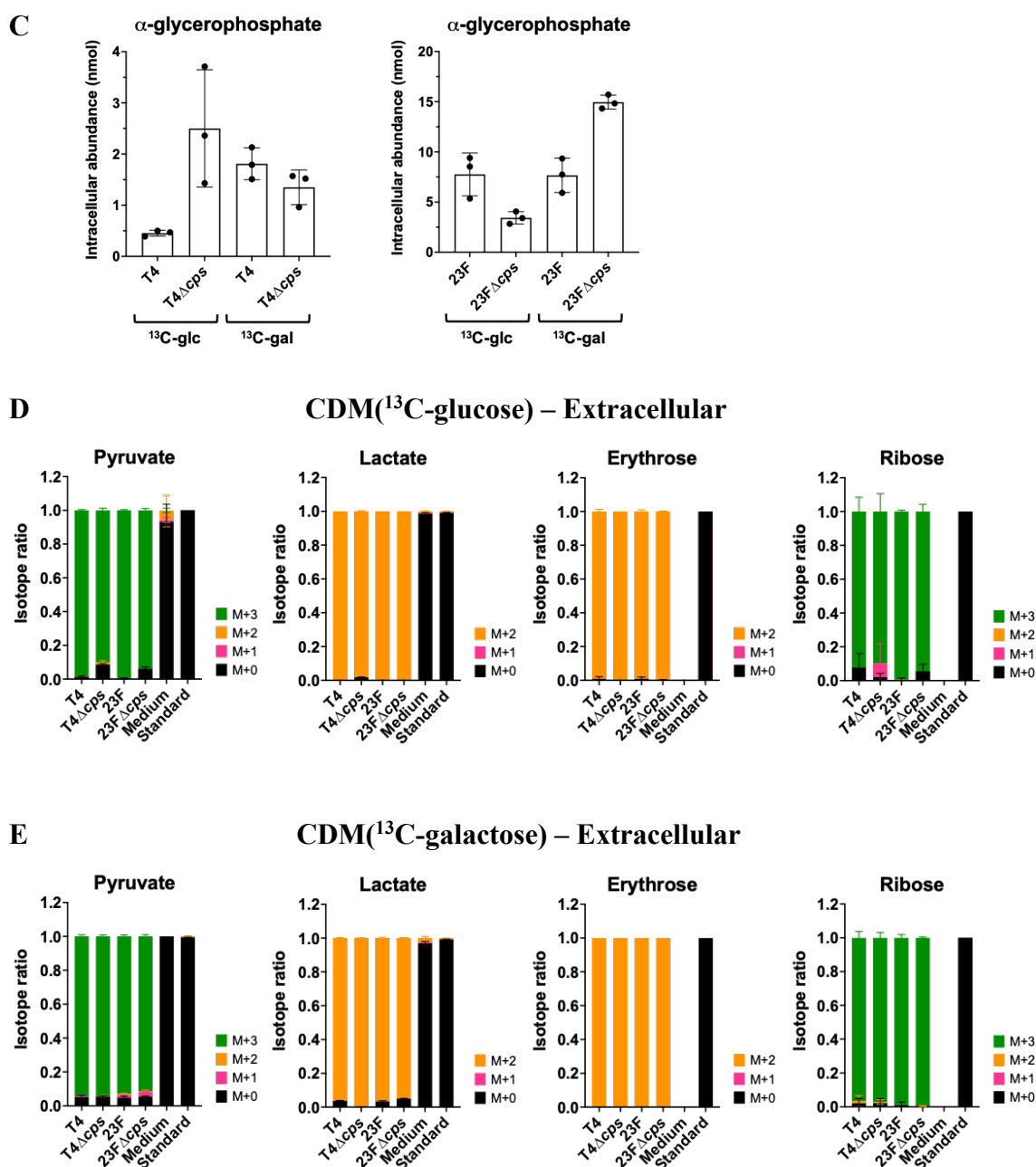


Figure 5.11 Intermediates of glycolysis and the pentose phosphate pathway – GC-MS.

Intracellular (A, B) and extracellular (D, E) isotope ratios denoting ^{13}C -label incorporation during *S. pneumoniae* culture in CDM(^{13}C -glucose) (A, D) and CDM(^{13}C -galactose) (B, E). M+[x] indicates [x] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). Medium = uninoculated culture medium; Standard = known amount of metabolite injected to visualise natural isotopic abundance. (C) Mean intracellular amounts of α -glycerophosphate per 2×10^8 CFU, with error bars denoting standard deviation. n = 3, with 2 technical replicates within each. DHAP = dihydroxyacetone phosphate; PEP = phosphoenolpyruvate.

5.5.2 Metabolites associated with the urea cycle

A diagram depicting the urea cycle is shown in **Figure 5.12A**. Aspartate, an intermediate of the urea cycle and generated from PEP-derived oxaloacetate, was only labelled (mainly [M+3]) in intracellular fractions (**Figure 5.13A, B**). Intracellular aspartate in TIGR4 showed 80-90% label, but was unlabelled and present at very low concentrations in TIGR4 Δ *cps* (**Figure 5.13A, C**). Aspartate showed similar labelling percentages (40-50% labelled) in 23F and 23F Δ *cps*. The carbohydrate source did not affect aspartate labelling patterns in any of the strains (**Figure 5.13A**).

Intracellular fumarate abundances were low, and most of these fumarate molecules were unlabelled. 25% labelling was seen in TIGR4 cultured in glucose, and 10% in galactose. TIGR4 Δ *cps* generated minimally labelled fumarate in both glucose and galactose (~5%). 23F and 23F Δ *cps* showed similar fumarate labelling in glucose (10-15%) and galactose (20%) (**Figure 5.13A**). Comparatively, fumarate present in the extracellular fractions exhibited higher [M+3] label percentages for all strains – 50-60% for TIGR4, 35-40% for 23F, ~ 60% for 23F Δ *cps* – except TIGR4 Δ *cps*, where extracellular fumarate was unlabelled and present at low concentrations (**Figure 5.13B, D**). Therefore, most of the fumarate containing ^{13}C -labelled carbon atoms was being secreted, and fumarate was derived from aspartate as seen by the corresponding [M+3] labelling of intracellular aspartate and extracellular fumarate (**Figure 5.13A, B**).

Intracellular ornithine and urea, the only other entities involved in the urea cycle which were detectable in this study, were generated but unlabelled (**Figure 5.13A, B**).

Unlabelled ornithine and urea would be expected during urea cycle activity, as the carbon atoms derived from aspartate are transferred to fumarate (**Figure 5.12B**).

Intriguingly, ornithine production was only observed in the 23F strain pair, and urea production only in the TIGR4 strain pair (**Figure 5.13A**).

Although citrulline and arginine were not detectable by the GC-MS derivatisation method used in this study, urea cycle activity in TIGR4 and 23F can be inferred from the labelling patterns of aspartate and fumarate, and the detection of ornithine and urea.

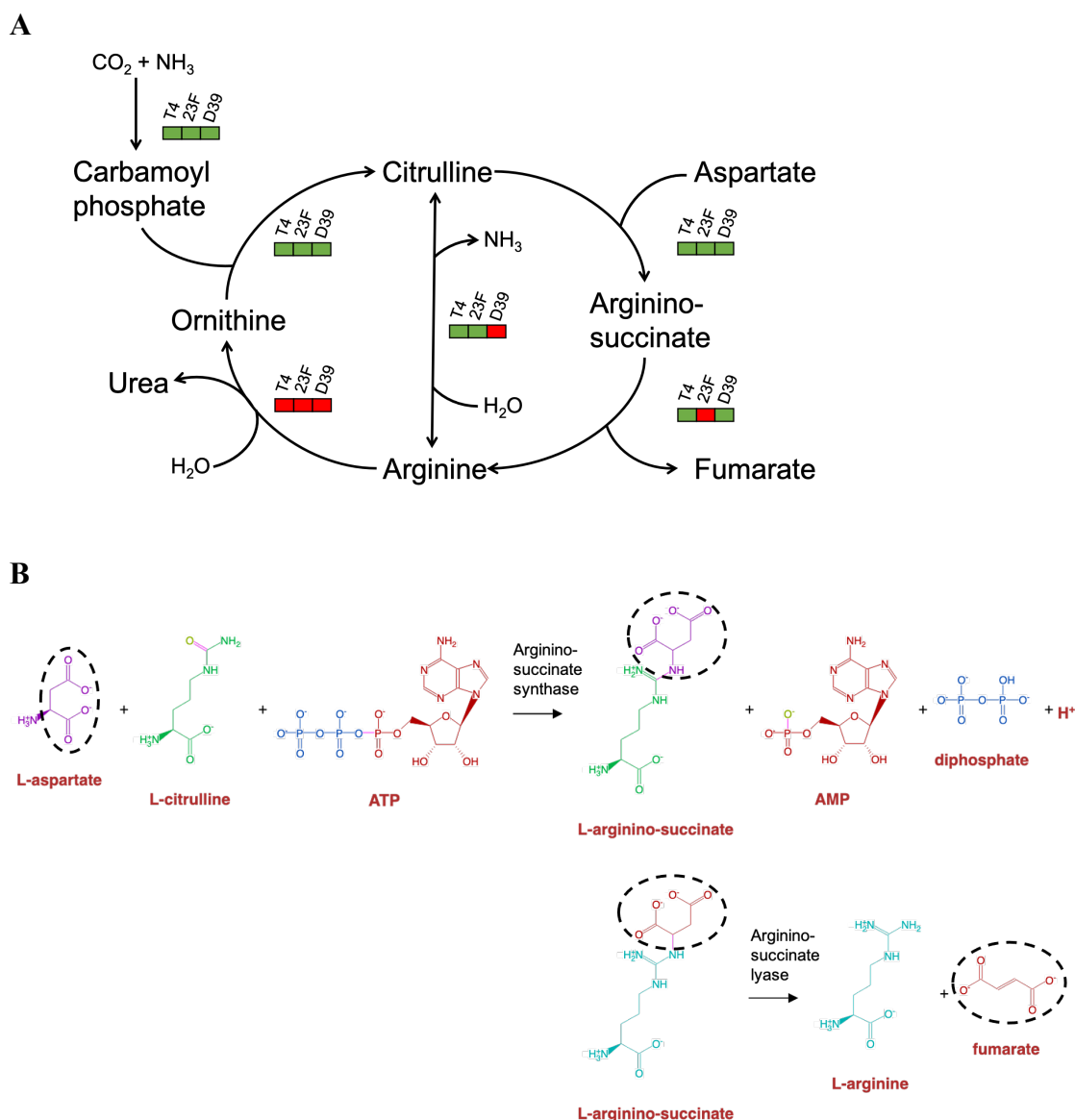
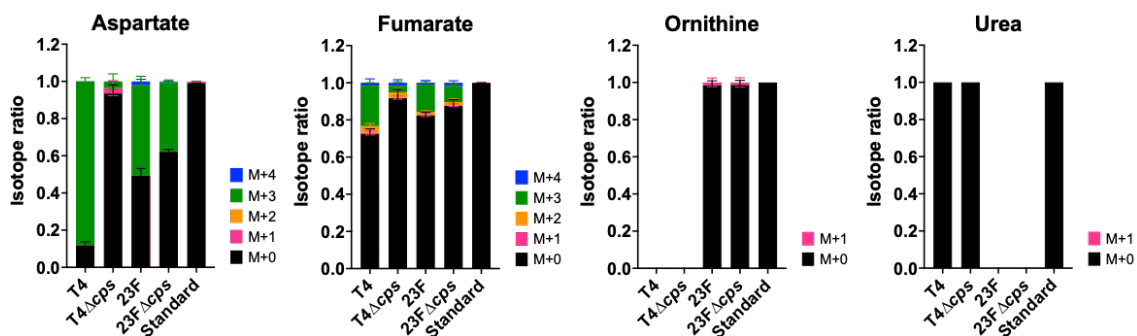
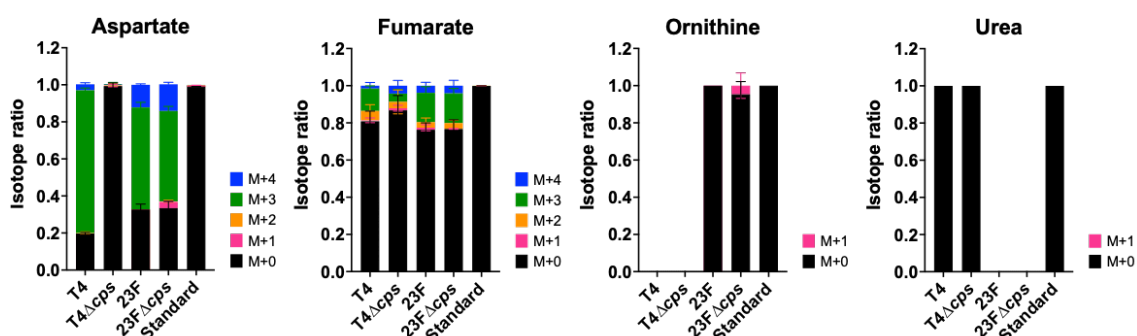


Figure 5.12 Urea cycle in *S. pneumoniae*

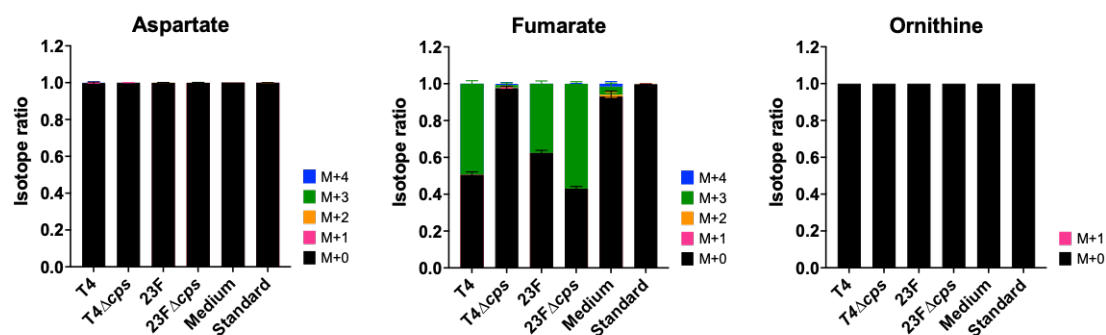
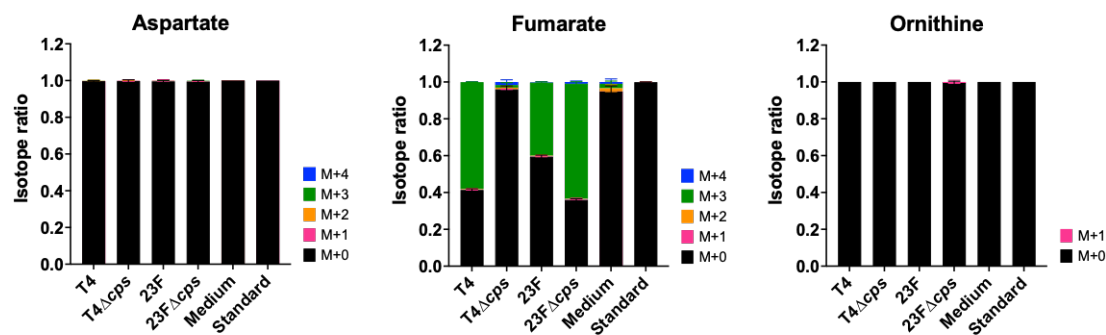
(A) Schematic depicting the urea cycle generated from up to date genome annotations in BioCyc (Karp et al., 2017) and KEGG (Kanehisa & Goto, 2000). Colour-coding represents the presence (green) or absence (red) of enzymes annotated in the genomes corresponding to *S. pneumoniae* TIGR4, 23F and D39.

(B) Figure generated in BioCyc (Karp et al., 2017). Dotted black circles indicate the same four carbon atoms from aspartate being transferred to fumarate via arginino-succinate as part of the urea cycle.

A

CDM(¹³C-glucose) – IntracellularCDM(¹³C-galactose) – Intracellular

B

CDM(¹³C-glucose) – ExtracellularCDM(¹³C-galactose) – Extracellular

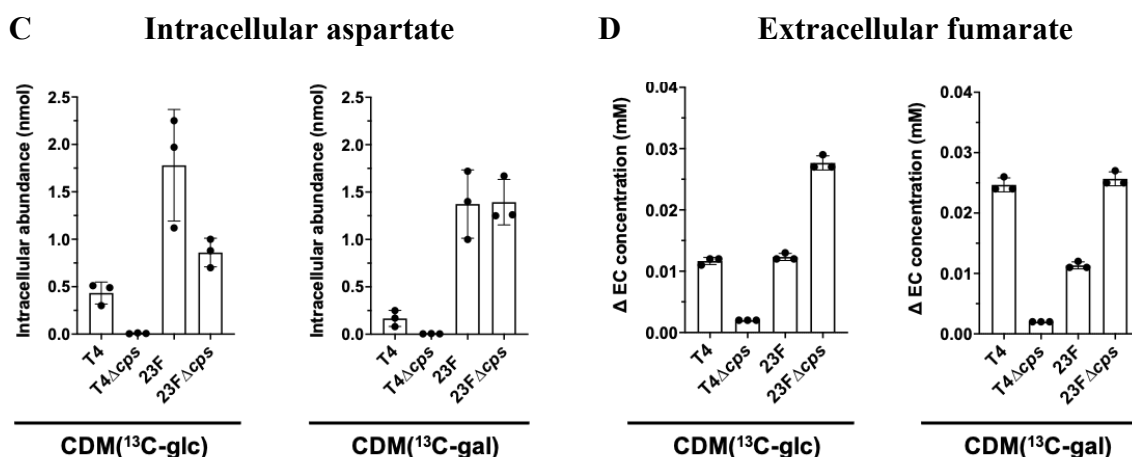


Figure 5.13 Metabolites involved in the urea cycle – GC-MS.

Intracellular (A) and extracellular (B) isotope ratios denoting ¹³C-label incorporation during *S. pneumoniae* culture in CDM(¹³C-glucose) and CDM(¹³C-galactose). M+[x] indicates [x] number of carbon atoms labelled with ¹³C carbons from ¹³C-sugar ('M' indicates the monoisotopic mass containing no ¹³C atoms). Medium = uninoculated culture medium; Standard = known amount of metabolite injected to visualise natural isotopic abundance. (C) Mean intracellular amounts of aspartate per 2x10⁸ CFU and (D) mean change in extracellular concentration of fumarate compared to uninoculated culture medium, with error bars denoting standard deviation. n = 3, with 2 technical replicates within each.

5.5.3 Succinate

Determining the source of pneumococcal succinate has been an aim throughout this study, because there are currently no enzyme annotations for succinate production in the *S. pneumoniae* genome. In both CDM(¹³C-glucose) and CDM(¹³C-galactose) cultures, intracellular succinate was detected but was entirely unlabelled (**Figure 5.14A, D**). Extracellular succinate labelling patterns were unaffected by the carbohydrate source in all strains – TIGR4 (~ 7% label), TIGR4Δcps (unlabelled), 23F (< 5% label), and 23FΔcps (15-20% label) (**Figure 5.14B, E**). Therefore, the highest amount of labelled succinate in the extracellular fraction was seen for 23FΔcps, which also showed 3-4x greater succinate abundance compared to the other strains. Extracellular succinate concentrations in TIGR4, TIGR4Δcps and 23F were similar (**Figure 5.14C, F**).

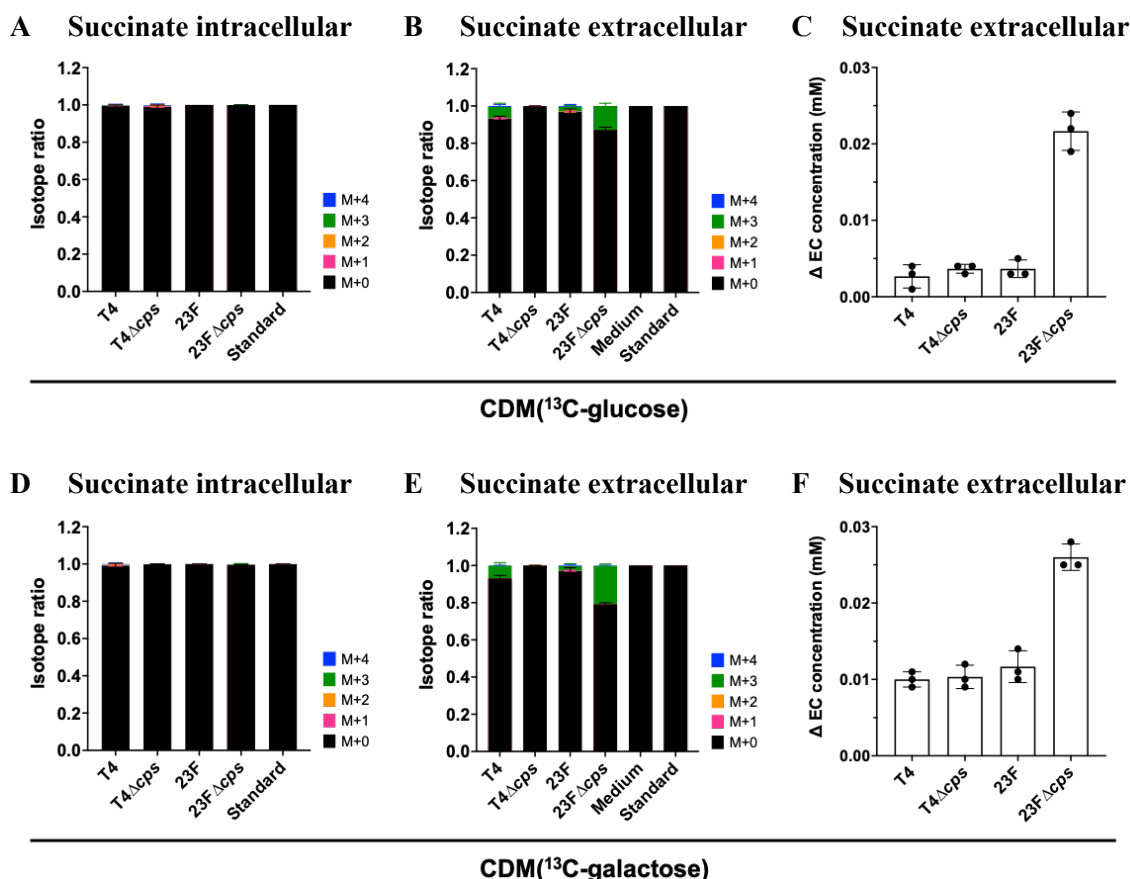


Figure 5.14 Succinate production and labelling in *S. pneumoniae* – GC-MS.

Intracellular (A, D) and extracellular (B, E) isotope ratios denoting ¹³C-label incorporation during *S. pneumoniae* culture in CDM(¹³C-glucose) and CDM(¹³C-galactose). M+[*x*] indicates [*x*] number of carbon atoms labelled with ¹³C carbons from ¹³C-sugar ('M' indicates the monoisotopic mass containing no ¹³C atoms). Medium = uninoculated culture medium; Standard = known amount of succinate injected to visualise natural isotopic abundance. (E, F) Mean change in extracellular concentration of succinate compared to uninoculated culture medium, with error bars denoting standard deviation. n = 3, with 2 technical replicates within each.

5.5.4 Glutamine and associated metabolites

No ^{13}C labelling was observed in glutamine and metabolites derived from glutamine – glutamate, 5-oxoproline, and α -ketoglutarate (**Figure 5.15**). Glutamine, an essential nitrogen source, was taken up by *S. pneumoniae* from the culture medium (**Figure 5.16A**). Glutamine and glutamate can interconvert via *glnA* (EC 6.3.1.2); I saw that extracellular glutamine uptake and glutamate accumulation patterns (not absolute concentrations) were similar across strains, showing that glutamine uptake does indeed correspond to glutamate production (**Figure 5.16A**). Glutamate is also known to rapidly interconvert with 5-oxoproline during GC-MS analysis, and similar intracellular abundance patterns for glutamate and 5-oxoproline were observed across strains (**Figure 5.16B**).

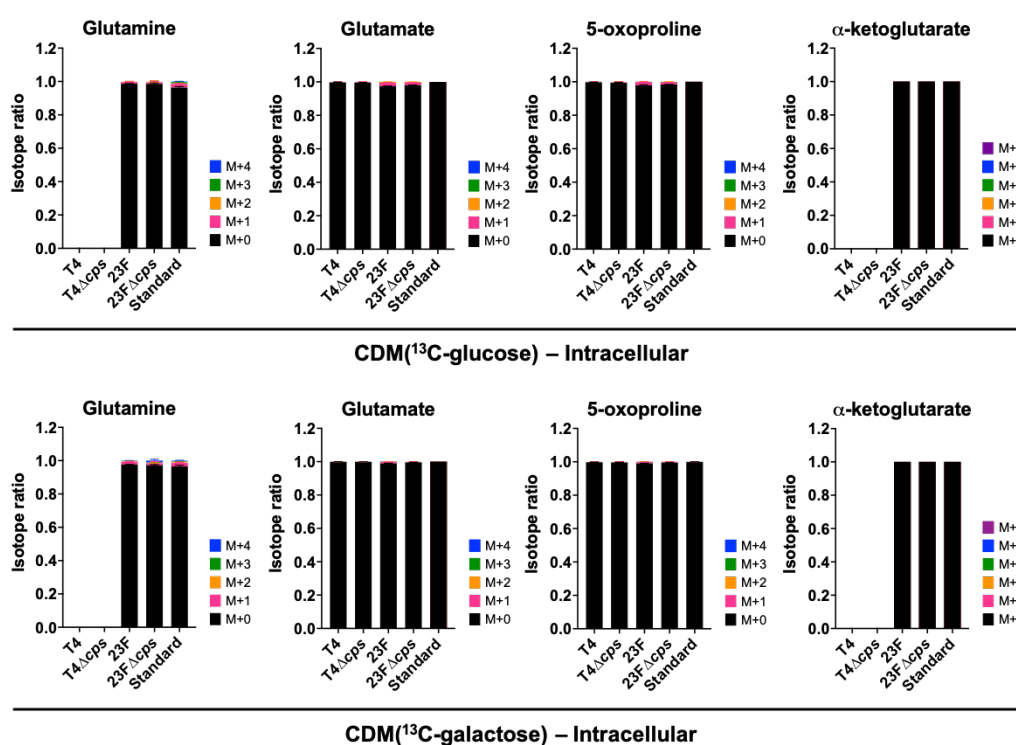


Figure 5.15 Isotope ratios of glutamine and glutamine-associated metabolites – GC-MS

Intracellular isotope ratios denoting ^{13}C -label incorporation during *S. pneumoniae* culture in CDM(^{13}C -glucose) and CDM(^{13}C -galactose). M+[x] indicates [x] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). Medium = uninoculated culture medium; Standard = known amount of metabolite injected to visualise natural isotopic abundance. $n = 3$, with 2 technical replicates within each.

