Elucidating Flux Regulation of the Fermentation Modes of Lactococcus lactis: A Multilevel Study

Chan, Siu Hung Joshua; Solem, Christian; Jensen, Peter Ruhdal

Link to article, DOI: 10.11581/DTU:00000016

Publication date: 2014

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Elucidating Flux Regulation of the Fermentation Modes of

*Lactococcus lactis*: A Multilevel Study

PhD Thesis

Siu Hung Joshua Chan

November, 2014

Supervisors:

Associate Professor Christian Solem

Professor Peter Ruudal Jensen

Systems Biotechnology and Biorefining

National Food Institute

Technical University of Denmark
Summary

The long history of application to the dairy industry has established *Lactococcus lactis* (*L. lactis*), the lactic acid bacterium, as one of the most extensively characterized low GC organisms. The relatively simple metabolism of *L. lactis* has also made it an attractive target for metabolic engineering for the production of non-food related chemicals. Moreover, the status of being the first genetically modified organism to deliver immunoproteins alive to human has brought *L. lactis* considerable fame in biomedical research.

Beside the exceptional industrial relevance of *L. lactis*, it is also an important subject for basic research in cellular metabolism because *L. lactis* exhibits an interesting metabolic shift. Under anaerobic conditions, on fast fermentable sugars, *L. lactis* produces lactate as the primary product, known as homolactic fermentation but on slowly fermentable sugars, significant amounts of formate, acetate and ethanol are formed, known as mixed-acid fermentation. This shift is termed the mixed-acid shift. This type of shift between a low-yield and a high-yield metabolism has drawn a lot of research focus and has similarly been observed in other bacteria, yeast and even tumor cells.

Efforts have been put to find out the mechanism regulating the mixed-acid shift as well as to answer questions such as why *L. lactis* prefers such a switch. Until now, some pieces of evidence have been reported and several factors and models have been proposed as the keys to regulating the shift, including the expression level of certain genes in glycolysis and fermentation pathways, the levels of the cofactors NADH, NAD\(^+\), ATP and ADP, the balance between catabolism and anabolism, etc.

In this project, we studied the mixed-acid fermentation of *L. lactis* by (i) examining the roles of the enzymes in the mixed-acid fermentation pathway under different growth conditions; (ii) testing the predicted effect of the cofactors NADH, NAD\(^+\) on the mixed-acid shift proposed in previous studies; (iii) looking into the connection between amino acid metabolism and the mixed-acid shift; and (iv) contrasting the difference regarding the mixed-acid shift between two widely studied laboratory strains of *L. lactis*, MG1363 that shifts significantly and IL1403 that does not shift.

We have measured the promoter activities of several mixed-acid genes which suggested that the regulatory elements governing the transcriptional regulation of the mixed-acid genes in MG1363 and IL1403 were different. This led us to performing experimental control analysis of the role of pyruvate formate-lyase (PFL) in MG1363 and IL1403. The expression of PFL in MG1363 appeared to be optimized for growth rate when growing on maltose whereas overexpressing PFL in IL1403 was probably detrimental.
The two homologous acetate kinases in MG1363 were also characterized with respect to the transcription and enzyme activities. The isozymes were found to have complementary physiological roles that became important in acetate-producing or acetate-assimilating conditions respectively.

The proposed roles of NADH and NAD$^+$ on the mixed-acid shift were tested by perturbation via introducing activities of 2,3-butanediol dehydrogenase and supplying extracellular acetoin as an oxidizing agent. The additional NAD$^+$-regenerating activities allowed a faster growth of MG1363 on maltose by shifting ethanol production into acetate production and also stimulated formate and acetate production in IL1403.

Dependance of the mixed-acid fermentation of MG1363 on amino acid availability was observed and the impact of individual amino acids could differ significantly. Meanwhile, a computational method for combining metabolic flux analysis and elementary mode analysis was developed and applied to analyse a case of amino acid metabolism of _L. lactis_.

iii
Dansk resumé

De mange års anvendelse af mælkesyrebakterien *Lactococcus lactis* (*L. lactis*) indenfor mejerindustrien, har været medvirkende til at *L. lactis* er blevet en af de mest velkarakteriserede bakterier. Denne Gram positive bakterie, som har et lavt GC indhold, har en relativt simpel metabolisme og er let at modificere genetisk. Dette har gjort den til et attraktivt mål for "metabolic engineering", bl.a. med henblik på produktion af non-food relaterede kemikalier. Derudover har den status som den første genetisk modificerede organismer der i levende form er blevet anvendt til at levere immunproteiner til mennesker, og *L. lactis* er velkendt inden for biomedicinsk forskning.

Ud over de vigtige roller som er blevet nævnt ovenfor, er *L. lactis* det også vigtig for grundforskning i cellulær metabolisme, bl.a. fordi *L. lactis* udfører et interessant metabolisk skift, som bevirker en ændring i produktdannelse. Let fermenterbare sukre giver anledning til hovedsagelig laktat (homolactic), medens sukre der omsættes langsommere giver anledning til andre organiske syrer og ethanol i tillæg til laktat (mixed-acid). Denne form for skift mellem en metabolisme med et lavt og et højt udbytte er der blevet forsket meget i, ikke kun i *L. lactis* men også i andre bakterier, gær og endda cancerceller.

Indtil nu har mange beviser blevet rapporteret, og flere faktorer og modeller er blevet foreslået som nøglerne til regulering af skift, herunder udtrykket af visse gener i glycolyse og fermentering veje, niveauet for cofaktorer NADH, NAD⁺, ATP og ADP, balancen mellem katabolisme og anabolisme, etc.

I dette projekt, vi studerede blandet syre fermentering af *L. lactis* ved (i) at undersøge de roller enzymer i blandet syre gæring vej under forskellige vækstbetingelser; (ii) testning af forudsagte virkning af cofaktorer NADH, NAD⁺ på blandede syre skift foreslået i tidligere undersøgelser; (iii) at se på sammenhængen mellem aminosyre metabolisme og det blandede syre skift; og (iv) kontrasterende forskellen med hensyn til blandet syre skift mellem to vidt studerede laboratoriestammer, *L. lactis* subsp. *cremoris* MG1363, der skifter betydeligt og *L. lactis* subsp. *lactis* IL1403, der ikke skifter.

Vi har målt promotoraktiviteter i flere blandet syre gener, som antydede, at de regulatoriske elementen for transkriptionel regulering af det blandede syre gener i MG1363 og IL1403 var anderledes. Dette førte os til at udføre eksperimentel kontrol analyse af den rolle, pyruvat format lyase (PFL) i MG1363 og IL1403. Ekspressionen af PFL i MG1363 syntes at være optimeret til vækst, når de vokser på maltose henviser overekspression PFL i IL1403 var sandsynligvis skadelige.
De to homologe acetat kinaser i MG1363 blev også characterized med hensyn til transkription og enzymaktiviteter. De isozymer viste sig at have komplementære fysiologiske roller, som fik betydning i acetat-producerende eller acetat-assimilere betingelser hhv.

De foreslåede roller af NADH og NAD⁺ på blandet syre skift blev testet ved perturbation via indføre aktiviteter af 2,3-butandiol dehydrogenase og leverer ekstracellulære acetoin som et oxidationsmiddel. De yderligere NAD⁺-regenererende aktiviteter have en hurtigere vækst i MG1363 på maltose ved at flytte ethanolproduktion i acetat produktion og også stimuleret formiat og acetat produktion i IL1403.

Acknowledgements

It is impossible to complete my PhD study without the help of many people. First of all, I must express my sincerest gratitude to my supervisors, Christian Solem and Peter Ruhdal Jensen.

Christian has taught me numerous principles, skills and knowledge in practical as well as theoretical molecular biology and cellular metabolism. From the very first day he has helped me, a person without any previous experience in biological experiments, to survive in the lab. He always came up with pragmatic solutions as well as interesting ideas that led to new directions. The most important thing of all that I have learnt from him, for a skeptical person like me, is his will power and faith to find out the truth through continuous efforts, as reflected by his passionate encouragement that usually put an end to our discussion: ‘Just do it!’ . It has powered me to continue the journey from time to time.

Peter has also been equally important to my PhD study who taught me a lot of knowledge in bacterial metabolism and how elegant it is. More importantly, I have learnt a lot from the perspective, insights and discernment he has demonstrated as a world-leading scientist when looking into biological problems. The close guidance, support and meanwhile flexibility and freedom he has given have allowed me to gradually become independent and better in productivity. In addition, the direction, spirit, morale and working environments he has built for the whole research group has propelled every one of us to move forward in a supportive and positive manner.

I must also thank people who have been working with me and helped me. Thank Jun Chen who from the very first day of my arrival in Copenhagen has helped me to settle my life here, provided countless useful tips and advices in my experiments and shared his extensive knowledge on *Lactococcus lactis*. This greatly accelerated my progress. Thank also Xiaoying Liu for her support and care in my beginning stage when she was in the group. Thank Lasse Nørregaard and Aria Aminzadeh for the good time that we worked together in their bachelor theses. Thank all the colleagues that have once helped me, including Anders Cai Holm Hansen and Tina Suhr for their help in using lab facilities from time to time; Kia Vest Petersen who always maintained a good working environment and taught me some good practices in the lab; professor Ole Michelsen who is very humorous and has shared some of his unimaginably vast knowledge with me; and many others who have been so nice to me.

Special thanks are given to Zhihao Wang who always has a lot of scientific and insightful discussion with me. He has personally become a good friend of mine from whom I got countless heartfelt physical and mental support and from whom I learnt a lot in improving my imperfect personality. I would also like to acknowledge the Danish Council for Independent Research in Natural Sciences for the financial support, Technical University of Denmark for the great research environment and the transnational SysMOLAB2 consortium for a fruitful discussion.

For my personal life in Denmark for these several years, I would like to thank my family and many friends in Hong Kong as well as in Denmark who show their love and care constantly which are vital for me to face the live and challenges in Denmark.

A person of paramount importance to me who I must thank most deeply is my wife, Sandy Wong, who has accompanied me to have a new life here and has always supported me and helped me in every aspect of my life. She deserves all the compliments that I receive (if any).

Finally, thank God whom I believe in. He made a way in the waste land, and rivers in the dry country.
Outline of the thesis

The thesis comprises seven chapters and three appendices.

Chapter 1 introduces the background of this study. Beginning with a general review of the industrial relevance, characteristics and available tools for research and engineering of lactic acid bacteria, the focus then turns to the metabolism of \textit{L. lactis}, especially the regulation of the glycolytic flux and the shift from homolactic to mixed-acid fermentation. The research objectives and strategies are defined at the end of the chapter.

Chapter 2 includes the growth characterization of the two widely studied laboratory strains, MG1363 and IL1403 as well as the transcriptional activity of the enzymes in the mixed-acid fermentation pathways in the two strains.

Chapter 3 reports the findings of the modulation of the expression of pyruvate formate-lyase and the control analysis of it in MG1363.

Chapter 4 presents the characterization of two genes encoding acetate kinases in MG1363 with respect to the transcription structure, enzyme kinetics and their physiological roles.

Chapter 5 describes the study of the effect of perturbing NADH/NAD⁺ for MG1363 and IL1403 growing on maltose or glucose respectively.

Chapter 6 briefly summarizes some results of a bachelor project initiated by this project regarding the effect of amino acid availability on the mixed-acid shift and at the same time reports some observations from a computational study in which a new method to analyse flux distributions by elementary modes was proposed. The derivation, validation and application of the method were presented in detail in the publication reprinted in Appendix A.

Chapter 7 concludes the results of the entire project and discusses some future directions.

Supplementary information for Chapter 4 and Chapter 5 are attached in Appendix B and Appendix C respectively.
# Table of Contents

Summary ............................................................................................................................ ii
Dansk resumé ...................................................................................................................... iv
Acknowledgements ........................................................................................................... vi
Outline of the thesis ......................................................................................................... vii
Table of Contents .............................................................................................................. viii
List of Figures ...................................................................................................................... xiii
List of Tables ..................................................................................................................... xv

## Chapter 1. Background ................................................................................................. 1

1.1 Lactic acid bacteria ................................................................................................... 1
1.2 Industrial relevance of lactic acid bacteria ............................................................... 2
  1.2.1 Starter cultures .................................................................................................... 2
  1.2.2 Probiotics .......................................................................................................... 2
  1.2.3 Cell factories .................................................................................................... 2
1.3 Some characteristics of lactic acid bacteria .............................................................. 3
  1.3.1 Genomics .......................................................................................................... 3
  1.3.2 Metabolism ...................................................................................................... 4
  1.3.3 Stress tolerance ............................................................................................... 8
1.4 Tools for studying lactic acid bacteria ..................................................................... 8
  1.4.1 Molecular tools ............................................................................................... 8
  1.4.2 Tools and computation models for systems biology .......................................... 11
1.5 *Lactococcus lactis* ................................................................................................... 14
  1.5.1 Glycolysis ........................................................................................................ 15
  1.5.2 Control of the glycolytic flux .......................................................................... 17
  1.5.3 Pyruvate metabolism .................................................................................... 23
  1.5.4 Flux regulation of mixed-acid shift ................................................................. 27
1.6 Objectives and strategies ......................................................................................... 33
  1.6.1 Genetic perturbation ....................................................................................... 33
  1.6.2 Cofactor perturbation .................................................................................... 33
  1.6.3 Nutrient availability ....................................................................................... 34
1.7 References ................................................................................................................. 34

## Chapter 2. Transcriptional activity of mixed-acid genes ............................................. 46

2.1 Introduction ............................................................................................................... 46
2.2 Materials and methods ............................................................................................ 46
  2.2.1 Bacteria strains and plasmids .......................................................................... 46
  2.2.2 Antibiotics ...................................................................................................... 48
  2.2.3 Culture media and conditions ....................................................................... 48
Chapter 3. Mixed-acid fermentation controlled by pyruvate-formate lyase .......................................................... 58

3.1 Introduction ..................................................................................... 58
3.2 Materials and methods ....................................................................... 58
  3.2.1 Bacteria strains and plasmids ............................................................ 58
  3.2.2 Antibiotics .................................................................................... 60
  3.2.3 Gene deletion and insertion .............................................................. 60
  3.2.4 Modulation of pfl expression in L. lactis ........................................ 60
  3.2.5 Others ......................................................................................... 61
3.3 Results .............................................................................................. 61
  3.3.1 PFLs of MG1363 and IL1403 expressed in MG1363Δpfl ................. 61
  3.3.2 Flux control by pfl expression level .................................................. 63
  3.3.3 Yield on sugars .............................................................................. 68
  3.3.4 Loss of formate-producing ability during the modulation of pfl in IL1403 ................................................................. 69
3.4 Discussion ......................................................................................... 70
  3.4.1 Similar behaviour of MG-PFL and IL-PFL when expressed in MG1363 ................................................................. 70
  3.4.2 Flux control by PFL in MG1363 ....................................................... 70
  3.4.3 Rate or yield optimization ................................................................. 71
  3.4.4 Comparison to previous results ...................................................... 71
  3.4.5 Unsuccessful modulation of pfl in IL1403 ....................................... 72
  3.4.6 Possible reasons for the low formate production in IL1403 ............. 72
3.5 Acknowledgement ............................................................................. 73
3.6 References ....................................................................................... 73

Chapter 4. Acetate kinase isozymes confer robustness in acetate metabolism ................................................................. 75

4.1 Abstract .......................................................................................... 75
4.2 Introduction ..................................................................................... 75
4.3 Materials and methods ....................................................................... 77
  4.3.1 Bacterial strains and plasmids ............................................................ 77
  4.3.2 Antibiotics .................................................................................... 77
  4.3.3 Culture media and growth conditions ............................................. 77
  4.3.4 Quantification of maltose and fermentation products ...................... 78
  4.3.5 DNA techniques .......................................................................... 78
4.3.6 Gene inactivation
4.3.7 Construction of gusA reporter strains
4.3.8 Rapid amplification of cDNA ends (RACE)
4.3.9 Overproduction of L. lactis ACK in E. coli
4.3.10 Measurement of ACK activities
4.3.11 Measurement of β-glucuronidase activity
4.3.12 Sequence analysis

4.4 Results
4.4.1 Homologous sequences of AckA1 and AckA2 in L. lactis MG1363
4.4.2 Multiple ackA genes existing in L. lactis and many other species
4.4.3 Two types of ACK conserved in Lactococcus
4.4.4 Predicted transcription terminator between the two ACK genes in L. lactis
4.4.5 Distinct transcription start sites for ackA1 and ackA2
4.4.6 Distinct transcription units and activities
4.4.7 Huge differences in $k_{\text{cat}}$ and $K_m$ for acetate
4.4.8 Additive cell extract activities of mutant strains
4.4.9 Slower acetate production by MG1363ΔackA1 at a high extracellular acetate concentration
4.4.10 Slower acetate uptake by MG1363ΔackA2Δpfl at low acetate concentrations

4.5 Discussion
4.5.1 Prevalence of multiple ACKs in bacteria
4.5.2 ackA1 and ackA2 in L. lactis probably resulted from gene duplication
4.5.3 AckA1 and AckA2 in MG1363 being isozymes rather than subunits
4.5.4 Possible different roles suggested by enzyme kinetics
4.5.5 Physiological roles of ACK reported in literature
4.5.6 Growth on maltose as a test of the physiological roles in L. lactis
4.5.7 Complementary roles in acetate metabolism
4.5.8 PTA-ACK as the only pathway interconverting Ac-CoA and acetate in L. lactis

4.6 Conclusions
4.7 Acknowledgments
4.8 Author contributions
4.9 References

Chapter 5. Influence of NADH/NAD$^+$ on mixed-acid fermentation

5.1 Introduction
5.2 Materials and methods
5.2.1 Bacteria strains and plasmids
5.2.2 Culture media and conditions
5.2.3 Quantification of sugar and fermentation products
5.2.4 Extraction and quantification of NADH and NAD$^+$
5.2.5 DNA techniques
5.2.6 Measurement of 2,3-butenediol dehydrogenase activity

5.3 Results
5.3.1 Introduction of 2,3-butenediol dehydrogenase activity
5.3.2 Faster growth of MG/SP1-bdh in the presence of acetoin in the media
5.3.3 Flux directed towards acetate production
5.3.4 No change in growth rate in media devoid of lipoic acid
5.3.5 Unexpected lower carbon recovery in MG/SP1-bdh at ≥ 10mM acetoin ........................................ 100
5.3.6 Growth of MG1363, MG/SP2-bdh and MG/SP1-bdh in MalSALN(BGP) ......................................... 100
5.3.7 Slower growth of MG/SP2-bdh and MG/SP1-bdh on glucose .................................................. 102
5.3.8 Increased NADH/NAD⁺ ratio ..................................................................................................... 103
5.4 Discussion ....................................................................................................................................... 103
5.4.1 Strategy to perturb NADH/NAD⁺ ............................................................................................ 103
5.4.2 Unexpected effect of extracellular acetoin on MG1363 .......................................................... 104
5.4.3 PDHc activity in anaerobic conditions ....................................................................................... 104
5.4.4 Shift towards acetate production in MG1363 and additionally formate production in IL1403 .......................................................... 105
5.4.5 Counterintuitive NADH/NAD⁺ ratio ......................................................................................... 106
5.5 References ....................................................................................................................................... 106

Chapter 6. Amino acid metabolism and fermentation modes ......................................................... 108