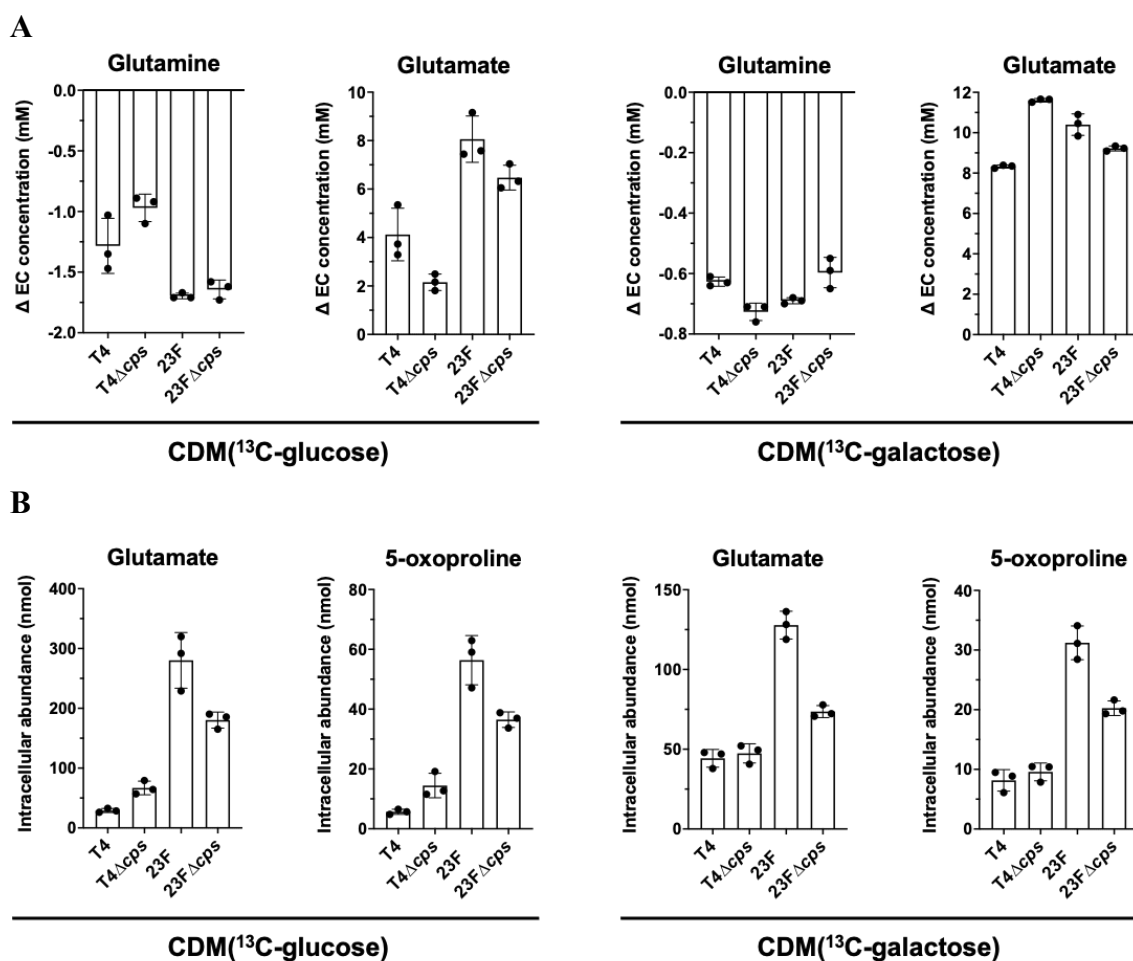


Figure 5.16 Amounts of glutamine, glutamate and 5-oxoproline – GC-MS.

(A) Mean change in extracellular concentration of glutamine and glutamate compared to uninoculated culture medium. (B) Mean intracellular amounts of glutamate and 5-oxoproline per 2×10^8 CFU. Error bars denote standard deviation. $n = 3$, with 2 technical replicates within each.

In *S. pneumoniae*, glutamate is converted to α -ketoglutarate by transamination via *aspC* which at the same time converts PEP-derived oxaloacetate to aspartate. Intracellular α -ketoglutarate was only detected in the 23F strain pair (**Figure 5.17A**), while α -ketoglutarate was present in the extracellular fractions of all strains. Abundance of extracellular α -ketoglutarate was higher in TIGR4 compared to TIGR4 Δ cps in both glucose and galactose, while similar amounts of extracellular α -ketoglutarate were observed in 23F and 23F Δ cps (**Figure 5.17B**).

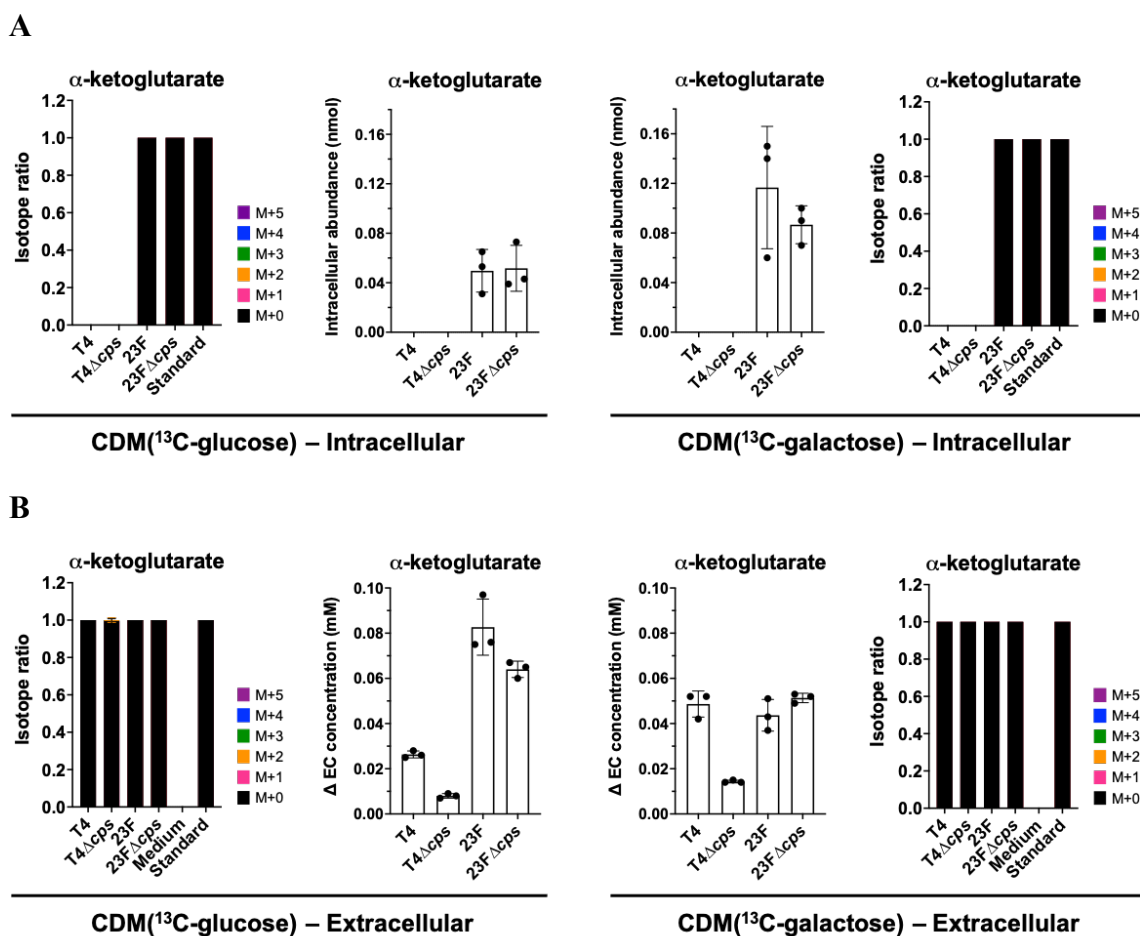


Figure 5.17 α -ketoglutarate production and labelling in *S. pneumoniae* – GC-MS.

Intracellular (A) and extracellular (B) isotope ratios (denoting ^{13}C -label incorporation) and amounts of α -ketoglutarate during *S. pneumoniae* culture in CDM(^{13}C -glucose) and CDM(^{13}C -galactose). M+[x] indicates [x] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). Medium = uninoculated culture medium; Standard = known amount of α -ketoglutarate injected to visualise natural isotopic abundance. Mean intracellular amount per 2×10^8 CFU and mean change in extracellular concentration of α -ketoglutarate compared to uninoculated culture medium are shown, with error bars denoting standard deviation. n = 3, with 2 technical replicates within each.

5.5.5 ^{13}C -label incorporation into amino acids

Amino acid metabolism in *S. pneumoniae* has been studied by others (Härtel et al., 2012; Leonard et al., 2018), albeit in a single strain (D39 Δcps or TIGR4 Δcps) cultured in excess glucose. In this section, I describe whether carbon flow through amino acids is affected by the presence of the capsular gene locus or during culture in a non-preferred sugar (galactose).

Alanine is produced from pyruvate via an alanine transaminase, and 80-90% labelling of intracellular alanine was observed in TIGR4 and TIGR4 Δcps cultured in both glucose and galactose (**Figure 5.18A, B**). 23F cultured in glucose also showed ~80% labelling in alanine, dropping to ~50% in galactose. The alanine labelling pattern was switched in 23F Δcps , i.e., ~50% label during glucose culture and ~70% label during galactose culture (**Figure 5.18A, B**). The unlabelled alanine observed was likely a by-product of selenocysteine metabolism (through the activity of *cdsB* (SPN23F07920) in *S. pneumoniae* 23F).

S. pneumoniae has previously been shown to be auxotrophic for the essential amino acid glycine (Härtel et al., 2012), and this is supported by our study showing the extracellular depletion of glycine (**Figure 5.19**). However, substantial intracellular glycine labelling was observed in TIGR4 (~50% in glucose and ~30% in galactose), which is consistent with the data in 7H9⁺(^{13}C -glucose) culture medium described in the preceding Chapter (4.7.2). Even though the glycine biosynthesis pathway is incomplete in *S. pneumoniae*, glycine may be generated by reverse action of the glycine cleavage system. It is puzzling that TIGR4 would utilise carbohydrate-derived carbons for glycine biosynthesis when it can be obtained from the culture medium, but may point to the role of glycine in purine biosynthesis. In contrast, intracellular glycine was unlabelled in TIGR4 Δcps . 23F generated ~15% labelled glycine in both glucose and galactose, while 23F Δcps only generated a similar level of labelled glycine when cultured in galactose (**Figure 5.18A, B**).

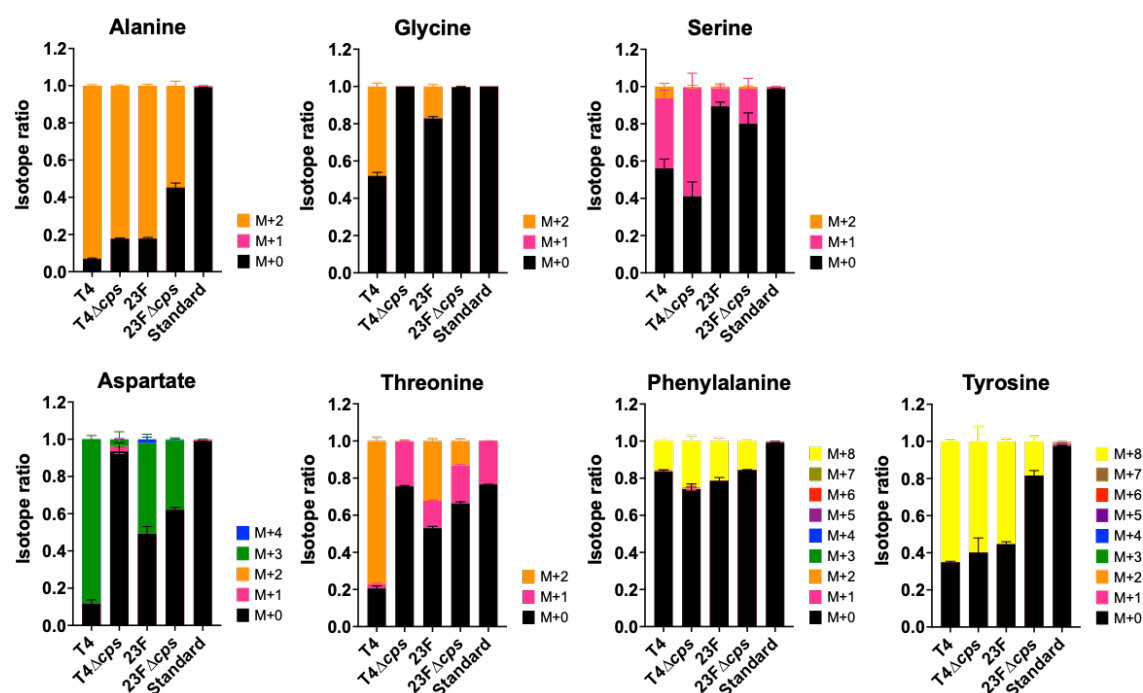
Serine can be generated in *S. pneumoniae* from formate and glycine via hydroxymethylation by *glyA*, present in both TIGR4 (gene SP_1024) and 23F (gene SPN23F09479). ^{13}C -labelled one-carbon formate (an end-product of anaerobic fermentation) and exogenous unlabelled glycine would then generate serine with one ^{13}C -labelled carbon atom ([M+1]). I observed differing proportions of intracellular [M+1] labelled serine in our strains, which was also affected by the carbohydrate source (**Figure 5.18A, B**). As no other pathways for serine biosynthesis are annotated in the pneumococcal genome, the presence of unlabelled serine cannot be explained by this study.

Asparagine was depleted from the culture medium and taken up by *S. pneumoniae* (**Figure 5.19**), and can interconvert with aspartate (SP_1970 and SPN23F19910 (EC 6.3.1.1); SP_1998 (EC 3.5.1.1)). Intracellular aspartate exhibited ^{13}C labelling in all strains except TIGR4 Δcps , and this labelling pattern is mirrored in intracellular threonine (**Figure 5.18A, B**), which is known to be generated from aspartate (Härtel et al., 2012).

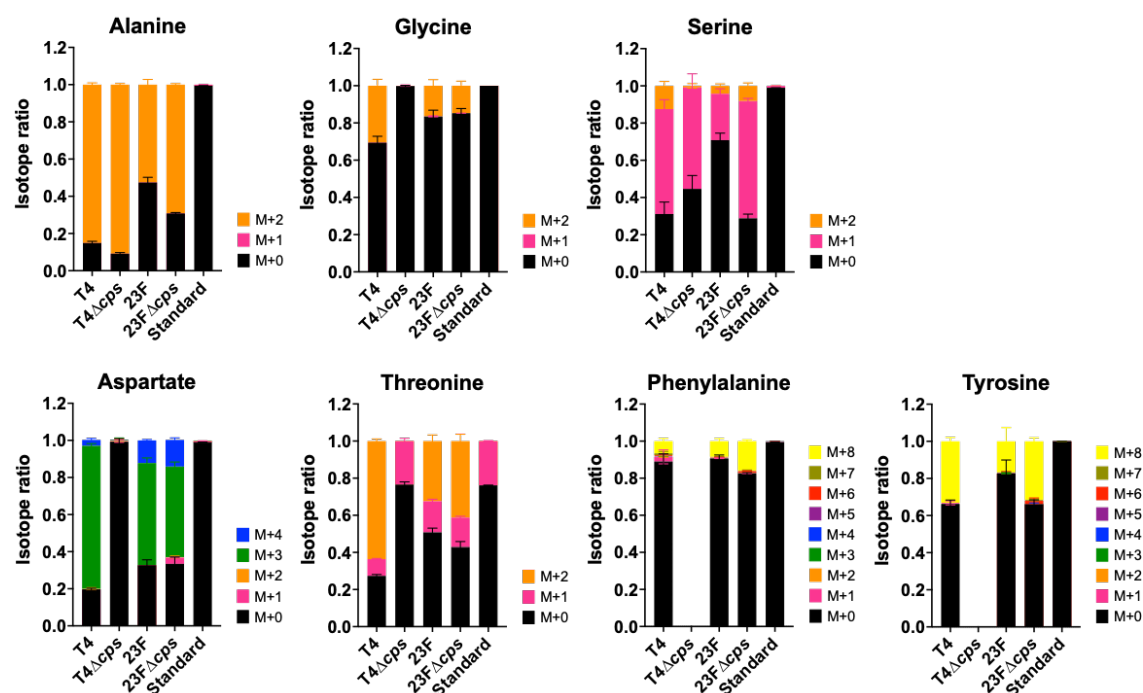
All the aforementioned amino acids were unlabelled in extracellular fractions (**Figure 5.18C, D**), so are not secreted outside *S. pneumoniae* cells.

The branched chain amino acids (BCAA) isoleucine, leucine and valine can be obtained by *S. pneumoniae* using an ABC transporter (LivJHMGF) (Basavanna et al., 2009). No labelling of intracellular leucine and isoleucine was observed, suggesting these amino acids are obtained from the culture medium. However, some [M+4] labelling of intracellular valine was observed, suggesting activity of the valine biosynthesis pathway (precursors pyruvate and threonine) along with potential uptake of exogenous valine (**Figure 5.20A, B**).

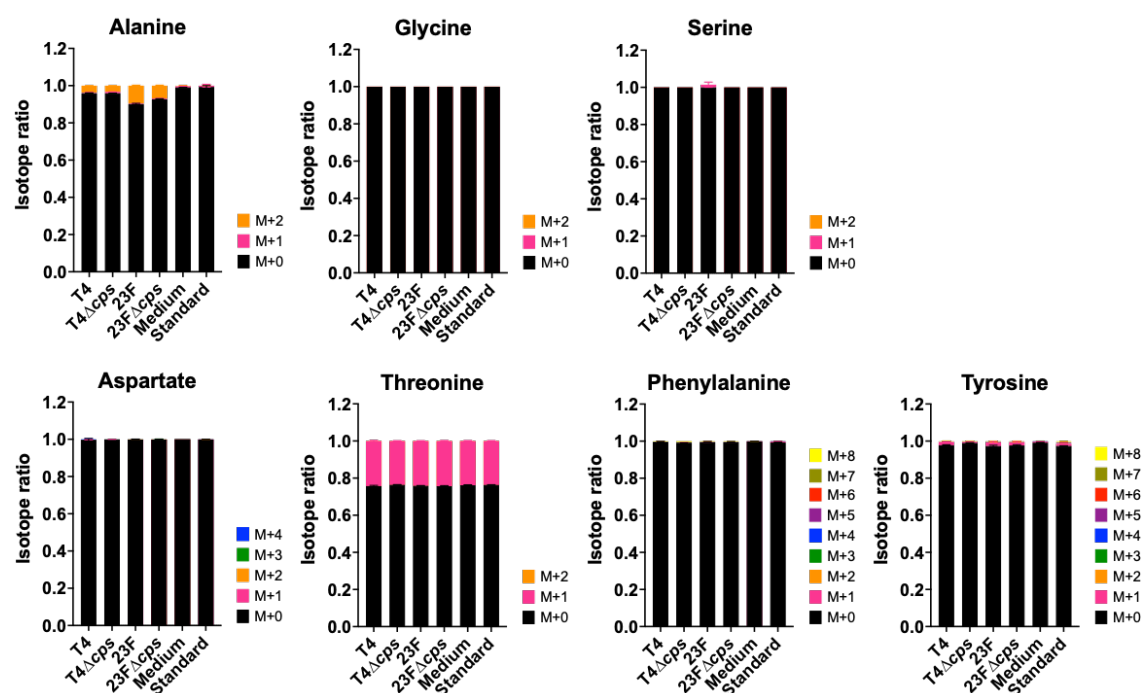
A

CDM(^{13}C -glucose) – Intracellular

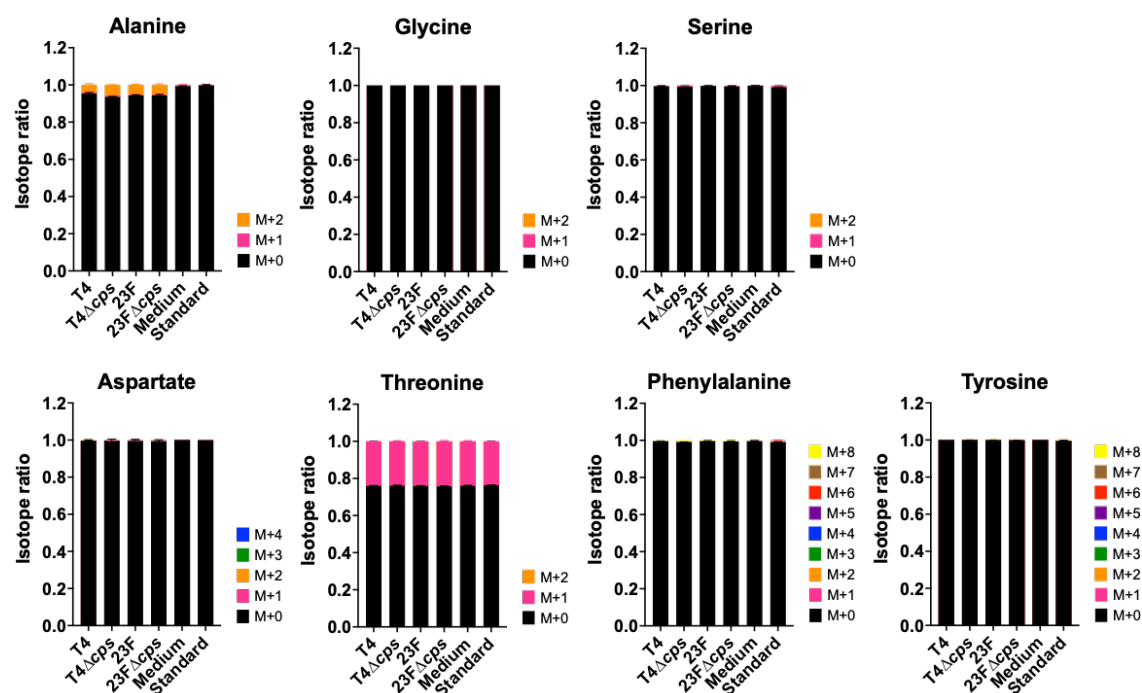
B

CDM(^{13}C -galactose) – Intracellular

C

CDM(^{13}C -glucose) – Extracellular

D

CDM(^{13}C -galactose) – ExtracellularFigure 5.18 Isotope ratios of amino acids in *S. pneumoniae* – GC-MS.

Intracellular (A, B) and extracellular (C, D) isotope ratios denoting ^{13}C -label incorporation during *S. pneumoniae* culture in CDM(^{13}C -glucose) (A, C) and CDM(^{13}C -galactose) (B, D). $\text{M}+[x]$ indicates $[x]$ number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). Medium = uninoculated culture medium; Standard = known amount of amino acid injected to visualise natural isotopic abundance. $n = 3$, with 2 technical replicates within each.

5.5.6 Unlabelled amino acids

Histidine can be synthesised from the PPP intermediate ribose-5-phosphate, but this biosynthetic pathway is absent in *S. pneumoniae* TIGR4 and 23F. The lysine biosynthesis pathway is also absent in *S. pneumoniae*. In agreement with this, *S. pneumoniae* in our system obtained histidine and lysine from the culture medium, as seen by the extracellular depletion of these amino acids (**Figure 5.19**). Proline, cysteine, and methionine were also unlabelled (**Figure 5.20**); proline and cysteine biosynthetic pathways have been postulated to be absent in *S. pneumoniae*, and methionine is generated from unlabelled intermediates (Leonard & Lalk, 2018).

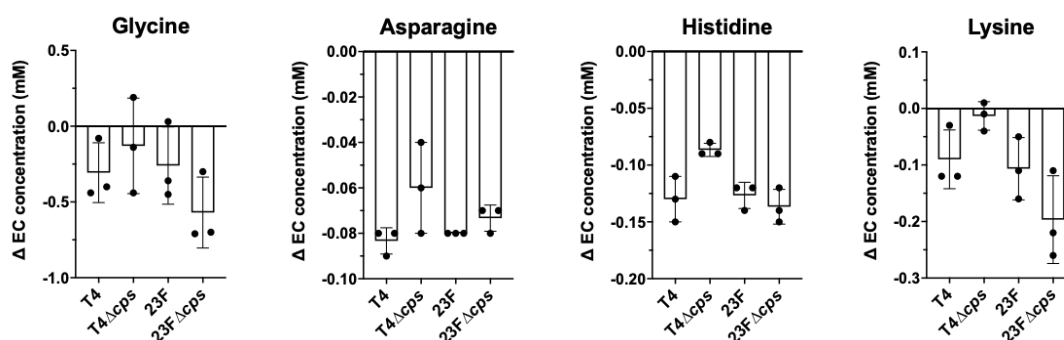
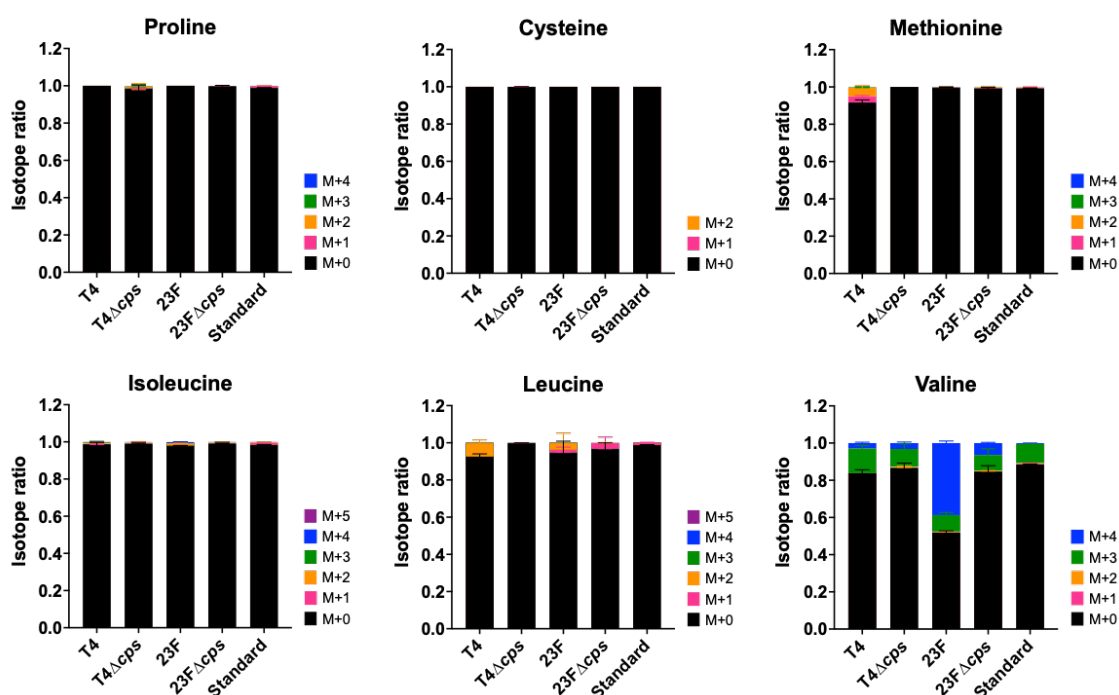


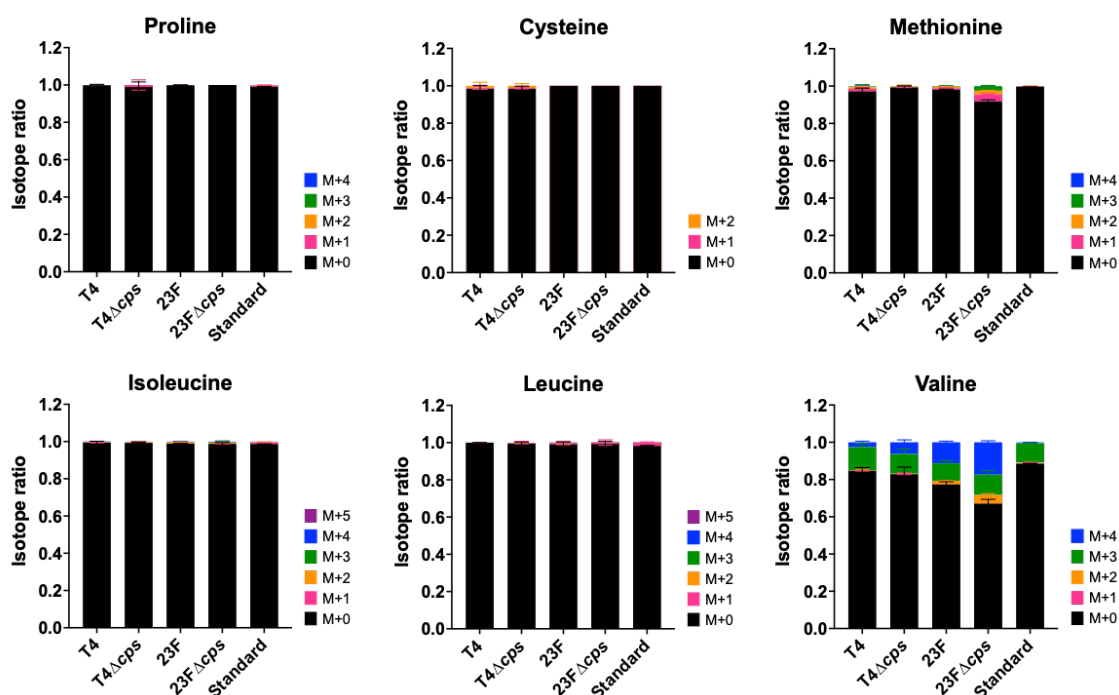
Figure 5.19 Amino acids depleted from the culture medium during *S. pneumoniae* growth.

The mean change in extracellular concentration of amino acids compared to uninoculated culture medium analysed by GC-MS is shown, with error bars denoting standard deviation. $n = 3$, with 2 technical replicates within each.

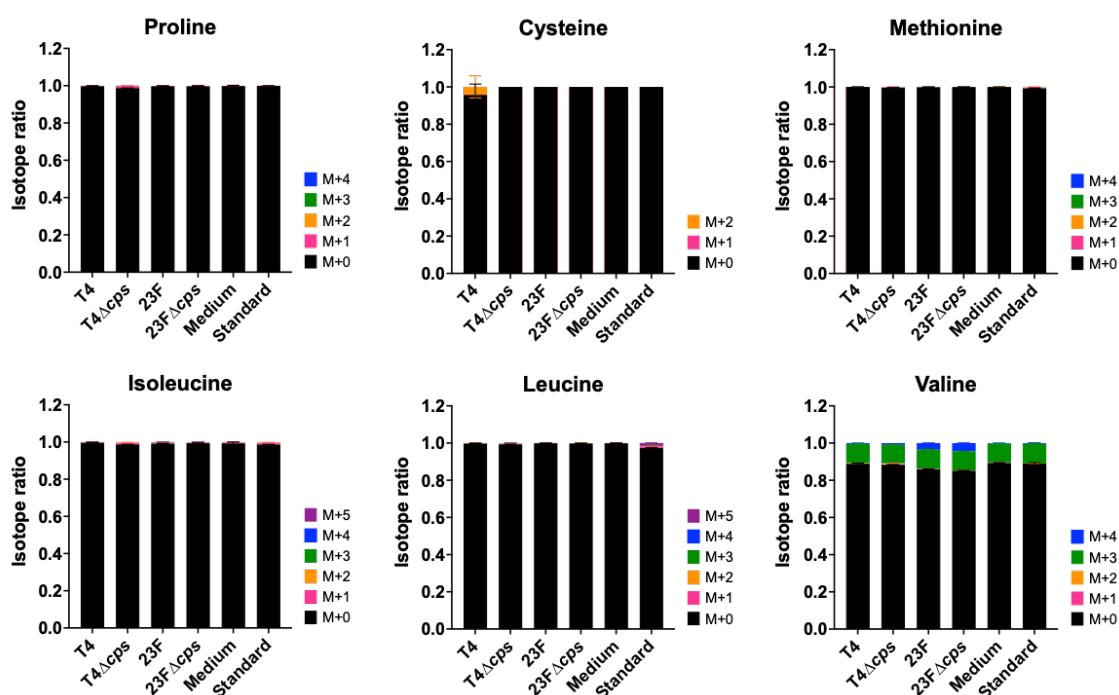
A

CDM(^{13}C -glucose) – Intracellular

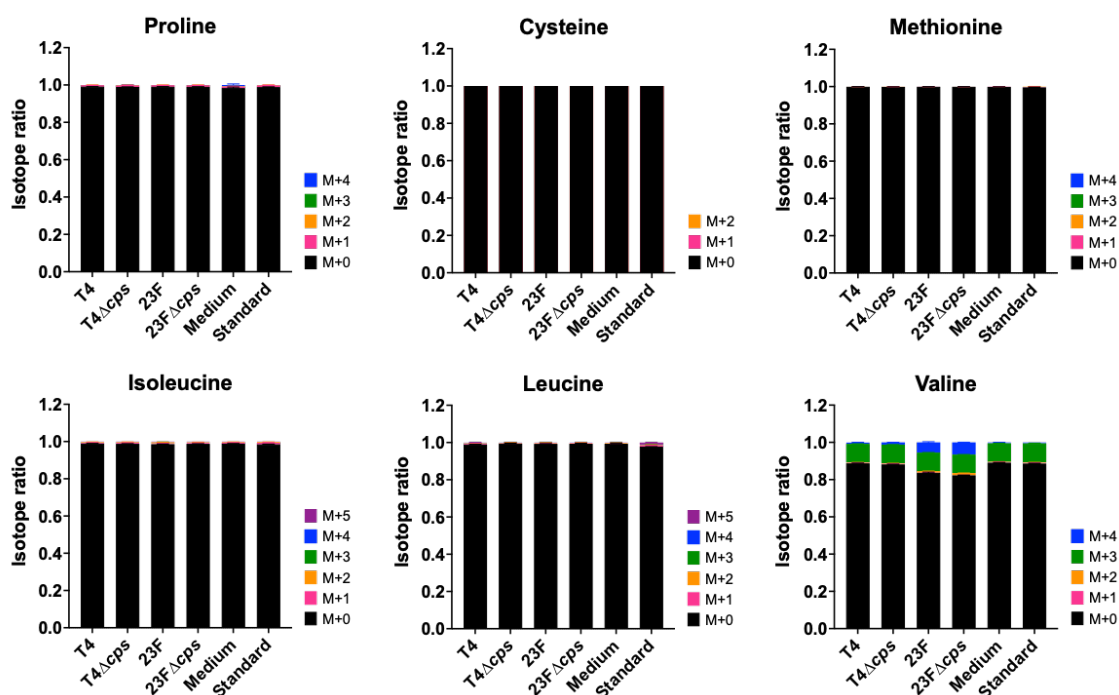
B

CDM(^{13}C -galactose) – Intracellular

C

CDM(^{13}C -glucose) – Extracellular

D

CDM(^{13}C -galactose) – ExtracellularFigure 5.20 Amino acids unlabelled during *S. pneumoniae* growth – GC-MS.

Intracellular (A, B) and extracellular (C, D) isotope ratios of amino acids denoting ^{13}C -label incorporation during *S. pneumoniae* culture in CDM(^{13}C -glucose) (A, C) and CDM(^{13}C -galactose) (B, D). M+[x] indicates [x] number of carbon atoms labelled with ^{13}C carbons from ^{13}C -sugar ('M' indicates the monoisotopic mass containing no ^{13}C atoms). Medium = uninoculated culture medium; Standard = known amount of amino acid injected to visualise natural isotopic abundance. n = 3, with 2 technical replicates within each.

5.6 Chapter 5 Discussion

In this Chapter, I first described the establishment of a method for the extraction of intracellular metabolites of *S. pneumoniae* for GC-MS analysis, showing that the most efficient extraction was obtained using the extraction solvent methanol + acetonitrile + water (2:2:1 volume) coupled with bead-assisted homogenisation of bacterial cells to release intracellular contents. I also described the process of standardising pneumococcal cell numbers and injection amounts of intracellular extracts for GC-MS analysis to enable the direct comparison of absolute intracellular amounts and isotope ratios of metabolites across the four *S. pneumoniae* strains studied – TIGR4, TIGR4 Δ *cps*, 23F and 23F Δ *cps*.

I then used [U- $^{13}\text{C}_6$]-glucose and [U- $^{13}\text{C}_6$]-galactose to trace carbon flow through pneumococcal metabolic pathways and study carbohydrate metabolism of *S. pneumoniae* cultured in the widely-used chemically defined medium CDM (adapted from van de Rijn & Kessler, 1980), with a particular focus on understanding the relationship between central carbon metabolism and capsule metabolism by using wild-type TIGR4 and 23F strains along with their capsule-deficient derivatives. ^1H -NMR and GC-MS were employed to detect a wide range of intracellular and extracellular metabolites, contributing to the most extensive metabolic profile of these *S. pneumoniae* strains generated to date.

5.6.1 Succinate and fumarate production in *S. pneumoniae*

In the previous Chapter, extracellular accumulation of succinate and fumarate was observed in *S. pneumoniae* TIGR4 corresponding to growth, and it was determined that they were not intermediates of a canonical TCA cycle. Here, I hypothesised that succinate is a by-product of anaerobic fermentation derived from formate, and fumarate is a by-product of the urea cycle and derived from aspartate.

Indeed, ^{13}C -label incorporation ([M+3]) was only observed in extracellular succinate and fumarate, suggesting these are secreted products of carbohydrate metabolism which

do not further participate in other metabolic or anabolic pathways. I also showed that fumarate is derived from aspartate, as corresponding [M+3] labelling patterns were observed for aspartate and fumarate in all strains, and that this occurs as part of the urea cycle – intracellular ornithine and urea production was observed in this GC-MS analysis, and the depletion of arginine corresponding to an increase in ornithine has been previously reported in *S. pneumoniae* (Leonard et al., 2018), suggesting that the genomic absence of the arginase converting arginine to urea and ornithine assumed in metabolic pathway reconstructions (KEGG and BioCyc) (**Figure 5.12A**) may be a result of incomplete annotation. Together, this data provides further evidence for the activity of a pneumococcal urea cycle.

As for succinate, appreciable extracellular ^{13}C -label incorporation ([M+3]) was only observed in 23F Δ *cps*, while the other strains showed no or negligible labelling. [M+3] labelled succinate may theoretically arise from formate ([M+1], released as $^{13}\text{CO}_2$) and fumarate ([M+3]), but this enzyme is not annotated in the *S. pneumoniae* genome. Moreover, the extracellular concentrations of succinate and formate did not correspond to each other, disproving the hypothesis that succinate is a by-product of anaerobic fermentation. The most likely explanation for [M+3] labelled succinate production based on the available data is through spontaneous fumarate conversion. Unlabelled intracellular and extracellular succinate observed is likely produced from α -ketoglutarate generated from exogenous (unlabelled) glutamate, but analysis with ^{13}C -glutamate would be necessary to confirm this.

5.6.2 The effect of carbohydrate source on *S. pneumoniae* metabolism

^1H -NMR and GC-MS analyses confirmed that carbon flow is mainly directed towards fermentation in all strains, as shown by label incorporation corresponding to carbohydrate uptake in extracellular end-products of fermentation (lactate, acetate, formate, and to a lesser extent, ethanol) via glycolysis, as shown by the complete labelling of glycolytic intermediates. The hypothesis that carbon flow through carbohydrate metabolic pathways besides fermentation would be unaffected during *S. pneumoniae* culture in ^{13}C -glucose or ^{13}C -galactose was found to be true for all strains

tested, exemplifying the metabolic adaptability of *S. pneumoniae* in the context of its demonstrated ability to utilise multiple carbohydrate sources.

Extracellular fermentation profiles of TIGR4 and 23F differed depending on the carbohydrate supplied – homolactic fermentation was observed during growth on glucose while *S. pneumoniae* grown on galactose produced a mixture of lactate, formate, acetate and ethanol (mixed acid fermentation). The differential fermentation pattern on glucose and galactose has been described in *S. pneumoniae* D39 by others, and its observation in TIGR4 and 23F here confirms that this is a conserved adaptation to growth in two carbon sources prevalent in different host sites. Acetate production results in the generation of ATP and H₂O₂, and other enzymes of mixed acid fermentation have been implicated in capsule production and virulence in mouse models of disease (Al-Bayati et al., 2017; Echlin et al., 2020, 2016; Yesilkaya et al., 2009).

5.6.3 The link between outer polysaccharide capsule and carbohydrate metabolism in *S. pneumoniae*

A key question I set out to address in this Chapter was the role of capsular polysaccharides within carbohydrate metabolism in *S. pneumoniae* – does the lack of a capsule in Δcps strains translate to a reduction in the energetic burden of synthesising capsule precursor polysaccharides, or does this remove the additional source of carbon that may be derived from capsule sugars in its role as a carbon reserve?

The TIGR4 capsule is made of repeating units of [pyruvate – galactose – *N*-acetylmannosamine – *N*-acetylfucosamine – *N*-acetylgalactosamine] and the 23F capsule is made of repeating units of [rhamnose – galactose – rhamnose – glucose] (Bentley et al., 2006) (Chapter 1 **Figure 1.1**). While capsule biosynthesis in TIGR4 and 23F comes at a similar energetic cost (as determined by the number of high-energy bonds per polysaccharide repeat unit) (Weinberger et al., 2009), the TIGR4 capsule incorporates a higher number of carbon atoms per repeat unit and can therefore be

presumed to have a higher nutritional cost for biosynthesis compared to the 23F capsule (Hathaway et al., 2012; Weinberger et al., 2009).

In this analysis, I found that the lack of a capsule generated different metabolic profiles in TIGR4 Δ *cps* and 23F Δ *cps* relative to their isogenic wild type strains. TIGR4 Δ *cps* generated a similar proportion of formate, acetate and ethanol when cultured in glucose or galactose (semi-mixed acid fermentation), contrasting with the clear homolactic (in glucose)/ mixed acid (in galactose) fermentation observed in TIGR4. 23F Δ *cps*, on the other hand, generated carbohydrate-dependent fermentation profiles identical to 23F.

During growth on galactose, TIGR4 Δ *cps* produced more lactate and less formate, acetate and ethanol compared to TIGR4, indicating increased lactate dehydrogenase (*ldh*) activity and reduced activity of pyruvate formate lyase (*pfl*) and acetate kinase (*ackA*). Intriguingly, the 23F/ 23F Δ *cps* strain pair displays the opposite fermentation trend to TIGR4/ TIGR4 Δ *cps* during galactose culture – 23F Δ *cps* produces more formate, acetate and ethanol and (marginally) less lactate than 23F. Genes of the *cps* locus are not known to affect pyruvate metabolism (which is the precursor for fermentation end-products), so this is unlikely to be an artefactual genetic effect of the *cps* deletion.

I hypothesised that in the absence of the metabolic burden of capsule biosynthesis, as modelled by the Δ *cps* strains, carbon flow would be directed downstream through glycolysis towards fermentation. This was found to be true for 23F/ 23F Δ *cps*, but false for TIGR4/ TIGR4 Δ *cps* where the overall concentration of fermentation end-products was lower in capsule-deficient TIGR4 compared to wild-type TIGR4 during galactose culture. This suggests that in the absence of external stressors (from immune cells, for example, which are not part of the experimental system described in this thesis), the balance between capsule production as a metabolic burden and capsular polysaccharides as a carbon reserve differs based on the strain background – when capsule biosynthesis is reduced, the existing TIGR4 capsule may be broken down to feed pneumococcal energetic metabolism, which does not seem to be a major factor in 23F.

I also observed reduced production of and ^{13}C -label incorporation into secondary (non-glycolytic) metabolites of TIGR4 Δcps , further suggesting that deletion of the *cps* gene locus, or indeed more likely the lack of the capsule itself, has an inhibitory effect on the ability of *S. pneumoniae* TIGR4 to assimilate carbohydrate-derived carbons. On the other hand, 23F Δcps showed similar ^{13}C -label incorporation into secondary metabolites as in 23F, suggesting that 23F is more metabolically resilient to the loss of capsule than TIGR4. This is an important finding that may be a contributing explanation for the greater nasopharyngeal carriage prevalence of 23F in the population, compared to TIGR4 which predominantly causes invasive pneumococcal disease (Brueggemann et al., 2003).

In summary, this Chapter establishes a link between the pneumococcal capsule and carbohydrate metabolism, and further shows that differences between *S. pneumoniae* strains in this context may inform carriage prevalence and invasive potential. This Chapter also provides evidence for the existence of a urea cycle in *S. pneumoniae*, despite a genomically incomplete urea cycle.

Chapter 6. Final Discussion

In recent years, there has been a growing appreciation of bacterial metabolism as a determinant of pathogenesis (Muñoz-Elías & McKinney, 2006; Richardson et al., 2015), as invading bacteria must be able to adapt their metabolism to the conditions of the host niche they inhabit. Some researchers have put forth the idea that the development of pathogenic characteristics may be informed by bacteria seeking to exploit nutritional sources within certain animal microenvironments (Brown et al., 2008; Rohmer et al., 2011). In the Introduction to this thesis (**Chapter 1**), I proposed that the description of *S. pneumoniae* virulence should be expanded from classical virulence factors to include its metabolism, considering its ability to survive in nutritionally and environmentally disparate host microenvironments encountered during colonisation and disease.

The primary aim of this thesis, therefore, was to characterise the central carbon metabolism of two *S. pneumoniae* strains (TIGR4 and 23F) by studying carbon flow in the context of two important aspects of its pathogenesis – the carbohydrate source available, and the outer polysaccharide capsule. It was interesting to see that there is historical precedence for these relationships, as early published studies aiming to understand *S. pneumoniae* metabolism used naturally encapsulated (“S form”) pneumococcal “types” (differentiated based on capsule composition, which later became classified as serotypes) and their spontaneous unencapsulated derivatives (“R form”), observing that respiration (amount of O₂ consumed) and glycolysis (amount of lactate produced) were affected by encapsulation, serotype, and the sugar supplied (Finkle, 1931; 1936).

6.1 Hypothesis-led metabolomics

6.1.1 *S. pneumoniae* displays conserved metabolic adaptation to carbohydrates

Glucose and galactose were used in this study to model the most abundant sugars available to the pneumococcus in the blood (invasive phenotype) and the upper respiratory tract (colonisation phenotype), respectively. Others have shown that the ratio of fermentation end-products (lactate, formate, acetate, and ethanol) generated in *S. pneumoniae* D39 differs during pneumococcal growth on glucose and galactose, with glucose catabolised predominantly to lactate and galactose catabolism generating a mix of all end-products of fermentation (Paixão et al., 2015a; 2015b; Yesilkaya et al., 2009).

In *S. pneumoniae* serotypes 4 and 23F, I observed the same differential fermentation profiles on glucose and galactose, showing that the adaptation to carbohydrate source is conserved in these, and possibly other, strains. These results also demonstrate that 23F, a serotype which is much more prevalent in carriage compared to invasive disease and thus would encounter minimal glucose in its predominant upper respiratory tract (URT) niche, retains the ability to switch its metabolism to survive in a different host site upon invasion into the blood or CSF where glucose is more abundant.

The fermentation profile is determined by enzymes involved in the metabolism of pyruvate, an intermediate of glycolytic carbohydrate breakdown, and the induction of these enzymes has been implicated in aspects of pneumococcal pathogenesis such as capsule production, hydrogen peroxide production, and bacterial burden in the nasopharynx and blood of infected mice (**Figure 6.1**).

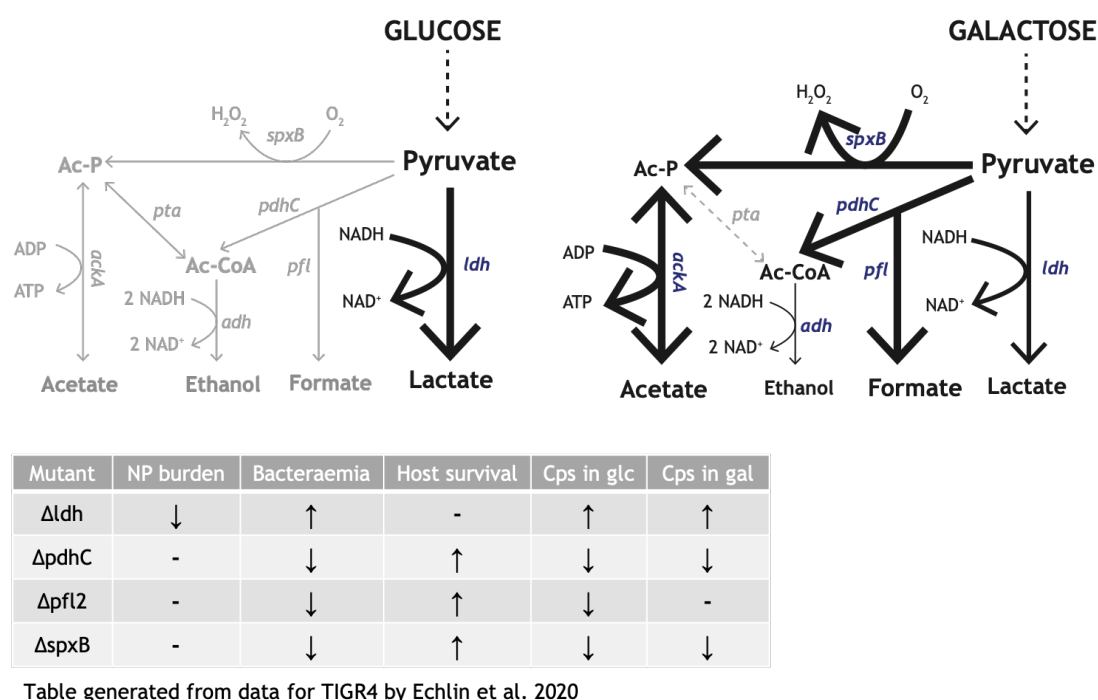


Figure 6.1 Summary schematic – Fermentation profiles of *S. pneumoniae* TIGR4 and 23F on glucose and galactose

Pathways illustrate results of extracellular concentrations of ^{13}C -labelled fermentation end-products (lactate, acetate, formate, ethanol) when *S. pneumoniae* TIGR4 or 23F was cultured in ^{13}C -glucose or ^{13}C -galactose, determined by ^1H -NMR (Chapter 5.4.1). Arrow sizes depict the relative amount of fermentation products generated. Grey = no/negligible pathway activity. The table has been generated from data for TIGR4 from another study using a mouse model (Echlin et al., 2020). Abbreviations: Ac-P = acetyl-phosphate; Ac-CoA = acetyl-CoA; NP = nasopharynx; Cps = capsule production; glc = glucose; gal = galactose.

I also showed that the difference in metabolic profiles during growth on glucose and galactose was limited to fermentation, as isotope tracing with $^{13}\text{C}_6$ -labelled carbohydrates showed identical patterns of label incorporation in intracellular and extracellular amino acids and CCM-associated metabolites between glucose and galactose for both TIGR4 and 23F. This shows that enzymes and products of pyruvate metabolism do not affect other metabolic and anabolic pathways, highlighting the ability of *S. pneumoniae* to effectively assimilate multiple carbohydrate sources to succeed in distinct host niches.

6.1.2 Capsule type of TIGR4 and 23F dictates differential carbon flow through pneumococcal metabolic pathways

The pneumococcal polysaccharide capsule exerts wide-ranging effects on growth; colonisation, through interactions with the epithelium and repelling respiratory mucus; immune evasion; and virulence (Hathaway et al., 2012; Hyams, 2009; Magee & Yother, 2001; Nahm et al., 2019; Nelson et al., 2007; Preston & Dockrell, 2008; Weight et al., 2019). Capsule biosynthesis is closely connected to carbohydrate metabolism, with capsule sugars being generated from intermediates of glycolysis, and the pyruvate metabolism enzymes *spxB* and *pdhC* affecting the transcriptional regulation of the *cps* biosynthetic locus and the availability of acetyl-CoA for the production of acetylated capsule sugars (Carvalho et al., 2013a; Echlin et al., 2020; 2016). Recent evidence has emerged for the breakdown of capsule sugars to supply carbohydrates for energetic metabolism during starvation, which may be a condition encountered during pneumococcal transmission between hosts (Hamaguchi et al., 2018), suggesting a role for the capsule in carbon storage.

However, the effect of encapsulation on global carbon metabolism has not yet been reported in the literature. In this thesis, I compared the metabolic profiles of wild-type encapsulated *S. pneumoniae* TIGR4 and 23F with their capsule-deficient derivatives (TIGR4 Δ *cps* and 23F Δ *cps*) to determine whether carbon flow through energetic and anabolic pathways would be affected by the lack of a capsule. The isogenic capsule-deficient mutants were constructed by replacing the *cps* biosynthetic gene locus with a Janus cassette (Trzciński et al., 2003), a highly specific recombination method with a low likelihood of off-target genetic effects due to the presence of conserved genes *dexB* and *aliA* flanking the *cps* locus.

Capsule biosynthesis is an additional metabolic cost for the pneumococcus. I found that 23F Δ *cps* generated a similar fermentation profile to 23F (and TIGR4), i.e., homolactic fermentation when supplied with glucose and mixed acid fermentation when supplied with galactose, but 23F Δ *cps* produced significantly more carbohydrate-derived formate, acetate, and ethanol during galactose culture than 23F. This suggests that lack of the

23F capsule removes the biosynthetic burden of capsule production, re-directing nutritional carbons towards growth and energetic metabolism via fermentation.