6.1 Introduction ...................................................................................................................................... 108
6.2 Results and discussion ................................................................................................................ 109
6.2.1 Computational method developed .......................................................................................... 109
6.2.2 Amino acid utilization and mixed-acid fermentation ............................................................. 109
6.3 References ....................................................................................................................................... 110

Chapter 7. Conclusion and future directions .................................................................................. 112

7.1 Conclusion ....................................................................................................................................... 112
7.1.1 Different promoter activities of mixed-acid genes ................................................................. 112
7.1.2 Contral on mixed-acid fermentation by pyruvate formate-lyase ............................................. 112
7.1.3 Complementary roles of acetate kinase isozymes .................................................................... 112
7.1.4 Acetate production induced by additional NAD⁺-regenerating activities ................................ 113
7.1.5 Connection between amino acid metabolism and fermentation modes .................................. 113
7.2 Future directions ............................................................................................................................ 113
7.2.1 Identify transcriptional regulators and cis-regulatory elements .............................................. 113
7.2.2 Restore the phenotype of mixed-acid fermentation for IL1403 ............................................... 113
7.2.3 Characterize acetate-dependent growth of IL1403 ............................................................... 114
7.2.4 Model the perturbation of NADH/NAD⁺ ratio ......................................................................... 114
7.2.5 Study the effect of another important cofactors, ATP/ADP ratio ............................................ 114
7.2.6 Comprehensive computational analysis on the connection between amino acid metabolism and fermentation modes .......................................................... 114

Appendix A Estimating biological elementary flux modes that decompose a flux distribution by the minimal branching property ........................................................................................................ 115

A.1 Abstract ....................................................................................................................................... 115
A.2 Introduction ................................................................................................................................. 115
A.3 Methods ....................................................................................................................................... 117
A.3.1 Optimization objective ............................................................................................................. 117
A.3.2 Minimal branching decomposition ........................................................................................... 119
A.3.3 Enumeration of alternative optima ........................................................................................... 119
A.3.4 Other decompositions for comparison ..................................................................................... 120
A.3.5 Implementations ....................................................................................................................... 120
A.4 Results .......................................................................................................................................... 120
  A.4.1 Core E. coli metabolic network ............................................................................................... 120
  A.4.2 Mouse myocardial metabolic network .................................................................................... 123
  A.4.3 Lactococcus lactis genome-scale metabolic network .............................................................. 126
A.5 Conclusion .................................................................................................................................. 130
A.6 Acknowledgements ....................................................................................................................... 130
A.7 References .................................................................................................................................. 130
A.8 Supplementary Material .............................................................................................................. 132
  A.8.1 Contributable EFMs .................................................................................................................. 132
  A.8.2 Decompositions by EFMs with unbalanced metabolites ......................................................... 132
  A.8.3 Test of uniqueness of MBDs ..................................................................................................... 133
  A.8.4 Mouse myocardial metabolic network .................................................................................... 134
  A.8.5 Lactococcus lactis’s metabolic network .................................................................................... 138
  A.8.6 Matlab script ............................................................................................................................. 144

Appendix B  Supplementary Material for Chapter 4 .................................................................... 145
  B.1. Supplementary tables .................................................................................................................. 145
  B.2. ACK entries in Uniprot ............................................................................................................... 147
  B.3. Supplementary figures ............................................................................................................... 147

Appendix C  Supplementary Material for Chapter 5 .................................................................... 150
List of Figures

Figure 1.1. Schematic phylogenetic tree of LAB. ................................................................. 1
Figure 1.2. Fermentation pathways in lactic acid bacteria. .................................................. 5
Figure 1.3. Carbon catabolite repression in Firmicutes including lactic acid bacteria .......... 7
Figure 1.4. Pathway for maltose uptake in L. lactis. ............................................................. 16
Figure 1.5. Two different scenarios leading to a zero FCC .................................................. 19
Figure 1.6. Mixed-acid fermentation pathway producing formate, acetate and ethanol ....... 24
Figure 1.7. Flavour compound production from pyruvate. .................................................. 26
Figure 2.1. Construction for measuring promoter activities. ................................................. 51
Figure 2.2. Identification of trans- and cis-regulation .......................................................... 52
Figure 2.3. pfl promoter activities. ......................................................................................... 52
Figure 2.4. eutD promoter activities. .................................................................................... 53
Figure 2.5. ackA1, ackA2 promoter activities ........................................................................ 53
Figure 2.6. adhE promoter activities in MG1363 ................................................................. 54
Figure 3.1. MG1363 derivatives with modulated PFL from MG1363 or IL1403 ................. 61
Figure 3.2. Screening of strains with modulated pfl ............................................................. 62
Figure 3.3. gusA activity of wild-type promoter and synthetic promoters ............................. 62
Figure 3.4. Dependence of formate flux on pfl expression ................................................... 64
Figure 3.5. Dependence of acetate flux on pfl expression .................................................... 65
Figure 3.6. Dependence of ethanol flux on pfl expression .................................................... 65
Figure 3.7. Dependence of lactate flux on pfl expression ..................................................... 66
Figure 3.8. Dependence of glycolytic flux on pfl expression ............................................... 67
Figure 3.9. Dependence of growth rate on pfl expression .................................................... 68
Figure 3.10. Biomass and product yield on sugars ............................................................... 69
Figure 4.1. Pyruvate and Ac-CoA metabolism ................................................................. 76
Figure 4.2. Phylogeny of acetate kinases from Lactococcus, Streptococcus and other species ... 81
Figure 4.3. Nucleotide sequence upstream of ackA1 and ackA2 ........................................ 82
Figure 4.4. Transcription activities of ackA genes .............................................................. 83
Figure 4.5. ACK activities in crude extracts of MG1363 and derived ackA deletion strains .... 85
Figure 4.6. Representative growth curves and growth rates of MG1363 and derived ackA deletion strains. ....... 85
Figure 4.7. Representative growth curves of MG1363Δpfl and derived ackA-pfl deletion strains. 87
Figure 5.1. Introducing NAD⁺-regenerating enzyme activities ............................................. 95
Figure 5.2. Products in overnight cultures in MalSALN media with/without acetoin ........... 98
Figure 5.3. Growth rates, fluxes of MG1363 and MG/SP1-bdh growing in MalSALN ....... 99
Figure 5.4. Growth rates in MalSALN or MalSALN(BGP) media ...................................... 101
Figure 5.5. Fluxes of MG1361, MG/SP2--bdh and MG/SP1--bdh plotted against 23BDH activity. .......................................................... 101
Figure A.1. The graphical meaning of NF ................................................................. 117
Figure A.2. Testing the ability of MBD to separate fluxes at branch points ......................... 122
Figure A.3. Reactions with non-zero fluxes in the mouse myocardial metabolic network .... 124
Figure A.4. (a) Number of uptake reactions, (b) number of excretion reactions and (c) length of EFM in MBD, MinEFMD and random decomposition ............. 127
Figure A.5. (a) Number of multiple MBDs and (b) time required for computing the first MBD for FDs randomly sampled from the core E. coli metabolic network .......... 134
Figure A.6. Reactions with non-zero fluxes in the mouse myocardial metabolic network (color)..............................................................................................................134
Figure A.7. Two sets of 16 EFMs composing (a) the unique MBD and (b) a randomly selected MinEFMD calculated from EFMs in which NADH and NAD$^+$ are left unbalanced.............................................................................................................137
Figure A.8. Branch points in the amino acid metabolism in the flux distribution in the L. lactis metabolic network.........................................................................................................................139
Figure A.9. MBD with high weights on amino acid branch point. .................................................................................................................................141
Figure A.10. A randomly selected EFM from a randomly selected random decomposition of the flux distribution in the L. lactis metabolic network..............142
Figure B.1. Multiple alignment of AckA1, AckA2 and ACKs from Salmonella typhimurium and Methanosarcina thermophila.............................................................................................................147
Figure B.2. Multiple alignment of all lactococcal ACKs.............................................................148
Figure B.3. Activities of purified acetate kinases. ...........................................................................149
List of Tables

Table 1.1. Summary of experimentally determined FCCs..................................................18
Table 2.1. Phenotype of L. lactis strains regarding shift in fermentation mode......................46
Table 2.2. Primers used in the study. ..................................................................................47
Table 2.3. Plasmid used or constructed in the study..........................................................47
Table 2.4. Strains used or constructed in the study............................................................48
Table 2.5. Specific rate of growth, sugar consumption and product formation......................50
Table 2.6. Derived growth statistics....................................................................................50
Table 3.1. Primers used in the study. ....................................................................................59
Table 3.2. Plasmid used or constructed in the study..........................................................59
Table 3.3. Strains used or constructed in the study............................................................60
Table 3.4. Growth data and gusA activities of SC105, SC106, SC100 and SC119..................63
Table 3.5. Estimated flux control coefficients (FCCs) at the wild-type pfl level....................71
Table 4.1. Estimated K_m and k_cat for AckA1 and AckA2..................................................84
Table 4.2. Average specific rates of consumption of maltose, production of lactate, formate and acetate of MG1363, MG1363ΔackA1 and MG1363ΔackA2 at 0 or 50 mM of extracellular acetate......................................................86
Table 4.3. Average specific rates of consumption of maltose, acetate and production of lactate of MG1363Δpfl, MG1363ΔackA1Δpfl and MG1363ΔackA2Δpfl........................................87
Table 5.1. Plasmid used in the study....................................................................................96
Table 5.2. Strains used or constructed in the study............................................................96
Table 5.3. MG1363 and MG/SP1−bdh growing in MalSALN with or without acetoin.........100
Table 5.4. Specific rates of MG1363, MG/SP2−bdh and MG/SP1−bdh growing in MalSALN(BGP)...............................................................................................................101
Table 5.5. Specific rates of MG1363, MG/SP2−bdh and MG/SP1−bdh growing in GluSALN(BGP)..................................................................................................................102
Table 5.6. Measurements of NADH/NAD⁺ ratio and their intracellular concentrations.......103
Table 5.7. Estimated net production of NADH from glycolysis and fermentation..............106
Table 6.1. Selected results about the impact of amino acid availability on mixed-acid fermentation.........................................................................................................................108
Table 6.2. Production and uptake fluxes of the estimated flux distribution and its decomposition by EFMs..................................................................................................................111
Table A.1. Summary of test results in the core E. coli metabolic network.........................121
Table A.2. Comparison of MBD, MinEFMD and random decomposition for the mouse myocardial flux distribution.................................................................125
Table A.3. AA branch points in the flux distribution in L. lactis network............................128
Table A.4. Summary about the tests of uniqueness of MBDs.............................................133
Table A.5. Comparison of the properties of MBD and MinEFMD for the mouse myocardial flux distribution calculated from the EFM matrix in which NADH and NAD⁺ are left unbalanced.........................................................135
Table A.6. Splitting the biomass reaction into groups of metabolites..................................143
Table B.1. Primers used in the study in Chapter 4.........................................................145
Table B.2. Plasmid used in Chapter 4.................................................................................146
Table B.3. Strains used in Chapter 4.................................................................................146
Table C.1. MG1363 and MG/SP1−bdh growing in MalSALN with or without acetoin.......150
Table C.2. Product yield, carbon recovery and biomass yield for growth in MalSALN......150
Table C.3. Product yield, carbon recovery and biomass yield for growth in MalSALN(BGP). .................................................................150
Table C.4. Product yield, carbon recovery and biomass yield for growth in GluSALN(BGP). .................................................................151
Table C.5. 23BDH activities and rates of acetoin consumption and 23BD production in SALN(BGP). .................................................................151
Chapter 1. Background

1.1 Lactic acid bacteria

At the beginning of the twentieth century, the concept of lactic acid bacteria (LAB) gradually emerged to describe a group of bacteria with similarities in morphology, metabolism and physiology that produces primarily lactic acid during fermentation of carbohydrates and whose first prominent role was recognized in milk fermentation [1]. LAB are generally Gram-positive, non-sporulating, non-respiring, nonmotile, low-GC, in the shape of rods or cocci [2, 3]. Mostly under the order Lactobacillales, LAB include genera Aerococcus, Carnobacterium, Enterococcus, Lactobacillus, Lactococcus, Leuconostoc, Oenococcus, Pediococcus, Streptococcus, Tetragenococcus, Vagococcus, Weissella, etc. [4, 5]. Figure 1.1 shows the evolutionary relationship between different genera of LAB [5].

![Figure 1.1. Schematic phylogenetic tree of LAB.](image)


The habitats of LAB range from plants, milk, meat and other different foods, to the mammalian intestine and mouth [1]. LAB also have diverse relationships with human, such as starter cultures in food fermentation [6], probiotics in human gut [7] and pathogens causing fatal diseases [8]. The ubiquitous presence of well-adapted LAB in different environments which still maintain central similarity, as well as their intimate relationship with human make them an extremely important subject in both basic and applied research in microbiology.
1.2 Industrial relevance of lactic acid bacteria

1.2.1 Starter cultures
LAB have been employed in the production of a wide variety of fermented food products. The ability to affect texture, flavour and prolonging shelf-life of the fermented product in combination with many years of safe use has led to the establishment of many LAB as workhorses within food fermentation. For instance, Lactococcus lactis (L. lactis) is famous for its role in producing dairy products including cheese, butter and sour cream [9]. When producing different types of cheese, other LAB species are present in the starter cultures used, e.g. Lactobacillus species, Streptococcus thermophilus, Leuconostoc mesenteroides and they all tend to have an effect on flavour and texture [10]. Flavour compounds of exceptional importance to cheese include various aldehydes, alcohols, carboxylic acids and esters, in particular those derived from methionine phenylalanine, threonine and the branched-chain amino acids (reviewed in [10, 11]). Lactobacillus delbrueckii (Lb. delbrueckii) subsp. bulgaricus and Streptococcus thermophilus are the essential bacteria in yogurt cultures [12]. Some Lactococcus, Lactobacillus, Pediococcus and Enterococcus species can be used in the preservation of fermented meats, fish, vegetables, soy sauce, wine etc. due to their bacteriocin production in addition to the acidification of the food (reviewed in [13–15]). Both of the effects can inhibit the growth of pathogenic bacteria and help LAB to dominate the microflora of food.

1.2.2 Probiotics
In addition to their essential roles in food fermentations, some LAB including some Lactobacillus, Enterococcus and Bifidobacterium species can also act as probiotics which exert extra health benefits more than the inherent nutrition [16]. They are naturally present in the human gut and together with other bacterial species form the so-called ‘microbiome’. They can survive in the acidic environment of the stomach and colonize the large intestine where they produce antioxidants that scavenge free radicals, antimicrobial compounds that inhibit the growth of intestinal and gastric pathogens and in addition occupy the intestinal surface and prevent other bacteria from invading the body [7]. For instance, Lb. acidophilus produces compounds called lactocidin, acidolin and acidophilin with a broad spectrum activity that targets both Gram-positive and negative bacteria [1]. Reuterin produced by Lb. reuteri has been reported to inhibit yeast, fungi, protozoa and bacteria including Salmonella, Clostridium, Staphylococcus, Listeria, etc. [1]. A number of LAB species have even been tested as live carriers of oral vaccine because of their safe status and mimicry of natural infection [17]. For example, recombinant L. lactis expressing tetanus toxin fragment C has been tested to successfully induce immune response when orally administered [18].

1.2.3 Cell factories
While the GRAS status of LAB establish their important position in food industry, the relatively simple carbon metabolism of LAB further makes them attractive targets for metabolic engineering for
production of food ingredients, nutraceuticals and even as cell factories for non-food related chemicals (reviewed in [19, 20]). By all means, the most natural product of all LAB is lactic acid. About 90% of the 80 000 tonnes of lactic acid produced worldwide were by LAB in year 2000 for application to food, pharmaceutical, leather and textile industries and for producing base chemicals and polylactic acid, an important precursor of biodegradable plastic [21]. The worldwide demand for lactic acid has increased to 130 000 – 150 000 tons in 2007 and 259 000 tons in 2012 [250]. Other compounds, for example, alanine, acetaldehyde, low-calorie sugars, vitamins, certain plant metabolites and polysaccharides can be produced by L. lactis and some Lactobacillus species (reviewed in [6, 20]). Streptococcus thermophilus, Lb. planetarium, Lb. acidophilus and Enterococcus faecalis have also been applied to reduce anti-nutritional compounds such as excessive sugars and amines [6]. LAB have also become popular hosts for overproduction of recombinant proteins, and this is largely due to available food-grade controlled gene expression systems, among which the most popular is the nisin controlled gene expression system [22]. For production of biofuels, L. lactis has been engineered to produce ethanol with a high yield by introducing genes from Zymomonas mobilis [23]. Production of butanol has been achieved in Lb. brevis by introducing the butanol pathway in Clostridium acetobutylicum [24]. 2,3-butanediol widely used as feedstock can be produced concomitant with a low-calorie sugar mannitol in L. lactis by cofactor engineering [25]. There have been many review articles on the metabolic engineering of LAB. Interested readers are referred to [19, 20, 26, 27].

1.3 Some characteristics of lactic acid bacteria

The exceptional relevance of LAB to food industry, human health and as microbial cell factories calls forth tremendous efforts to study their genetics, molecular biology, metabolism and biochemistry in order to optimize fermentation yield, to increase the health benefits of products, to better engineer LAB for production of other chemicals, as well as to learn the interactions with other bacterial communities in gut and the host. In this section, some of the biological knowledge on LAB is reviewed.

1.3.1 Genomics

Genome sequences

The genome of a laboratory strain L. lactis subsp. lactis IL1403 was the first completely sequenced genome of LAB and was published in 2001 [28]. Since then, more and more LAB have been genome-sequenced. Until 2009, there were 31 complete genome sequences of LAB [29]. Thanks to the popularization of next generation sequencing, up to September 2014, search in the NCBI database showed that there are 232 species containing thousands of strains under the order Lactobacillales with complete genome sequences. Almost all of them are considered as members of LAB. It is a dramatic increase and highlights the growing attention received by LAB from the scientific community as well.
The typical size of the genome of LAB ranges from 1.7 Mbp to 3.4 Mbp with the number of genes lying between 1,500 to 3,300 and GC content between 35% and 50% [30].

**Reductive genomes**
From evolutionary genomic analysis, LAB shows a reductive genomic evolution [29, 30]. It has been estimated that the common ancestor of *Lactobacillales* had lost 600 to 1,200 genes but gained less than 100 genes since the divergence from class *Bacilli* [31]. The reduction in genome size was probably due to adaptation to environments with rich nutrient availability [29, 32], for instance, from plant to milk, which has recently been elegantly demonstrated by subjecting *L. lactis* isolated from plant to adaptive evolution in milk [33]. Another example is the smaller genome sizes and lower GC content of *Lactobacillus* species found in vaginal cavity among all LAB which was suggested to be a result of vaginal adaptation [30, 34].

**Pseudogenes**
The genomic reduction is also evidenced by the pervasive presence of pseudogenes (genes that have lost their functions during the course of evolution), which in LAB generally are more frequent than in other bacteria [4, 29]. This in general has been compensated by enrichment for genes encoding transporters for amino acids, sugars, ions, peptides, etc. to take up the nutrients required for growth from the rich environments instead of synthesizing them *de novo* [29, 30].