TIGR4 Δcps , on the other hand, did not exhibit the characteristic homolactic/ mixed acid fermentation on glucose/ galactose, but rather displayed a “semi-mixed acid” fermentation profile regardless of the carbohydrate supplied – small and similar amounts of formate, acetate, and ethanol produced on glucose and galactose, with a shift towards increased lactate production on galactose. The implications of this are uncertain but, taken together with the reduced ^{13}C -label incorporation into some amino acids by TIGR4 Δcps compared to the wild-type, suggests that the lack of a capsule in TIGR4 may inhibit growth and the assimilation of carbohydrate-derived carbons. By comparison, 23F Δcps showed similar ^{13}C -label incorporation into CCM-associated metabolites as 23F, suggesting that 23F may be more metabolically resilient to the lack of capsule than TIGR4.

In summary, I saw divergent carbohydrate-derived carbon flow between TIGR4 Δcps and 23F Δcps in a manner suggesting that the capsule is essential for normal metabolism of TIGR4 but is much less crucial for 23F. These observations may explain the greater prevalence of TIGR4 in invasive disease compared to carriage (Brueggemann et al., 2003), as the capsule is a requirement for immune evasion during invasion of host sites such as the blood. Whether this divergent carbon flow has been a driver or effect of invasive potential remains uncertain. Additionally, the serotype 4 and 23F capsules are biochemically distinct (Chapter 1 **Figure 1.1**) and are likely to have different biosynthetic costs (Hathaway et al., 2012; Weinberger et al., 2009). Therefore, further work is required to delineate whether these observations are serotype-dependent or dependent on the strain genetic background, and to prove a link with carriage prevalence or invasive potential. To this end, similar isotope labelling studies with capsule-switch mutants (same strain with capsules of different serotypes) and additional serotypes showing high (e.g., 6B, 9V, 19F) and low (e.g., 7F, 14, 15, 18C) carriage prevalence would be useful.

6.2 Discoveries from unbiased metabolomics

A major contribution of this thesis has been the development and validation of a GC-MS stable isotope labelling method in *S. pneumoniae* to investigate the activity of central carbon metabolism pathways. The use of metabolomics in the pneumococcal research field has been relatively limited, and I hope that the characterisation and simplification of isotope tracing using GC-MS described here will encourage its adoption in the streptococcal research community to help further our understanding of the metabolic adaptations and associated pathogenic mechanisms of streptococci. The main discoveries made in this thesis as a consequence of employing unbiased metabolomics are summarised below.

6.2.1 TCA cycle activity in *S. pneumoniae* was hypothesised and disproved, but discovery of succinate production raises the possibility of alternative pathways

Extracellular metabolite analysis of *S. pneumoniae*-conditioned culture medium over time showed an accumulation of succinate, fumarate, and α -ketoglutarate (α -KG) corresponding to growth. These may be connected via the TCA cycle, but the pneumococcal genome does not encode the necessary genes to operate a classical TCA cycle (Hoskins et al., 2001; Lanie et al., 2007; Tettelin et al., 2001). However, some other members of the *Streptococcus* spp. operate a fragmentary TCA cycle (Willenborg & Goethe, 2016) and variant branched, non-cyclic, and reductive TCA cycles have been described in various organisms (Guest, 1995; Huynen et al., 1999; Kwong et al., 2017; Pitson et al., 1999; Steffens et al., 2021; Tian et al., 2005), as the normal condition or as an adaptive mechanism. Therefore, we hypothesised that a variation of the TCA cycle may be operational in *S. pneumoniae*, which could be tested by intracellular ^{13}C -label incorporation into these and other intermediates of the TCA cycle.

Classical TCA cycle activity leads to [M+2] labelling of TCA intermediates when cultured in uniformly ^{13}C -labelled hexose sugars (e.g., glucose and galactose) ([U- $^{13}\text{C}_6$]-hexose). This means that 2 carbon atoms from [U- $^{13}\text{C}_2$]-Ac-CoA derived from

glycolysis would be incorporated into TCA cycle intermediates (**Figure 6.2A**). However, I did not find labelling patterns indicative of classical or variant TCA cycle activity in *S. pneumoniae* – Malate and citrate were not detected; oxaloacetate is undetectable by the GC-MS method; α -KG was unlabelled; fumarate exhibited [M+3] labelling; and succinate was unlabelled in the intracellular fraction and minimally [M+3] labelled in the extracellular fraction (**Figure 6.2B**). Pneumococcal production of α -KG from exogenous glutamate has been described, explaining the lack of label incorporation in α -KG (Härtel et al., 2012). The observed [M+3] labelling of fumarate is discussed in the following section (6.6.2).

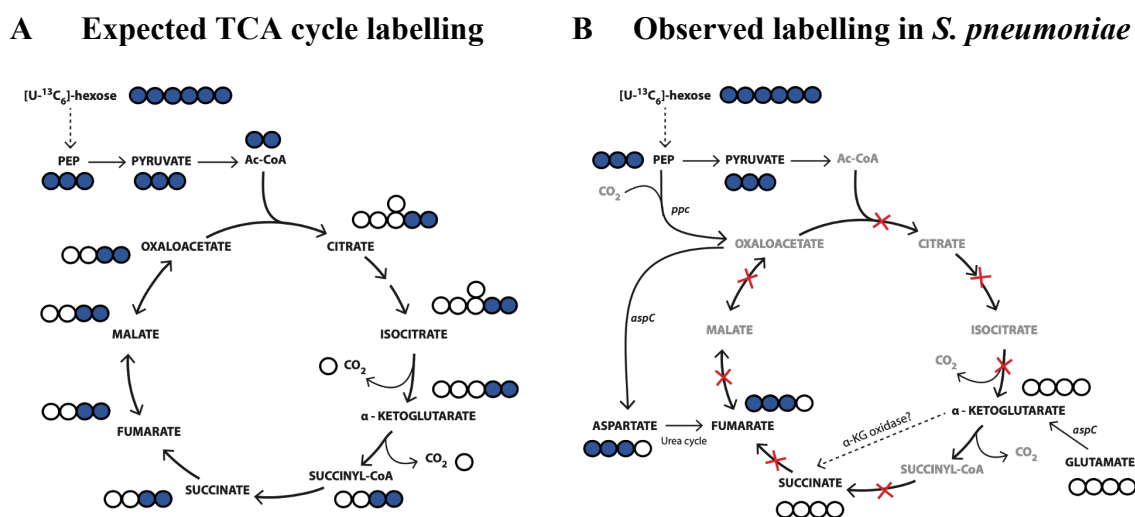


Figure 6.2 Intracellular carbon labelling patterns in *S. pneumoniae* metabolites do not correspond to TCA cycle activity

(A) ^{13}C carbon labelling indicative of one turn of the TCA cycle, adapted from (Metallo et al., 2012). (B) Observed carbon labelling in *S. pneumoniae* in this study. Circles depict individual carbon atoms. Open circle = ^{12}C (unlabelled) carbon atom; filled circle = ^{13}C carbon atom. Intermediates in grey were not detected. Red crosses indicate enzymes unannotated in the *S. pneumoniae* genome (Tettelin et al., 2001).

However, the substrates and reaction(s) involved in succinate production remain uncertain. I hypothesise that unlabelled succinate may be generated from α -KG, as they were both unlabelled in the current study, and because one of the two enzymes catalysing the transformation of α -KG to succinate (α -ketoglutarate dehydrogenase, but not succinyl-CoA synthetase – **Figure 6.2B**) has been putatively annotated in the pneumococcal core genome (https://www.genome.jp/entry/spn:SP_1161; **Chapter 3.3**

Nucleotide sequence similarity analysis of metabolism-associated genes). Additionally, the one-step conversion of α -KG to succinate by α -ketoglutarate oxidase is reported in some anaerobic *Helicobacterium* spp species and potentially also other obligate anaerobes (Allison et al., 1979; Pickett et al., 1994; Pitson et al., 1999).

To test the hypothesis that succinate is generated from glutamate-derived α -KG, ^{13}C -glutamine labelling experiments will be necessary and form part of the future work proposed. I would also perform ^{13}C -hexose and ^{13}C -glutamine labelling experiments during anaerobic pneumococcal culture to determine whether the reaction(s) involved in succinate production are regulated by oxygen, because succinate production is observed in obligate anaerobes such as *Helicobacter pylori*, and in *Escherichia coli* undergoing anaerobic respiration. Studying pneumococcal metabolism under varying environmental oxygen concentrations also has physiological relevance, as variable oxygen concentrations are encountered by the pneumococcus in different host sites during colonisation and disease (Nagaoka et al., 2018; Weiser et al., 2001).

If the hypothesis is proven, functional enzyme assays can be performed to assess activities of succinyl-CoA synthetase and α -ketoglutarate oxidase by ^1H -NMR spectroscopy, because even though these enzymes are not annotated in the pneumococcal genome, experimental elucidation is critical to determine whether they are in fact absent in *S. pneumoniae*, or whether alternative enzymes are present.

The current data clearly show that succinate is produced in *S. pneumoniae* and perhaps crucially, secreted. Without knowing the enzymatic reactions involved in generating succinate, which require further work as described, it is difficult to determine the relevance of succinate production. However, I believe that attempting to delineate this is a worthwhile endeavour, particularly considering the emerging evidence for succinate involvement in broad-ranging processes beyond metabolism, such as ROS detoxification and tumour progression. More relevant to bacteria, succinate accumulating in the gut (via activity of succinate-producing gut microbes) can act as a cross-feeding metabolite, and additional roles for succinate as a pro-inflammatory signalling molecule have also been described (Guo et al., 2020; Murphy & O'Neill,

2018; Rosenberg et al., 2021; Tretter et al., 2016). We speculate that succinate in *S. pneumoniae* may act as a signalling molecule with potential involvement in quorum sensing or induction of host inflammation.

6.2.2 Evidence for a pneumococcal urea cycle

As mentioned in the previous section (6.2.1), the [M+3] labelled fumarate observed in stable isotope labelling GC-MS experiments cannot be generated through action of a classical TCA cycle. However, the labelling pattern of fumarate corresponded closely to that of aspartate across *S. pneumoniae* strains. Aspartate and fumarate may be connected via the urea cycle, but the urea cycle is incomplete in *S. pneumoniae* genome annotations (Chapter 5 **Figure 5.12**).

Leonard and colleagues (Leonard et al., 2018) first hypothesised the potential for a urea cycle in *S. pneumoniae* upon observing a decrease in extracellular arginine levels corresponding to an increase in extracellular ornithine and fumarate. Employing carbon isotope tracing, I not only saw the production of ornithine and urea, but also determined their lack of ¹³C-labelling to be indicative of a urea cycle (**Figure 6.3**). Although other intermediates of the urea cycle – argininosuccinate, arginine and citrulline – are not detectable using this GC-MS method, the abundance and labelling patterns of aspartate, fumarate, ornithine, and urea provide further evidence for operation of a urea cycle in *S. pneumoniae*.

However, more work is required to definitively prove this. *S. pneumoniae* is auxotrophic for arginine and arginine uptake is essential for pneumococcal growth (Härtel et al., 2012); therefore, isotopically-labelled arginine could be used to determine label incorporation into urea cycle metabolites. This will be especially important in unpicking urea cycle activity from that of the *S. pneumoniae* arginine deiminase system (ADS), which generates ornithine from arginine and in the process produces ATP, ammonia and carbon dioxide (Abdelal, 1979; Gupta et al., 2013; Schulz et al., 2014).

The urea cycle may be an important method to assimilate nitrogen from ammonia, and/or excrete nitrogenous waste generated during amino acid catabolism.

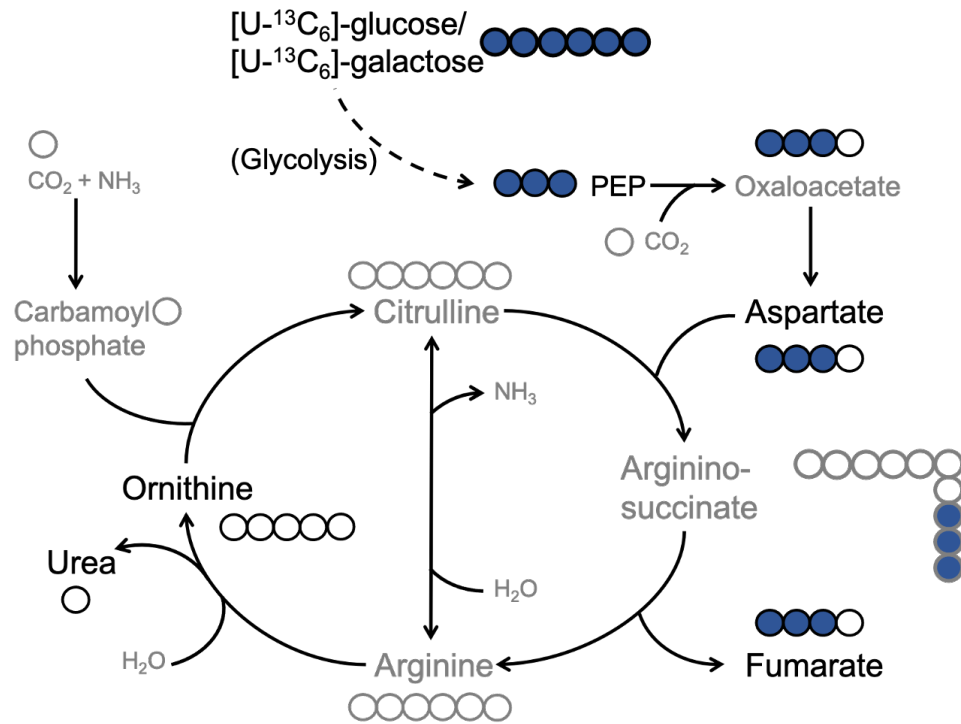


Figure 6.3 Stable isotope labelling of urea cycle intermediates in *S. pneumoniae*

Circles depict individual carbon atoms. Open circle = ^{12}C (unlabelled) carbon atom; filled circle = ^{13}C carbon atom. Black font indicates intermediates detected by the GC-MS method. Intermediates in grey were not detected by the GC-MS method, but their expected labelling during urea cycle operation is shown.

6.2.3 Potential biosynthesis of proline and glycine in *S. pneumoniae*

S. pneumoniae was previously shown to be auxotrophic for the essential amino acid glycine (Härtel et al., 2012), and I observed the extracellular depletion of glycine (indicating glycine uptake) in all *S. pneumoniae* strains tested. Härtel and colleagues (Härtel et al., 2012) have also described the conversion of glycine to serine via a unique serine hydroxymethyltransferase (SHMT) enzyme *glyA*, which uses 5,10-methylenetetrahydrofolate (5,10-MTHF) (generated from formate so $[M+1]$ labelled) and glycine (exogenous unlabelled) to generate $[M+1]$ serine. In agreement, *glyA*

activity could be inferred from the isotope labelling experiments in this thesis, indicated by the production of [M+1] serine.

Unexpectedly, I also observed the production of uniformly labelled [U- $^{13}\text{C}_2$]-glycine, suggesting pneumococcal glycine biosynthesis even in the presence of utilisable exogenous glycine. Glycine is commonly generated in bacteria from the glycolytic intermediate 3-phosphoglycerate during serine catabolism, which would lead to [U- $^{13}\text{C}_2$]-glycine when cultured on [U- $^{13}\text{C}_6$]-sugar (Eisenreich et al., 2006; Pizer, 1965; Vidal et al., 2005). This common pathway is considered absent in pneumococci, on the basis of stable isotope labelling experiments performed by Härtel and colleagues (2012) using *S. pneumoniae* D39 Δ *cps* showing no ^{13}C -label incorporation into glycine. Their use of D39 Δ *cps* provides a critical hint, because in my experiments [U- $^{13}\text{C}_2$]-glycine was only generated in the wild-type encapsulated TIGR4 (and to a lesser extent, 23F), but not in the Δ *cps* mutants.

Therefore, this data suggests a complex glycine metabolism system involving import for serine biosynthesis, as well as a route for intracellular glycine production derived from central carbon metabolism which is abrogated in the absence of the capsule. The involvement of capsule could hint at adaptive metabolism and virulence. Pneumococcal glycine metabolism requires significant characterisation and contextualisation and should be investigated further using different strains of *S. pneumoniae* along with their capsule-deficient derivatives.

Finally, my data also suggests *de novo* proline biosynthesis in *S. pneumoniae*, even though the proline biosynthetic pathway (from glutamate) is genomically incomplete (Hoskins et al., 2001; Tettelin et al., 2001). I observed very high levels of extracellular proline accumulation during *S. pneumoniae* growth, and found no label incorporation in intracellular- or extracellular-detected proline in the presence of $^{13}\text{C}_6$ -glucose. This points to proline synthesis in *S. pneumoniae* from glutamate – experiments with isotopically-labelled glutamine/ glutamate will be necessary – and highlights the power of unbiased metabolomics to uncover new aspects of microbial metabolism.

6.3 Summary, limitations, and outlook

The results described in this thesis have demonstrated that stable isotope profiling of global pneumococcal metabolic pathways can uncover novel aspects of metabolism and generate hypotheses for further investigation. In this study, bacterial cells were cultured in a ^{13}C -enriched carbon source and the incorporation of ^{13}C isotope into various metabolites inside and outside pneumococcal cells was determined at isotopic steady state using GC-MS. Intracellular metabolite concentrations are constant at metabolic steady state, allowing us to trace carbohydrate-derived ^{13}C carbon atoms through metabolic networks to understand energetic and anabolic metabolism as well as highlight missing or variant reactions. However, this does not provide information about metabolic flux and enzyme kinetics in a dynamic system (Sauer, 2006).

Broad metabolite coverage requires the use of multiple metabolomics technologies as they have variable sensitivities and specificities for different types of compounds, such as GC-MS, LC-MS, and NMR. In this study, GC-MS was the primary method employed, with ^1H -NMR used to detect a targeted set of 5 metabolites, and LC-MS used only for methodological validation. Therefore, the metabolite coverage in this study was limited to those detectable by GC-MS; the amino acid and urea cycle intermediate arginine, for example, is not detectable by the method used. Additionally, achieving a truly global understanding of bacterial metabolism and the various ways in which pathways are connected and regulated requires the integration of techniques beyond isotope labelling metabolomics, including genomic analyses, transcriptomics, enzyme kinetic assays, metabolic flux analyses, and functional characterisation (Feist et al., 2008).

I have shown that the capsules of two different pneumococcal serotypes, 4 and 23F, have divergent effects on carbon assimilation into anabolic pathways. The strain and serotype repertoire must be expanded to validate the conclusions derived from these studies and determine whether differences observed are dependent on capsule features (polysaccharide composition, energetic burden, thickness), the strain genetic background, or a combination of both.

Carbohydrate and central carbon metabolism were studied in this thesis using glucose and galactose as niche-specific carbon sources, but a mixture of sugars (such as galactose, *N*-acetylglucosamine, mannose) could be used to better model the nasopharyngeal mucus environment (Paixão et al., 2015a). There is also broad scope for the investigation of other nutritional sources by isotope labelling and related techniques. Studies with ^{13}C - or ^{15}N -labelled glutamine will help understand nitrogen metabolism in *S. pneumoniae*, and specifically answer some of the questions generated during the course of this work and discussed earlier – can proline be generated from glutamate? Is succinate a product of novel succinyl-CoA synthetase or α -KG oxidase activity derived from glutamate?

Metabolic adaptability to other variable components of host microenvironments, such as low oxygen tension in deeper host sites and changing temperatures, can also be probed by stable isotope labelling and metabolic flux analyses. This will complement transcriptomic studies showing differential gene expression during pneumococcal colonisation and the invasive disease at different sites (Aprianto et al., 2018; D’Mello et al., 2020), eventually improving our understanding of a key determinant of pneumococcal pathogenesis – metabolic adaptation.

In conclusion, I have demonstrated that the application of cutting-edge technologies to probe bacterial metabolic networks can reveal novel aspects of metabolic adaptation and nutrient assimilation. I have also shown that the pneumococcal polysaccharide capsule can influence central and associated carbon metabolism in a serotype-dependent manner, and this may be connected to factors determining the carriage and invasive potential of the highly diverse population of *S. pneumoniae* strains. These findings open a wide-ranging set of possibilities to fully delineate the relationship between virulence and carbohydrate metabolism, and possibly also understand the basis of non-vaccine serotype replacement in vaccinated individuals as well as identify targets for broad-spectrum vaccines and more effective drugs.

Chapter 7. Appendix

Chapters 2 and 3

Table 7.1 List of *S. pneumoniae* isolates included in nucleotide sequence similarity analysis of metabolic genes

Table 7.2 List of metabolic genes included in nucleotide sequence similarity analysis across 120 *S. pneumoniae* isolates

Chapter 4

Table 7.3 Extracellular metabolites of *S. pneumoniae* TIGR4 cultured in THY

Figure 7.1 Extracellular metabolites of *S. pneumoniae* TIGR4 cultured in 7H9+(glucose) and 7H9+(galactose)

Table 7.4 List of nucleotide variants in two TIGR4 strains compared to reference genome

Table 7.1 List of *S. pneumoniae* isolates included in nucleotide sequence similarity analysis of metabolic genes

The isolates were obtained from pubmlst.org.

¹ VT = Vaccine type; NVT = Non vaccine type. VT isolates include serotypes found in the pneumococcal conjugate vaccines (PCV) 10 and 13, and the pneumococcal polysaccharide vaccine (PPSV) 23.

² One contig ideal, as complete sequence available.

³ L50 = **position** of the shortest contig in a set that together represents 50% of the total assembly length, when all contigs in the assembly are aligned in descending order of length. Smaller the N50 contig number, the more 'complete' the assembly.

⁴ N50 = **length** of the shortest contig that encompasses 50% of the total length, when all contigs in the assembly are aligned in descending order of length. Larger this value, the more 'complete' the assembly.

Serotype	Isolate	VT/ NVT ¹	MLST	Contig(s) ²	L50 ³	N50 ⁴	Length (bp)
1	NCTC7465	VT	615	26	2	410559	2084938
	P1031	VT	303	1	1	2111882	2111882
	SPnINV104B	VT	227	1	1	2142122	2142122
2	D39	VT	595	1	1	2046115	2046115
	2/2	VT	574	32	5	174677	2049712
	2/3	VT	3744	28	3	268043	2011336
3	SPnOXC141	VT	180	1	1	2034829	2034829
	Sp03-4183	VT	180	1	1	2037254	2037254
	A45	VT	6934	4	1	1212544	2124944
4	TIGR4	VT	205	1	1	2160842	2160842
	GA07643	VT	695	6	1	1168444	2127739
5	5/4	VT	7222	38	5	198717	2183089
	5/3	VT	289	35	5	162021	2183256
	70585	VT	289	1	1	2184682	2184682
8	8/2	VT	53	31	5	97352	2032551
	8/3	VT	53	23	4	196630	2031147
13	13/2	NVT	7173	27	5	186424	2057476
	s 1°_Hakenbeck	NVT	5647	29	3	269960	2057185
14	JJA	VT	66	1	1	2120234	2120234
	England14-9	VT	9	29	5	144499	2067848
	IS7_Hakenbeck	VT	230	26	3	286744	2082067
20	20/4	VT	5977	29	6	106943	2017121
	20/3	VT	235	34	5	130045	2093105
21	21/2		7207	38	5	142842	2091517

	21/3	NVT	193	35	5	173390	2110480
34	34/2	NVT	7218	25	4	199413	2040428
36	36/2	NVT	7237	33	5	131201	2031262
	36/5	NVT	4031	27	3	280189	2014778
39	s 7°_Hakenbeck	NVT	13149	34	5	129995	2051504
	s 10°_Hakenbeck	NVT	13149	30	4	174408	2049736
40	40/2	NVT	7185	25	4	258088	2030457
43	43/4	NVT	7193	30	4	162112	2016820
	43/3	NVT	7197	18	4	226163	2084514
44	44/3	NVT	7176	34	5	129652	2127229
45	45/3	NVT	7239	25	2	518591	2142743
	45/4	NVT	7191	24	4	173092	2019503
48	48/3	NVT	7232	26	3	301638	2043332
	48/5	NVT	7232	22	3	232548	2079434
10A	10A/1	VT	97	27	3	196680	2072630
10C	10C/2	NVT	2862	39	7	92011	2098901
10F	10F/2	NVT	7186	20	3	378766	2048181
11A	AP200	VT	62	1	1	2130580	2130580
	SP11-BS70	VT	62	25	4	206311	2059769
	VICE1758	VT	62	28	5	132675	2032993
11D	VICE1791a	NVT	10345	30	6	94171	2024602
	11D/1	NVT	62	30	5	179941	2011523
12F	CDC0288-04	VT	220	38	5	208663	2049808
	12F/6	VT	218	28	5	186921	2023118
	s 8°_Hakenbeck	VT	13148	33	5	158885	2144919
15A	VICE1642	NVT	63	27	5	148128	2057189
15B/C	VICE1726	VT	199	26	4	164064	2049971
	VICE1774	VT	1262	28	6	130402	2083538
	VICE1779a	VT	1262	25	5	164084	2076296
16A	16A/1	NVT	6543	20	3	321529	2088451
16F	16F/2	NVT	7167	32	5	136301	2134100
17F	17F/3	VT	392	24	4	171292	2125590
	17F/4	VT	392	28	5	147002	2124813
18A	18A/2	NVT	241	26	3	303452	2039841
18B	18B/2	NVT	4706	31	5	177362	2141199
18C	18C/3	VT	7195	32	5	155644	2101943
	Netherlands18C-36	VT	113	28	5	163353	2093092

	GA13637	VT	3060	21	3	242831	2128292
18F	18F/1	NVT	7181	20	3	330156	2047749
	18F/2	NVT	4863	27	3	214545	2016793
19A	Hungary19A-6	VT	268	1	1	2245615	2245615
	TCH8431	VT	320	1	1	2088772	2088772
	GA54644	VT	156	7	1	1313102	2128321
19F	Taiwan19F-14	VT	236	338	12	57345	2057950
	G54	VT		1	1	2078953	2078953
	ST556	VT	1392	1	1	2145902	2145902
22A	22A/3	NVT	7181	29	5	143010	2055285
	22A/2	NVT	7182	27	4	239438	2057334
22F	VICE1717	VT	433	22	6	125840	2070440
	VICE1743	VT	433	23	6	127502	2070742
	A_Hakenbeck	VT	13152	24	3	306136	2066762
23A	VICE1820	NVT	190	17	4	214964	2050169
	2301_Hakenbeck	NVT	524	19	2	587312	2047400
23B	VICE1712b	NVT	439	20	4	222553	2063496
	VICE1750	NVT	439	19	4	190143	2048795
23F	Spain23F-1	VT	81	1	1	2221315	2221315
	23F/5	VT	439	18	2	491607	2065101
	2306_Hakenbeck	VT	13141	18	3	258436	2043993
24F	VICE1987	NVT	13131	38	4	132401	2090805
25A	25A/2	NVT	5407	36	4	211292	2154220
32F	32F/2	NVT	7164	31	3	301358	2093989
33A	33A/3	NVT	1012	30	5	182669	2036546
	33A/2	NVT	1012	26	4	179342	2003668
33B	33B/2	NVT	2864	30	5	94524	2115280
33F	33F/2	VT	7192	28	3	290639	2045606
	GA41317	VT	2705	6	2	418734	2090801
	VICE1761b	VT	100	9	3	320168	2043286
35B	35B/2	NVT	7234	24	2	412458	2029839
	35B/4	NVT	198	34	6	130649	2059216
35F	VICE1667	NVT	1635	22	4	174445	2106387
	VICE1840	NVT	1635	21	4	139811	2107766
41A	41A/2	NVT	7238	24	4	199053	2065894
41F	41F/2	NVT	7213	20	3	387224	2107691
47A	47A/1	NVT	7220	34	4	155961	2110699

6A	SMRU1447	VT	4936	179	22	31331	2065313
	GA17971	VT	1175	10	2	441236	2122606
	VICE1690	VT	2756	21	5	173830	2033735
6B	VICE1047	VT	138	25	3	274273	2204584
	VICE1153	VT	176	25	4	204093	2161573
	U26_Hakenbeck	VT	473	36	3	320994	2155576
6C	GA47033	NVT	4150	10	2	540003	2106476
	VICE1632	NVT	1379	19	3	262979	2048598
6E	670	NVT	90	1	1	2240045	2240045
	NP112	NVT	1536	9	2	582976	2072320
7B	7B/3	NVT	230	30	3	197103	2123191
7F	Netherlands7F-39	VT	191	36	8	76326	2005102
	GA47439	VT	1176	11	3	337258	2019801
	GA47283	VT	191	16	4	170377	2061946
9L	9L/3	NVT	7204	28	5	152453	2059114
	9L/4	NVT	7240	29	4	173479	2088807
9N	9N/2	VT	7205	16	2	399019	2043388
	9N/3	VT	3983	17	2	422980	2063167
	VICE1770	VT	66	30	6	119730	2121179
9V	9V/4	VT	123	40	5	195336	2154347
	665_Hakenbeck	VT	156	32	4	126976	2089978
	676_Hakenbeck	VT	156	33	4	133819	2089698

Table 7.2 List of metabolic genes included in nucleotide sequence similarity analysis across 120 *S. pneumoniae* isolates

Gene list obtained from selected metabolic pathways of *S. pneumoniae* TIGR4 on the KEGG database.

Gene ID	Gene name	Annotated/ putative function
<i>Glycolysis/ gluconeogenesis</i>		
SP_0066	galM	aldose 1-epimerase
SP_0240		phosphoglycerate mutase
SP_0265		glycosyl hydrolase
SP_0285		alcohol dehydrogenase
SP_0303	bglA-1	6-phospho- β -glucosidase
SP_0499	pgk	phosphoglycerate kinase
SP_0578	bglA-2	6-phospho- β -glucosidase
SP_0605	fba	fructose-bisphosphate aldolase, class II
SP_0668	gki	glucokinase
SP_0758		PTS system, IIBC components
SP_0896	pfk	6-phosphofructokinase
SP_0897	pyk	pyruvate kinase
SP_0984		putative phosphoglycerate mutase family
SP_1119	gapN	glyceraldehyde-3-phosphate dehydrogenase
SP_1128	eno	enolase
SP_1161		putative acetoin dehydrogenase
SP_1162		putative acetoin dehydrogenase
SP_1163		putative acetoin dehydrogenase
SP_1164		putative acetoin dehydrogenase
SP_1220	ldh	L-lactate dehydrogenase
SP_1498	pgm	phosphoglucomutase
SP_1574	tpi	triosephosphate isomerase
SP_1655	gpmA	phosphoglycerate mutase
SP_1684		PTS system, IIBC components
SP_1855		alcohol dehydrogenase
SP_2012	gap	glyceraldehyde 3-phosphate dehydrogenase
SP_2026		alcohol dehydrogenase
SP_2070	pgi	glucose-6-phosphate isomerase

<i>Pentose phosphate pathway</i>		
SP_0027	prsA-1	ribose-phosphate pyrophosphokinase
SP_0317		4-hydroxy-2-oxoglutarate aldolase
SP_0318		carbohydrate kinase, PfkB family
SP_0375	gnd	6-phosphogluconate dehydrogenase
SP_0605	fba	fructose-bisphosphate aldolase, class II
SP_0828	rpiA	ribose 5-phosphate isomerase
SP_0829	deoB	phosphopentomutase
SP_0843	deoC	deoxyribose-phosphate aldolase
SP_0896	pfk	6-phosphofructokinase
SP_1095	prsA-2	ribose-phosphate pyrophosphokinase
SP_1119	gapN	glyceraldehyde-3-phosphate dehydrogenase
SP_1243	zwf	glucose-6-phosphate 1-dehydrogenase
SP_1498	pgm	phosphoglucomutase
SP_1506		conserved hypothetical protein
SP_1615	recP-1	transketolase
SP_1983	rpe	ribulose-phosphate 3-epimerase
SP_2030	recP-2	transketolase
SP_2070	pgi	glucose-6-phosphate isomerase
SP_2127		transketolase
SP_2128		transketolase
<i>Galactose metabolism</i>		
SP_0060		glycosyl hydrolase
SP_0066	galM	aldose 1-epimerase
SP_0321		PTS system, IIA component
SP_0323		PTS system, IIB component
SP_0324		PTS system, IIC component
SP_0325		PTS system, IID component
SP_0358	cap4J	UDP-glucose 4-epimerase
SP_0476	lacF-1	PTS system, lactose-specific IIA component
SP_0477	lacG-1	6-phospho-beta-galactosidase
SP_0478	lacE-1	PTS system, lactose-specific IIBC components
SP_0648	bgaA	β -galactosidase
SP_0668	gki	glucokinase
SP_0896	pfk	6-phosphofructokinase
SP_1184	lacG-2	6-phospho- β -galactosidase

SP_1185	lacE-2	PTS system, lactose-specific IIBC components
SP_1186	lacF-2	PTS system, lactose-specific IIA component
SP_1190	lacD	tagatose 1,6-diphosphate aldolase
SP_1191	lacC	tagatose-6-phosphate kinase
SP_1192	lacB	galactose-6-phosphate isomerase, LacB subunit
SP_1193	lacA	galactose-6-phosphate isomerase, LacA subunit
SP_1498	pgm	phosphoglucomutase
SP_1607	galE-1	UDP-glucose 4-epimerase
SP_1724	scrB	sucrose-6-phosphate hydrolase
SP_1795		putative sucrose-6-phosphate hydrolase
SP_1828	galE-2	UDP-glucose 4-epimerase
SP_1829	galT-1	galactose-1-phosphate uridylyltransferase
SP_1852	galT-2	galactose-1-phosphate uridylyltransferase
SP_1853	galK	galactokinase
SP_1867		NAD-dependent epimerase
SP_1898	aga	α -galactosidase
SP_2092	galU	UTP-glucose-1-phosphate uridylyltransferase
<i>Pyruvate metabolism</i>		
SP_0251		putative formate acetyltransferase
SP_0423	accB	acetyl-CoA carboxylase, biotin carboxyl carrier protein
SP_0425	accC	acetyl-CoA carboxylase, biotin carboxylase
SP_0426	accD	acetyl-CoA carboxylase, carboxyl transferase, β -subunit
SP_0427	accA	acetyl-CoA carboxylase, carboxyl transferase, α -subunit
SP_0459	pfl	pyruvate formate lyase
SP_0730	spxB	pyruvate oxidase
SP_0897	pyk	pyruvate kinase
SP_0962	gloA	lactoylglutathione lyase
SP_1068	ppc	phosphoenolpyruvate carboxylase
SP_1100	pta	phosphate acetyltransferase
SP_1161		putative acetoin dehydrogenase complex, E3 component, dihydrolipoamide dehydrogenase
SP_1162		putative acetoin dehydrogenase complex, E2 component, dihydrolipoamide acetyltransferase
SP_1163		putative acetoin dehydrogenase, E1 component, β -subunit
SP_1164		putative acetoin dehydrogenase, E1 component, α -subunit
SP_1220	ldh	L-lactate dehydrogenase

SP_1258		putative 2-isopropylmalate synthase
SP_1974		acylphosphatase
SP_2026		alcohol dehydrogenase
SP_2044	ackA	acetate kinase
<i>Additional genes involved in carbon metabolism</i>		
SP_0105	sdhA	L-serine dehydratase, iron-sulfur-dependent, α -subunit
SP_0106	sdhB	L-serine dehydratase, iron-sulfur-dependent, β -subunit
SP_0450	ilvA	threonine dehydratase
SP_0586		putative 5,10-methylenetetrahydrofolate reductase
SP_0589	cysE	serine acetyltransferase
SP_0825	folD	methylenetetrahydrofolate dehydrogenase
SP_1024	glyA	serine hydroxymethyltransferase
SP_1229	fhs	formate-tetrahydrofolate ligase
SP_2151	arcC	carbamate kinase
SP_2210	cysM	cysteine synthase

Table 7.3 Extracellular metabolites of *S. pneumoniae* TIGR4 cultured in THY

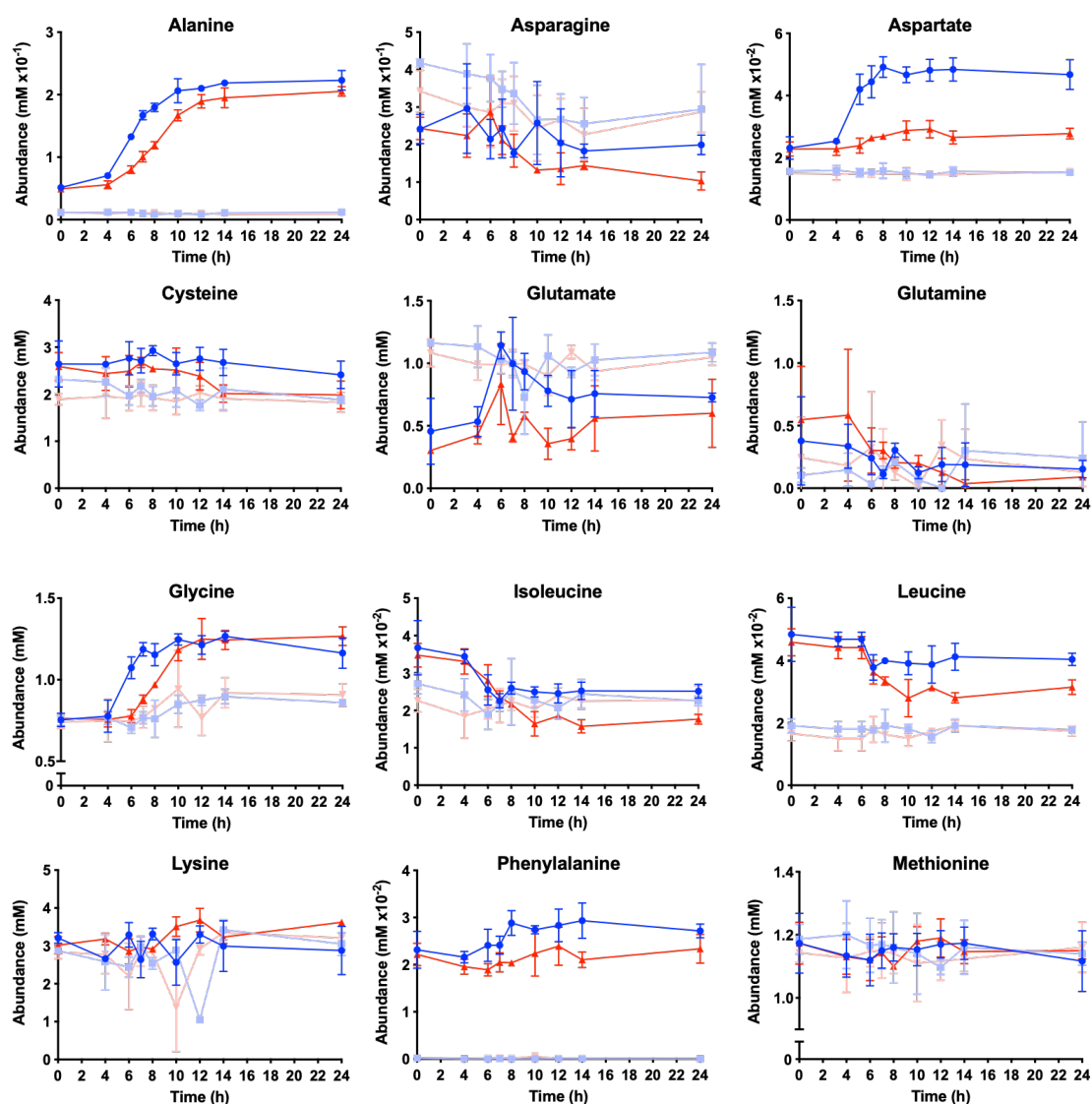
Extracellular concentrations (mM) of metabolites over time relative to the internal standard are reported. The experiment was performed once, with no technical replicates.

Time (h)	0	1	2	3	4	5	6	7	8	10	14
Glucose	4.27	4.29	4.27	4.28	3.86	2.44	0.21	0.11	0.07	0.03	0.02
Fructose	2.66	2.80	2.67	2.60	1.86	0.17	0.16	0.16	0.16	0.17	0.16
Pyruvate	2.26	2.07	1.76	1.37	1.39	0.44	0.30	0.23	0.17	0.17	0.15
Lactate	0.74	0.71	0.72	0.82	0.99	1.50	1.85	1.84	1.85	1.89	1.83
Ala	0.59	0.75	0.71	0.76	0.77	0.71	0.65	0.74	1.00	0.90	0.71
Asn	1.68	1.79	1.67	1.71	1.54	1.28	1.11	1.14	1.04	1.26	1.40
Asp	0.76	0.33	0.06	0.80	0.76	0.94	1.06	1.13	1.09	1.25	1.23
Glu	47.1	46.4	43.7	48.4	47.5	47.3	48.3	49.6	49.5	52.2	53.2
Gly	5.46	5.04	4.69	5.65	5.25	5.29	5.54	5.69	5.70	5.98	6.04
Ile	0.84	0.84	0.84	0.85	0.87	1.04	1.21	1.23	1.28	1.49	1.67
Leu	0.88	0.86	0.75	0.93	1.09	1.43	1.59	1.50	1.53	1.64	1.84
Lys	7.52	5.92	5.10	7.44	7.56	7.44	7.56	7.72	7.64	7.97	7.93
Met	1.39	0.78	0.38	1.39	1.37	1.47	1.52	1.55	1.32	1.45	1.59
Pro	0.08	0.09	0.08	0.10	0.11	0.15	0.23	0.25	0.32	0.36	0.35
Ser	0.89	0.38	0.27	0.80	0.82	1.53	1.68	1.67	1.81	2.02	2.18
Thr	1.07	0.31	0.19	0.99	0.97	1.35	1.45	1.50	1.57	1.70	1.86
Trp	0.45	0.50	0.77	0.54	0.52	0.56	0.56	0.49	0.48	0.49	0.61
Tyr	3.55	3.63	3.37	3.64	3.61	3.68	3.86	4.01	3.97	4.18	4.25
Val	0.32	0.34	0.30	0.38	0.35	0.35	0.39	0.42	0.60	0.52	0.45
5-oxopro	14.5	15.3	15.0	15.0	15.0	14.5	15.0	16.1	16.1	16.9	16.4
Citrate	0.11	0.12	0.12	0.12	0.11	0.11	0.11	0.11	0.11	0.11	0.11
Fumarate	0.10	0.09	0.08	0.10	0.12	0.12	0.15	0.15	0.13	0.13	0.15
Malate	3.00	3.02	3.03	3.12	3.06	2.95	2.94	2.94	2.89	2.98	3.00
Ornithine	0.28	0.21	0.18	0.32	0.30	0.29	0.29	0.30	0.34	0.52	0.77
Succinate	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.11	0.11	0.11
Thymine	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Uracil	0.13	0.11	0.11	0.14	0.13	0.10	0.08	0.09	0.09	0.10	0.11

Figure 7.1 Extracellular metabolites of *S. pneumoniae* TIGR4 cultured in 7H9⁺(glucose) and 7H9⁺(galactose)

Extracellular concentrations (mM) of metabolites over time relative to the internal standard are plotted. Data points show the mean of three replicates and error bars represent the standard deviation.

Blue dots = TIGR4 in 7H9⁺(glucose); pale blue squares = 7H9⁺(glucose) only;
Red triangles = TIGR4 in 7H9⁺(galactose); pale red inverted triangles = 7H9⁺(galactose) only.



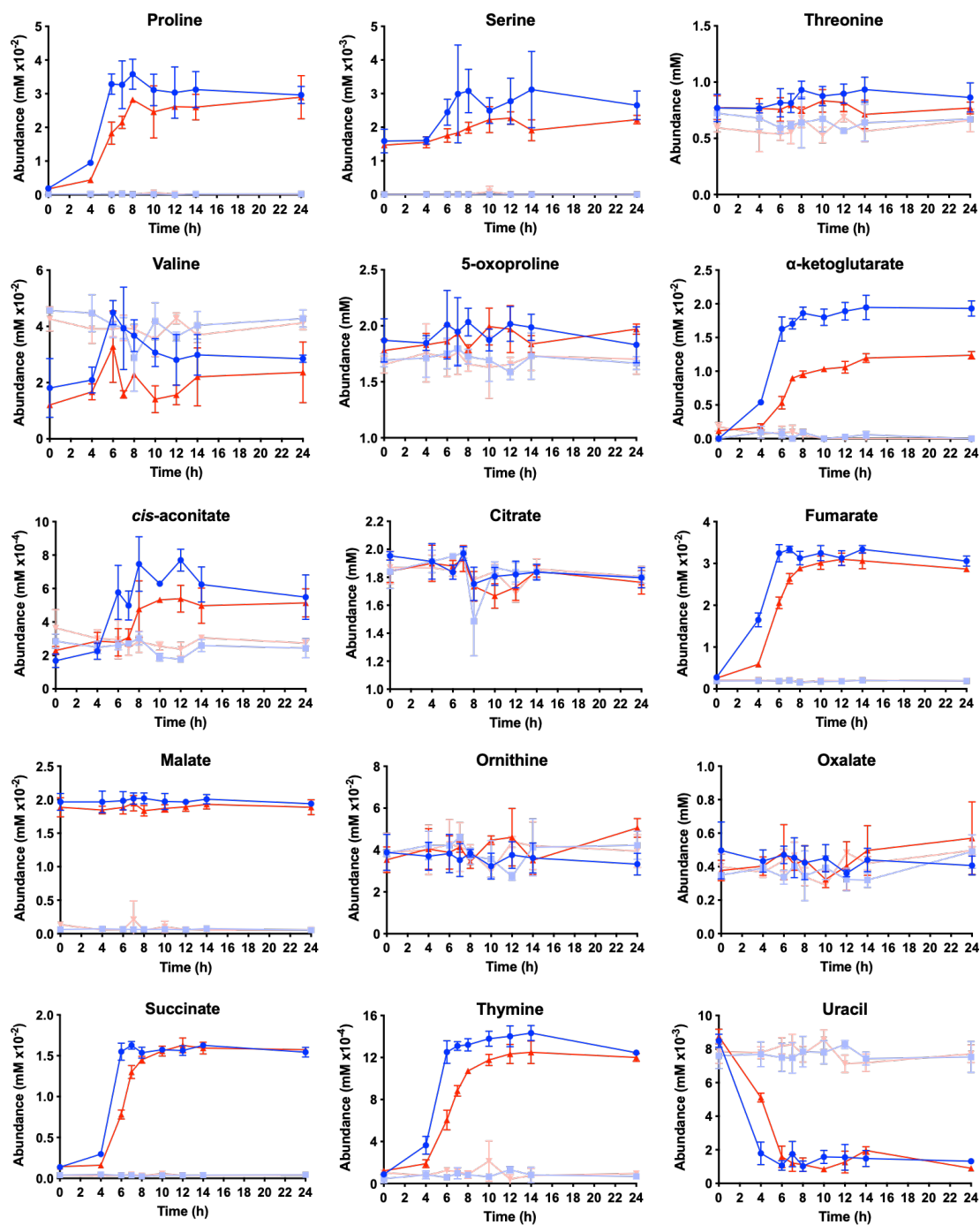


Table 7.4 List of nucleotide variants in two TIGR4 strains compared to reference genome

Whole genomes of TIGR4-lab (strain P1542) and ATCC-TIGR4 (type strain BAA-334) were sequenced and aligned to the reference TIGR4 genome AE05672.3. The variant list is reported here.

Position	Reference	Alternative	TIGR-lab (P1542)	ATCC-TIGR4 (BAA-334)
146055	G	C	1	1
188292	TC	T	1	1
192435	G	GC	1	1
247805	A	G	1	1
273377	CG	C	1	1
463630	G	A	1	1
463631	A	G	1	1
469287	GC	G	1	1
476406	G	T	1	1
597325	A	AG	1	1
656387	G	GA	1	1
730678	C	T	1	1
737108	A	G	1	1
834922	A	AC	1	1
972509	CT	C	1	1
985055	GT	G	1	1
991106	G	GC	1	1
1101585	C	G	1	1
1101586	G	C	1	1
1113285	G	GC	1	1
1127020	C	G	1	1
1127021	G	C	1	1
1132391	A	AG	1	1
1132435	G	A	1	1
1132436	A	G	1	1
1267219	C	G	1	1
1267220	G	C	1	1
1282010	A	AG	1	1
1618223	A	G	1	1

1618514	A	AG	1	1
1622021	G	A	1	1
1622022	A	G	1	1
1622066	G	A	1	1
1622067	A	G	1	1
1695199	G	GT	1	1
1696085	A	AC	1	1
1759030	GA	G	1	1
1985217	A	AG	1	1
2016330	A	AG	1	1
1132343	A	AG	1	1
204679	C	T	0	1
417383	T	C	1	0
426530	C	A	1	0
762843	C	A	1	0
807984	G	A	0	1
846780	CAAAAAAAAA	CAAAAAAAAA	1	0
879981	A	C	1	0
947167	G	C	0	1
967257	G	T	0	1
1001704	T	G	0	1
1096817	T	G	0	1
1257186	G	T	0	1
1320529	C	A	0	1
1350621	G	T	0	1
1543556	C	A	0	1
1593439	CT	C	1	0
1634975	A	T	1	0
1634977	G	C	1	0
1634978	T	C	1	0
1634981	A	G	1	0
1634986	C	A	1	0
1634987	A	G	1	0
1634993	A	G	1	0
1634999	T	A	1	0
1635159	G	A	1	0

1635179	A	G	1	0
1796273	G	A	1	0
1828206	G	T	0	1
1857996	GAAAAAAAAA	GAAAAAAAAA	1	0
1862985	C	T	0	1
1901723	G	T	1	0
2042324	G	A	0	1
2050193	CGTTTTT	C	0	1
2092182	T	C	1	0
2127997	G	T	0	1
257296	C	A	1	0
324506	G	A	0	1
324848	G	GA	1	0
395317	G	T	1	0
395471	G	A	1	0
422834	C	A	0	1
488215	C	T	0	1
745861	G	T	0	1
1256183	C	CTT	0	1
1635002	G	T	1	0
1674302	GC	G	0	1
1787558	G	T	0	1
2091412	G	A	0	1

Reference List

- Abbott, D. W., Higgins, M. A., Hyrnuik, S., Pluvinae, B., Lammerts van Bueren, A., & Boraston, A. B. (2010). The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae*. *Molecular Microbiology*, 77(1), 183–199. <https://doi.org/10.1111/j.1365-2958.2010.07199.x>
- Abdelal, A. T. (1979). Arginine catabolism by microorganisms. *Annual Review of Microbiology*, 33, 139–168. <https://doi.org/10.1146/annurev.mi.33.100179.001035>
- Agarwal, V., Hammerschmidt, S., Malm, S., Bergmann, S., Riesbeck, K., & Blom, A. M. (2012). Enolase of *Streptococcus pneumoniae* binds human complement inhibitor C4b-binding protein and contributes to complement evasion. *Journal of Immunology*, 189, 3575–3584. <https://doi.org/10.4049/jimmunol.1102934>
- Al-Bayati, F. A. Y., Kahya, H. F. H., Damianou, A., Shafeeq, S., Kuipers, O. P., Andrew, P. W., & Yesilkaya, H. (2017). Pneumococcal galactose catabolism is controlled by multiple regulators acting on pyruvate formate lyase. *Scientific Reports*, 7. <https://doi.org/10.1038/srep43587>
- Ali, F., Lee, M. E., Iannelli, F., Pozzi, G., Mitchell, T. J., Read, R. C., & Dockrell, D. H. (2003). *Streptococcus pneumoniae*–associated human macrophage apoptosis after bacterial internalization via complement and Fcγ receptors correlates with intracellular bacterial load. *The Journal of Infectious Diseases*, 188(8), 1119–1131. <https://doi.org/10.1086/378675>
- Allan, R. N., Skipp, P., Jefferies, J., Clarke, S. C., Faust, S. N., Hall-Stoodley, L., & Webb, J. (2014). Pronounced metabolic changes in adaptation to biofilm growth by *Streptococcus pneumoniae*. *PLoS ONE*, 9(9). <https://doi.org/10.1371/journal.pone.0107015>
- Allison, M. J., Robinson, I. M., & Baetz, A. L. (1979). Synthesis of α-ketoglutarate by reductive carboxylation of succinate in *Veillonella*, *Selenomonas*, and *Bacteroides* species. *Journal of Bacteriology*, 140(3), 980–986. <https://doi.org/10.1128/jb.140.3.980-986.1979>

- Alvarez-Dominguez, C., Salcines-Cuevas, D., Teran-Navarro, H., Calderon-Gonzalez, R., Tobes, R., Garcia, I., ... Yañez-Díaz, S. (2020). Epitopes for multivalent vaccines against *Listeria*, *Mycobacterium* and *Streptococcus* spp: A novel role for glyceraldehyde-3-phosphate dehydrogenase. *Frontiers in Cellular and Infection Microbiology*, 10.
<https://doi.org/10.3389/fcimb.2020.573348>
- Andersson, B., Dahmén, J., Frejd, T., Leffler, H., Magnusson, G., Noori, G., & Edén, C. S. (1983). Identification of an active disaccharide unit of a glycoconjugate receptor for pneumococci attaching to human pharyngeal epithelial cells. *The Journal of Experimental Medicine*, 158(2), 559–570.
<https://doi.org/10.1084/JEM.158.2.559>
- Aprianto, R., Slager, J., Holsappel, S., & Veening, J. W. (2018). High-resolution analysis of the pneumococcal transcriptome under a wide range of infection-relevant conditions. *Nucleic Acids Research*, 46(19), 9990–10006.
<https://doi.org/10.1093/nar/gky750>
- Argimón, S., Abudahab, K., Goater, R. J. E., Fedosejev, A., Bhai, J., Glasner, C., ... Aanensen, D. M. (2016). Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microbial Genomics*, 2(11), e000093.
<https://doi.org/10.1099/mgen.0.000093>
- Attali, C., Durmort, C., Vernet, T., & Di Guilmi, A. M. (2008). The interaction of *Streptococcus pneumoniae* with plasmin mediates transmigration across endothelial and epithelial monolayers by intercellular junction cleavage. *Infection and Immunity*, 76(11), 5350–5356.
<https://doi.org/10.1128/IAI.00184-08>
- Ayoola, M. B., Shack, L. A., Nakamya, M. F., Thornton, J. A., Swiatlo, E., & Nanduri, B. (2019). Polyamine synthesis effects capsule expression by reduction of precursors in *Streptococcus pneumoniae*. *Frontiers in Microbiology*, 10.
<https://doi.org/10.3389/fmicb.2019.01996>
- Bansil, R., & Turner, B. S. (2006). Mucin structure, aggregation, physiological functions and biomedical applications. *Current Opinion in Colloid and Interface Science*, 11(2–3), 164–170.
<https://doi.org/10.1016/j.cocis.2005.11.001>