**Mobile genetic elements**
Plasmids and other mobile genetic elements are also important factors that have shaped the evolution of LAB [30], by facilitating horizontal gene transfer. This has enabled many LAB to utilize lactose, galactose, oligo-peptide and has also provided them with various phage resistance mechanism, etc. (reviewed in [29, 35]). For example, the megaplasmid in a *Lb. salivarius* strain contains genes for bacteriocin production and bile salt hydrolase which make the strain probiotic [36]. Plasmids responsible for traits critical for dairy fermentation were also found in a dairy *L. lactis* strain [37] but not present in plant isolates [30].

### 1.3.2 Metabolism

**Fermentation of carbohydrates**
LAB do not have an electron transport chain and thus are unable to perform oxidative phosphorylation for ATP synthesis. Instead, LAB rely on substrate-level phosphorylation to generate energy from sugars. Sugars are taken up by the phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS), a common sugar transport system of bacteria in which extracellular sugar is directly phosphorylated to become intracellular sugar phosphate by receiving the phosphate group from PEP [38]. PTS is also the main system for carbon catabolite repression, a major way of regulating carbohydrate metabolism in bacteria, which is discussed in a later subsection. An alternative way of
Chapter 1. Background

Sugar transport is uptake by sugar-specific permeases. The ability to utilize different sugars by different species of LAB is reviewed in [39] in great detail. Usually, over 95% of sugar is converted into lactic acid by LAB [20]. Generically, there are two fermentation pathways for this in LAB [5].

**Glycolysis**

The first is glycolysis which is the typical Embden-Meyerhof-Parnas (EMP) pathway in which one mole of glucose is oxidized to form two moles of pyruvate concomitant with two moles of ADP phosphorylated into two moles of ATP (Figure 1.2). Lactate dehydrogenase (LDH) then reduces the two moles of pyruvate into two moles of lactate, balancing the redox. Other sugars such as galactose can also be metabolized through glycolysis by conversion into glucose 6-phosphate (G6P) through the Leloir pathway or conversion into triosephosphate through the tagatose 6-phosphate pathway (Figure 1.2) [5]. This pathway has sometimes been called homofermentative pathway in the literature.

![Figure 1.2. Fermentation pathways in lactic acid bacteria.](image)

(A) Homofermentative pathway including tagatose 6-phosphate pathway and Lelior pathway for galactose uptake and mixed-acid fermentation. (B) Heterofermentative pathways. PEP, phosphoenolpyruvate; GAP, glyceraldehyde 3-phosphate; DHAP, dihydroxyacetone phosphate. Lumped reactions are in dotted lines. Fermentation end products are highlighted.

**Mixed-acid fermentation and flavour compound production**

Under certain environmental conditions, the pyruvate formed in the homofermentative pathway is not totally converted to lactate, but goes into the other two branches. The first branch is the formation of acetyl-CoA by pyruvate-formate lyase or pyruvate dehydrogenase, followed by acetate and ethanol production [40, 41]. When a significant amount of these end products is produced in addition to lactate during fermentation, it is called ‘mixed-acid fermentation’. This is also the main subject of study in this thesis which is discussed in section 1.5.3. The second is the conversion of pyruvate into
α-acetolactate, followed by diacetyl, acetoin and 2,3-butanediol production which are important flavour compounds in dairy industry [5].

**Phosphoketolase pathway**
The second fermentation pathway is the phosphoketolase pathway (Figure 1.2) [1, 5]. In this pathway, glucose 6-phosphate or pentose enters pentose phosphate pathway rather than EMP pathway. Xylulose 5-phosphate is then converted by phosphoketolase into glyceraldehyde 3-phosphate (GAP) entering glycolysis, and acetyl phosphate from which acetate or ethanol is produced [1, 5, 41]. This pathway has also been called heterofermentative/heterolactic pathway in the literature.

The fermentation type is an important characteristic of LAB. For instance, *Leuconostoc, Oenococcus, Weissella, Lb. brevis, Lb. buchneri, Lb. fermentum*, and *Lb. reuteri* are obligatory heterolactic; *Lb. acidophilus, Lb. delbrueckii, Lb. helveticus* and *Lb. salivarius* cannot catabolize pentose and are obligatory homolactic; many other LAB are facultatively heterolactic [5, 42].

**Carbon catabolite repression**
Most of the bacteria including LAB are known to have the ability to selectively consume carbon sources for fastest immediate growth by the carbon catabolite repression (CCR) system (reviewed in [43, 44]). The most classical example must be the diauxic growth of *Escherichia coli* in the presence of both glucose and lactose discovered by Monod [45]. In Firmicutes, which consists of many low-GC Gram-positive bacteria including LAB and *Bacillus*, CCR is exerted through the concerted effort of the histidine protein (HPr), which is one of the component proteins of PTS and the global regulator called catabolite control protein (CcpA) [46].

When a sugar allowing fast growth such as glucose is present, the glycolytic flux is relatively high and so are the levels of ATP and fructose 1,6-bisphosphate (FBP). This activates the HPr kinase/phosphorylase which phosphorylates HPr at Ser46 (HPr-Ser-P). HPr-Ser-P on the one hand triggers inducer exclusion which inhibits the activities of permeases of other sugars by allosteric regulation and on the other hand forms a complex with CcpA which serves as a global transcriptional regulator binding to cre sites on DNA to exert transcriptional regulation [46]. In the absence of glucose or when the levels of glycolytic intermediates are low, HPr is mostly not Ser46-phosphorylated for catabolization of carbon sources. In particular, in *B. subtilis*, induction of specific catabolic operons by glycerol kinase and other substrate-specific regulators containing PTS-regulatory domains (PRDs) is prevented by the low level of HPr(His-P) in the presence of glucose. The mechanism is visualized in Figure 1.3 (reprinted from [43] with permission).
Chapter 1. Background

Figure 1.3. Carbon catabolite repression in Firmicutes including lactic acid bacteria. EI, enzyme I of PTS; EII, enzyme II of PTS; HPr, histidine protein; HPrK, HPr kinase/phosphorylase; P, phosphate; CcpA, catabolite control protein; Glu, glucose; Glu-6-P, glucose 6-phosphate; FBP, fructose 1,6-bisphosphate; PRD, PTS-regulatory domain; GlpK, glycerol kinase. Reprinted from [43] with permission.

Nutritional auxotrophy

As discussed in section 1.3.1, LAB have evolved reductively to adapt to nutritionally rich environments. Many of the lost genes encode enzymes that are involved in biosynthetic pathways. LAB are thus in general auxotrophic for many nutrients such as some amino acids and vitamins. There is however not a clear pattern of the auxotrophy depending on species. Some strains of L. lactis, Lb. plantarum, Lb. casei, Lb. helveticus were tested for the amino acid requirements for growth and it was found that even different strains under the same species can have different amino acid requirements [47]. Recent results from comparative genomics have shown that, for instance, the families of Lactobacillaceae and Leuconostocaceae lost the genes for synthesizing serine and glycine; Pediococcus pentosaceus, Oenococcus oeni and some Lb. species lost the genes for arginine and aromatic amino acid biosynthesis; some genes for biosynthesis of fatty acid were not found in Lb. gasseri and Lb. johnsonii [4]. Most of these results, however, have not yet been confirmed by physiological characterization. A lot of effort is required to characterize the different nutritional requirements of LAB.

Proteolytic system

To adapt to nutritionally rich environments, the loss of ability to synthesize nutrients necessary for growth usually comes along with gain of ability to take up nutrients from environments quickly [29, 30]. The proteolytic ability of LAB, especially those adapted to the dairy environment, is a very important trait for LAB to prevail. LAB can express cell-envelope proteinases to break down extracellular casein, the protein of largest fraction in milk, into peptides primarily ranging from 4 to
10 residues which are then taken up by LAB with specific transport systems followed by further hydrolysis by peptidases into single amino acids [48]. The characterized proteinases, peptide transport systems and peptidases have been extensively reviewed (e.g. [48–50]).

1.3.3 Stress tolerance
LAB encounter various stressful conditions when used in food processing, and stress tolerance is important to consider if a consistent product quality is to be obtained. The quality of the final products may not be assured if the culture dies too early or produces different compounds under stress condition. There are several types of stress, including acid, oxidative, osmotic, heat, cold and starvation stress [51]. Tolerance to these stresses is also an important phenotype that distinguishes different species of LAB. Examples include: *Streptococcus* is unable to grow at 10 °C; *Enterococcus* is able to grow at 45 °C but most other species are unable; *Aerococcus*, *Enterococcus* and *Tetrageonococcus* are able to grow under significant osmotic stress (6.5% sodium chloride, and up to 18% for *Enterococcus*) or at pH as high as 9.6; *Aerococcus* and *Streptococcus* are unable to grow at pH 4.4 [5]. It should be noted that sometimes the tolerance can differ from strain to strain and exceptions can exist. The species-specific phenotypes under various stress conditions of many LAB are reviewed in [39] in great detail. The response to stress and its mechanism in LAB are extensively reviewed in [51].

1.4 Tools for studying lactic acid bacteria
The advance in science is deeply influenced by the technology available to study it. It is exceptionally conspicuous in the course of development of biological science in the last century. Since the discovery of the molecular model of DNA by Watson and Crick and the later proposed central dogma of molecular biology by Crick in the 1950s, much of the biological research has turned the focus to biomolecules and their interactions. Many molecular tools have been developed to manipulate cells. The different ‘–omics’ techniques developed in the last two decades also gave birth to the systems biology paradigm which studies biology at a system level. A lot of bioinformatics and computational tools have thus been developed for this purpose. The study of LAB, as an important branch of microbiology, has undergone the same transformation. In this section, the molecular tools, computational models and tools for systems biology developed for or applied to LAB are briefly reviewed.

1.4.1 Molecular tools
Different genetic tools, with applications within strain optimization or basic research, have been developed for LAB. Beside the general tools designed for DNA manipulation, there are two important tools specifically for modulating gene expression, namely inducible promoters for controlled gene expression and synthetic promoter libraries. Many of the tools were first engineered for *L. lactis* and later applied to other LAB.
**Controlled gene expression systems**

The most influential controlled gene expression system invented for LAB must be the nisin controlled gene expression (NICE) system [22, 52]. In this system, the histidine-protein kinase NisK acts as a sensor which autophosphorylates in the presence of nisin, a broad spectrum bacteriocin, and the phosphate group is transferred to the response regulator NisR which then activates the specific promoter which is originally responsible for initiating the transcription of the gene cluster for nisin biosynthesis in the nisin-producing *L. lactis* strain NZ9700 [53]. There are a number of advantages of the NICE system. First, nisin is known to be safe so the overexpression of recombinant proteins by this system can be considered food-grade [22]. Second, it has a high sensitivity in the sense that only a tiny amount of nisin is required for induction (0.1 – 5 ng/ml) which normally does not cause growth inhibition to other microorganisms in a starter culture during fermentation [22, 54]. Third, the induced expression level is linearly correlated with the level of inducer, namely nisin, in a dynamic range which can be more than 1,000-fold [53, 55]. This property enables the study of hierarchical flux control by the inductively expressed enzyme, for instance, how the pyruvate distribution is controlled by the activity of NADH oxidase in *L. lactis* [56], the conversion of *L. lactis* from homolactic to homoalanine fermentation [57], high-level production of diacetyl [58]. The NICE system has been applied in other LAB such as *Streptococci* and *Enterococci* (reviewed in [22]).

Later another similar system using another bacteriocin sakacin as inducer was developed for *Lb. sakei* and *Lb. plantarum* [59–61]. There are other attempts to develop other expression systems, e.g. expression controlled by zinc availability in *L. lactis* [62]. The controlled gene expression systems naturally existing in or developed for LAB have been reviewed in [54, 63, 64].

**Synthetic promoter libraries**

A synthetic promoter library (SPL) is a set of promoters with the consensus sequences (e.g. -10 element and -35 element for bacteria) fixed and sequence of spacers in between randomized (reviewed in [65–67]). Such a library of promoters, essentially equivalent to saturation mutagenesis, has been shown to be capable of modulating gene expression in a dynamic range of up to thousand-fold [68, 69]. Traditional approaches usually either knock out a gene or overexpress a gene by a strong promoter. In comparison, SPL results in a continuous range of activity which allows the optimization of the desired metabolic phenotype as well as the estimation of flux control coefficient in the theory of metabolic control which requires only a slight perturbation from the original level of enzyme activity [65–67, 69].

Several strategies for practical construction of a collection of strains with a SPL have been proposed later [70]. One strategy is a single integration of a plasmid with a truncated part of the gene of interest following a SPL inserted into the plasmid. A marker for selection will reside in the resultant strains in this case. This approach has been applied to study glycolytic flux control in *L. lactis* [71–73]. Another
approach is to remove the marker and the rest of the plasmid by a double-crossover event but then there should be a homologous region upstream of the native promoter inserted prior to the SPL in the plasmid. These two methods are suitable for modulating the promoter strength of a natively existing gene. A third way is to make use of the site specific integration system to directly introduce a complete gene of interest following an SPL into the chromosome. For example, a site-specific integration system requiring a resolvase-like integrase derived from the temperate lactococcal bacteriophage TP901-1 [74–76] has been used together with SPLs in a number of studies on glycolytic flux control in *L. lactis* (e.g. [77, 78]). A fourth way is to express the gene of interest by a SPL on a replicating plasmid directly. It has been applied to study, e.g. the dependence of glycolytic flux on ATP/ADP level in *E. coli* and *L. lactis* [79–81]. The method of SPL has also been used in other LAB such as *Lb. plantarum* [82].

**Other genetic tools**

In addition to inducible controlled gene expression systems and the method of synthetic promoter libraries, other basic genetic tools for DNA manipulation are also indispensable for studying LAB.

**DNA modifications by homologous recombination**

In *L. lactis*, several plasmids have been created for insertion, deletion or mutation of chromosomal DNA by homologous recombination. An earlier one is pINT1 which is non-replicable and was derived from the plasmid pWV01 isolated from *L. lactis* subsp. *cremoris* Wg2 [83] by removing the repA gene [84]. Another plasmid system pGhost also derived from pWV01 is replication-thermosensitive [85, 86]. The property that the plasmid is replicable below 35°C but not above eases the requirement of high transformation frequency by uncoupling the transformation step and the recombination step but the higher temperature required for selection can meanwhile pose other disadvantages to the bacteria [87]. Later the pORI series was developed to make use of the two systems [87, 88]. Another plasmid, pCS1966 derived from the pBluescript [89], was made to manipulate chromosomal DNA [90]. It makes use of the gene *oroP* encoding an orotate transporter as a counterselection tool by adding 5-fluoroorotate, a toxic analogue of orotate. pSEUDO, a derivative of pCS1966, was then designed for chromosomal integration into a pseudo gene in *L. lactis* [91]. The similar idea of adding the gene *upp* encoding uracil phosphoribosyltransferase to the pORI vector for counterselection has also been applied to *Lb. acidophilus* [92, 93]. Systems for integration by homologous recombination in other LAB include, for example, integration into a number of LAB species with the transposon Tn916 on chromosomes [94], other thermosensitive plasmids in *Streptococcus thermophilus* [95], *Lb. helveticus* [96], *Enterococcus faecalis* [97] (reviewed in [98]).

**DNA insertion by site-specific integration**

Site-specific integration is another popular tool for introducing a single copy of gene into the chromosome. An obvious example is the system derived from the temperate lactococcal
bacteriophage TP901-1 in *L. lactis* mentioned above [74]. Cre-*lox*-based system, which has been used in many organisms, has recently been used to develop a multiple-gene-deletion system in *Lb. plantarum* [99]. Recently, based on these two systems, a new system for repetitive, marker-free, site-specific integration has even been developed [100]. In fact, similar site-specific integration systems originating from temperate bacteriophage naturally exist in many LAB such as *Lactobacillus* [101, 102] and *Streptococcus thermophilus* [103]. Other examples include the pTRK system developed for *Lb. acidophilus* and *Lb. gasseri* adapted from the pORI system [104].

**Random mutagenesis**

Transposition-based mutagenesis is also an effective method for both targeted and random DNA disruption. While targeted mutagenesis aims to study the effect of disrupting a specific gene which can also be achieved by the aforementioned chromosomal integration, random mutagenesis is an elegant method to match genotype to an observed phenotype. Building a genomic library for random mutagenesis based on homologous recombination is a feasible way, e.g. [105], but transposons and insertion sequences are particular convenient tools for this purpose. For instance, an efficient system using lactococcal ISS1 concomitant with pGhost has been developed for *L. lactis* [106] and applied to discover regulators of the operon of cystathionine lyase and cysteine synthase [107], as well as to identify a gene important for protein secretion [108]. Transposase IS1223 and its target *Pjunc* sequence have been used for a random mutagenic system in *Lb. casei* [109] and finding critical genes for the growth of *Lb. pentosus* in olive brine [110]. Systems for insertional mutagenesis were reviewed in [98]. Commonly used mobile and other genetic elements among LAB can be found in [35].

It should be noted that there are molecular tools for LAB other than the aforementioned, e.g. the ssDNA recombineering that has become very popular recently [111]. Some commonly used LAB strains for molecular cloning and their characteristics have been summarized in [35].

1.4.2 **Tools and computation models for systems biology**

Different -omics and high-throughput experimental techniques have been applied to the research on LAB. They are able to generate data which give a holistic view of cells. Several transcriptional regulatory networks, genome-scale metabolic networks of LAB have been reconstructed based on these data. Meanwhile, some refined kinetic models of LAB incorporating enzyme kinetics and enzyme assay data were also built as bottom-up approaches to study cellular metabolism.

**Transcriptional regulatory networks**

Transcriptional regulatory networks (TRNs) are networks that describe the interactions between genes and transcription factors (TFs). It is able to reveal changes at transcriptional level in response to a stimulus. The whole set of genes and operon regulated by a particular TF is called a regulon. The regulon of a given TF can be found experimentally by first comparing the transcriptomes of the wild-
type strain and the TF-knockout strain under a condition in which the TF is expressed. The binding of the TF to the upstream region of a differentially expressed gene can then be confirmed by other assays. Many of the known regulons in LAB were characterized by this method. A more direct but more expensive method is to combine chromatin immunoprecipitation (ChIP) and transcriptomic techniques which target DNA bound to the TF only (ChIP-chip or ChIP-seq). A general review on the construction of TRNs can be found in [112].

Many of the regulons of TFs that are common in many LAB have been characterized, such as CcpA for global catabolite repression [113, 114], CodY for amino acid metabolism [115–118], CopR for copper homeostasis [119, 120], FabT for fatty acid biosynthesis [121] and many others (e.g. [122–124]). Many evidences of genes regulated by TFs were also reviewed by transcriptomic profiling of the same strain in different growth conditions, e.g. changing growth rate [125, 126], changing temperature [127].

Based on the huge amount of experimental data, several attempts have been made to construct regulatory networks for LAB recently. Regulatory network of a *Lb. plantarum* strain has been reconstructed based on correlation of gene expression in more than 70 experimental conditions and the binding motifs predicted from it [128]. Although it was not a TRN, it clustered genes into different transcription units that were co-expressed. A TRN for *L. lactis* growing in milk has been reconstructed by time-resolved transcriptomic analysis as well [129]. Since the proposal of the comparative genomic approach for reconstructing TRNs which is a very powerful tool to make use of genome sequences and biological knowledge of closely related species [130], it has been applied to reconstruct TRNs for 30 LAB strains with available genome sequences in a single study [131]. The results are available on RegPrecise which is a database of both manually curated and automatically propagated regulons with detailed description and interactive graphics [132]. The approach has later been applied for the reconstruction of a TRN of *Enterococcus faecalis* [120]. In this study, sub-networks of the TRN induced by varying copper levels were identified and the authors showed that the TRN of *Enterococcus faecalis* can be refugured to maintain copper homeostasis under genetic perturbation.