- Basavanna, S., Khandavilli, S., Yuste, J., Cohen, J. M., Hosie, A. H. F., Webb, A. J., ... Brown, J. S. (2009). Screening of *Streptococcus pneumoniae* ABC transporter mutants demonstrates that LivJHMGF, a branched-chain amino acid ABC transporter, is necessary for disease pathogenesis. *Infection and Immunity*, 77(8), 3412–3423. <https://doi.org/10.1128/IAI.01543-08>
- Beckonert, O., Keun, H. C., Ebbels, T. M. D., Bundy, J., Holmes, E., Lindon, J. C., & Nicholson, J. K. (2007). Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols*, 2, 2692–2703. <https://doi.org/10.1038/nprot.2007.376>
- Behrends, V., Tredwell, G. D., & Bundy, J. G. (2011). A software complement to AMDIS for processing GC-MS metabolomic data. *Analytical Biochemistry*, 415(2), 206–208. <https://doi.org/10.1016/J.AB.2011.04.009>
- Bennett, N. M., Whitney, C. G., Moore, M., Pilishvili, T., & Dooling, K. L. (2013). Use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine for adults with immunocompromising conditions: Recommendations of the Advisory Committee on Immunization Practices (ACIP). *Morbidity & Mortality Weekly Report*, 13(1), 232–235. <https://doi.org/10.1111/ajt.12073>
- Bentley, S. D., Aanensen, D. M., Mavroidi, A., Saunders, D., Rabinowitsch, E., Collins, M., ... Spratt, B. G. (2006). Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genetics*, 2(3). <https://doi.org/10.1371/journal.pgen.0020031>
- Berry, A. M., & Paton, J. C. (2000). Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infection and Immunity*, 68(1), 133–140. <https://doi.org/10.1128/IAI.68.1.133-140.2000>
- Bidossi, A., Mulas, L., Decorosi, F., Colomba, L., Ricci, S., Pozzi, G., ... Oggioni, M. R. (2012). A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. *PLoS ONE*, 7(3). <https://doi.org/10.1371/journal.pone.0033320>
- Bigogo, G. M., Audi, A., Auko, J., Aol, G. O., Ochieng, B. J., Odiembo, H., ... Verani,

- J. R. (2019). Indirect effects of 10-valent pneumococcal conjugate vaccine against adult pneumococcal pneumonia in rural western Kenya. *Clinical Infectious Diseases*, 69(12), 2177–2184.
<https://doi.org/10.1093/cid/ciz139>
- Blanchette-Cain, K., Hinojosa, C. A., Akula Suresh Babu, R., Lizcano, A., Gonzalez-Juarbe, N., Munoz-Almagro, C., ... Orihuela, C. J. (2013). *Streptococcus pneumoniae* biofilm formation is strain dependent, multifactorial, and associated with reduced invasiveness and immunoreactivity during colonization. *mBio*, 4(5).
<https://doi.org/10.1128/mBio.00745-13>
- Blanchette, K. A., Shenoy, A. T., Milner, J., Gilley, R. P., McClure, E., Hinojosa, C. A., ... Orihuela, C. J. (2016). Neuraminidase A-exposed galactose promotes *Streptococcus pneumoniae* biofilm formation during colonization. *Infection and Immunity*, 84(10), 2922–2932.
<https://doi.org/10.1128/IAI.00277-16>
- Blau, K., Portnoi, M., Shagan, M., Kaganovich, A., Rom, S., Kafka, D., ... Mizrachi Nebenzahl, Y. (2007). Flamingo Cadherin: A putative host receptor for *Streptococcus pneumoniae*. *The Journal of Infectious Diseases*, 195(12), 1828–1837. <https://doi.org/10.1086/518038>
- Blom, A. M., Bergmann, S., Fulde, M., Riesbeck, K., & Agarwal, V. (2014). *Streptococcus pneumoniae* phosphoglycerate kinase is a novel complement inhibitor affecting the Membrane Attack Complex formation. *Journal of Biological Chemistry*, 289(47), 32499–32511.
<https://doi.org/10.1074/jbc.M114.610212>
- Bogaert, D., de Groot, R., & Hermans, P. W. M. (2004). *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infectious Diseases*, 4, 144–154. [https://doi.org/10.1016/S1473-3099\(04\)00938-7](https://doi.org/10.1016/S1473-3099(04)00938-7)
- Bonofiglio, L., García, E., & Mollerach, M. (2012). The *galU* gene expression in *Streptococcus pneumoniae*. *FEMS Microbiology Letters*, 332, 47–53.
<https://doi.org/10.1111/j.1574-6968.2012.02572.x>
- Brissac, T., Shenoy, A. T., Patterson, L. D. A., & Orihuela, C. J. (2018). Cell invasion and pyruvate oxidase-derived H₂O₂ are critical for *Streptococcus pneumoniae*-mediated cardiomyocyte killing. *Infection and Immunity*, 86(1).

<https://doi.org/10.1128/IAI.00569-17>

- Brittan, J. L., Buckeridge, T. J., Finn, A., Kadioglu, A., & Jenkinson, H. F. (2012). Pneumococcal neuraminidase A: An essential upper airway colonization factor for *Streptococcus pneumoniae*. *Molecular Oral Microbiology*, 27(4), 270–283.
<https://doi.org/10.1111/J.2041-1014.2012.00658.X>
- Brown, A. O., Mann, B., Gao, G., Hankins, J. S., Humann, J., Giardina, J., ... Orihuela, C. J. (2014). *Streptococcus pneumoniae* translocates into the myocardium and forms unique microlesions that disrupt cardiac function. *PLoS Pathogens*, 10(9).
<https://doi.org/10.1371/journal.ppat.1004383>
- Brown, S. A., Palmer, K. L., & Whiteley, M. (2008). Revisiting the host as a growth medium. *Nature Reviews Microbiology*, 6, 657–666.
<https://doi.org/10.1038/nrmicro1955>
- Brueggemann, A. B., Griffiths, D. T., Meats, E., Peto, T., Crook, D. W., & Spratt, B. G. (2003). Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. *The Journal of Infectious Diseases*, 187, 1424–1432.
<https://doi.org/10.1086/374624>
- Bryant, J. C., Dabbs, R. C., Oswalt, K. L., Brown, L. R., Rosch, J. W., Seo, K. S., ... Thornton, J. A. (2016). Pyruvate oxidase of *Streptococcus pneumoniae* contributes to pneumolysin release. *BMC Microbiology*, 16.
<https://doi.org/10.1186/s12866-016-0881-6>
- Caceres-Cortes, J., & Reily, M. D. (2010). NMR spectroscopy as a tool to close the gap on metabolite characterization under MIST. *Bioanalysis*, 2(7), 1263–1276.
<https://doi.org/10.4155/bio.10.77>
- Cámara, M., Boulnois, G. J., Andrew, P. W., & Mitchell, T. J. (1994). A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. *Infection and Immunity*, 62(9), 3688–3695.
<https://doi.org/10.1128/iai.62.9.3688-3695.1994>
- Camilli, A., & van Opijnen, T. (2012). A fine scale phenotype-genotype virulence map of a bacterial pathogen. *Genome Research*, 2541–2551.
<https://doi.org/10.1101/gr.137430.112>
- Carvalho, S. M., Farshchi Andisi, V., Gradstedt, H., Neef, J., Kuipers, O. P., Neves, A.

- R., & Bijlsma, J. J. E. (2013a). Pyruvate oxidase influences the sugar utilization pattern and capsule production in *Streptococcus pneumoniae*. *PLoS ONE*, 8(7).
<https://doi.org/10.1371/journal.pone.0068277>
- Carvalho, S. M., Kloosterman, T. G., Kuipers, O. P., & Neves, A. R. (2011). CcpA ensures optimal metabolic fitness of *Streptococcus pneumoniae*. *PLoS ONE*, 6(10).
<https://doi.org/10.1371/journal.pone.0026707>
- Carvalho, S. M., Kuipers, O. P., & Neves, A. R. (2013b). Environmental and nutritional factors that affect growth and metabolism of the pneumococcal serotype 2 strain D39 and its nonencapsulated derivative strain R6. *PLoS ONE*, 8(3).
<https://doi.org/10.1371/journal.pone.0058492>
- CDC, NCIRD, & WHO. (2011). Identification and Characterization of *Streptococcus pneumoniae*. In *Laboratory Methods for the Diagnosis of Meningitis Caused by Neisseria meningitides, Streptococcus pneumoniae, and Haemophilus influenzae* (2nd ed.).
<http://www.cdc.gov/ncidod/biotech/strep/strep-doc/index.htm>.
- Centers for Disease Control and Prevention. (2013). *Streptococcus pneumoniae* 2013. Active Bacterial Core Surveillance Report, Emerging Infections Program Network.
<https://www.cdc.gov/abcs/reports-findings/survreports/spneu13.html>
- Chaguza, C., Cornick, J. E., & Everett, D. B. (2015). Mechanisms and impact of genetic recombination in the evolution of *Streptococcus pneumoniae*. *Computational and Structural Biotechnology Journal*, 13, 241–247.
<https://doi.org/10.1016/J.CSBJ.2015.03.007>
- Chang, J. C., Lasarre, B., Jimenez, J. C., Aggarwal, C., & Federle, M. J. (2011). Two Group A streptococcal peptide pheromones act through opposing Rgg regulators to control biofilm development. *PLoS Pathogens*, 7(8), 1002190.
<https://doi.org/10.1371/journal.ppat.1002190>
- Chanotiya, C., Pragadhessh, V., & Uniyal, G. (n.d.). *Avoiding interferences and contaminants using Eppendorf Safe-Lock Tubes in mass spectrometry studies of natural products*.
https://www.eppendorf.com/uploads/media/Application_282_Safe-Lock-Tubes_DNA-LoBind-Tubes_Avoiding-interferenc_01.pdf
- Chao, Y., Marks, L. R., Pettigrew, M. M., & Hakansson, A. P. (2015). *Streptococcus*

- pneumoniae* biofilm formation and dispersion during colonization and disease. *Frontiers in Cellular and Infection Microbiology*, 4.
<https://doi.org/10.3389/FCIMB.2014.00194>
- Chi, F., Nolte, O., Bergmann, C., Ip, M., & Hakenbeck, R. (2007). Crossing the barrier: Evolution and spread of a major class of mosaic *pbp2x* in *Streptococcus pneumoniae*, *S. mitis* and *S. oralis*. *International Journal of Medical Microbiology*, 297, 503–512.
<https://doi.org/10.1016/J.IJMM.2007.02.009>
- Chokkathukalam, A., Kim, D.-H., Barrett, M. P., Breitling, R., & Creek, D. J. (2014). Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks. *Bioanalysis*, 6(4), 511–524.
<https://doi.org/10.4155/bio.13.348>
- Cooper, V. S., Honsa, E., Rowe, H., Deitrick, C., Iverson, A. R., Whittall, J. J., ... Rosch, J. W. (2020). Experimental evolution *in vivo* to identify selective pressures during pneumococcal colonization. *mSystems*, 5(3).
<https://doi.org/10.1128/mSystems.00352-20>
- Corander, J., Fraser, C., Gutmann, M. U., Arnold, B., Hanage, W. P., Bentley, S. D., ... Croucher, N. J. (2017). Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nature Ecology and Evolution*, 1(12), 1950–1960. <https://doi.org/10.1038/s41559-017-0337-x>
- Corfield, A. P. (2015). Mucins: A biologically relevant glycan barrier in mucosal protection. *Biochimica et Biophysica Acta*, 1850, 236–252.
<https://doi.org/10.1016/j.bbagen.2014.05.003>
- Crain, M. J., Waltman, W. D., Turner, J. S., Yother, J., Talkington, D. F., McDaniel, L. S., ... Briles, D. E. (1990). Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. *Infection and Immunity*, 58(10), 3293–3299.
<https://doi.org/10.1128/IAI.58.10.3293-3299.1990>
- Croucher, N. J., Harris, S. R., Fraser, C., Quail, M. A., Burton, J., Van Der Linden, M., ... Bentley, S. D. (2011). Rapid pneumococcal evolution in response to clinical interventions. *Science*, 331.
<https://doi.org/https://www.science.org/doi/10.1126/science.1198545>

- D'Mello, A., Riegler, A. N., Martínez, E., Beno, S. M., Ricketts, T. D., Foxman, E. F., ... Tettelin, H. (2020). An in vivo atlas of host-pathogen transcriptomes during *Streptococcus pneumoniae* colonization and disease. *PNAS*, *117*(52), 33507–33518. <https://doi.org/10.1073/pnas.2010428117>
- Dalia, A. B., Standish, A. J., & Weiser, J. N. (2010). Three surface exoglycosidases from *Streptococcus pneumoniae*, NanA, BgaA, and StrH, promote resistance to opsonophagocytic killing by human neutrophils. *Infection and Immunity*, *78*(5), 2108–2116. <https://doi.org/10.1128/IAI.01125-09>
- Dananché, C., Paranhos-Baccalà, G., Messaoudi, M., Sylla, M., Awasthi, S., Bavdekar, A., ... Picot, V. S. (2020). Serotypes of *Streptococcus pneumoniae* in children aged <5 years hospitalized with or without pneumonia in developing and emerging countries: a descriptive, multicenter study. *Clinical Infectious Diseases*, *70*(5), 83. <https://doi.org/10.1093/cid/ciz277>
- Daniels, C. C., Coan, P., King, J., Hale, J., Benton, K. A., Briles, D. E., & Hollingshead, S. K. (2010). The proline-rich region of pneumococcal surface proteins A and C contains surface-accessible epitopes common to all pneumococci and elicits antibody-mediated protection against sepsis. *Infection and Immunity*, *78*(5), 72. <https://doi.org/10.1128/IAI.01199-09>
- Dave, S., Brooks-Walter, A., Pangburn, M. K., & McDaniel, L. S. (2001). PspC, a pneumococcal surface protein, binds human factor H. *Infection and Immunity*, *69*(5), 3435–3437. <https://doi.org/10.1128/IAI.69.5.3435-3437.2001>
- de Carvalho, L. P. S., Fischer, S. M., Marrero, J., Nathan, C., Ehrt, S., & Rhee, K. Y. (2010). Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chemistry and Biology*, *17*, 1122–1131. <https://doi.org/10.1016/j.chembiol.2010.08.009>
- Diven, W. F., Doyle, W. J., & Vietmeier, B. (1988). Hydrolytic enzymes in otitis media pathogenesis. *Annals of Otology, Rhinology and Laryngology*, *97*, 6–9. <https://doi.org/10.1177/00034894880970s303>
- Dockrell, D. H., Marriott, H. M., Prince, L. R., Ridger, V. C., Ince, P. G., Hellewell, P. G., & Whyte, M. K. B. (2003). Alveolar macrophage apoptosis contributes to pneumococcal clearance in a resolving model of pulmonary infection. *The Journal*

- of Immunology*, 171, 5380–5388.
<https://doi.org/10.4049/jimmunol.171.10.5380>
- Donati, C., Hiller, N. L., Tettelin, H., Muzzi, A., Croucher, N. J., Angiuoli, S. V., ... Masignani, V. (2010). Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biology*, 11(10).
<https://doi.org/10.1186/gb-2010-11-10-r107>
- Douglas, A. E. (2020). The microbial exometabolome: ecological resource and architect of microbial communities. *Philosophical Transactions of the Royal Society*, 375.
<https://doi.org/10.1098/rstb.2019.0250>
- Echlin, H., Frank, M., Rock, C., & Rosch, J. W. (2020). Role of the pyruvate metabolic network on carbohydrate metabolism and virulence in *Streptococcus pneumoniae*. *Molecular Microbiology*, 114, 536–552.
<https://doi.org/10.1111/mmi.14557>
- Echlin, H., Frank, M. W., Iverson, A., Chang, T. C., Johnson, M. D. L., Rock, C. O., & Rosch, J. W. (2016). Pyruvate oxidase as a critical link between metabolism and capsule biosynthesis in *Streptococcus pneumoniae*. *PLoS Pathogens*, 12(10).
<https://doi.org/10.1371/journal.ppat.1005951>
- Ehrt, S., & Rhee, K. Y. (2012). *Mycobacterium tuberculosis* metabolism and host interaction: mysteries and paradoxes. *Current Topics in Microbiology and Immunology*.
https://doi.org/10.1007/82_2012_299
- Eisenreich, W., Slaghuis, J., Laupitz, R., Bussemer, J., Stritzker, J., Schwarz, C., ... Bacher, A. (2006). ¹³C isotopologue perturbation studies of *Listeria monocytogenes* carbon metabolism and its modulation by the virulence regulator PrfA. *PNAS*, 103(7), 2040–2045.
<https://doi.org/10.1073/pnas.0507580103>
- Enright, M. C., & Spratt, B. G. (1998). A multilocus sequence typing scheme for *Streptococcus pneumoniae*: Identification of clones associated with serious invasive disease. *Microbiology*, 144(11), 3049–3060.
<https://doi.org/10.1099/00221287-144-11-3049>
- Eoh, H. (2014). Metabolomics: A window into the adaptive physiology of *Mycobacterium tuberculosis*. *Tuberculosis*, 94(6), 538–543.

<https://doi.org/10.1016/j.tube.2014.08.002>

Ercoli, G., Fernandes, V. E., Chung, W. Y., Wanford, J. J., Thomson, S., Bayliss, C. D., ... Oggioni, M. R. (2018). Intracellular replication of *Streptococcus pneumoniae* inside splenic macrophages serves as a reservoir for septicaemia. *Nature Microbiology*, 3(5), 600–610.

<https://doi.org/10.1038/s41564-018-0147-1>

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048.

<https://doi.org/10.1093/bioinformatics/btw354>

Fahy, J. V., & Dickey, B. F. (2010). Airway mucus function and dysfunction. *New England Journal of Medicine*, 363(23), 2233–2247.

<https://doi.org/10.1056/NEJMr0910061>

Fan, R. R., Howard, L. M., Griffin, M. R., Edwards, K. M., Zhu, Y., Williams, J. V., ... Grijalva, C. G. (2016). Nasopharyngeal pneumococcal density and evolution of acute respiratory illnesses in young children, Peru, 2009–2011. *Emerging Infectious Diseases*, 22(11), 1996–1999.

<https://doi.org/10.3201/eid2211.160902>

Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., & Palsson, B. Ø. (2008). Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, 7(2), 129–143. <https://doi.org/10.1038/nrmicro1949>

Finkle, P. (1931). Metabolism of S and R forms of pneumococcus. *Journal of Experimental Medicine*, 53(5), 661–676.

Finkle, P. (1936). Metabolism of various types of sugars by S and R forms of pneumococcus. *Journal of Bacteriology*, 32(5), 473–483.

Fleming, E. (2016). *Mechanisms of carbon catabolite repression in Streptococcus pneumoniae*. Tufts University (PhD Thesis).

Fleming, E., & Camilli, A. (2016). ManLMN is a glucose transporter and central metabolic regulator in *Streptococcus pneumoniae*. *Molecular Microbiology*, 102(3), 467–487.

<https://doi.org/10.1111/mmi.13473>

Fleming, E., Lazinski, D. W., & Camilli, A. (2015). Carbon catabolite repression by

- seryl phosphorylated HPr is essential to *Streptococcus pneumoniae* in carbohydrate-rich environments. *Molecular Microbiology*, 97(2), 360–380.
<https://doi.org/10.1111/mmi.13033>
- Fulde, M., Bernardo-García, N., Rohde, M., Nachtigall, N., Frank, R., Preissner, K. T., ... Bergmann, S. (2014). Pneumococcal phosphoglycerate kinase interacts with plasminogen and its tissue activator. *Thrombosis and Haemostasis*, 111, 401–416.
<https://doi.org/10.1160/TH13-05-0421>
- Gaspar, P., Al-Bayati, F. A. Y., Andrew, P. W., Neves, A. R., & Yesilkaya, H. (2014). Lactate dehydrogenase is the key enzyme for pneumococcal pyruvate metabolism and pneumococcal survival in blood. *Infection and Immunity*, 82(12), 5099–5109.
<https://doi.org/10.1128/IAI.02005-14>
- Giammarinaro, P., & Paton, J. C. (2002). Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*. *Infection and Immunity*, 70(10), 5454–5461.
<https://doi.org/10.1128/IAI.70.10.5454-5461.2002>
- Gilley, R. P., & Orihuela, C. J. (2014). Pneumococci in biofilms are non-invasive: implications on nasopharyngeal colonization. *Frontiers in Cellular and Infection Microbiology*, 4.
<https://doi.org/10.3389/FCIMB.2014.00163>
- Gladstone, R. A., Lo, S. W., Lees, J. A., Croucher, N. J., van Tonder, A. J., Corander, J., ... The Global Pneumococcal Sequencing Consortium. (2019). International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine*, 43, 338–346.
<https://doi.org/10.1016/j.ebiom.2019.04.021>
- Goldblatt, D., Hussain, M., Andrews, N., Ashton, L., Virta, C., Melegaro, A., ... Miller, E. (2005). Antibody responses to nasopharyngeal carriage of *Streptococcus pneumoniae* in adults: A longitudinal household study. *The Journal of Infectious Diseases*, 192(3), 93.
<https://doi.org/10.1086/431524>
- Gosink, K. K., Mann, E. R., Guglielmo, C., Tuomanen, E. I., & Masure, H. R. (2000). Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infection and Immunity*, 68(10), 5690–5695.

- <https://doi.org/10.1128/IAI.68.10.5690-5695.2000>
- Gratz, N., Loh, L. N., Mann, B., Gao, G., Carter, R., Rosch, J., & Tuomanen, E. I. (2017). Pneumococcal neuraminidase activates TGF- β signalling. *Microbiology*, 163. <https://doi.org/10.1099/mic.0.000511>
- Guest, J. R. (1995). The Leeuwenhoek Lecture, 1995: Adaptation to Life without Oxygen. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 350(1332), 189–202. <https://doi.org/10.1098/rstb.1995.0152>
- Guo, Y., Cho, S. W., Saxena, D., & Li, X. (2020). Multifaceted actions of succinate as a signaling transmitter vary with its cellular locations. *Endocrinology and Metabolism*, 35(1), 36–43. <https://doi.org/10.3803/EnM.2020.35.1.36>
- Gupta, R., Yang, J., Dong, Y., Swiatlo, E., Zhang, J. R., Metzger, D. W., & Bai, G. (2013). Deletion of *arcD* in *Streptococcus pneumoniae* D39 impairs its capsule and attenuates virulence. *Infection and Immunity*, 81(10), 3903–3911. <https://doi.org/10.1128/IAI.00778-13>
- Hajaj, B., Yesilkaya, H., Shafeeq, S., Zhi, X., Benisty, R., Tchalal, S., ... Porat, N. (2017). CodY regulates thiol peroxidase expression as part of the pneumococcal defense mechanism against H₂O₂ stress. *Frontiers in Cellular and Infection Microbiology*, 7. <https://doi.org/10.3389/fcimb.2017.00210>
- Hakenbeck, R., Balmelle, N., Weber, B., Gardès, C., Keck, W., & De Saizieu, A. (2001). Mosaic genes and mosaic chromosomes: Intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infection and Immunity*, 69(4), 2477–2486. <https://doi.org/10.1128/IAI.69.4.2477-2486.2001>
- Hamaguchi, S., Zafar, M. A., Cammer, M., & Weiser, J. N. (2018). Capsule prolongs survival of *Streptococcus pneumoniae* during starvation. *Infection and Immunity*, 86(3). <https://doi.org/10.1128/IAI.00802-17>
- Hammerschmidt, S., Wolff, S., Hocke, A., Rosseau, S., Müller, E., & Rohde, M. (2005). Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infection and Immunity*, 73(8), 4653–4667. <https://doi.org/10.1128/IAI.73.8.4653-4667.2005>

- Hardy, G. G., Caimano, M. J., & Yother, J. (2000). Capsule biosynthesis and basic metabolism in *Streptococcus pneumoniae* are linked through the cellular phosphoglucosyltransferase. *Journal of Bacteriology*, 182(7), 1854–1863.
<https://journals.asm.org/doi/10.1128/JB.182.7.1854-1863.2000>
- Härtel, T., Eylert, E., Schulz, C., Petruschka, L., Gierok, P., Grubmüller, S., ... Hammerschmidt, S. (2012). Characterization of central carbon metabolism of *Streptococcus pneumoniae* by isotopologue profiling. *Journal of Biological Chemistry*, 287(6), 4260–4274.
<https://doi.org/10.1074/jbc.M111.304311>
- Hathaway, L. J., Brugger, S. D., Morand, B., Bangert, M., Rotzetter, J. U., Hauser, C., ... Mühlemann, K. (2012). Capsule type of *Streptococcus pneumoniae* determines growth phenotype. *PLoS Pathogens*, 8(3).
<https://doi.org/10.1371/journal.ppat.1002574>
- Hausdorff, W. P., Feikin, D. R., & Klugman, K. P. (2005). Epidemiological differences among pneumococcal serotypes. *Lancet Infectious Diseases*, 5, 83–93.
[https://doi.org/10.1016/S1473-3099\(05\)01280-6](https://doi.org/10.1016/S1473-3099(05)01280-6)
- Hava, D. L., & Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Molecular Microbiology*, 45(5), 1389–1406. <https://doi.org/10.1046/j.1365-2958.2002.03106.x>
- Hendriksen, W. T., Bootsma, H. J., Estevão, S., Hoogenboezem, T., de Jong, A., de Groot, R., ... Hermans, P. W. M. (2008). CodY of *Streptococcus pneumoniae*: link between nutritional gene regulation and colonization. *Journal of Bacteriology*, 190(2), 590–601.
<https://doi.org/10.1128/JB.00917-07>
- Hendriksen, W. T., Kloosterman, T. G., Bootsma, H. J., Estevão, S., de Groot, R., Kuipers, O. P., & Hermans, P. W. M. (2008). Site-specific contributions of glutamine-dependent regulator GlnR and GlnR-regulated genes to virulence of *Streptococcus pneumoniae*. *Infection and Immunity*, 76(3), 1230–1238.
<https://doi.org/10.1128/IAI.01004-07>
- Henriques-Normark, B., & Tuomanen, E. I. (2013). The Pneumococcus: Epidemiology, Microbiology, and Pathogenesis. *Cold Spring Harbor Perspectives in Medicine*.
<https://dx.doi.org/10.1101%2Fcshperspect.a010215>

- Hiller, N. L., & Sá-Leão, R. (2018). Puzzling over the pneumococcal pangenome. *Frontiers in Microbiology*, 9.
<https://doi.org/10.3389/fmicb.2018.02580>
- Hobbs, J. K., Meier, E. P. W., Pluvinau, B., Mey, M. A., & Boraston, A. B. (2019). Molecular analysis of an enigmatic *Streptococcus pneumoniae* virulence factor: The raffinose-family oligosaccharide utilization system. *Journal of Biological Chemistry*, 294(46), 17197–17208.
<https://doi.org/10.1074/jbc.RA119.010280>
- Hoskins, J., Alborn, W. E., Arnold, J., Blaszcak, L. C., Burgett, S., Estrem, S. T., ... Glass, J. I. (2001). Genome of the bacterium *Streptococcus pneumoniae* strain R6. *Journal of Bacteriology*, 183(19), 5709–5717.
<https://doi.org/10.1128/JB.183.19.5709>
- Huynen, M. A., Dandekar, T., & Bork, P. (1999). Variation and evolution of the citric-acid cycle: A genomic perspective. *Trends in Microbiology*, 7(7), 281–291.
[https://doi.org/10.1016/S0966-842X\(99\)01539-5](https://doi.org/10.1016/S0966-842X(99)01539-5)
- Hyams, C., Camberlein, E., Cohen, J. M., Bax, K., & Brown, J. S. (2010). The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. *Infection and Immunity*.
<https://doi.org/10.1128/IAI.00881-09>
- Hyams, C. J. (2009). The role of the *Streptococcus pneumoniae* capsule in interactions with complement and phagocytes. University College London (PhD Thesis).
- Hyams, C., Opel, S., Hanage, W., Yuste, J., Bax, K., Henriques-Normark, B., ... Brown, J. S. (2011). Effects of *Streptococcus pneumoniae* strain background on complement resistance. *PLoS ONE*, 6(10).
<https://doi.org/10.1371/journal.pone.0024581>
- Iyer, R., Baliga, N. S., & Camilli, A. (2005). Catabolite control protein A (CcpA) contributes to virulence and regulation of sugar metabolism in *Streptococcus pneumoniae*. *Journal of Bacteriology*, 187(24), 8340–8349.
<https://doi.org/10.1128/JB.187.24.8340-8349.2005>
- Iyer, R., & Camilli, A. (2007). Sucrose metabolism contributes to *in vivo* fitness of *Streptococcus pneumoniae*. *Molecular Microbiology*, 66(1).
<https://doi.org/10.1111/j.1365-2958.2007.05878.x>

- Jang, C., Chen, L., & Rabinowitz, J. D. (2018). Metabolomics and isotope tracing. *Cell*, 173, 822–837. <https://doi.org/10.1016/j.cell.2018.03.055>
- Jolley, K. A., Bray, J. E., & Maiden, M. C. J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Research*, 3. <https://doi.org/10.12688/wellcomeopenres.14826.1>
- Kadioglu, A., & Andrew, P. W. (2004). The innate immune response to pneumococcal lung infection: The untold story. *Trends in Immunology*, 25(3), 143–149. <https://doi.org/10.1016/j.it.2003.12.006>
- Kahya, H. F., Andrew, P. W., & Yesilkaya, H. (2017). Deacetylation of sialic acid by esterases potentiates pneumococcal neuraminidase activity for mucin utilization, colonization and virulence. *PLoS Pathogens*, 13(3), 1–21. <https://doi.org/10.1371/journal.ppat.1006263>
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1), 27–30. <https://doi.org/10.1093/nar/28.1.27>
- Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., ... Subhraveti, P. (2017). The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics*, 20(4). <https://doi.org/10.1093/bib/bbx085>
- Keller, L. E., Robinson, D. A., & McDaniel, L. S. (2016). Nonencapsulated *Streptococcus pneumoniae*: Emergence and pathogenesis. *mBio*, 7(2). <https://doi.org/10.1128/mBio.01792-15>
- Kelly, R. T., Farmer, S., & Greiff, D. (1967). Neuraminidase activities of clinical isolates of *Diplococcus pneumoniae*. *Journal of Bacteriology*, 94(1), 272–273. <https://doi.org/10.1128/JB.94.1.272-273.1967>
- Kent, A., Makwana, A., Sheppard, C. L., Collins, S., Fry, N. K., Heath, P. T., ... Ladhani, S. N. (2019). Invasive pneumococcal disease in UK children <1 year of age in the post-13-valent pneumococcal conjugate vaccine era: What are the risks now? *Clinical Infectious Diseases*, 69(1), 84–90. <https://doi.org/10.1093/cid/ciy842>
- Kerr, A. R., Paterson, G. K., McCluskey, J., Iannelli, F., Oggioni, M. R., Pozzi, G., &

- Mitchell, T. J. (2006). The contribution of PspC to pneumococcal virulence varies between strains and is accomplished by both complement evasion and complement-independent mechanisms. *Infection and Immunity*, 74(9), 5324. <https://doi.org/10.1128/IAI.00543-06>
- Kilian, M., Poulsen, K., Blomqvist, T., Håvarstein, L. S., Bek-Thomsen, M., Tettelin, H., & Sørensen, U. B. S. (2008). Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS ONE*, 3(7), e2683. <https://doi.org/10.1371/JOURNAL.PONE.0002683>
- Kim, G.-L., Lee, S., Luong, T. T., Nguyen, C. T., Park, S.-S., Pyo, S., & Rhee, D.-K. (2017). Effect of decreased BCAA synthesis through disruption of *ilvC* gene on the virulence of *Streptococcus pneumoniae*. *Archives of Pharmacal Research*, 40, 921–932. <https://doi.org/10.1007/s12272-017-0931-0>
- Kim, J. O., Romero-Steiner, S., Skov Sørensen, U. B., Blom, J., Carvalho, M., Barnard, S., ... Weiser, J. N. (1999). Relationship between cell surface carbohydrates and intrastrain variation on opsonophagocytosis of *Streptococcus pneumoniae*. *Infection and Immunity*, 67(5), 2327–2333. <https://doi.org/10.1128/IAI.67.5.2327-2333.1999>
- Kim, K. C. (2012). Role of epithelial mucins during airway infection. *Pulmonary Pharmacology and Therapeutics*, 25(6), 415–419. <https://doi.org/10.1016/j.pupt.2011.12.003>
- Kind, T., Tsugawa, H., Cajka, T., Ma, Y., Lai, Z., Mehta, S. S., ... Fiehn, O. (2018). Identification of small molecules using accurate mass MS/MS search. *Mass Spectrometry Reviews*, 37(4), 513–532. <https://doi.org/10.1002/mas.21535>
- King, S. J., Whatmore, A. M., & Dowson, C. G. (2005). NanA, a neuraminidase from *Streptococcus pneumoniae*, shows high levels of sequence diversity, at least in part through recombination with *Streptococcus oralis*. *Journal of Bacteriology*, 187(15), 5376–5386. <https://doi.org/10.1128/JB.187.15.5376-5386.2005>
- King, Samantha J., Hippe, K. R., & Weiser, J. N. (2006). Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Molecular Microbiology*, 59(3), 961–974.

- <https://doi.org/10.1111/j.1365-2958.2005.04984.x>
- Klugman, K. P., Chien, Y. W., & Madhi, S. A. (2009). Pneumococcal pneumonia and influenza: A deadly combination. *Vaccine*, 27(SUPPL. 3), C9–C14.
<https://doi.org/10.1016/j.vaccine.2009.06.007>
- Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., ... Wilson, R. K. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, 22(3), 568–576.
<https://doi.org/10.1101/gr.129684.111>
- Kohlstedt, M., & Wittmann, C. (2019). GC-MS-based ¹³C metabolic flux analysis resolves the parallel and cyclic glucose metabolism of *Pseudomonas putida* KT2440 and *Pseudomonas aeruginosa* PAO1. *Metabolic Engineering*, 54, 35–53.
<https://doi.org/10.1016/j.ymben.2019.01.008>
- Kwong, W. K., Zheng, H., & Moran, N. A. (2017). Convergent evolution of a modified, acetate-driven TCA cycle in bacteria. *Nature Microbiology*, 2.
<https://doi.org/10.1038/nmicrobiol.2017.67>
- Lai, Z., & Fiehn, O. (2016). Mass spectral fragmentation of trimethylsilylated small molecules. *Mass Spectrometry Reviews*.
<https://doi.org/10.1002/mas.21518>
- Lanie, J. A., Ng, W. L., Kazmierczak, K. M., Andrzejewski, T. M., Davidsen, T. M., Wayne, K. J., ... Winkler, M. E. (2007). Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. *Journal of Bacteriology*, 189(1), 38–51.
<https://doi.org/10.1128/JB.01148-06>
- Lau, G. W., Haataja, S., Lonetto, M., Kensit, S. E., Marra, A., Bryant, a P., ... Holden, D. W. (2001). A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. *Molecular Microbiology*, 40(3), 555–571.
<https://doi.org/mmi2335> [pii]
- Leen, W. G., Willemsen, M. A., Wevers, R. A., & Verbeek, M. M. (2012). Cerebrospinal fluid glucose and lactate: Age-specific reference values and implications for clinical practice. *PLoS ONE*, 7(8).
<https://doi.org/10.1371/journal.pone.0042745>
- Leonard, A., & Lalk, M. (2018). Infection and metabolism – *Streptococcus pneumoniae*

- metabolism facing the host environment. *Cytokine*, 112, 75–86.
<https://doi.org/10.1016/j.cyto.2018.07.021>
- Leonard, A., Gierok, P., Methling, K., Gómez-Mejia, A., Hammerschmidt, S., & Lalk, M. (2018). Metabolic inventory of *Streptococcus pneumoniae* growing in a chemical defined environment. *International Journal of Medical Microbiology*, 308, 705–712.
<https://doi.org/10.1016/j.ijmm.2018.01.001>
- Leonard, A., Möhlis, K., Schlüter, R., Taylor, E., Lalk, M., & Methling, K. (2020). Exploring metabolic adaptation of *Streptococcus pneumoniae* to antibiotics. *Journal of Antibiotics*, 73, 441–454.
<https://doi.org/10.1038/s41429-020-0296-3>
- Leventer-Roberts, M., Feldman, B. S., Brufman, I., Cohen-Stavi, C. J., Hoshen, M., & Balicer, R. D. (2015). Effectiveness of 23-valent pneumococcal polysaccharide vaccine against invasive disease and hospital-treated pneumonia among people aged ≤ 65 years: A retrospective case-control study. *Clinical Infectious Diseases*, 60(10), 1472–1480.
<https://doi.org/10.1093/cid/civ096>
- Levy, C., Ouldali, N., Caeymaex, L., Angoulvant, F., Varon, E., & Cohen, R. (2019). Diversity of serotype replacement after pneumococcal conjugate vaccine implementation in Europe. *Journal of Pediatrics*, 213.
<https://doi.org/10.1016/j.jpeds.2019.07.057>
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595.
<https://doi.org/10.1093/bioinformatics/btp698>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
<https://doi.org/10.1093/bioinformatics/btp352>
- Ling, E., Feldman, G., Portnoi, M., Dagan, R., Overweg, K., Mulholland, F., ... Mizrachi-Nebenzahl, Y. (2004). Glycolytic enzymes associated with the cell surface of *Streptococcus pneumoniae* are antigenic in humans and elicit protective immune responses in the mouse. *Clinical and Experimental Immunology*, 138,

- 290–298. <https://doi.org/10.1111/j.1365-2249.2004.02628.x>
- Lisher, J. P., Tsui, H.-C. T., Ramos-Montañez, S., Hentchel, K. L., Martin, J. E., Trinidad, J. C., ... Giedroc, D. P. (2017). Biological and chemical adaptation to endogenous hydrogen peroxide production in *Streptococcus pneumoniae* D39. *mSphere*, 2(1).
<https://doi.org/10.1128/msphere.00291-16>
- Liu, X., Gallay, C., Kjos, M., Domenech, A., Slager, J., Van Kessel, S. P., ... Veening, J.-W. (2017). High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Molecular Systems Biology*, 13, 931.
<https://doi.org/10.15252/msb.20167449>
- Lo, S. W., Gladstone, R. A., van Tonder, A. J., Lees, J. A., du Plessis, M., Benisty, R., ... The Global Pneumococcal Sequencing Consortium. (2019). Pneumococcal lineages associated with serotype replacement and antibiotic resistance in childhood invasive pneumococcal disease in the post-PCV13 era: an international whole-genome sequencing study. *The Lancet Infectious Diseases*, 19(7), 759–769.
[https://doi.org/10.1016/S1473-3099\(19\)30297-X](https://doi.org/10.1016/S1473-3099(19)30297-X)
- Lopatkin, A. J., Bening, S. C., Manson, A. L., Stokes, J. M., Kohanski, M. A., Badran, A. H., ... Collins, J. J. (2021). Clinically relevant mutations in core metabolic genes confer antibiotic resistance. *Science*, 371.
<https://doi.org/10.1126/science.aba0862>
- Lopatkin, A. J., Stokes, J. M., Zheng, E. J., Yang, J. H., Takahashi, M. K., You, L., & Collins, J. J. (2019). Bacterial metabolic state more accurately predicts antibiotic lethality than growth rate. *Nature Microbiology*, 4, 2109–2117.
<https://doi.org/10.1038/s41564-019-0536-0>
- Loughran, A. J., Orihuela, C. J., & Tuomanen, E. I. (2019). *Streptococcus pneumoniae*: Invasion and Inflammation. *Microbiology Spectrum*, 7(2).
<https://doi.org/10.1128/microbiolspec>
- Luong, T. T., Kim, E.-H., Bak, J. P., Nguyen, C. T., Choi, S., Briles, D. E., ... Rhee, D.-K. (2015). Ethanol-induced alcohol dehydrogenase E (AdhE) potentiates pneumolysin in *Streptococcus pneumoniae*. *Infection and Immunity*, 83, 108–119.
<https://doi.org/10.1128/IAI.02434-14>
- MacLeod, C. M., & Krauss, M. R. (1950). Relation of virulence of pneumococcal

- strains for mice to the quantity of capsular polysaccharide formed *in vitro*. *The Journal of Experimental Medicine*, 92(1), 1–9.
<https://doi.org/10.1084/jem.92.1.1>
- MacRae, J. I., Dixon, M. W. A., Dearnley, M. K., Chua, H. H., Chambers, J. M., Kenny, S., ... McConville, M. J. (2013). Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite *Plasmodium falciparum*. *BMC Biology*, 11, 67.
<https://doi.org/10.1186/1741-7007-11-67>
- Magee, A. D., & Yother, J. (2001). Requirement for capsule in colonization by *Streptococcus pneumoniae*. *Infection and Immunity*, 69(6), 3755–3761.
<https://doi.org/10.1128/IAI.69.6.3755-3761.2001>
- Mahdi, L. K., Higgins, M. A., Day, C. J., Tiralongo, J., Hartley-Tassell, L. E., Jennings, M. P., ... Ogunniyi, A. D. (2017). The pneumococcal alpha-glycerophosphate oxidase enhances nasopharyngeal colonization through binding to host glycoconjugates. *EBioMedicine*, 18, 236–243.
<https://doi.org/10.1016/j.ebiom.2017.03.002>
- Mahdi, L. K., Wang, H., van der Hoek, M. B., Paton, J. C., & Ogunniyi, A. D. (2012). Identification of a novel pneumococcal vaccine antigen preferentially expressed during meningitis in mice. *Journal of Clinical Investigation*, 122(6), 20.
<https://doi.org/10.1172/JCI45850>
- Manso, A. S., Chai, M. H., Atack, J. M., Furi, L., De Ste Croix, M., Haigh, R., ... Oggioni, M. R. (2014). A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nature Communications*, 5.
<https://doi.org/10.1038/ncomms6055>
- Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalian, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43, 34–40.
<https://doi.org/10.1016/j.copbio.2016.08.001>
- Marriott, H. M., Gascoyne, K. A., Gowda, R., Geary, I., Nicklin, M. J. H., Iannelli, F., ... Dockrell, D. H. (2012). Interleukin-1 β regulates CXCL8 release and influences disease outcome in response to *Streptococcus pneumoniae*, defining intercellular cooperation between pulmonary epithelial cells and macrophages. *Infection and*

- Immunity*, 80(3), 1140–1149.
<https://doi.org/10.1128/IAI.05697-11>
- Marriott, H. M., Hellewell, P. G., Cross, S. S., Ince, P. G., Whyte, M. K. B., & Dockrell, D. H. (2006). Decreased alveolar macrophage apoptosis is associated with increased pulmonary inflammation in a murine model of pneumococcal pneumonia. *Journal of Immunology*, 177, 6480–6488.
<https://doi.org/10.4049/jimmunol.177.9.6480>
- McCullers, J. A., & Tuomanen, E. I. (2001). Molecular pathogenesis of pneumococcal pneumonia. *Journal of Experimental Medicine*, 877–889.
<https://doi.org/10.2741/mccullers>
- McDaniel, L. S., Yother, J., Vijayakumar, M., McGarry, L., Guild, W. R., & Briles, D. E. (1987). Use of insertional inactivation to facilitate studies of biological properties of pneumococcal surface protein A (PspA). *Journal of Experimental Medicine*, 165(2), 381–394.
<https://doi.org/10.1084/JEM.165.2.381>
- McDevitt, E., Khan, F., Scasny, A., Thompson, C. D., Eichenbaum, Z., McDaniel, L. S., & Vidal, J. E. (2020). Hydrogen peroxide production by *Streptococcus pneumoniae* results in alpha-hemolysis by oxidation of oxy-hemoglobin to met-hemoglobin. *mSphere*, 5(6).
<https://doi.org/10.1128/msphere.01117-20>
- Metallo, C. M., Gameiro, P. A., Bell, E. L., Mattaini, K. R., Yang, J., Hiller, K., ... Stephanopoulos, G. (2012). Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature*, 481(7381), 380–384.
<https://doi.org/10.1038/nature10602>
- Middleton, D. R., Paschall, A. V., Duke, J. A., & Avci, F. Y. (2018). Enzymatic hydrolysis of pneumococcal capsular polysaccharide renders the bacterium vulnerable to host defense. *Infection and Immunity*, 86(8).
<https://doi.org/10.1128/IAI.00316-18>
- Minhas, V., Aprianto, R., McAllister, L. J., Wang, H., David, S. C., McLean, K. T., ... Trappetti, C. (2020). In vivo dual RNA-seq reveals that neutrophil recruitment underlies differential tissue tropism of *Streptococcus pneumoniae*. *Nature Communications Biology*, 3(293).

<https://doi.org/10.1038/s42003-020-1018-x>

Minhas, V., Harvey, R. M., McAllister, L. J., Seemann, T., Syme, A. E., Baines, S. L., ... Trappetti, C. (2019). Capacity to utilize raffinose dictates pneumococcal disease phenotype. *mBio*, 10(1).

<https://doi.org/10.1128/mBio.02596-18>

Mitchell, A. M., & Mitchell, T. J. (2010). *Streptococcus pneumoniae*: virulence factors and variation. *Clinical Microbiology and Infection*, 16(5), 411–418.

<https://doi.org/10.1111/j.1469-0691.2010.03183.x>

Mitchell, T. J. (2000). Virulence factors and the pathogenesis of disease caused by *Streptococcus pneumoniae*. *Research in Microbiology*, 151, 413–419.

[https://doi.org/10.1016/S0923-2508\(00\)00175-3](https://doi.org/10.1016/S0923-2508(00)00175-3)

Moberley, S., Holden, J., Tatham, D. P., Andrews, R. M., & Cochrane Acute Respiratory Infections Group. (2013). Vaccines for preventing pneumococcal infection in adults. *Cochrane Database of Systematic Reviews*.

<https://doi.org/10.1002/14651858.CD000422.pub3>

Morona, J. K., Miller, D. C., Morona, R., & Paton, J. C. (2004). The effect that mutations in the conserved capsular polysaccharide biosynthesis genes *cpsA*, *cpsB*, and *cpsD* have on virulence of *Streptococcus pneumoniae*. *The Journal of Infectious Diseases*, 189, 1905–1913.

<https://doi.org/10.1086/383352>

Moscoso, M., García, E., & López, R. (2006). Biofilm formation by *Streptococcus pneumoniae*: Role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. *Journal of Bacteriology*, 188(22), 7785–7795.

<https://doi.org/10.1128/JB.00673-06>

Muchnik, L., Adawi, A., Ohayon, A., Dotan, S., Malka, I., Azriel, S., ... Nebenzahl, Y. M. (2013). NADH oxidase functions as an adhesin in *Streptococcus pneumoniae* and elicits a protective immune response in mice. *PLoS ONE*, 8(4).

<https://doi.org/10.1371/journal.pone.0061128>

Mukerji, R., Hendrickson, C., Genschmer, K. R., Park, S. S., Bouchet, V., Goldstein, R., ... Briles, D. E. (2018). The diversity of the proline-rich domain of pneumococcal surface protein A (PspA): Potential relevance to a broad-spectrum vaccine. *Vaccine*, 36(45), 43.

- <https://doi.org/10.1016/j.vaccine.2018.08.045>
- Muñoz-Elías, E. J., & McKinney, J. D. (2006). Carbon metabolism of intracellular bacteria. *Cellular Microbiology*, 8(1), 10–22.
<https://doi.org/10.1111/j.1462-5822.2005.00648.x>
- Murphy, M. P., & O'Neill, L. A. J. (2018). Krebs Cycle reimaged: The emerging roles of succinate and itaconate as signal transducers. *Cell*, 174, 780–784.
<https://doi.org/10.1016/j.cell.2018.07.030>
- Musher, D. M. (2021). Pneumococcal vaccination in adults - UpToDate.
 Retrieved July 3, 2021, from <https://www.uptodate.com/contents/pneumococcal-vaccination-in-adults>
- Nagaoka, K., Yamashita, Y., Kimura, H., Suzuki, M., Konno, S., Fukumoto, T., ... Nishimura, M. (2018). Effects of anaerobic culturing on pathogenicity and virulence-related gene expression in pneumococcal pneumonia. *The Journal of Infectious Diseases*, 219, 1545–1553.
<https://doi.org/10.1093/infdis/jiy718>
- Nahm, M. H., Brissac, T., Kilian, M., Vlach, J., Orihuela, C. J., Saad, J. S., & Ganaie, F. (2019). Pneumococci can become virulent by acquiring a new capsule from oral streptococci. *The Journal of Infectious Diseases*, 222(3), 372–380.
<https://doi.org/https://academic.oup.com/jid/article/222/3/372/5564845>
- Neijssel, O. M., Snoep, J. L., & Teixeira de Mattos, M. J. (1997). Regulation of energy source metabolism in streptococci. *Journal of Applied Microbiology*, 83, 12S-19S.
<https://doi.org/10.1046/j.1365-2672.83.s1.2.x>
- Nelson, A. L., Roche, A. M., Gould, J. M., Chim, K., Ratner, A. J., & Weiser, J. N. (2007). Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infection and Immunity*, 75(1), 83–90.
<https://doi.org/10.1128/IAI.01475-06>
- O'Toole, R. D., Goode, L., & Howe, C. (1971). Neuraminidase activity in bacterial meningitis. *The Journal of Clinical Investigation*, 50(5), 979–985.
<https://doi.org/10.1172/JCI106591>
- Oggioni, M. R., Trappetti, C., Kadioglu, A., Cassone, M., Iannelli, F., Ricci, S., ... Pozzi, G. (2006). Switch from planktonic to sessile life: a major event in pneumococcal pathogenesis. *Molecular Microbiology*, 61(5), 1196–1210.

<https://doi.org/10.1111/j.1365-2958.2006.05310.x>

- Ogunniyi, A. D., Giammarinaro, P., & Paton, J. C. (2002). The genes encoding virulence-associated proteins and the capsule of *Streptococcus pneumoniae* are upregulated and differentially expressed *in vivo*. *Microbiology*, 148, 2045–2053. <https://doi.org/https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-148-7-2045>
- Orihuela, C. J., Radin, J. N., Sublett, J. E., Gao, G., Kaushal, D., & Tuomanen, E. I. (2004). Microarray analysis of pneumococcal gene expression during invasive disease. *Infection and Immunity*. <https://doi.org/10.1128/IAI.72.10.5582-5596.2004>
- Paixão, L., Caldas, J., Kloosterman, T. G., Kuipers, O. P., Vinga, S., & Neves, A. R. (2015a). Transcriptional and metabolic effects of glucose on *Streptococcus pneumoniae* sugar metabolism. *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.01041>
- Paixão, L., Oliveira, J., Veríssimo, A., Vinga, S., Lourenço, E. C., Ventura, M. R., ... Neves, A. R. (2015b). Host glycan sugar-specific pathways in *Streptococcus pneumoniae*: galactose as a key sugar in colonisation and infection. *PLoS ONE*, 10(3). <https://doi.org/10.1371/journal.pone.0121042>
- Parker, D., Soong, G., Planet, P., Brower, J., Ratner, A. J., & Prince, A. (2009). The NanA neuraminidase of *Streptococcus pneumoniae* is involved in biofilm formation. *Infection and Immunity*, 77(9), 3730. <https://doi.org/10.1128/IAI.00228-09>
- Pedram, N., Rashedi, H., & Motamedian, E. (2020). A systematic strategy using a reconstructed genome-scale metabolic network for pathogen *Streptococcus pneumoniae* D39 to find novel potential drug targets. *Pathogens and Disease*, 78. <https://doi.org/10.1093/femspd/ftaa051/5900975>
- Perez Diaz, J., Martin Requero, A., Ayuso Parrilla, M. S., & Parrilla, R. (1977). Metabolic features of isolated rat lung cells. I. Factors controlling glucose utilization. *American Journal of Physiology*, 232(4). <https://doi.org/10.1152/ajpendo.1977.232.4.E394>
- Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J. C., & Vorholt, J. A. (2009). Demonstration of the ethylmalonyl-CoA pathway by using ¹³C metabolomics.