**Genome-scale metabolic networks**

Thanks to popularization of whole-genome sequencing technologies, genome-scale metabolic networks (GSMNs) of a cell containing all the metabolic reactions catalysed by genes with known functions that are present in the genome can be reconstructed. Roughly speaking, the reconstruction involves an iterative procedure, starting from a drafted list of reactions, to reconcile known experimental data with computational feasibility of the generated model. A detailed guide for a high-quality reconstruction can be found in [133]. A huge and expanding computational toolbox for GSMNs, the so-called constraint-based reconstruction and analysis (COBRA) approach, has also been
developed (reviewed in [134, 135]). It depends on the quasi-steady-state assumption of metabolism so that the kinetic details can be neglected and fluxes can be estimated by solving linear constraints. Among the many methods, flux balance analysis (FBA) is one of the most important techniques to predict optimal flux distributions [136]. GSMNs which represent the metabolic capability reflected by the genome can have several useful functions in systems biology. A very crucial one is the ‘context’ provided by GSMNs to understand the content of –omics data. Integrating –omics data (e.g. genome, transcriptome, proteome, metabolome) into GSMNs offers a way to link those data to the metabolic phenotype (reviewed in [137]). As the embodiment of genotype-phenotype relations, GSMNs can guide metabolic engineering by suggesting strategies of gene perturbation (reviewed in [138]) and even predict evolution (reviewed in [139]).

In LAB, the application of GSMNs is becoming more and more popular, too. Since the reconstruction of the first GSMN among LAB (L. lactis subsp. lactis IL1403) in 2005 for studying aerobic and anaerobic growth [140], models have been constructed for several species of LAB. They include Lb. plantarum for studying growth on complex media [141]; Streptococcus thermophilus for comparing the amino acid auxotrophy and the capability to produce volatiles with other LAB models [142]; Lb. reuteri to look into the metabolic mechanism of probiotic features [143]; an updated model of L. lactis IL1403 for increasing the production of recombinant proteins using dynamic FBA [144]; and L. lactis subsp. cremoris MG1363 for analysing the flavour-forming ability [145] (reviewed in [146]). As a good example to follow, the practical reconstruction of Lb. plantarum’s model has been described in detail [147]. It is believed that ‘metagenome-scale’ models of LAB would be the next step in response to metagenomics to push forward the frontiers of the study on microbial communities such as gut microbiota [146].

**Kinetic models**

As opposed to GSMNs, kinetic modelling is a bottom-up approach which includes only a limited number (usually dozens) of reactions of interest but the kinetics of each reaction is provided. The change in metabolite concentrations depends on fluxes of reactions which in turn depend on metabolite concentrations. This forms a dynamical system which can simulate system response to dynamic changes and well-established techniques are available for system analysis. Kinetic models however usually suffer from the lack of kinetic parameters which can sometimes be unidentifiable even through data fitting.

For LAB, many attempts have been made to model the kinetics of L. lactis’s glycolysis because of its intriguing regulation. Kinetics of individual enzymes, in particular those involving the conversion of global cofactors NADH/NAD’ and ATP/ADP have been studied and modelled because of their high relevance to the regulation of glycolytic flux, including lactate dehydrogenase (LDH) [148, 149], glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [149], pyruvate kinase (PYK) [150, 151],
alcohol dehydrogenase [149], acetate kinase [152] etc. The first kinetic model that gathered different enzyme kinetic data for L. lactis’s glycolysis was constructed to analyse the effect of different metabolic engineering strategies on acetoin and diacetyl formation [153]. It has also been used to simulate glucose run-out experiments [154] and later adapted to incorporate the effect of pH level [155]. Recently kinetic models for L. lactis and Streptococcus pyogenes taking extracellular phosphate concentration into considerations have been built and successfully described the different characteristics of the two species [156]. The model of Streptococcus pyogenes has later been shown to possess the characteristic of glycolytic oscillation [157].

Another type of models was constructed by assigning a generic rate equation to each reaction and obtaining parameters by fitting time-series data, such as a rational function of substrate and product concentrations [158], the generalized mass action [159] and its special case S-system [160]. They are quite capable of capturing the kinetic features but usually do not lead to mechanistic insights.

For these two types of kinetic models, many of the data of metabolite concentrations used for fitting were generated from nuclear magnetic resonance (NMR) applied to live cells consuming labelled substrates (e.g. [158, 161–163]. Detailed reviews on modelling LAB can be found in [146, 164]

**Metabolic control analysis**

Metabolic control analysis (MCA) aims to study how steady-state fluxes change in response to small changes in enzyme activities [165]. One application is the identification of targets for drugs [166] and metabolic engineering [167]. The flux control coefficient (FCC) is one of the central objects in MCA defined by the rate of fractional change of a steady-state flux with respect to the (infinitesimal) fractional change of an enzyme activity. It can be computed given a kinetic model by perturbing the $V_{\text{max}}$ value of an enzyme or determined experimentally by modulating the expression level of an enzyme using methods like synthetic promoter libraries described in section 1.4.1. It has been widely employed to study the glycolysis of L. lactis which is reviewed and discussed later in section 1.5.1. The kinetic model mentioned previously which analysed the engineering strategies for increasing the production of acetoin and diacetyl actually used MCA for prediction [153]. A method combining MCA and kinetic modelling has also been proposed to predict flux distributions from the proteome in L. lactis [168].

### 1.5 Lactococcus lactis

*L. lactis* is a species of LAB in the family of streptococcaceae. It is in the shape of cocci and mesophilic with maximum growth rate at 30°C and pH 7.0. *L. lactis* can be isolated from dairy environments or non-dairy environments such as plants [169].

There are two subspecies under *L. lactis*, namely subsp. *lactis* and subsp. *cremoris*. The subspecies were originally differentiated by phenotypic differences which include the ability of subsp. *lactis* to
consume arginine, to grow at higher temperatures or higher salt concentrations [170]. There is one biovar in subsp. *lactis* which can utilize citrate, namely subsp. *lactis* biovar *diacetylactis* [170, 171]. Two main genome homology groups were also found by nucleotide sequence comparison. They do, however, not perfectly match the two phenotypic subspecies [172]. An obvious example is the two most widely studied *L. lactis* laboratory strains, IL1403 and MG1363. IL1403 was derived from *L. lactis* subsp. *lactis* biovar *diacetylactis* CNRZ157 (or IL594) by curing the citrate plasmid [171]. It has a *lactis* genotype and a *lactis* phenotype. In contrast, MG1363, a plasmid-free derivative of the dairy strain NCDO712 [173], was considered to have a *cremoris* genotype but a *lactis* phenotype.

As one may notice from the review of LAB in the previous sections, among all LAB, *L. lactis* is one of the most studied species and has become a model organism of the group. Much pioneering research on LAB was first conducted in *L. lactis*. One of the reasons is its extreme industrial importance. *L. lactis* is involved in the fermentation of a wide range of dairy products [170], primarily soft and hard cheeses. It can also be engineered to become a cell factory for chemicals of different categories, including recombinant protein [174], therapeutic proteins, DNA and vaccine antigens [175], flavour ingredients [19], nutraceuticals [6] etc.

Another reason for the huge curiosity aroused by *L. lactis* is its theoretical importance in cellular metabolism because it exhibits very significant metabolic shift from homolactic fermentation to mixed-acid fermentation as sugar concentration decreases, termed the mixed-acid shift or switch [40]. The control and regulation of the glycolytic flux is also a related and interesting object of study because of its intrinsic theoretical value [164] as well as the potential application to industrial production by increasing the glycolytic flux. In this section, the glycolysis and pyruvate metabolism which includes the well-known shift between fermentation modes in *L. lactis* and the regulation of fluxes in these pathways are reviewed.

### 1.5.1 Glycolysis

The glycolysis of *L. lactis* comprises the typical EMP pathway with different carbohydrates entering at different points. This subsection describes the uptake of glucose and maltose in *L. lactis* which are used extensively in our study. The transcriptional regulation of glycolytic enzymes is also briefly reviewed. The roles of glycolytic enzymes and metabolic regulation in determining glycolytic flux are reviewed in the next subsection.

#### Sugar transport

Sugars enter cells of *L. lactis* by either the PEP-dependent PTSs described in section 1.3.2 or sugar-specific permeases. For glucose uptake, there are two distinct PTSs, mannose PTS and cellobiose PTS, and a proton-motive force dependent permease in *L. lactis* and the kinetic properties of these transport systems in MG1363 have been characterized recently [176]. Mannose PTS has a much higher affinity
of micro-molar than the permease of milli-molar. PTSs should be responsible for the majority of the glucose uptake especially for cells starved for sugar for a period of time [176]. Many other monosaccharides like galactose can be transported into *L. lactis* similarly and enter glycolysis via G6P or GAP [177].

For the disaccharide maltose, the pathway for its uptake and entry into glycolysis is shown in Figure 1.4. Maltose is taken up into *L. lactis* cells by an ABC transporter ATP-binding protein [178, 179]. Intracellular maltose is split by maltose phosphorylase (MP) into one glucose and one β-glucose 1-phosphate which is then converted into glucose 6-phosphate by β-phosphoglucomutase (β-PGM) [180, 181]. Through this and the EMP pathway, one maltose is converted into four pyruvate with the formation of four ATP and four NADH from four ADP and four NAD⁺. A gene, predicted to encode for an α-glucosidase (or called maltase) which hydrolyzes one unit of maltose into two units of glucose, exists in the genome of some *L. lactis* strains, such as MG1363 [182, 183]. The physiological significance of the gene product, however, has not yet been demonstrated.

**Figure 1.4. Pathway for maltose uptake in *L. lactis***

The gene for α-glucosidase is only predicted and the function has not been confirmed.

**Transcriptional regulation**

Glycolytic enzymes are in general highly expressed to sustain the huge glycolytic flux in *L. lactis*. A recent study suggested that about 20% of total soluble protein is glycolytic enzymes [184]. Their genes are in general located close to the origin of replication for higher expression [177]. Transcriptional regulation of the expression of glycolytic genes in *L. lactis* primarily lies in the carbon catabolite repression (CCR) as described in 1.3.2 which has been verified experimentally in *L. lactis* [185]. The HPr protein, at high levels of FBP and ATP which reflect high glycolytic flux, is phosphorylated and then forms a complex with CcpA acting as a global regulator binding to the cre sites on chromosome. The consensus well-conserved in Gram-positive bacteria for cre binding site is TGNANCGNTNNCA which is roughly palindromic with the central CG base always present [186]. A study on CcpA regulon in MG1363 proposed the consensus WGWAARCGBTYWWMMA which is specific for *L. lactis* [113].
The binding of CcpA to cre sites can lead to either activation or repression, depending on the position relative to promoters as in the case of Bacillus [187]. When a cre site is upstream of, inside, or downstream of the promoter of a gene, transcription is activated, repressed, or aborted accordingly [177]. Results also indicated that the interaction between CcpA and the transcription machinery may be dependent on the helix side of CcpA binding because strongest repression was observed for cre sites that were consecutively separated by around 10.5 bp, equal to a full helical turn of double-stranded DNA [113]. The expression of a number of glycolytic enzymes was found to be up-regulated by CcpA, including PFK, PYK and LDH (the members of the las operon) [188], phosphoglucone isomerase (PGI), GAPDH, enolase [186].

1.5.2 Control of the glycolytic flux

In the intracellular environment, assuming constant environmental factors, such as pH, temperature, viscosity etc., the reaction rate of an enzyme depends on concentrations of the enzyme, substrates, products and other effectors modulating the activity. Reaction rates would in turn change metabolite concentrations, forming a dynamical system. The metabolic flux of a pathway is the overall conversion rate of metabolites by the pathway resulting from the dynamical interactions between all metabolic enzymes and metabolites. The aforementioned factors controlling flux can be divided into two levels. The first level is the enzyme level, which accounts for changes in fluxes caused by changes in gene expression level. It is sometimes called hierarchical regulation in the literature. The second level is the metabolic level referring to changes in fluxes which are not caused by altered gene expression but by changes in metabolite concentrations and the inherent kinetic properties of enzymes such as maximum velocity and substrate affinity. It is called metabolic regulation. Here the distinction between ‘control’ and ‘regulation’ suggested in [164] is worth mentioned. The statement that an enzyme has ‘control’ on a flux should refer to the phenomenon that change in the enzyme level leads to change in the flux but not the direct regulatory mechanism. ‘Regulation’ should refer to the exact mechanism causing the change. These two terms have occasionally been used interchangeably though.

Control by individual enzyme levels

Control of fluxes by individual enzymes can be quantified by FCCs in MCA (which is defined by the rate of fractional change of a steady-state flux with respect to the fractional change of an enzyme activity). Finding out the so-called ‘rate-limiting’ step or an enzyme with high flux control that can lead to increase in glycolytic flux in L. lactis can have direct industrial relevance, e.g. speeding up the production. An earlier study trying to inhibit the activity of GAPDH by the specific inhibitor iodoacetate indicated that GAPDH had a high FCC of about 0.9 on glycolytic flux in non-growing cells of L. lactis subsp. cremoris Wg2 [189]. Similar results were obtained in another strain NCDO2118 with GAPDH having a FCC equal to 0.7 [190]. Later, the control of glycolytic flux by
glycolytic enzymes has been extensively studied by Jensen’s group by experimental estimation of FCCs in the laboratory strains *L. lactis* IL1403 and MG1363. The results are summarized in Table 1.1.

### Table 1.1. Summary of experimentally determined FCCs.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>FCC at wild-type level</th>
<th>Min. fraction of wild-type enzyme level for maximal glycolytic flux</th>
<th>Definite expression level for optimality*</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>growth rate</td>
<td>glycolytic flux</td>
<td>lactate flux</td>
<td>formate flux</td>
</tr>
<tr>
<td><strong>MG1363</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDH</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>-1.45 ~ -1.27</td>
</tr>
<tr>
<td>TPI</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ -0.25</td>
</tr>
<tr>
<td>GAPDH</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
</tr>
<tr>
<td>PFK</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
</tr>
<tr>
<td>PYK</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>0.9 ~ 1.1</td>
</tr>
<tr>
<td><strong>IL1403</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPI</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
</tr>
<tr>
<td>PGM</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
</tr>
<tr>
<td>ENO</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
<td>≈ 0</td>
</tr>
</tbody>
</table>

*Definite expression level for optimality* refers to a unique maximum of growth rate and glycolytic flux at the wild-type enzyme level. See Figure 1.5 for illustration. LDH, lactate dehydrogenase; TPI, triosephosphate isomerase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; PFK, phosphofructokinase; PYK, pyruvate kinase; PGM, phosphoglycerate mutase; ENO, enolase.

Interestingly, usually each individual enzyme appears to have no control on growth rate and glycolytic flux at the wild-type enzyme level, including LDH [191], GAPDH [78], phosphofructokinase (PFK), PYK [77] and triosephosphate isomerase (TPI) [72] for MG1363; TPI, enolase [71] and phosphoglycerate mutase (PGM) [73] for IL1403. Among these enzymes, some are present in the wild type in significant excess for attaining maximum glycolytic flux, such as LDH, TPI and GAPDH whereas some enzymes appear to be optimally expressed in the wild type for maximum glycolytic flux, such as PFK, PYK, PGM and ENO. This means that when the level of any of the latter set of enzymes is perturbed slightly, the growth rate and glycolytic flux do decrease no matter for over- or under-expression. This property of maximum growth rate and glycolytic flux in the wild type also leads to a zero FCC at the wild-type level. Figure 1.5 illustrates the two different scenarios leading to zero FCCs in the wild type encountered in the experimental studies of glycolytic flux control by glycolytic enzymes. In the literature, nonetheless, the possible consequence and interpretation of this observation have not been discussed thoroughly.
Figure 1.5. Two different scenarios leading to a zero FCC.
(a) An enzyme is present in excess and perturbing its level slightly does not change the flux. (b) The level of an enzyme allows maximum flux and perturbing the level slightly always decreases the flux.

The reason for the zero flux control for these several enzymes also remains elusive. One possible explanation is that glycolysis is already running at its maximum possible rate or the control is distributed over many enzymes [193]. Another conjecture is that the glycolysis is so optimized throughout evolution that the true FCCs were not measured due to optimal regulation of protein expression which somehow counteracted the effect of modulating an enzyme by reallocating the protein expression profile [164]. If this is true, the calculated rate is not the defined partial derivative because the concentrations of other enzymes are also functions of the concentration of the perturbed enzyme.

The conflicting results on the role of GAPDH from different studies also highlight the difficulty of studying flux control. In an earlier study, nearly full control of glycolytic flux by GAPDH in L. lactis Wg2 and NCDO2118 was found [189, 190] but zero control was found in L. lactis MG1363 [78]. One possible explanation is the intrinsic difference between the two strains as the GAPDH level was found to be two-fold higher in MG1363 compared to Wg2 [78]. Another possibility is the difference in experimental methods. GAPDH’s activity in Wg2 was only inhibited but not increased whereas both under- and over-expression of GAPDH were included in the study of MG1363. This can lead to contradictory estimation of FCCs.

Hierarchical regulation under different growth conditions
Another approach used to distinguish between hierarchical regulation and metabolic regulation was proposed by Ter Kuile and Westerhoff [194] by estimating coefficients for the two types of regulation. This approach is to a certain extent working in a reverse sense to the aforementioned approach which estimates the FCC of an enzyme by growing strains with different activities of the enzyme under the same condition and then measuring the changes in activities and fluxes. In contrast, the same strain is cultured under different growth conditions, for example, in chemostat at different dilution rates [194], or different starvation conditions [195]. Then the relative change in enzyme level, measured by enzyme assay, is divided by the relative change in fluxes to define the ‘hierarchical regulation.
coefficient’ (HRC). It is equal to one in the ideal case of pure hierarchical regulation. The ‘metabolic regulation coefficient’ (MRC) can then be computed by $1 - HRC$ to account for the change in fluxes not accountable by hierarchical regulation.

In *L. lactis*, several studies have been conducted using this approach. For instance, MG1363 has been grown in chemostat at a dilution rate of 0.1 h\(^{-1}\) at different pH, from 4.7 to 6.6 [196]. It was found that when taking the inhibitory effect of pH on enzyme activities into account, metabolic regulation was the dominant force controlling glycolytic flux. Also among the hierarchical regulation, post-transcriptional regulation of gene expression was found to be more prominent than transcriptional regulation by comparing change in mRNA transcript level with change in enzyme activity. Recently, MG1363 has been grown in chemostat at different dilution rates from 0.15 to 0.6 h\(^{-1}\) and transcriptomes, proteomes, enzyme activities have been quantified simultaneously [184]. Similar conclusions were reached. For dilution rates between 0.15 to 0.5 h\(^{-1}\), the changes in flux through most enzymes were predominantly caused by metabolic regulation instead of hierarchical regulation except for alcohol dehydrogenase (ADH) and possible pyruvate formate lyase (PFL) whose concentrations decreased as the dilution rate increased and the flux through mixed-acid fermentation pathway decreased. So these two enzymes probably controlled the switch between fermentation modes but not the glycolytic flux. Significant hierarchical regulation only occurred during transition from 0.5 to 0.6 h\(^{-1}\) in which the expression of several enzymes were found to have changed, probably due to the effect of carbon catabolite repression by the regulatory protein CcpA. Indeed, similar results of the lack of significant change in expression of glycolytic enzymes have also been observed in an accelerostat study on IL1403 in which the dilution rate increased very slowly from 0.1 to 0.6 h\(^{-1}\) to obtain different steady states [126]. Only PGM was found to show changes in expression.

**Metabolic regulation**

Metabolic regulation is not easy to discover because it usually involves interactions between an enzyme and metabolites other than substrates and products of that enzyme. Extensive *in vitro* enzyme characterization is required to identify possible effector metabolites and experiments for confirmation of *in vivo* regulatory roles can even be more difficult to design. As mentioned above some studies indicated that metabolic regulation was the main driving force for flux regulation and meanwhile many pieces of knowledge on particular regulatory relationships are available, nonetheless, a clear and integrative picture of how different types of metabolic regulation work together to explain most of the known experimental results still remains elusive.

**Negative feedback on PTS by FBP and inorganic phosphate**

One example of the metabolic regulation of glycolysis is the regulation of the phosphorylation of HPr protein (HPr/HPr-Ser-P) by FBP, ATP and inorganic phosphate (Pi) mentioned previously. Since HPr helps sugar uptake through PTS but HPr-Ser-P does not, high level of FBP due to high rate of sugar
uptake causing more HPr phosphorylated into HPr-Ser-P eventually slow down the sugar uptake, forming a negative feedback loop. This loop may help to stabilize the glycolytic flux, especially against sudden change in sugar availability [164]. The question of whether this negative feedback loop poses a bottleneck on maximum glycolytic flux, nevertheless, remains unanswered.

Feedforward on PYK by FBP
Beside the role in PTS, FBP has also been known to be an activator for PYK [197]. A kinetic study of glycolytic intermediates in glucose-pulse experiments using NMR found that FBP level rose to the peak during glucose uptake and started to drop after glucose was exhausted and until a certain low FBP level, PEP started to accumulate and remained at a high level during glucose starvation [159]. The authors proposed that the low FBP level reflecting low supply of glucose could serve as a way to preserve high PEP pool by inhibiting PYK which consumes PEP during sugar starvation for future rapid sugar uptake through PTS. Others, nonetheless, observed that such an activation relationship was also preserved in other organisms including those without PTS and remained conservative about the role of this FBP-PYK relation in glycolysis [164]. They suggested another possible role in which high FBP level could serve as a signal for PYK to remove the phosphoglycerate compounds in favour of a high flux through GAPDH which operated close to thermodynamic equilibrium and was thus sensitive to mass action.

Global cofactors: NADH/NAD⁺ ratio
Another interesting example is the control of glycolytic flux by cofactors levels NADH/NAD⁺ and ATP/ADP. NADH/NAD⁺ ratio was first proposed by Garrigues et al. to be the controlling factor of the glycolytic flux (as well as the mixed-acid shift, which is reviewed in section 1.5.4) in a study on L. lactis NCDO2118 exponentially growing on three sugars, glucose, galactose and lactose, with decreasing glycolytic fluxes [198]. This was based on several observations: (i) in vitro activity of GAPDH almost completely inhibited by NADH/NAD⁺ ratio higher than 0.05; (ii) NADH/NAD⁺ ratio positively correlated with glycolytic flux and as high as 0.08 on glucose (severe inhibition of GAPDH expected); (iii) high pools of metabolites upstream of GAPDH including FBP, GAP and dihydroxyacetone phosphate (DHAP) (suggesting the insufficient GAPDH activity to metabolize GAP). A later study by the same group on MG1363 found the same correlation between NADH/NAD⁺ ratio and glycolytic flux but the factors determining the ratio remained unknown [199].

Global cofactors: ATP/ADP ratio
Glycolytic kinetics in non-growing cells of L. lactis had then been studied using in vivo NMR by Neves et al. [158]. The kinetic model built in the study fitted with experimental data predicted that conversion of PEP into pyruvate by PYK is inhibited by high ATP surplus, i.e. high ATP/ADP ratio, which in turn inhibits NAD⁺ regeneration by LDH and bounds the glycolytic flux. This somehow provided an explanation for the positive correlation between NADH/NAD⁺ ratio and glycolytic flux.
observed in other studies [198]. Later NMR study by the same group focusing on the role of NADH and NAD⁺ found that GAPDH was able to sustain a flux as high as in the wild-type MG1363 in a LDH-knockout strain in which the NADH concentration was 1.5 mM while the inhibitory constant of NADH for GAPDH was found to be 0.4 mM [162]. The authors to a certain extent dismissed the control by GAPDH and NADH and proposed ATP, ADP and Pi as important regulatory metabolites in glycolysis. When interpreting these results, however, one should bear in mind that the cases that in vivo NMR studies were dealing with were non-growing L. lactis so it may not be directly applicable to the case where L. lactis is exponentially growing.

To test the in vivo role of ATP/ADP ratio in growing L. lactis, Koebmann et al. decreased the intracellular ATP/ADP ratio in MG1363 by expressing ATPase using a synthetic promoter library [79]. Surprisingly, for these strains growing exponentially on glucose, the glycolytic flux showed no significant change over a large range of ATP/ADP ratio from around 9 in the wild type to around 5 in the strain with the highest expression of ATPase. When these strains were in a non-growing state (achieved by resuspending cells in media without amino acids and vitamins), however, the glycolytic flux increased with ATPase activities until a level close to that in the growing wild type. These observations suggested the possibility that glycolysis was already at its maximal rate in the wild type. Another possible situation suggested by the authors was that although lowering the ATP/ADP ratio might stimulate glycolysis (e.g. by increasing the activities of kinases in the pay-off phase), it might eventually reduce the activity of PFK which becomes a bottleneck countering the effect, known as the risk of a ‘turbo design’ in which part of the desirable products are invested in the first place as input [200].

Other factors
Other sources of metabolic regulation thought to be important for regulating glycolytic flux include the inhibition of PYK by Pi [201], inhibition of PFK by PEP [202] and inhibition of GAPDH by NADH [198, 199]. Hoefnagel et al. integrated these three inhibitive relation and the activation of PYK by FBP in a single kinetic model with all rate equations and parameters adopted from literature without fitting [154]. The model succeeded in simulating the observed kinetic behaviour during glucose run-out experiments in other studies ([158]) including the rapid increase in PEP and Pi, decrease in ATP and slow depletions of FBP. The authors reasoned the following sequence of kinetic responses upon glucose depletion: (1) less Pi is retained in G6P, F6P and FBP; (2) Pi increases and inhibits PK; (3) PEP increases due to inhibited PK and no glucose for PTS and thus less pyruvate available; (4) Less substrate for LDH and thus NADH accumulates; (5) GAPDH is inhibited by NADH.
1.5.3 Pyruvate metabolism

Since *L. lactis* does not have a complete and functional tricarboxylic acid (TCA) cycle and cannot perform oxidative phosphorylation due to the inability to synthesize heme [28, 182], glycolysis as the major energy generating pathway must operate at a high speed. Since only a small fraction of the sugar metabolized is required for biomass production (usually <10%, e.g. [203]), most of the pyruvate flux ends up as fermentation products through different pathways. *L. lactis* is usually a homolactic fermenting LAB. As described in section 1.3.2, there are two branches of pathways producing mixed acids and flavour compounds respectively from pyruvate other than homolactic fermentation in *L. lactis* (Figure 1.2A).

**Homolactic fermentation**

In the metabolism of *L. lactis*, at the pyruvate branch point, lactate is produced via a single reaction catalysed by lactate dehydrogenase (LDH). LDH converts pyruvate into lactate and generates NAD\(^+\) from NADH at the same time. Several studies on the kinetics of LDH from *L. lactis* have been conducted [198, 199, 204]. From these studies, the K\(_m\) for pyruvate and NADH were found to be 1.5–3 mM and around 0.1 mM respectively. Also, FBP was found to be an activator of LDH with a micro-molar activation constant. Pi was considered to be an inhibition of LDH [204] but the inhibition was confirmed later to be negligible under physiological conditions [198, 199].

The expression of LDH is in general very robust as reflected by a fairly constant in vitro specific activity under different growth conditions, from strictly anaerobic, microaerobic to aerobic conditions [205, 206], or growth on different sugars [207]. LDH also appeared to be in huge excess compared to the in vivo flux it had in some studies (e.g. [207, 208]) and was shown to be able to sustain the same flux as in the wild type after about 40% reduction of its expression [191] (Table 1.1).

There are three genes in most genome-sequenced *L. lactis* strains, as observed from the UniProt database (http://www.uniprot.org), annotated as genes for LDH, named as *ldh*, *ldhB* and *ldhX* in both MG1363 and IL1403. *ldh* is the gene lying in the *las* operon following *pfk* and *pyk* and deemed to be the canonical gene for the LDH responsible for producing lactate in most cases [177]. The *las* operan has a cre-containing promoter and hence LDH, together with PFK and PYK, is transcriptionally activated by CcpA [188]. Whilst the role of *ldhX* remains unknown, *ldhB* has been found to be transcriptionally inactive when *ldh* was present and the transcription of *ldhB* was recovered with a significant capability of lactate production in an *ldh*-knockout strain after adaptive evolution [209]. The LDH synthesized from *ldhB* has later been characterized to have kinetic properties that depend heavily on the pH level and are quite different from the canonical LDH, e.g. activation by Pi at low pH [210].
**Mixed-acid fermentation**

Under anaerobic conditions, when the sugar uptake rate is constrained to be slow, for instance, by a low dilution rate in chemostat experiment (e.g. [40]), or by growth on slowly fermentable sugars such as maltose or galactose (e.g. [211, 212]), a significant fraction of pyruvate flux would be directed to the production of formate, acetate and ethanol, termed mixed-acid fermentation (Figure 1.6). The first step is the conversion of pyruvate into acetyl-coenzyme A (Ac-CoA) by either pyruvate formate lyase (PFL) or pyruvate dehydrogenase complex (PDHc). Ac-CoA can then be converted into acetyl-phosphate (Ac-P) by phosphotransacetylase (PTA) followed by acetate production by acetate kinase (ACK) which produces one ATP. Alternatively, Ac-CoA can be reduced into first acetaldehyde and then ethanol by a single enzyme alcohol dehydrogenase (ADH).

![Figure 1.6. Mixed-acid fermentation pathway producing formate, acetate and ethanol.](image)

Products excreted are shaded in green. Enzymes are in bold and italic style. Pi, inorganic phosphate; CoA, coenzyme A; PDHc, pyruvate dehydrogenase complex; PFL, pyruvate formate lyase; ADH, alcohol dehydrogenase; PTA, phosphotransacetylase; ACK, acetate kinase.

Of all these enzymes, two possible elementary routes are available for converting pyruvate into end products with proper redox balance. The generation of one pyruvate is accompanied by one NADH being reduced from NAD\(^+\). Redox balance can therefore be kept by consuming one pyruvate with simultaneous oxidation of one NADH. The first route is one unit of pyruvate converted by PFL into Ac-CoA, followed by half of the Ac-CoA entering PTA-ACK pathway to produce acetate and the other half forming ethanol via ADH. The formate-acetate-ethanol ratio in this case is 2:1:1, which is the most common case observed in many studies under anaerobic conditions. This route generates 0.5 more ATP per pyruvate being catabolized. The second possible route is the production of one unit of CO2 and ethanol by PDHc and ADH but it has not been reported under anaerobic conditions due to low level and low activity of PDHc. On the contrary, under aerobic conditions, due to the expression of the NADH oxidase (NOX) and its activation by oxygen, PDHc activity coupled with acetate production with the aid of NOX for regeneration of NAD\(^+\) has been observed in different studies [163, 205, 206].

**Pyruvate formate-lyase**

PFL in *L. lactis* has been manifested to be extremely vulnerable to exposure to oxygen as in *E. coli* [213]. It has been reported that the *K*\(_m\) for pyruvate of PFL in MG1363 was 0.4 mM, lower than that
of LDH and PFL was inhibited by triosephosphate pool with inhibitory constant equal to 0.7 mM and 1 mM for GAP and DHAP respectively [199]. The expression of PFL was usually found to be significantly induced for growth under anaerobic conditions at a low dilution rate in chemostat [206, 208] or on slowly fermentable sugars such as galactose [213].

The expression of pfl of MG1363, the gene encoding PFL, has been characterized [214]. A monocistronic mRNA for pfl was found. There are two sequences upstream of pfl with significant homology to the consensus of the global regulator FNR of E. coli which activates the transcription of pfl in E. coli anaerobically [215]. One of them even overlaps with the -35 promoter region. The genes of two proteins in L. lactis MG1363 homologous to FNR, FlpA and FlpB, were studied [216]. An flpA mutant was found to produce significantly more formate when growing on glucose or galactose but much less when growing on maltose compared to the wild type whereas no significant effect was observed for an flpB mutant and a double mutant. FlpA and FlpB have later been purified for studying their binding to DNA in vitro and FlpA was confirmed to recognize FNR site but appeared to work in the opposite direction to FNR because opposite to FNR, the binding was abolished by an iron-sulphur cluster and restored by exposure to air [217]. A later study on the expression of FlpA and FlpB indicated that transcription of flpB was much lower than flpA and was induced under anaerobic conditions while the transcription of flpA was not oxygen-responsive but required the presence of both FlpA and FlpB. Until now, there is no clear conclusion on the transcriptional regulation of pfl at the molecular level.

Pyruvate dehydrogenase complex

PDHc, another route for Ac-CoA formation from pyruvate, unlike PFL, produces one CO₂ and meanwhile generates one NADH from NAD⁺. Kinetic studies on the kinetics of PDHc indicated that it was highly sensitive to high NADH/NAD⁺ ratio [218, 219]. 50% and complete inhibition were observed at ratios of about 0.033 and 0.135 respectively in the studies, suggesting complete inhibition under strict anaerobiosis.

The expression of PDHc has usually been associated to growth under aerobic conditions [198, 199]. Under different growth conditions, however, conflicting results have been reported. PDH activity in MG1363 under anaerobic conditions has been found to be absent when growing on lactose [199]; or present at low but observable levels when growing either at 0.1 h⁻¹ dilution rate or on glucose in batch (growth rate up to 0.9 h⁻¹) [205, 206]. It has also been found to increase >10-fold with aeration under microaerobic condition (air-to-nitrogen ratio from 0 to around 2) at a constant dilution rate of 0.1 h⁻¹ [206] or without significant change in anaerobic, microaerobic or aerobic conditions for batch growth on glucose with high growth rates (>0.78 h⁻¹ for all three cases) [205]. The well-known requirement of lipoate as a cofactor for PDHc is worthwhile to mention because L. lactis and some other LAB such
as *Enterococcus faecalis* lack the ability to synthesize lipoate and thus addition of lipoic acid into the growth media is necessary for PDHc activation [220].

**Phosphotransacetylase, acetate kinase and alcohol dehydrogenase**

For *L. lactis*’s PTA, ACK and ADH, little kinetic characterization has been performed except a recent study on PTA and ACK [221] and a very recent study on ACK isozymes [152]. PTA has been found to be inhibited by FBP and diacetyl [221]. In addition to these two compounds, ACK has also been found to be inhibited by GAP and PEP at concentrations within physiological ranges [152, 221]. The expression of these three enzymes was found to be in generally higher at lower growth rates in growth experiments on galactose, maltose or at low dilution rates in chemostat [179, 184, 207, 221].

**Flavour compound production**

Production of diacetyl, acetoin and 2,3-butanediol (2,3BTD) by *L. lactis*, all being C₄ flavour compounds, is very important in food industry. It has been associated with citrate fermentation [222, 223] as well as aerobic growth on sugars [224, 225]. In this pathway, first α-acetolactate synthase (ALS) catalyses the decarboxylation of two pyruvate to form one α-acetolactate, which upon further decarboxylation becomes one acetoin. 2,3BTD is produced from acetoin by 2,3BTD dehydrogenase (BTDDH). α-Acetolactate also reacts non-enzymatically with oxygen to form diacetyl which can be reduced to acetoin by acetoin dehydrogenase (ACETDH).

![Figure 1.7. Flavour compound production from pyruvate.](image)

Products excreted are shaded in green. Enzymes are in bold and italic style. The reaction producing diacetyl from α-acetolactate is non-enzymatic. ALS, α-acetolactate synthase; ALDC, α-acetolactate decarboxylase; ACETDH, acetoin dehydrogenase; BTDDH, 2,3-butanediol dehydrogenase.

The kinetics of ALS has been characterized for a *L. lactis* strain and showed strong positive cooperativity for pyruvate (Hill coefficient = 2.4) and a very high $K_m$ for pyruvate equal to 50 mM [219]. Another study on another strain found a similar value of 30 mM [226]. In the study ALDC was also tested. Positive cooperativity was again observed with Hill coefficient equal to 1.89. The $K_m$ for α-acetolactate estimated was 60 mM. For ACETDH and BTDDH, in a *L. lactis* strain, an enzyme capable of catalysing both reactions was isolated and $K_m$ for acetoin and diacetyl were estimated to be 60 $\mu$M and 3.6 mM respectively [227].
Beside the precursor of several flavour compounds, α-acetolactate is also a metabolite connecting carbon and amino acid metabolism since it is a precursor of the branched-chain amino acids (BCAAs) leucine, isoleucine and valine (provided that a particular *L. lactis* strain is able to synthesize any of them). The expression of ALDC was found to be induced by the presence of leucine in media and inhibit growth of *L. lactis* NCDO2118 under valine starvation while disrupting the gene for ALDC succeeded in growth recovery under such a condition [228]. ALDC was attributed to the exhaustion of intracellular α-acetolactate so that valine could not be synthesized. The regulation of ALDC expression has later been investigated at transcriptional and translational levels [229].

In many studies, Fluxes through this flavour-producing pathway and acetate production increased concomitantly when *L. lactis* was granted extra ability to grow aerobically by either overexpressing NOX [56, 153, 163, 221] or supplementing heme or its precursor such as protoporphyrin IX for respiration [225, 230–232]. The reason for the redirection has not yet been fully elucidated. Comparison at the protein level between heme-dependent respiratory and static conditions found that among glycolytic enzymes, the expression of phosphoglycerate mutase (PGM) and one isoform of GAPDH decreased markedly during respiration whereas the expression of PDHc and ALS was upregulated [233]. This may have some control on the distribution of pyruvate flux. For the case of NOX overexpression, the lower $K_m$ for NADH of NOX compared to that of LDH and ADH has been ascribed to the phenomenon observed [163]. NOX outperforms LDH and ADH regarding NADH competition. This is compatible with the observation that other than the flavour-producing pathway, only acetate is produced via PDHc, PTA and ACK without formate and much ethanol production. PDHc may have high activity due to the expected low NADH/NAD$^+$ ratio. Meanwhile, PFL is inactivated under aerobic conditions [213].