- PNAS*, 106(12), 4846–4851.
<https://doi.org/10.1073/pnas.0810932106>
- Pezzulo, A. A., Gutiérrez, J., Duschner, K. S., McConnell, K. S., Taft, P. J., Ernst, S. E., ... Zabner, J. (2011). Glucose depletion in the airway surface liquid is essential for sterility of the airways. *PLoS ONE*, 6(1).
<https://doi.org/10.1371/journal.pone.0016166>
- Philips, B. J., Meguer, J.-X., Redman, J., & Baker, E. H. (2003). Factors determining the appearance of glucose in upper and lower respiratory tract secretions. *Intensive Care Medicine*, 29(12), 2204–2210.
<https://doi.org/10.1007/s00134-003-1961-2>
- Pickett, M. W., Williamson, M. P., & Kelly, D. J. (1994). An enzyme and (13)C-NMR study of carbon metabolism in heliobacteria. *Photosynthesis Research*, 41(1), 75–88. <https://doi.org/10.1007/BF02184147>
- Pitson, S. M., Mendz, G. L., Srinivasan, S., & Hazell, S. L. (1999). The tricarboxylic acid cycle of *Helicobacter pylori*. *European Journal of Biochemistry*, 260(1), 258–267. <https://doi.org/10.1046/j.1432-1327.1999.00153.x>
- Pizer, L. I. (1965). Glycine synthesis and metabolism in *Escherichia coli*. *Journal of Bacteriology*, 89(4), 1145–1150.
<https://doi.org/10.1128/jb.89.4.1145-1150.1965>
- Polissi, A., Pontiggia, A., Feger, G., Altieri, M., Mottl, H., Ferrari, L., & Simon, D. (1998). Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infection and Immunity*.
<https://doi.org/10.1128/iai.66.12.5620-5629.1998>
- Preston, J. A., & Dockrell, D. H. (2008). Virulence factors in pneumococcal respiratory pathogenesis. *Future Microbiology*, 3(2), 205–221.
<https://doi.org/10.2217/17460913.3.2.205>
- Price, C. E., Zeyniyev, A., Kuipers, O. P., & Kok, J. (2012). From meadows to milk to mucosa – adaptation of *Streptococcus* and *Lactococcus* species to their nutritional environments. *FEMS Microbiology Reviews*, 36, 949–971.
<https://doi.org/10.1111/j.1574-6976.2011.00323.x>
- Ram, Y., Dellus-gur, E., Bibi, M., Karkare, K., Obolski, U., Feldman, M. W., ... Hadany, L. (2019). Predicting microbial growth in a mixed culture from growth

- curve data. *PNAS*, 116(29).
<https://doi.org/10.1073/pnas.2010013117>
- Ramos-Montañez, S., Kazmierczak, K. M., Hentchel, K. L., & Winkler, M. E. (2010). Instability of ackA (acetate kinase) mutations and their effects on acetyl phosphate and ATP amounts in *Streptococcus pneumoniae* D39. *Journal of Bacteriology*, 192(24), 6390–6400.
<https://doi.org/10.1128/JB.00995-10>
- Regev-Yochay, G., Raz, M., Dagan, R., Porat, N., Shainberg, B., Pinco, E., ... Rubinstein, E. (2004). Nasopharyngeal carriage of *Streptococcus pneumoniae* by adults and children in community and family settings. *Clinical Infectious Diseases*, 38(5), 9.
<https://doi.org/10.1086/381547>
- Ren, B., Szalai, A. J., Thomas, O., Hollingshead, S. K., & Briles, D. E. (2003). Both family 1 and family 2 PspA proteins can inhibit complement deposition and confer virulence to a capsular serotype 3 strain of *Streptococcus pneumoniae*. *Infection and Immunity*, 71(1), 75–85.
<https://doi.org/10.1128/IAI.71.1.75-85.2003>
- Richardson, A. R., Somerville, G. A., & Sonenshein, A. L. (2015). Regulating the intersection of metabolism and pathogenesis in gram-positive bacteria. *Microbiology Spectrum*, 3.
<https://doi.org/10.1128/microbiolspec.MBP-0004-2014>
- Ridsdale, R., & Post, M. (2004). Surfactant lipid synthesis and lamellar body formation in glycogen-laden type II cells. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 287(4), L743–L751.
<https://doi.org/10.1152/ajplung.00146.2004>
- Robb, M., Hobbs, J. K., Woodiga, S. A., Shapiro-Ward, S., Suits, M. D. L., McGregor, N., ... Boraston, A. B. (2017). Molecular characterization of N-glycan degradation and transport in *Streptococcus pneumoniae* and its contribution to virulence. *PLOS Pathogens*, 13(1).
<https://doi.org/10.1371/journal.ppat.1006090>
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*,

29(1), 24–26.

<https://doi.org/10.1038/nbt.1754>

Rodrigues, F., Foster, D., Nicoli, E., Trotter, C., Vipond, B., Muir, P., ... Finn, A. (2013). Relationships between rhinitis symptoms, respiratory viral infections and nasopharyngeal colonization with *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus* in children attending daycare. *Pediatric Infectious Disease Journal*, 32(3), 227–232.

<https://doi.org/10.1097/INF.0b013e31827687fc>

Rohmer, L., Hocquet, D., & Miller, S. I. (2011). Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. *Trends in Microbiology*, 19(7), 341–348.

<https://doi.org/10.1016/j.tim.2011.04.003>

Rooney, S. A. (2001). Regulation of surfactant secretion. In *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology* (Vol. 129, pp. 233–243).

[https://doi.org/10.1016/S1095-6433\(01\)00320-8](https://doi.org/10.1016/S1095-6433(01)00320-8)

Rose, M. C., & Voynow, J. A. (2006). Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiological Reviews*, 86(1), 245–278.

<https://doi.org/10.1152/physrev.00010.2005>

Rosenberg, G., Yehezkel, D., Hoffman, D., Mattioli, C. C., Fremder, M., Ben-Arosh, H., ... Avraham, R. (2021). Host succinate is an activation signal for *Salmonella* virulence during intracellular infection. *Science*, 371, 400–405.

<https://doi.org/10.1126/science.aba8026>

Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A., & Masure, H. R. (1997). Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Molecular Microbiology*, 25(5), 819–829.

<https://doi.org/10.1111/J.1365-2958.1997.MMI494.X>

Roth, G. A., Abate, D., Hassen Abate, K., Abay, S. M., Abbafati, C., Abbasi, N., ... GBD 2017 Causes of Death Collaborators (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392, 1736–1788.

- [https://doi.org/10.1016/S0140-6736\(18\)32203-7](https://doi.org/10.1016/S0140-6736(18)32203-7)
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., ... Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, 12(1), 87.
<https://doi.org/10.1186/s12915-014-0087-z>
- Salvadori, G., Junges, R., Morrison, D. A., & Petersen, F. C. (2019). Competence in *Streptococcus pneumoniae* and close commensal relatives: Mechanisms and implications. *Frontiers in Cellular and Infection Microbiology*, 9.
<https://doi.org/10.3389/fcimb.2019.00094>
- Sanchez-Rosario, Y., & Johnson, M. D. L. (2021). Media Matters, examining historical and modern *Streptococcus pneumoniae* growth media and the experiments they affect. *Frontiers in Cellular and Infection Microbiology*, 11.
<https://doi.org/10.3389/fcimb.2021.613623>
- Sauer, U. (2006). Metabolic networks in motion: 13C-based flux analysis. *Molecular Systems Biology*.
<https://doi.org/10.1038/msb4100109>
- Schulz, C., Gierok, P., Petruschka, L., Lalk, M., Mäder, U., & Hammerschmidt, S. (2014). Regulation of the arginine deiminase system by ArgR2 interferes with arginine metabolism and fitness of *Streptococcus pneumoniae*. *mBio*, 5(6).
<https://doi.org/10.1128/mBio.01858-14>
- Schulz, C., & Hammerschmidt, S. (2014). Exploitation of physiology and metabolomics to identify pneumococcal vaccine candidates. *Expert Review of Vaccines*, 12(9), 1061–1075. <https://doi.org/10.1586/14760584.2013.824708>
- Shakhnovich, E. A., King, S. J., & Weiser, J. N. (2002). Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of *Neisseria meningitides* and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract. *Infection and Immunity*, 70(12), 7161–7164.
<https://doi.org/10.1128/IAI.70.12.7161-7164.2002>
- Sicard, A. M. (1964). A new synthetic medium for *Diplococcus pneumoniae*, and its use for the study of reciprocal transformations at the *amiA* locus. *Genetics*, 50, 31–44.
- Siegel, S. J., & Weiser, J. N. (2015). Mechanisms of bacterial colonization of the

- respiratory tract. *Annual Review of Microbiology*, 69, 425–444.
<https://doi.org/10.1146/annurev-micro-091014-104209>
- Sjöström, K., Spindler, C., Ortqvist, A., Kalin, M., Sandgren, A., Kühlmann-Berenzon, S., & Henriques-Normark, B. (2006). Clonal and capsular types decide whether pneumococci will act as a primary or opportunistic pathogen. *Clinical Infectious Diseases*, 42(4), 9.
<https://doi.org/10.1086/499242>
- Smith, A. W., Roche, H., Trombe, M.-C., Briles, D. E., & Håkansson, A. (2002). Characterization of the dihydrolipoamide dehydrogenase from *Streptococcus pneumoniae* and its role in pneumococcal infection. *Molecular Microbiology*, 44(2), 431–448.
<https://doi.org/10.1046/j.1365-2958.2002.02883.x>
- Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., ... Siuzdak, G. (2005). METLIN: A metabolite mass spectral database. *Therapeutic Drug Monitoring*, 27(6), 747–751.
<https://doi.org/10.1097/01.ftd.0000179845.53213.39>
- Spellerberg, B., Cundell, D. R., Sandros, J., Pearce, B. J., Idänpään-Heikkilä, I., Rosenow, C., & Masure, H. R. (1996). Pyruvate oxidase, as a determinant of virulence in *Streptococcus pneumoniae*. *Molecular Microbiology*, 19(4), 803–813.
<https://doi.org/10.1046/j.1365-2958.1996.425954.x>
- Stahl, W. L., & O'Toole, R. D. (1972). Pneumococcal neuraminidase: purification and properties. *Biochimica et Biophysica Acta*, 268(2), 480–487.
[https://doi.org/10.1016/0005-2744\(72\)90343-9](https://doi.org/10.1016/0005-2744(72)90343-9)
- Steffens, L., Pettinato, E., Steiner, T. M., Mall, A., König, S., Eisenreich, W., & Berg, I. A. (2021). High CO₂ levels drive the TCA cycle backwards towards autotrophy. *Nature*, 592.
<https://doi.org/10.1038/s41586-021-03456-9>
- Subramanian, K., Henriques-Normark, B., & Normark, S. (2019). Emerging concepts in the pathogenesis of the *Streptococcus pneumoniae*: From nasopharyngeal colonizer to intracellular pathogen. *Cellular Microbiology*, 21(11).
<https://doi.org/10.1111/cmi.13077>
- Subramanian, K., Neill, D. R., Malak, H. A., Spelmink, L., Khandaker, S., Dalla Libera

- Marchiori, G., ... Henriques-Normark, B. (2019). Pneumolysin binds to the mannose receptor C type 1 (MRC-1) leading to anti-inflammatory responses and enhanced pneumococcal survival. *Nature Microbiology*, 4(1), 62–70.
<https://doi.org/10.1038/s41564-018-0280-x>
- Tang, J. (2011). Microbial metabolomics. *Current Genomics*, 12(6), 391–403.
<https://doi.org/10.2174/138920211797248619>
- Taniai, H., Iida, K. I., Seki, M., Saito, M., Shiota, S., Nakayama, H., & Yoshida, S. I. (2008). Concerted action of lactate oxidase and pyruvate oxidase in aerobic growth of *Streptococcus pneumoniae*: Role of lactate as an energy source. *Journal of Bacteriology*, 190(10), 3572–3579.
<https://doi.org/10.1128/JB.01882-07>
- Teran-Navarro, H., Salcines-Cuevas, D., Calderon-Gonzalez, R., Tobes, R., Calvo-Montes, J., Pérez-Del Molino Bernal, I. C., ... Alvarez-Dominguez, C. (2021). A comparison between recombinant *Listeria* GAPDH proteins and GAPDH encoding mRNA conjugated to lipids as cross-reactive vaccines for *Listeria*, *Mycobacterium*, and *Streptococcus*. *Frontiers in Immunology*, 12.
<https://doi.org/10.3389/fimmu.2021.632304>
- Terra, V. S., Homer, K. A., Rao, S. G., Andrew, P. W., & Yesilkaya, H. (2010). Characterization of novel beta-galactosidase activity that contributes to glycoprotein degradation and virulence in *Streptococcus pneumoniae*. *Infection and Immunity*, 78, 348–357.
<https://doi.org/10.1128/IAI.00721-09>
- Terra, V. S., Zhi, X., Kahya, H. F., Andrew, P. W., & Yesilkaya, H. (2016). Pneumococcal 6-Phospho- β -Glucosidase (BglA3) is involved in virulence and nutrient metabolism. *Infection and Immunity*, 84(1), 286–292.
<https://doi.org/10.1128/IAI.01108-15>
- Terrasse, R., Tacnet-Delorme, P., Moriscot, C., Pérard, J., Schoehn, G., Vernet, T., ... Frachet, P. (2012). Human and pneumococcal cell surface glyceraldehyde-3-phosphate dehydrogenase (GAPDH) proteins are both ligands of human C1q protein. *Journal of Biological Chemistry*, 287(51), 42620–42633.
<https://doi.org/10.1074/jbc.M112.423731>
- Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., ...

- Fraser, C. M. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science*, 293(5529), 498–506.
<https://doi.org/10.1126/science.1061217>
- The Global Pneumococcal Sequencing Project. (n.d.).
<https://www.pneumogen.net/gps/>
- Tian, J., Bryk, R., Itoh, M., Suematsu, M., & Nathan, C. (2005). Variant tricarboxylic acid cycle in *Mycobacterium tuberculosis*: Identification of α -ketoglutarate decarboxylase. *PNAS*, 102(30), 10670–10675.
<https://doi.org/10.1073/pnas.0501605102>
- Titgemeyer, F., & Hillen, W. (2002). Global control of sugar metabolism: A Gram-positive solution. *Antonie van Leeuwenhoek*, 82, 59–71.
<https://doi.org/10.1023/A:1020628909429>
- Todd, E. W., & Hewitt, L. F. (1932). A new culture medium for the production of antigenic streptococcal haemolysin. *The Journal of Pathology and Bacteriology*, 35(6), 74. <https://doi.org/10.1002/path.1700350614>
- Tong, H. H., Blue, L. E., James, M. A., & Demaria, T. F. (2000). Evaluation of the virulence of a *Streptococcus pneumoniae* neuraminidase-deficient mutant in nasopharyngeal colonisation and development of otitis media in the chinchilla model. *Infection and Immunity*, 68(2), 921–924.
<https://doi.org/10.1128/iai.68.2.921-924.2000>
- Tong, H. H., Liu, X., Chen, Y., James, M., & Demaria, T. (2002). Effect of neuraminidase on receptor-mediated adherence of *Streptococcus pneumoniae* to chinchilla tracheal epithelium. *Acta Oto-Laryngologica*, 122(4), 413–419.
<https://doi.org/10.1080/00016480260000111>
- Tounta, V., Liu, Y., Cheyne, A., & Larrouy-Maumus, G. (2021). Metabolomics in infectious diseases and drug discovery. *Molecular Omics*, 17(3).
<https://doi.org/10.1039/d1mo00017a>
- Trappetti, C., McAllister, L. J., Chen, A., Wang, H., Paton, A. W., Oggioni, M. R., ... Paton, J. C. (2017). Autoinducer 2 signaling via the phosphotransferase FruA drives galactose utilization by *Streptococcus pneumoniae*, resulting in hypervirulence. *mBio*, 8(1).
<https://doi.org/10.1128/mBio.02269-16>

- Trappetti, C., Potter, A. J., Paton, A. W., Oggioni, M. R., & Paton, J. C. (2011). LuxS mediates iron-dependent biofilm formation, competence, and fratricide in *Streptococcus pneumoniae*. *Infection and Immunity*, 79(11), 4550–4558.
<https://doi.org/10.1128/IAI.05644-11>
- Trappetti, C., van der Maten, E., Amin, Z., Potter, A. J., Chen, A. Y., van Mourik, P. M., ... Paton, J. C. (2013). Site of isolation determines biofilm formation and virulence phenotypes of *Streptococcus pneumoniae* serotype 3 clinical isolates. *Infection and Immunity*, 81(2), 505–513.
<https://doi.org/10.1128/IAI.01033-12>
- Tretter, L., Patocs, A., & Chinopoulos, C. (2016). Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis. *Biochimica et Biophysica Acta - Bioenergetics*, 1857(8), 1086–1101.
<https://doi.org/10.1016/j.bbabo.2016.03.012>
- Trimble, A., Connor, V., Robinson, R. E., McLenaghan, D., Hancock, C. A., Wang, D., ... Collins, A. M. (2020). Pneumococcal colonisation is an asymptomatic event in healthy adults using an experimental human colonisation model. *PLoS ONE*, 15(3).
<https://doi.org/10.1371/journal.pone.0229558>
- Troxler, L. J., Werren, J. P., Schaffner, T. O., Mostacci, N., Vermathen, P., Vermathen, M., ... Hilty, M. (2019). Carbon source regulates polysaccharide capsule biosynthesis in *Streptococcus pneumoniae*. *Journal of Biological Chemistry*, 294(46). <https://doi.org/10.1074/jbc.RA119.010764>
- Trzciński, K., Thompson, C. M., & Lipsitch, M. (2003). Construction of otherwise isogenic serotype 6B, 7F, 14, and 19F capsular variants of *Streptococcus pneumoniae* strain TIGR4. *Applied and Environmental Microbiology*, 69(12), 7364–7370.
<https://doi.org/10.1128/AEM.69.12.7364-7370.2003>
- Tuomanen, E. I. (2021). Pneumococcal vaccination in children - UpToDate.
Retrieved July 4, 2021, from https://www.uptodate.com/contents/pneumococcal-vaccination-in-children?topicRef=7021&source=see_link
- Usuf, E., Bottomley, C., Adegbola, R. A., & Hall, A. (2014). Pneumococcal carriage in sub-Saharan Africa—A systematic review. *PLoS ONE*, 9(1).
<https://doi.org/10.1371/journal.pone.0085001>

- Vadlamudi, N. K., Chen, A., & Marra, F. (2019). Impact of the 13-valent pneumococcal conjugate vaccine among adults: A systematic review and meta-analysis. *Clinical Infectious Diseases*, 69(1), 34–49.
<https://doi.org/10.1093/cid/ciy872>
- van Beelen, P., van der Hoeven, J. S., de Jong, M. H., & Hoogendoorn, H. (1986). The effect of oxygen on the growth and acid production of *Streptococcus mutans* and *Streptococcus sanguis*. *FEMS Microbiology Ecology*, 38, 25–30.
- van Bueren, A. L., Higgins, M., Wang, D., Burke, R. D., & Boraston, A. B. (2007). Identification and structural basis of binding to host lung glycogen by streptococcal virulence factors. *Nature Structural and Molecular Biology*, 14(1), 76–84. <https://doi.org/10.1038/nsmb1187>
- van de Rijn, I., & Kessler, R. E. (1980). Growth characteristics of group A streptococci in a new chemically defined medium. *Infection and Immunity*, 27(2), 444–448.
<https://doi.org/10.1128/iai.27.2.444-448.1980>
- van Opijnen, T., Bodi, K. L., & Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature Methods*, 6(10), 767–772.
<https://doi.org/10.1038/nmeth.1377>
- Vernocchi, P., Del Chierico, F., & Putignani, L. (2016). Gut microbiota profiling: Metabolomics based approach to unravel compounds affecting human health. *Frontiers in Microbiology*.
<https://doi.org/10.3389/fmicb.2016.01144>
- Vidal, L., Calveras, J., Clapés, P., Ferrer, P., & Caminal, G. (2005). Recombinant production of serine hydroxymethyl transferase from *Streptococcus thermophilus* and its preliminary evaluation as a biocatalyst. *Applied Microbiology and Biotechnology*, 68(4), 489–497.
<https://doi.org/10.1007/s00253-005-1934-1>
- Wahl, B., O'Brien, K. L., Greenbaum, A., Majumder, A., Liu, L., Chu, Y., ... Knoll, M. D. (2018). Burden of *Streptococcus pneumoniae* and *Haemophilus influenzae* type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000–15. *The Lancet Global Health*, 6(7).

- [https://doi.org/10.1016/S2214-109X\(18\)30247-X](https://doi.org/10.1016/S2214-109X(18)30247-X)
- Watkins, E. R., Penman, B. S., Lourenço, J., Buckee, C. O., Maiden, M. C. J., & Gupta, S. (2015). Vaccination drives changes in metabolic and virulence profiles of *Streptococcus pneumoniae*. *PLoS Pathogens*, 11(7).
<https://doi.org/10.1371/journal.ppat.1005034>
- Watson, D. A., & Musher, D. M. (1990). Interruption of capsule production in *Streptococcus pneumoniae* serotype 3 by insertion of transposon Tn916. *Infection and Immunity*, 58(9), 8.
<https://doi.org/10.1128/iai.58.9.3135-3138.1990>
- Weight, C. M., Venturini, C., Pojar, S., Jochems, S. P., Reiné, J., Nikolaou, E., ... Heyderman, R. S. (2019). Microinvasion by *Streptococcus pneumoniae* induces epithelial innate immunity during colonisation at the human mucosal surface. *Nature Communications*, 10.
<https://doi.org/10.1038/s41467-019-11005-2>
- Weinberger, D. M., Malley, R., & Lipsitch, M. (2011). Serotype replacement in disease after pneumococcal vaccination. *The Lancet*, 378(9807), 1962–1973.
[https://doi.org/10.1016/S0140-6736\(10\)62225-8](https://doi.org/10.1016/S0140-6736(10)62225-8)
- Weinberger, D. M., Trzciński, K., Lu, Y.-J., Bogaert, D., Brandes, A., Galagan, J., ... Lipsitch, M. (2009). Pneumococcal capsular polysaccharide structure predicts serotype prevalence. *PLoS Pathogens*, 5(6).
<https://doi.org/10.1371/journal.ppat.1000476>
- Weiser, J. N., Bae, D., Epino, H., Gordon, S. B., Kapoor, M., Zenewicz, L. A., & Shchepetov, M. (2001). Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. *Infection and Immunity*, 69(9), 5430–5439.
<https://doi.org/10.1128/IAI.69.9.5430-5439.2001>
- Weiser, Jeffrey N. (2010). The pneumococcus: why a commensal misbehaves. *Journal of Molecular Medicine*, 88(2), 97–102.
<https://doi.org/10.1007/s00109-009-0557-x>
- Werren, J. P., Troxler, L. J., Oyewole, O. R. A., Ramette, A., Brugger, S. D., Bruggmann, R., ... Hilty, M. (2021). Carbon source-dependent changes of the structure of *Streptococcus pneumoniae* capsular polysaccharide with serotype 6F.

- International Journal of Molecular Sciences*, 22(9).
<https://doi.org/10.3390/ijms22094580>
- WHO. (2020). Children: improving survival and well-being.
<https://www.who.int/news-room/fact-sheets/detail/children-reducing-mortality>
- WHO. (2021). *Global Priority List of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics*.
<http://www.cdc.gov/drugresistance/threat-report-2013/>
- Willenborg, J., & Goethe, R. (2016). Metabolic traits of pathogenic streptococci. *FEBS Letters*, 590, 3905–3919.
<https://doi.org/10.1002/1873-3468.12317>
- Wilson, R., Cohen, J. M., Jose, R. J., de Vogel, C., Baxendale, H., & Brown, J. S. (2015). Protection against *Streptococcus pneumoniae* lung infection after nasopharyngeal colonization requires both humoral and cellular immune responses. *Mucosal Immunology*, 8(3), 627–639.
<https://doi.org/10.1038/mi.2014.95>
- Wong, H. S., Maker, G. L., Trengove, R. D., & O’Handley, R. M. (2015). Gas chromatography-mass spectrometry-based metabolite profiling of *Salmonella enterica* serovar typhimurium differentiates between biofilm and planktonic phenotypes. *Applied and Environmental Microbiology*, 81(8), 2660–2666.
<https://doi.org/10.1128/AEM.03658-14>
- Wood, D. E., & Salzberg, S. L. (2014). Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, 15(3), R46.
<https://doi.org/10.1186/gb-2014-15-3-r46>
- Yadav, M. K., Vidal, J. E., Go, Y. Y., Kim, S. H., Chae, S.-W., & Song, J.-J. (2018). The LuxS/AI-2 quorum-sensing system of *Streptococcus pneumoniae* is required to cause disease, and to regulate virulence- and metabolism-related genes in a rat model of middle ear infection. *Frontiers in Cellular and Infection Microbiology*, 8.
<https://doi.org/10.3389/fcimb.2018.00138>
- Yesilkaya, H., Manco, S., Kadioglu, A., Terra, V. S., & Andrew, P. W. (2008). The ability to utilize mucin affects the regulation of virulence gene expression in *Streptococcus pneumoniae*. *FEMS Microbiology Letters*, 278, 231–235.
<https://doi.org/10.1111/j.1574-6968.2007.01003.x>

- Yesilkaya, H., Spissu, F., Carvalho, S. M., Terra, V. S., Homer, K. A., Benisty, R., ... Andrew, P. W. (2009). Pyruvate formate lyase is required for pneumococcal fermentative metabolism and virulence. *Infection and Immunity*, 77(12), 5418–5427. <https://doi.org/10.1128/IAI.00178-09>
- Yother, J. (2011). Capsules of *Streptococcus pneumoniae* and other bacteria: Paradigms for polysaccharide biosynthesis and regulation. *Annual Review of Microbiology*, 65, 563–581. <https://doi.org/10.1146/annurev.micro.62.081307.162944>
- Zakrzewicz, D., Bergmann, S., Didiasova, M., Giaimo, B. D., Borggreffe, T., Mieth, M., ... Wygrecka, M. (2016). Host-derived extracellular RNA promotes adhesion of *Streptococcus pneumoniae* to endothelial and epithelial cells. *Scientific Reports*, 6. <https://doi.org/10.1038/srep37758>
- Zhang, J.-R., Mostov, K. E., Lamm, M. E., Nanno, M., Shimida, S., Ohwaki, M., & Tuomanen, E. (2000). The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells. *Cell*, 102(6), 827–837. [https://doi.org/10.1016/S0092-8674\(00\)00071-4](https://doi.org/10.1016/S0092-8674(00)00071-4)
- Zhang, Y. Y., Ji, J., Lan, M. Z., Wang, T. W., & Sun, X. L. (2020). Development of a non-targeted high-coverage microbial metabolomics pretreatment method and its application to drug resistant *Salmonella*. *Analytical Methods*, 12(11), 1449–1459. <https://doi.org/10.1039/c9ay02579k>

Conference Abstracts

1. **L. Mane**, N. Legrave, C. Weight, J. I. MacRae, R.J. Wilkinson, R.S. Heyderman.
Metabolomic analyses of *Streptococcus pneumoniae* to understand niche adaptation. Gordon Research Conference Streptococcal Biology, August 2018.
2. **L. Mane**, N. Legrave, C. Weight, J. I. MacRae, R.J. Wilkinson, R.S. Heyderman.
Metabolomic analyses of *Streptococcus pneumoniae* reveal novel metabolic routes that may facilitate niche adaptation. 14th European Meeting on the Molecular Biology of the Pneumococcus, June 2019.
3. **L. Mane**, N. Legrave, C. Weight, J. I. MacRae, R.J. Wilkinson, R.S. Heyderman.
Nutritional and environmental regulation of pneumococcal central carbon metabolism at the commensal-pathogen interface. 12th International Symposium on Pneumococci and Pneumococcal Disease (ISPPD-12), accepted for 2020 but postponed to 2022 due to the COVID-19 pandemic.

*The road to wisdom? — Well, it's plain
and simple to express:*

Err

and err

and err again,

but less

and less

and less.

— Piet Hein; poet, designer, inventor, and scientist