### 1.5.4 Flux regulation of mixed-acid shift

The metabolic shift between homolactic fermentation and mixed-acid fermentation is intriguing in the sense that *L. lactis* prefers production of lactate with two ATP generated per glucose to production of mixed acids with three ATP generated per glucose. Such a shift from high-yield to low-yield metabolism, or equivalently use of energetically inefficient pathway, has been called overflow metabolism (usually for bacteria) [234, 235], or ‘crabtree effect’ for yeast and other organisms [236, 237], or ‘Warburg effect’ for tumor cells [238]. Unravelling the mechanism and the evolutionary advantage of this phenomenon is thus a very important biological question. In *L. lactis*, from the large amount of experimental evidence accumulated over the years, different factors and mechanisms have been proposed. They are divided into three aspects, protein level, metabolic level and environmental conditions. The possible evolutionary advantages are also discussed along the subsection.
**Protein level**

Elevated expression levels of enzymes in the mixed-acid branch have been reported when *L. lactis* grows at low dilution rates or on slowly fermentable sugars as described in the previous subsection. A study comparing the proteomes of *L. lactis* growing on glucose, growing on maltose and resting on maltose also revealed that expression of glycolytic enzymes was in general slightly down-regulated (<2-fold) for growth on maltose compared with growth on glucose while proteomes for *L. lactis* growing or resting on maltose were in general very similar [179].

Control analysis on different enzymes further emphasized the role of protein expression on the shift. PFL level in MG1363 has been modulated and strong control on formate flux was observed for growth on galactose (FCC > 1, estimated from the data in the article) but only weak control was found for growth on glucose (FCC ≈ 0.26) [239]. Modulation of LDH and PYK levels showed that there are a strong negative and positive control of formate flux and acetate flux by LDH and PYK respectively (FCCs close to -1 and 1) when growing on glucose [77, 191]. This means decreasing and increasing the expression of LDH and PYK respectively did increase the pyruvate flux distributed to mixed-acid fermentation. The same directions of control by LDH and PYK remained for growth on maltose, however, with much smaller FCCs (around -0.2 and 0.2) [240]. In another study, inhibition of the activity of GAPDH of *L. lactis* growing on lactose to different extents was achieved by adding increasing concentrations of iodoacetate [190]. A marked shift from mixed-acid to homolactate fermentation accompanied by decreasing growth rates was observed.

**Beyond protein level**

From the above experimental evidence, it is tempting to conclude that the mixed-acid shift is caused by the change in gene expression level but this was generally not the conclusion of most studies because there are other experimental results in which the shift could not be ascribed to change in gene expression level. For example, in the modulation of LDH and PYK levels [77, 191], the flux through the two reactions did not change apparently but they changed the formate flux, so it is reasonable to hypothesize that the change was caused by interactions between enzymes and metabolite, i.e. events at the metabolic level.

As aforementioned, proteomes for *L. lactis* growing or resting on maltose were in general very similar and no significant change was observed for glycolytic enzymes and enzymes in pyruvate metabolism [179]. However, while the growing cells produced mixed acids from maltose, the resting cells were found to produce lactate only. The glycolytic flux in resting cells was about 3-fold lower than that in growing cells, similar to the case of growth on glucose [79]. This study suggested that the flux distribution did not solely depend on the proteome. A similar conclusion has been drawn in the aforementioned study which calculated hierarchical and metabolic regulation coefficients by growing
MG1363 at different dilution rates [184]. Only the flux of ADH and possibly PFL appeared to be regulated by the protein expression level in conditions with different fermentation profiles.

A final example is a set of shift-up and shift-down experiments [208]. In the shift-up experiment, *L. lactis* MG1363 had been cultured to reach a steady state at a low dilution rate (0.1 h⁻¹) which produced mixed acid and the dilution rate was then instantaneously elevated to 0.5 h⁻¹. Switch to homolactic fermentation almost occurred at once, contrary to the hypothetical case of sole dependence on gene expression level in which a delayed switch should be observed due to the time required for regulation of expression. The shift-down experiment was conducted in a similar fashion but a reverse change of dilution rates. Unlike shifting up, in the shift-down experiment a profile of mixed-acid fermentation was gradually built up, suggesting the plausible dependence of the switch on gene expression. The results indicated the existence of a rapid regulatory mechanism for the switch from mixed-acid to homolactic fermentation, which is probably able to sense the glycolytic flux.

**Metabolic level**

**LDH activation by FBP**

In fact, allosteric regulation of the enzymes in pyruvate metabolism of *L. lactis* has long been proposed as a mechanism for the mixed-acid shift. Activation of LDH by FBP has first been proposed to regulate the shift since FBP is the only activator of LDH for growth on glucose and FBP level was found to be correlated with glycolytic flux [40]. The theory was however turned down convincingly by two later studies conducted by Garrigues *et al* on *L. lactis* NCDO2118 and MG1363 respectively [198, 199]. In these studies, full activation concentrations of LDH by FBP were measured to be 2 – 4 μM whereas the intracellular FBP concentration measured ranged from 0.3 mM to 118 mM, >50-fold higher than the activation concentration even for the case of the lowest FBP level.

**NADH/NAD⁺ ratio**

Garrigues *et al*, in the same studies, instead proposed another mechanism for the switch which is caused by the change in NADH/NAD⁺ ratio [198, 199]. Since GAPDH was found to be very sensitive to NADH/NAD⁺ ratio, triose-phosphate (DHAP and GAP) accumulates when the ratio are high. Triose-phosphate then inhibits the activity of PFL and meanwhile LDH has a high activity because of high NADH/NAD⁺ ratio, causing homolactic fermentation. When NADH/NAD⁺ ratio is low, PFL is therefore not inhibited and the activity of LDH is lower. Combining with the higher affinity of PFL for pyruvate, mixed-acid fermentation is achieved. In the studies, positive correlation between NADH/NAD⁺ ratio and glycolytic flux was observed. The model can therefore describe the homolactic and mixed-acid fermentation occurring at high and low glycolytic flux respectively. The inhibition of PFL by DHAP and GAP (first observed in *Streptococcus mutans* [241]) was confirmed *in vitro*. Their intracellular levels measured for exponentially growing cells in different conditions
also matched the hypothesis: being lower than the inhibitory constants in mixed-acid fermenting cells and being significantly higher than the constants in homolactic fermenting cells. This proposed mechanism is able to explain the rapid switch from mixed-acid to homolactic fermentation during the aforementioned shift-up experiment [208].

**ATP/ADP surplus**

The role of NADH/NAD$^+$ ratio was consistent with experimental evidences. How this ratio depends on substrate availability or glycolytic flux, however, was left unexplained. In fact, in the second of the two studies mentioned in the last paragraph, Garrigues *et al* proposed a more generalized theory stating that the switch depends on the balance of anabolic demand of ATP which is generated by catabolism [199]. It was based on the additional experiments in the study that compared the growth and fermentation pattern of *L. lactis* NCDO 2118 on galactose in media with different amino acid and vitamin contents. It was found that with fewer amino acid and vitamin contents, growth and glycolytic flux decreased as expected whereas fermentation switched from mixed-acid to homolactic unexpectedly. The authors thus proposed that under conditions of catabolic excess (growth on glucose) or anabolic limitation (growth in media with fewer amino acid and vitamin contents) in which ATP/ADP was predicted to be high, homolactic fermentation prevailed whereas under conditions of catabolic limitation and anabolic excess (growth on galactose in rich media), mixed-acid fermentation took place to maximize ATP. The concrete link between this theory and NADH/NAD$^+$ ratio was however not provided.

The proposed role of ATP/ADP surplus in the mixed-acid shift actually coincides with its role in glycolytic flux as discussed in section 1.5.2. The glycolytic flux in resting cells could be stimulated by lowering ATP/ADP ratio using ATPase, showing that high ATP/ADP level limited the flux [79]. A later study provided evidences supporting this assumed role of ATP/ADP [242]. For resting cells of *L. lactis* fermenting maltose which produced lactate only, reducing ATP/ADP ratio by adding monensin or valinomycin (ionophores known to induce ATPase activity) switched the fermentation pattern to mixed acids. Inhibition of ADH by the ATP and ADP pool was proposed as the caused since the measured level of ATP and ADP in resting cells brought 95% inhibition of ADH activity *in vitro*. Inhibition to a less extent by ATP and ADP was also observed for GAPDH and LDH.

The authors in the same study also compared growth on maltose with nitrogen limitation or growth arrested by chloramphenicol. In either case the decreasing growth rate was negatively correlated to increasing homolactic fermentation and the authors claimed that the change in fermentation pattern was due to change in growth rate instead of nitrogen content available. We think, however, this can be overly interpreted because chloramphenicol has long been known to inhibit bacterial growth by specifically inhibiting the assimilation of amino acids and therefore protein synthesis [243]. Thus
chloramphenicol addition and nitrogen limitation may actually have very similar effects on metabolism.

There is a final remark about the hypotheses of NADH/NAD$^+$ ratio and ATP/ADP surplus. The former has experimental support of the proposed mechanism but how it is related to the environment to which a cell is subject remains unclear. The latter, namely ATP/ADP surplus, provides a rationale of the switch and is also supported by some experimental evidence, but how it is related to the exact mechanism like the NADH/NAD$^+$ ratio remains unknown. Moreover, there have not yet been convincing experiments carried out as a proof of the proposed mechanism.

**Environmental conditions**

The most well-established environmental condition causing the occurrence of the mixed-acid shift is the catabolic limitation by either growth on slowly fermentable sugars or low availability of sugars in situation such as chemostat at low dilution rates. For growth under these conditions, when other anabolic limitations are active, such as less amino acid or vitamin availability as described previously, fermentation appears to shift back from mixed-acid to homolactic fermentation.

Environmental factors other than substrate availability including pH and temperature have also been found to influence the fermentation pattern. A study on *L. lactis* growing on maltose at different pH and temperatures revealed that *L. lactis* became more homolactic at lower pH compared to pH 6.5 and at higher temperature compared to 30 °C [244]. The authors suggested the inactivation of PFL at low pH found in *Streptococcus mutans* [241] and the increased ATP spent on maintenance at low pH or high temperatures as the reasons for the observed shift.

**Evolutionary advantage**

Despite the wealth of the knowledge regarding the metabolic shift (though an integrated consensus of the relative importance is still not available), the evolutionary advantage, as an 'ultimate regulatory force’, of switching between a low-yield (homolactic fermentation) and high-yield (mixed-acid fermentation) metabolism remains far from clear. An advantage that has been sometimes cited in the literature is to inhibit the growth of other organisms by acidification of the environment due to fast lactate production but Teusink et al. argued that it may not be valid [164]. They reasoned that ‘cheater’ mutants that can grow faster by switching to high-yield metabolism and meanwhile survive the acidic environment should have evolved according to this hypothesis but it has never been observed for *L. lactis*.

Another hypothesis is the tradeoff between protein cost and ATP generation. On the one hand, mixed-acid fermentation generates one more ATP per hexose catabolized. On the other hand, mixed-acid fermentation requires three enzymes to be expressed while homolactic fermentation requires one only. Since protein expression also costs ATP and the rate of the cost should increase with growth rate due
to dilution by cell division, different strategies at different growth rates can be a result of growth rate optimization considering the ATP gain and protein cost of a pathway. This cannot be confirmed without a quantitative model. A very first model addressing this type of shift was the self-replicator model consisting of minimal elements for self-replication including one transport protein, ribosome, one anabolic pathway and two metabolic pathways, one with lower ATP yield but higher catalytic efficiency $k_{\text{cat}}$ and one with higher ATP yield but lower $k_{\text{cat}}$ [245]. The model did succeed in predicting a shift from high-yield metabolism to low-yield metabolism upon the increase of growth rate. A later attempt based on a similar rationale employed the genome-scale metabolic network of *L. lactis* IL1403 and applied flux balance analysis (FBA) with the constraints of molecular crowding to model the mixed-acid shift successfully [246]. Molecular crowding here refers to explicit consideration of the distribution of protein with the total amount being finite over all metabolic reactions. Each reaction has a maximum velocity equal to $k_{\text{cat}}$ multiplied by the protein concentration assigned to that reaction.

This hypothesis, though logically sound, is yet to be verified experimentally. The very recent study mentioned several times previously conducted on MG1363 growing at different dilution rates with the aid of multi-omics techniques has been designed to test this hypothesis [184]. It was found that only the flux of ADH and possibly PFL was controlled at gene expression level and many other proteins appeared to be excessive than the necessary amounts to carry fluxes. This to a certain extent suggested that there was not a case in which tight control of gene expression was exerted to satisfy metabolic requirement minimally for protein-cost saving. One may however argue that while the excessive amount of most enzymes may bring other important evolutionary advantage, the gain from saving the cost of one or two enzymes, though little, can become significant in evolutionary sense. For one point that can be sure about, nonetheless, this hypothesis cannot explain the existence of a rapid shift mechanism because it is not related to the protein expression level.

Another evolutionary point of view that can be taken into consideration is the evolutionary game theory that has drawn increasing interest in the last decade [247–249]. It advocates defining objective function for optimization by considering the interactions between the organism, environment and other competing organisms (other players) as in game theory instead of simply maximizing growth rate. In this way of thinking, the objective function may no longer be the growth rate defined in genome-scale metabolic networks calculated by FBA which is basically an yield optimization but could be true rate optimization including the use of metabolically inefficient pathway. Or it is even possible to maximize the substrate depletion rate, regardless of the growth yield that such a strategy brings, to minimize the growth of other players.

A final remark about inferring the evolutionary advantage is that although the same phenomenon of the shift of fermentation modes was observed at different condition, one should not simply draw the
same conclusion regarding the cause and effect. Under the conditions that \textit{L. lactis} has long been adapted to, it is reasonable to think that the shift represents a way to optimize the growth rate or other fitness function. But under other conditions, the observed shift may be an effect of the evolved mechanism and advantage of the shift under those conditions may not be found. To properly study the evolutionary advantage of the shift, the evolutionary history, habitats and the ecological community of \textit{L. lactis} should be put into considerations together.

\textbf{1.6 Objectives and strategies}

In spite of the many efforts put on investigating the glycolysis and pyruvate metabolism of \textit{L. lactis} in the past, especially about the mixed-acid shift, many elements in the proposed theory regarding the shift are required to be elucidated for a more comprehensive picture. Also, the widely studied laboratory strain \textit{L. lactis} subsp. \textit{lactis} IL1403 was found to remain homolactic even when growing on slowly fermentable sugars such as galactose [207]. By studying the fermentation behaviour of the strain and comparison with strains able to switch, further insights can be gained.

In this study, in response to the factors affecting the switching behaviour as reviewed in the previous section, namely gene expression level, metabolic level and environmental conditions, we aimed to study the mixed-acid fermentation of \textit{L. lactis} subsp. \textit{cremoris} MG1363 using the following strategies: (i) genetic perturbation, (ii) cofactor perturbation and (iii) growth at varying nutrient availability, in terms of amino acids. Growth behaviour on glucose and maltose of MG1363 respectively is usually compared and it is occasionally contrasted with IL1403. Maltose is chosen because it was observed to cause the most prominent mixed-acid shift in \textit{L. lactis} MG1363.

\textbf{1.6.1 Genetic perturbation}

The roles of some mixed-acid-fermenting enzymes on mixed-acid fermentation remain to be elucidated, such as PFL, PTA, ACK and ADH, especially ACK which appears to be encoded by two different uncharacterized genes in many \textit{L. lactis} strains. 1.7 reports the study of transcriptional activity of these enzymes and the modes of transcriptional regulation in the two widely studied laboratory strains, MG1363 and IL1403. The former exhibits mixed-acid shift to a large extent while the latter does not. Chapter 3 reports the results of modulation of PFL levels in MG1363 and IL1403 and the comparison. In Chapter 3, characterization of the two genes encoding ACK in \textit{L. lactis} MG1363 is presented, including the transcription structure, enzyme kinetics and their physiological roles.

\textbf{1.6.2 Cofactor perturbation}

The role of NADH/NAD$^+$ in anaerobic growth of \textit{L. lactis} was supported by comparing growth on different sugars or in different media [198, 199]. The question, however, that whether it truly controls
the pyruvate metabolism, remains unexamined to a large extent. Chapter 5 attempts to study the effect of perturbing NADH/NAD⁺ under different conditions.

1.6.3 Nutrient availability

The effect of availability of nutrients other than sugars, especially amino acids, on the mixed-acid shift has only been studied roughly in [198, 199, 242]. Chapter 6 briefly summarizes some experimental findings in this respect from a bachelor project initiated by us and especially a general computational method developed and applied to look into the amino acid metabolism in the genome-scale metabolic network of L. lactis MG1363 [145] which has been published and reprinted in Appendix A.

1.7 References

3. Orla-Jensen S (1919) Lactic Acid Bacteria (Fred Host and Son, Copenhagen) Available at: http://books.google.dk/books?id=mxOmgECAAAJ.


178. Neves AR et al. (1995) A system to generate chromosomal mutations in genetic regulation and enzyme con


Chapter 1. Background


Chapter 2. Transcriptional activity of mixed-acid genes

2.1 Introduction

Many strains of Lactococcus lactis (L. lactis), including both subsp. cremoris and subsp. lactis, have been found to exhibit a metabolic shift from homolactic to mixed-acid fermentation when transiting from a fast fermentable sugar to one that is more slowly fermentable. In the former fermentation mode lactate is the main product whereas in the latter mode significant amounts of formate, acetate and ethanol are also produced. Table 2.1 summarizes the ability to exhibit a shift in fermentation mode for different strains that have been studied in the past. Most of the studied strains were found to exhibit the shift but two strains, L. lactis subsp. lactis ML8 and the widely studied laboratory strain L. lactis subsp. lactis IL1403, appeared to be unable to shift, though there has once been a study reporting the mixed-acid fermenting behavior of IL1403 when growing on maltose [1].

In this study, IL1403 and another extensively studied strain L. lactis subsp. cremoris MG1363, which exhibits the mixed-acid shift, were compared with respect to growth physiology, relative transcriptional activities of mixed-acid genes and elements of transcription regulation.

Table 2.1. Phenotype of L. lactis strains regarding shift in fermentation mode.

<table>
<thead>
<tr>
<th>Subspecies</th>
<th>Strain</th>
<th>Shift to mixed-acid fermentation at low sugar uptake rates</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>lactis</td>
<td>ML₁</td>
<td>Yes</td>
<td>[2, 3]</td>
</tr>
<tr>
<td>lactis</td>
<td>7962</td>
<td>Yes</td>
<td>[2]</td>
</tr>
<tr>
<td>lactis</td>
<td>ML₈</td>
<td>No</td>
<td>[2]</td>
</tr>
<tr>
<td>cremoris</td>
<td>Eₚ</td>
<td>Yes</td>
<td>[2]</td>
</tr>
<tr>
<td>cremoris</td>
<td>HP</td>
<td>Yes</td>
<td>[2]</td>
</tr>
<tr>
<td>lactis</td>
<td>65.1</td>
<td>Yes</td>
<td>[4, 5]</td>
</tr>
<tr>
<td>lactis</td>
<td>NCD0 2118</td>
<td></td>
<td>[6]</td>
</tr>
<tr>
<td>lactis</td>
<td>ATCC 19435</td>
<td></td>
<td>[7]</td>
</tr>
<tr>
<td>cremoris*</td>
<td>MG1363</td>
<td>Yes</td>
<td>[8]</td>
</tr>
<tr>
<td>lactis</td>
<td>IL1403</td>
<td>No</td>
<td>[9]</td>
</tr>
</tbody>
</table>

*MG1363 is known to have cremoris genotype but lactis phenotype.

2.2 Materials and methods

2.2.1 Bacteria strains and plasmids

The two L. lactis strains, IL1403 and MG1363, were plasmid-free derivatives of L. lactis CNRZ157 (or IL594) [10] and L. lactis NCD0712 [11] respectively. Other L. lactis strains used were derived from the two strains. The plasmids pLB65 expressing the T901-1 integrase was used for mediating the site-specific integration into the chromosomal attachment site, attB, of pLB85, which contains a promoterless gusA reporter gene and a gene encoding erythromycin resistance [12]. E. coli ABLE-K (Stratagene) [13] was used for cloning purposes. Primers, plasmids and strains used or constructed in this study are listed in Table 2.2, Table 2.3 and Table 2.4 respectively.
Table 2.2. Primers used in the study.

<table>
<thead>
<tr>
<th>Primer</th>
<th>RS</th>
<th>Amplified region</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3f</td>
<td>SpeI</td>
<td>upstream region of pfl in IL1403</td>
<td>atcgaacttgaGAGAAGATGAGAAGAC</td>
</tr>
<tr>
<td>3r</td>
<td>SalI</td>
<td>upstream region of pfl in IL1403</td>
<td>atcgaactcgacTAAATGTTCAAGATATTTTCGTT</td>
</tr>
<tr>
<td>4f</td>
<td>XbaI</td>
<td>upstream region of adhE in IL1403</td>
<td>atcgaacttgaTGCTGAGCAAAATACGA</td>
</tr>
<tr>
<td>4r</td>
<td>PstI</td>
<td>upstream region of adhE in IL1403</td>
<td>atcgaactcgacTAAACGGGCTTTTATTTG</td>
</tr>
<tr>
<td>5f</td>
<td>XbaI</td>
<td>upstream region of eutD in IL1403</td>
<td>atcgaacttgaTGGTTGATACAGATGACGA</td>
</tr>
<tr>
<td>5r</td>
<td>PstI</td>
<td>upstream region of eutD in IL1403</td>
<td>atcgaactcgacTTACAGTGATCACAAGATTTCC</td>
</tr>
<tr>
<td>6f</td>
<td>XbaI</td>
<td>upstream region of ackA1 in IL1403</td>
<td>atcgaacttgaTTTTCCTACATTTGTC</td>
</tr>
<tr>
<td>6r</td>
<td>PstI</td>
<td>upstream region of ackA1 in IL1403</td>
<td>atcgaactcgacTTATTTTATGAGAGGAGCC</td>
</tr>
<tr>
<td>7f</td>
<td>XbaI</td>
<td>upstream region of ackA2 in IL1403</td>
<td>atcgaacttgaCCTAGCAACATTTGTC</td>
</tr>
<tr>
<td>7r</td>
<td>PstI</td>
<td>upstream region of ackA2 in IL1403</td>
<td>atcgaactcgacTTATTTTATGAGAGGAGCC</td>
</tr>
<tr>
<td>8f</td>
<td>XbaI</td>
<td>upstream region of pfl in MG1363</td>
<td>atcgaacttgaAGATGTATGGACCAAGAGAG</td>
</tr>
<tr>
<td>8r</td>
<td>PstI</td>
<td>upstream region of pfl in MG1363</td>
<td>atcgaactcgacTTATTAAACCATTTCAAGC</td>
</tr>
<tr>
<td>9f</td>
<td>SpeI</td>
<td>upstream region of adhE in MG1363</td>
<td>atcgaacttgaTGATATTACGTACGGCTTT</td>
</tr>
<tr>
<td>9r</td>
<td>SalI</td>
<td>upstream region of adhE in MG1363</td>
<td>atcgaactcgacTTATGAGAGGCTTTTATTTG</td>
</tr>
<tr>
<td>10f</td>
<td>SpeI</td>
<td>upstream region of eutD in MG1363</td>
<td>atcgaacttgaAGATATTACGTACGGCTTT</td>
</tr>
<tr>
<td>10r</td>
<td>PstI</td>
<td>upstream region of eutD in MG1363</td>
<td>atcgaactcgacTTACAGTGATCACAAGATTTCC</td>
</tr>
<tr>
<td>11f</td>
<td>XbaI</td>
<td>upstream region of ackA1 in MG1363</td>
<td>atcgaacttgaAGATATTACGTACGGCTTT</td>
</tr>
<tr>
<td>11r</td>
<td>PstI</td>
<td>upstream region of ackA1 in MG1363</td>
<td>atcgaactcgacTTATGAGAGGCTTTTATTTG</td>
</tr>
<tr>
<td>12f</td>
<td>XbaI</td>
<td>upstream region of ackA2 in MG1363</td>
<td>atcgaacttgaAGATATTACGTACGGCTTT</td>
</tr>
<tr>
<td>12r</td>
<td>PstI</td>
<td>upstream region of ackA2 in MG1363</td>
<td>atcgaactcgacTTATGAGAGGCTTTTATTTG</td>
</tr>
<tr>
<td>CSO50</td>
<td>BamHI</td>
<td>verify chromosomal integration of pLB85</td>
<td>ggaaggattcCCTAGATCTACAGTTAYC</td>
</tr>
<tr>
<td>CSO263</td>
<td>/</td>
<td>verify chromosomal integration of pLB85</td>
<td>CGCGTACAGAATATT</td>
</tr>
</tbody>
</table>

Each amplified region includes a short coding sequence (CDS) (20 – 50 bp) from the gene and a stop codon.

RS, restriction site.

Table 2.3. Plasmid used or constructed in the study.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Antibiotic resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>pLB65 [12]</td>
<td>Containing orf1 expressing phage TP901-1 integrase</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85 [12]</td>
<td>Reporter vector containing a promoterless gusA gene and an attP site for integration into attB site on chromosomes in the presence of the TP901-1 integrase</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-Ilpfl</td>
<td>pLB85 containing pfl upstream region of IL1403 obtained with primer 3f, 3r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-IladhE</td>
<td>pLB85 containing adhE upstream region of IL1403 obtained with primer 4f, 4r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-IleutD</td>
<td>pLB85 containing eutD upstream region of IL1403 obtained with primer 5f, 5r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-IlackA1</td>
<td>pLB85 containing ackA1 upstream region of IL1403 obtained with primer 6f, 6r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-IlackA2</td>
<td>pLB85 containing ackA2 upstream region of IL1403 obtained with primer 7f, 7r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-MGpfl</td>
<td>pLB85 containing pfl upstream region of MG1363 obtained with primer 8f, 8r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-MGadhE</td>
<td>pLB85 containing adhE upstream region of MG1363 obtained with primer 9f, 9r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-MGeutD</td>
<td>pLB85 containing eutD upstream region of MG1363 obtained with primer 10f, 10r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-MGackA1</td>
<td>pLB85 containing ackA1 upstream region of MG1363 obtained with primer 11f, 11r</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85-MGackA2</td>
<td>pLB85 containing ackA2 upstream region of MG1363 obtained with primer 12f, 12r</td>
<td>Cam</td>
</tr>
</tbody>
</table>

Cam, chloramphenicol; Amp, ampicillin; Erm, erythromycin.
Table 2.4. Strains used or constructed in the study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL1403</td>
<td>A plasmid-free strain derived from L. lactis subsp. diacetylactis CNRZ157</td>
</tr>
<tr>
<td>SC072</td>
<td>IL1403 containing pLB65</td>
</tr>
<tr>
<td>SC072/ILpfl</td>
<td>SCO72 with pLB85-ILpfl integrated</td>
</tr>
<tr>
<td>SC072/ILeutD</td>
<td>SCO72 with pLB85-ILeutD integrated</td>
</tr>
<tr>
<td>SC072/ILackA1</td>
<td>SCO72 with pLB85-ILackA1 integrated</td>
</tr>
<tr>
<td>SC072/ILackA2</td>
<td>SCO72 with pLB85-ILackA2 integrated</td>
</tr>
<tr>
<td>SC072/MGpfl</td>
<td>SCO72 with pLB85-MGpfl integrated</td>
</tr>
<tr>
<td>SC072/MGeutD</td>
<td>SCO72 with pLB85-MGeutD integrated</td>
</tr>
<tr>
<td>SC072/MGackA1</td>
<td>SCO72 with pLB85-MGackA1 integrated</td>
</tr>
<tr>
<td>SC072/MGackA2</td>
<td>SCO72 with pLB85-MGackA2 integrated</td>
</tr>
<tr>
<td>MG1363</td>
<td>A plasmid-free strain derived from L. lactis subsp. cremoris NCDO 712</td>
</tr>
<tr>
<td>LB436</td>
<td>MG1363 containing pLB65</td>
</tr>
<tr>
<td>LB436/ILpfl</td>
<td>SCO72 with pLB85-ILpfl integrated</td>
</tr>
<tr>
<td>LB436/ILadhE</td>
<td>SCO72 with pLB85-ILadhE integrated</td>
</tr>
<tr>
<td>LB436/ILeutD</td>
<td>SCO72 with pLB85-ILeutD integrated</td>
</tr>
<tr>
<td>LB436/ILackA1</td>
<td>SCO72 with pLB85-ILackA1 integrated</td>
</tr>
<tr>
<td>LB436/ILackA2</td>
<td>SCO72 with pLB85-ILackA2 integrated</td>
</tr>
<tr>
<td>LB436/MGpfl</td>
<td>SCO72 with pLB85-MGpfl integrated</td>
</tr>
<tr>
<td>LB436/MGadhe</td>
<td>SCO72 with pLB85-MGadhe integrated</td>
</tr>
<tr>
<td>LB436/MGeutD</td>
<td>SCO72 with pLB85-MGeutD integrated</td>
</tr>
<tr>
<td>LB436/MGackA1</td>
<td>SCO72 with pLB85-MGackA1 integrated</td>
</tr>
<tr>
<td>LB436/MGackA2</td>
<td>SCO72 with pLB85-MGackA2 integrated</td>
</tr>
<tr>
<td>ABLE-K</td>
<td>An E. coli strain reducing the copy number of common cloning vectors</td>
</tr>
</tbody>
</table>

2.2.2 Antibiotics

When needed ampicillin, erythromycin or chloramphenicol were added at 100, 5 and 5 μg ml$^{-1}$ for L. lactis respectively. Tetracycline and kanamycin were applied at 15 and 8 μg ml$^{-1}$ for E. coli ABLE-K.

2.2.3 Culture media and conditions

E. coli strains were cultivated aerobically at 28 °C in lysogeny broth (LB) [14]. L. lactis was cultivated in M17 media or chemically defined SA media [15] at 30 °C without aeration, supplemented with 0.2% (w/v) glucose. For growth experiments, defined SA medium [15] devoid of sodium acetate and supplemented with nucleosides (adenosine, cytidine, guanosine, thymidine, uridine and inosine, 20 mg/L), α-lipoic acid (2 mg/L) and either glucose or maltose (0.2%) were used, abbreviated as GluSALN or MalSALN respectively. Growth experiments were conducted in flasks at 30 °C under static conditions with slow stirring. Optical density at 600 nm (OD$_{600}$) was regularly measured. To prepare a preculture, a single colony from an agar plate was inoculated into GluSALN or MalSALN media in dilution series. Growth experiments were started by inoculating an exponentially growing overnight culture selected from the dilution series into a flask containing the same medium, up to OD$_{600} = 0.001$ or 0.02 for growth on glucose or maltose respectively. Biological
triplicates for each experiment were performed. The previously found cell density equal to 0.36 g (dry weight) (gdw) per liter of SALN medium at OD\textsubscript{600} = 1 was used for the calculation of specific rates [16].

2.2.4 Quantification of sugar and fermentation products

High-performance liquid chromatography (HPLC) was employed to measure the concentration of glucose, maltose, lactate, formate and acetate in the samples taken during the growth experiments as previously described [16]. The biomass yield on sugar \( Y_{X/S} \) (gdw/C6-mol) was estimated by:

\[
Y_{X/S} = \frac{1}{n_{c_6}} \frac{dX}{dS} = \frac{d(X/V)}{d(S/V)} = \frac{(\rho/OD_{600}) d(OD_{600})}{n_{c_6}} = \frac{0.36 d(OD_{600})}{n_{c_6} dC}
\]

where \( S \) is the substrate (sugar), \( X \) is the biomass, \( V \) the culture volume, \( \rho \) the cell density and \( C \) the substrate concentration. \( n_{c_6} \) = 1 for glucose and 2 for maltose. It is therefore proportional to the slope of OD\textsubscript{600} plotted against C6 concentration. Under the assumption of exponential growth (\( dX/dt = \mu X \) where \( \mu \) is the specific growth rate), the specific rate of consumption or production of substrate or product, respectively, or simply the substrate or product flux, were calculated as:

\[
v = \frac{1}{X} \frac{dS}{dt} = \frac{dS}{dX} = \frac{\mu}{0.36 d(OD_{600})} \frac{dC}{dC}
\]

2.2.5 DNA techniques

The method used to isolate the chromosomal DNA from \( L. \) lactis was modified from a previous method [17]. PCR amplification, restriction, ligation, transformation and plasmid purification from \( E. \) coli were performed following procedures described in Sambrook and Russell [18] and the description from the manufacturer of the enzymes used. Electrocompetent cells of \( L. \) lactis were grown in M17 broth supplemented with 10 g L\textsuperscript{-1} glucose and 10 g L\textsuperscript{-1} glycine and transformed by electroporation as described previously [19].

2.2.6 Construction of gusA reporter strains

The promoter containing region upstream a specific gene (≈500-bp) was PCR amplified and inserted into plasmid pLB85 and transformed into LB436 which expresses phage TP901-1 integrase as described previously [12] to construct strains reporting the promoter activity of mixed-acid genes. For SC072 which was derived from IL1403, since the attB site in IL1403 does not perfectly match the consensus for high-frequency integration, vectors derived from pLB85 were first transformed into and then purified from \( E. \) coli. It is noted that ampicillin instead of erythromycin was used as selection marker for \( E. \) coli transformation to avoid abnormal forms of plasmids observed (possibly multimer). \( L. \) lactis transformants were selected on GM17 with erythromycin and verified by sequencing using primers CSO50 and CSO263 (Table 2.2).
2.2.7 Measurement of β-glucuronidase activity

Exponentially growing culture at \( \text{OD}_{600} = 0.6 \) was quenched on ice and harvested for \( \beta \)-glucuronidase assay. The procedure used for measuring \( \beta \)-glucuronidase activities was modified from Miller [20] and Israelsen et al. [21].

2.3 Results

2.3.1 Growth of MG1363 and IL1403

Before studying the difference between the two strains, the growth behavior and fermentation pattern of the strains should be confirmed. For this purpose, \( L. \ lactis \) MG1363 and IL1403 were cultivated on GluSALN or MalSALN media. Sugars were verified to be the limiting substrate for growth by the absence of sugar in cultures where growth had ceased. The growth rate and the fermentation pattern during exponential growth phase for each experiment were analyzed. Table 2.5 shows the specific growth rate and specific rates of consumption of sugars and production of lactate, formate, acetate and ethanol. The growth rate of both strains on glucose was indistinguishable from each other, around 1.15 h\(^{-1}\), equivalent to a generation time of 36 minutes. Product formation as revealed by HPLC measurements was found to be homolactic, and only 5% (molar) of the pyruvate ended up as alternative products (Table 2.6).

<table>
<thead>
<tr>
<th>Growth rate ((\text{h}^{-1}))</th>
<th>Specific rate of consumption/production ((\text{mmol h}^{-1} \text{gdw}^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sugar</td>
</tr>
<tr>
<td>GluSALN</td>
<td></td>
</tr>
<tr>
<td>MG1363</td>
<td>1.15 ± 0.04</td>
</tr>
<tr>
<td>IL1403</td>
<td>1.15 ± 0.07</td>
</tr>
<tr>
<td>MalSALN</td>
<td></td>
</tr>
<tr>
<td>MG1363</td>
<td>0.61 ± 0.03</td>
</tr>
<tr>
<td>IL1403</td>
<td>0.43 ± 0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carbon recovery</th>
<th>Formate %</th>
<th>Acetate + ethanol %</th>
<th>Biomass yield ((\text{gdw}/ \text{C}_5\text{-mol}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>GluSALN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG1363</td>
<td>92% ± 3%</td>
<td>4% ± 1%</td>
<td>5% ± 1%</td>
</tr>
<tr>
<td>IL1403</td>
<td>88% ± 5%</td>
<td>4% ± 1%</td>
<td>5% ± 1%</td>
</tr>
<tr>
<td>MalSALN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG1363</td>
<td>92% ± 7%</td>
<td>57% ± 2%</td>
<td>55% ± 2%</td>
</tr>
<tr>
<td>IL1403</td>
<td>97% ± 6%</td>
<td>10% ± 2%</td>
<td>10% ± 2%</td>
</tr>
</tbody>
</table>

Carbon recovery was calculated as the sum of acetate, ethanol and lactate fluxes divided by the sugar uptake flux in \( \text{C}_5\)-mole. Formate percentage was the formate flux divided by the sum of formate and lactate fluxes. Acetate and ethanol percentage was the sum of acetate and ethanol fluxes divided by the sum of the two and lactate fluxes.
When growing on maltose, however, a notable higher growth rate of MG1363 (0.61 h\(^{-1}\)) than that of IL1403 (0.43 h\(^{-1}\)) was observed. Calculation of product fluxes indicated very different fermentation patterns between the two strains, prominently mixed-acid for MG1363 (≈55% mixed acids) and homolactic for IL1403 (10% mixed acids). Especially low in production was ethanol, basically close to zero production as in the case of growth on glucose. It is also worth noting that the maltose flux in IL1403 is slightly but systematically higher than that in MG1363 (8 replications throughout 1 year) (Table 2.5). Another observation is the significantly lower growth yield of IL1403 on maltose.

### 2.3.2 Strategy to identify trans- and cis-regulation of transcription

We aimed to study the regulation of the promoter activities of the mixed-acid genes in both MG1363 and IL1403. The genes include *pfl*, *eutD*, *ackA1*, *ackA2* and *adhE*, respectively encoding for pyruvate formate-lyase (PFL), phosphotransacetylase (PTA), two acetate kinases (ACKs) and alcohol dehydrogenase (ADH). For the activity of promoter of gene *G* of strain *A* expressed in strain *B*, a single copy of a 500-bp upstream region of *G* from strain *A* followed by the gene *gusA*, encoding β-glucuronidase as a reporter, was created on the chromosome of strain *B* (see Figure 2.1 and ‘construction of gusA reporter strains’ in the method section for construction details).

For each gene, strains were constructed for measuring the promoter activity from MG1363 or IL1403 expressed in MG1363 or IL1403 respectively (Figure 2.2). The rationale behind this was to identify possible trans- and cis-regulatory elements. If the promoter activities of the two promoters are similar within a strain but the values differ between the two strains, trans-regulation specific to a strain is suggested. In contrast, if the activity of the promoter from a strain remains unchanged when expressed in both strains but is persistently different from that of the promoter from the other strain, cis-regulatory elements on either promoter region subject to regulation in both strains is suggested.

![Diagram](image.png)

*Figure 2.1. Construction for measuring promoter activities.*
Chapter 2. Transcriptional activity of mixed-acid genes

Figure 2.2. Identification of trans- and cis-regulation.

Comparing activity of promoters from the two strains expressed in the two strains respectively.

PA, promoter activity.

**pfl promoters**

The activities of the *pfl* promoters of both strains were assayed (Figure 2.3). An obvious observation is the 10-fold induction of *pfl* promoters in all constructs for growth on maltose compared to growth on glucose. More interestingly, a surprisingly clear pattern resembling *trans*-regulation was seen. *pfl* promoters of both strains had very similar activities and they were 3-fold higher when expressed in MG1363 than in IL1403.

Figure 2.3. *pfl* promoter activities.

**eutD promoters**

For *eutD* promoters, similarly, induction for growth on maltose was observed (Figure 2.4). A sharp contrast to the case of *pfl* is a clear pattern resembling *cis*-regulation on *eutD* promoters. The *eutD* promoter of MG1363 always has an activity about 2 times higher than that of IL1403 and the levels are similar when expressed in both strains.
Chapter 2. Transcriptional activity of mixed-acid genes

Promoter activities of *ackA1* and *ackA2*

For the two *ackA* genes, no clear pattern could be observed. For *ackA1*, only a higher activity on maltose compared to glucose (~2-fold) for all cases could be concluded. For *ackA2*, more significant induction on maltose was observed (~10-fold). The promoter of IL1403 appeared to be more active than that of MG1363 and was particularly active in its original host IL1403.
Promoter activities of adhE

For adhE promoter activity, only the promoter activities in MG1363 as the background were measured (Figure 2.6). The induction on maltose was very strong for MG1363’s promoter (75-fold). The activity of IL1403’s adhE promoter was two-fold to three-fold higher on maltose than that of MG1363’s.

![Figure 2.6. adhE promoter activities in MG1363.](image)

2.4 Discussion

2.4.1 Homolactic fermentation in IL1403 confirmed

The growth experiments of IL1403 in MalSALN media supported two observations in Even et al. (2001) [9]. First, IL1403 largely remained homolactic even when growing on a slowly fermentable sugar (galactose in that study). Second, alcohol dehydrogenase (ADH) was probably not sufficiently functional as it was found in the previous study that ADH had an extremely low activity in IL1403 ($10^2$- to $10^3$-fold lower than other fermentation enzymes) and in this study the ethanol production was very low, not comparable to the acetate production. In usual mixed-acid fermentation the production of the two products should be in the same order of magnitude.

The lower growth rate and biomass yield (as well as the final OD$_{600}$, data not shown) of IL1403 growing on maltose are probably results of decreased glycolytic flux limited by maltose uptake and the inability of IL1403 to switch into mixed-acid fermentation to get one more ATP per acetate produced. The effect of the lowered ATP production rate can be explained by the simplistic Pirt’s equation:

\[ \nu_{ATP} = \mu Y_{ATP/X} + \nu_{ATPm} \]

where $\nu_{ATP}$ is the ATP production rate, $\mu$ the growth rate, $Y_{ATP/X}$ the ATP required for forming one unit of biomass, $\nu_{ATPm}$ the rate of ATP needed for maintenance. Assume a constant maintenance and constant ATP requirement for biomass production, the growth rate decreases with the ATP production
rate. Then, during exponential growth, \( v_{ATP} \) is related to the ATP yield \( Y_{ATP/S} \) and the biomass yield on substrate \( Y_{X/S} \) as follows:

\[
v_{ATP} = \frac{1}{X} \frac{d(ATP)}{dt} = \frac{Y_{ATP/S}}{X} \frac{dS}{dt} = Y_{ATP/S} \frac{dS}{\mu dX} = \mu \frac{Y_{ATP/S}}{Y_{X/S}}
\]

Rearranging, we have:

\[
Y_{X/S} = \frac{Y_{ATP/S}}{Y_{ATP/X}} \left(1 - \frac{v_{ATPm}}{v_{ATP}}\right)
\]

Thus the biomass yield decreases with the ATP production rate and increases with the ATP yield on substrate.

Other factors could also contribute to the decreased ATP production in addition to the limited maltose uptake rate. (i) In the upstream maltose pathway, if more intracellular maltose is hydrolyzed by \( \alpha \)-glucosidase than by maltose phosphorylase, the outcome is more glucose which needs ATP in order to be converted into G6P (Figure 1.4). The ATP yield is then lower than 2 moles per mole of C6. But again the activity of \( \alpha \)-glucosidase in L. lactis has not been determined. (ii) Compared to the PTS converting extracellular glucose to intracellular G6P in a single step, both ABC transporter transporting maltose and glucokinase produce extra protons while L. lactis maintains a high intracellular pH [22]. The extra protons produced may need to be pumped out of the cell by ATPase, using extra ATP. (iii) More ATP may be consumed in the whole cellular program for growth on maltose than on glucose, for instance, expression of additional enzymes for maltose assimilation.

### 2.4.2 Regulation of promoter activities

The activities of \( pfl \) promoters in different backgrounds suggested regulation by a trans-element in either strain. It could be a repressor in IL1403 or an activator in MG1363 not present in the other strain. A motif well matching the cre consensus at -114 (relative to the transcription start site) is present in both strains but it is not likely to be the reason since there have been in vitro and in vivo evidences showing the functionality of CcpA in IL1403 [23, 24]. Another candidate with higher possibility is FlpA in MG1363. It has been found to recognize the FNR consensus [25] which overlaps with the -35 promoter of \( pfl \) for both of the strains and once reported to affect the formate production in MG1363 [26]. No homolog of FlpA could be found in IL1403 but there is another FNR-like protein, RcfA, that however was unable to substitute FlpA in MG1363 [27].

For \( eutD \), the regulation of promoter activity was more likely to be governed by cis-elements present in the promoter region. They could be sequence motifs matching the consensus of a certain activator targeting MG1363’s promoter or a repressor targeting IL1403’s promoter that exists in both strains. One possible candidate is a motif roughly matching the cre consensus present on MG1363
Chapter 2. Transcriptional activity of mixed-acid genes

(TGGGGTCTGATATTA), 257 bp upstream of the start codon and 165 bp upstream of a putative TGn-10 promoter (TGGTATACT). In IL1403, the corresponding sequence has lost the central CG element (TGGGGTGATATTA) which is essential for CcpA recognition.

For the two ackA genes, the repression of ackA2 on glucose is expected because it has been found to be subject to carbon catabolite repression and a cre consensus 22 bp upstream of the start codon is present just after a putative transcription start site [28]. Besides, the quite different activity profiles of the two promoters are interesting and worth further studying in the sense that the proteins encoded by the two genes (ACK) should be responsible for the same reaction. One hypothetical possible situation is that the ACK from ackA1 produces acetate from acetyl-phosphate while that from ackA2 is for the reverse direction. It was hypothesized because IL1403 does not exhibit significant mixed-acid fermentation, so acetyl-CoA is expected to be insufficient and acetate is one of the precursors. This coincides with the higher promoter activity of ackA2 than ackA1 for IL1403 growing on maltose.

Despite the insights provided by these results, a factor contributing to the expression level of a gene at the transcription level not taken into account using the present approach is the possibly varying gene dosage along the position of a gene in the genome. All transcriptional fusions were integrated into the same site on the chromosome, 18 kbp away from the origin of replication. pfl, for example, is located 618 and 657 kbp away from the origin in MG1363 and IL1403 respectively. The actual expression level may be different because the closer to the origin a gene is, the higher the expression level one can expect [29].

For the gene adhE, only the promoter activities in MG1363 were measured as the construction for IL1403 was not successful initially. Further attempts have not been made since later the focus had been shifted to PFL and ACKs, as reported in Chapter 3 and Chapter 1.

2.5 References


20. Miller JG (1972) Experiments in molecular genetics (Cold Spring Harbor Laboratory, New York (US)).


Chapter 3. Mixed-acid fermentation controlled by pyruvate-formate lyase

3.1 Introduction

*L. lactis* subsp. *cremoris* MG1363 and subsp. *lactis* IL1403 have been confirmed to exhibit different extents of mixed-acid fermentation when growing on maltose in SALN media. In MG1363, 60% (in molar) of pyruvate ended up in mixed acids (formate, acetate and ethanol) whereas mixed acids accounted for only 10% of pyruvate distribution in IL1403 (1.7). Meanwhile, the transcriptional activity induced by *pfl* promoters was found to be 3-fold higher in MG1363 than in IL1403. *pfl* encodes for pyruvate formate-lyase (PFL), an oxygen sensitive enzyme producing formate and acetyl-CoA from pyruvate, and is the first enzyme of the mixed-acid pathway (see section 1.5.3). The lower transcription level of *pfl* could be a reason for the inability to shift from homolactic to mixed-acid fermentation in IL1403 as in the case of MG1363, for which PFL has been found to control the mixed-acid shift for growth on galactose [1].

In this study, to examine the role and functionality of PFL in IL1403, first PFL activities from IL1403 were introduced into a PFL-knockout strain derived from MG1363. Second, an attempt was made to modulate the level of PFL in IL1403.

3.2 Materials and methods

3.2.1 Bacteria strains and plasmids

All *L. lactis* strains were derived from MG1363 and IL1403. pCS1966 containing genes for erythromycin resistance and orotate transport was used for markerless *pfl* deletion in MG1363 [2]. pLB65 and pLB85-derived plasmids (described in section 2.2) were used for inserting a chromosomal copy of *pfl* preceded with a synthetic promoter library (SPL) in MG1363. The thermosensitive plasmid pGhost8 containing a gene encoding for tetracycline resistance was used for *gusA* insertion into IL1403’s chromosome [3]. pRC1, non-replicating in *L. lactis*, was used for integrating synthetic promoters upstream of *pfl* in IL1403 by a single cross-over event [4]. *E. coli* MC1000 was used for cloning pCS1966- and pGhost8-derived plasmids [5]. *E. coli* ABLE-C (Stratagene) was used for cloning pRC1 [6]. Primers, plasmids and strains used or constructed in this study are listed in Table 3.1, Table 3.2 and Table 3.3 respectively.
Table 3.1. Primers used in the study.

<table>
<thead>
<tr>
<th>Primer</th>
<th>RS</th>
<th>Amplified region / Description</th>
<th>Sequence (5’ - 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSO834</td>
<td>XbaI</td>
<td>pf upstream</td>
<td>ctagctagCAAGTGATGTACCAAATGAC</td>
</tr>
<tr>
<td>CSO835</td>
<td>BamHI</td>
<td>in MG1363</td>
<td>cgccgtacctTTGAAAATCTCCTTTGTTCT</td>
</tr>
<tr>
<td>CSO836</td>
<td>BamHI</td>
<td>pf downstream</td>
<td>cgccgtacctTCTTAGATTTAAAAAATATAAAG</td>
</tr>
<tr>
<td>CSO837</td>
<td>XhoI</td>
<td>in MG1363</td>
<td>ggtactcgagTGTGATTCACCCCTATTTTCT</td>
</tr>
<tr>
<td>CSO838 /</td>
<td></td>
<td>Verify pf deletion</td>
<td>CTTGATTCGATTTGCTATTATC</td>
</tr>
<tr>
<td>CSO839 /</td>
<td></td>
<td></td>
<td>CTGGTACGATCAGATTACTCG</td>
</tr>
<tr>
<td>76f</td>
<td>XhoI</td>
<td>≈900bp upstream region of pfI</td>
<td>actgactcgagGACTGATACCTTTGTTCT</td>
</tr>
<tr>
<td>76r</td>
<td>HindIII in IL1403</td>
<td></td>
<td>actgaaagcttTAAGAATTAGATGGTTGAAGTATGC</td>
</tr>
<tr>
<td>CSO380</td>
<td>HindIII</td>
<td></td>
<td>gcataagctaggtGACCAATTTATTATC</td>
</tr>
<tr>
<td>CSO381</td>
<td>PstI</td>
<td></td>
<td>tggactcgagAGGAGAGTTGTTGATTCATTG</td>
</tr>
<tr>
<td>77f</td>
<td>PstI</td>
<td>≈900bp downstream region of pfI</td>
<td>actgactcgaaATTTAATGGAATATCTCGGTCTCAGT</td>
</tr>
<tr>
<td>77r</td>
<td>XbaI</td>
<td></td>
<td>actgactagaaCAGATTACTCGGTCTCAGT</td>
</tr>
<tr>
<td>16f</td>
<td>XbaI</td>
<td>Synthetic promoter library for</td>
<td>actgactgtcagATCTTTGAAAAGGGAAGG</td>
</tr>
<tr>
<td>16r</td>
<td>SalI</td>
<td>pfI in MG1363</td>
<td>tnnnTTGAAATTTTGAACAAAGAAAGAC</td>
</tr>
<tr>
<td>8f</td>
<td>XbaI</td>
<td>upstream region of pfI in MG1363</td>
<td>actgactgtcagACTTTGAAAAGGGAAGG</td>
</tr>
<tr>
<td>54r</td>
<td>SalI</td>
<td>Reverse primer for IL1403’s pfI</td>
<td>actgactgacAAAGAATTAGATGGTTGAAGTATGC</td>
</tr>
<tr>
<td>78r</td>
<td>PstI</td>
<td>Reverse primer for truncated IL1403’s pfI</td>
<td>actgactgactgaGAATTTAATGGAAGGGAAGG</td>
</tr>
</tbody>
</table>

RS, restriction site.

Table 3.2. Plasmid used or constructed in the study.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCS1966</td>
<td>A vector containing oroP for orotate transporter as a counterselection tool for deletion.</td>
<td>Erm</td>
</tr>
<tr>
<td>pCS1966-pfl</td>
<td>pCS1966 containing the upstream and downstream regions of pfI in MG1363 obtained with Erm primers CSO834-837.</td>
<td>Erm</td>
</tr>
<tr>
<td>pGhost8</td>
<td>Thermosensitive plasmid, non-replicating above 35 °C.</td>
<td>Tet</td>
</tr>
<tr>
<td>pGhost8-gusA</td>
<td>pGhost8 containing the upstream and downstream regions of pfI stop codon in IL1403, separated by E. coli’s gusA, PCR amplified with primers 76f, r, CSO380, 381 and 77f, r.</td>
<td></td>
</tr>
<tr>
<td>pLB65</td>
<td>Containing orfI expressing phage TP901-1 integrase</td>
<td>Cam</td>
</tr>
<tr>
<td>pLB85 (7)</td>
<td>Reporter vector containing a promoterless gusA gene and an attP site for integration into attB Amp, Erm site on chromosomes in the presence of the TP901-1 integrase.</td>
<td></td>
</tr>
<tr>
<td>pLB85-SPL-MGpf</td>
<td>pLB85 containing a synthetic promoter library preceding pfI of MG1363 PCR amplified with Amp, Erm primer 16f, r.</td>
<td>Erm</td>
</tr>
<tr>
<td>pLB85-SPL-ILpf</td>
<td>pLB85 containing a synthetic promoter library preceding pfI of IL1403 PCR amplified with Amp, Erm primer 16f, 54r.</td>
<td></td>
</tr>
<tr>
<td>pLB85-WTP-MGpf</td>
<td>pLB85 containing the wild-type promoter region and pfI of MG1363 PCR amplified with Amp, Erm primers 8f, 16r</td>
<td></td>
</tr>
<tr>
<td>pRC1 (6)</td>
<td>A pBlue-script-derived vector, non-replicating in L. lactis.</td>
<td>Erm</td>
</tr>
<tr>
<td>pRC1-SPL-ILpf</td>
<td>pRC1 containing a synthetic library preceding a truncated version of pfI of IL1403 (≈900 bp)</td>
<td>Erm</td>
</tr>
</tbody>
</table>

Erm, erythromycin; Tet, tetracycline; Cam, chloramphenicol; Amp, ampicillin.
Table 3.3. Strains used or constructed in the study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG1363 [8]</td>
<td>A plasmid-free strain derived from <em>L. lactis</em> subsp. <em>cremoris</em> NCDO 712</td>
</tr>
<tr>
<td>IL1403 [9]</td>
<td>A plasmid-free strain derived from <em>L. lactis</em> subsp. <em>lactis</em> biovar <em>diacetylactis</em> CNRZ157</td>
</tr>
<tr>
<td>MG1363Δpfl</td>
<td>pfl-knockout derived from MG1363 using pCS1966-pfl</td>
</tr>
<tr>
<td>SC100</td>
<td>MG1363Δpfl containing pLB65</td>
</tr>
<tr>
<td>SC105 a – c</td>
<td>Strains with modulated expression of <em>pfl</em> from MG1363 by integration of pLB85-SPL-MGpfl into SC100</td>
</tr>
<tr>
<td>SC106 a – c</td>
<td>Strains with modulated expression of <em>pfl</em> from IL1403 by integration of pLB85-SPL-ILpfl into SC100</td>
</tr>
<tr>
<td>SC119</td>
<td>A strain with the wild-type promoter region and <em>pfl</em> of MG1363 by integration of pLB85-WTP-MGpfl into SC100</td>
</tr>
<tr>
<td>SC183</td>
<td>IL1403 with <em>gusA</em> inserted at <em>pfl</em> downstream, constructed using pGhost8-<em>gusA</em></td>
</tr>
<tr>
<td>SC193 a – c</td>
<td>Strains with synthetic promoters preceding the native <em>pfl</em> in SC183 resulting from integrating pRC1-SPL-ILpfl into SC183’s chromosome</td>
</tr>
<tr>
<td>MC1000 [5]</td>
<td>An <em>E. coli</em> strain used for cloning purposes</td>
</tr>
<tr>
<td>ABLE-K [6]</td>
<td>An <em>E. coli</em> strain reducing the copy number of common cloning vectors</td>
</tr>
</tbody>
</table>

3.2.2 Antibiotics

When needed, erythromycin, chloramphenicol, tetracycline were added at 5 μg ml\(^{-1}\) for *L. lactis* respectively. Erythromycin, tetracycline and kanamycin were applied at 150, 15 and 8 μg ml\(^{-1}\) for *E. coli*.

3.2.3 Gene deletion and insertion

*pfl* deletion was achieved by using the plasmid pCS1966 [2]. \(\approx900\)-bp regions upstream and downstream of the target to be deleted were PCR amplified (using primers CSO834 – 837) and inserted into pCS1966. The resulting plasmids were used as previously described [2]. For insertion of *gusA*, the sequence to be inserted was also PCR amplified (using primers CSO380, 381) and cloned in between the upstream and downstream regions (amplified using primers 76, 77) in pGhost8 [3].

3.2.4 Modulation of *pfl* expression in *L. lactis*

For modulating *pfl* expression in MG1363, first the native *pfl* was deleted followed by the transformation of pLB65 into the deletion strain, resulting in SC100. The vectors for strain construction were pLB85-SPL-MGpfl and pLB85-SPL-ILpfl containing a SPL preceding *pfl* of MG1363 and IL1403 amplified using primers 16f, 16r and 16f, 54r respectively. The vectors were directly transformed into SC100. The transformants became the SC105 and SC106 series of strains with modulated expression from MG1363 and IL1403 respectively via site-specific integration of the vectors at *attP, attB* sites in the presence of the integrase encoded from pLB65 [7]. For control, pLB85-WTP-MGpfl containing the wild-type promoter followed MG1363’s *pfl* was also prepared. Transforming the vector yielded the strain.
For modulation of \( pfl \) in IL1403, the \textit{gusA} gene was inserted just downstream of \( pfl \) and the resultant strain was named SC183. For the integration vector, the fragment of a truncated version of \( pfl \) with a SPL was amplified using primers 16f and 78r. It was digested and inserted into pRC1 to become pRC1-SPL-IL\textit{pfl}. It was then cloned through \textit{E. coli} ABLE-C which reduced the copy number of plasmids to compensate the potential selective disadvantage of strong synthetic promoters. The transformants of the purified plasmids become the SC193 series of strains with different expression of \( pfl \).

For the calculation of flux control coefficients (FCCs), a suitable curve \( y = f(x) \) for flux as a function of \textit{gusA} activity was fitted using Matlab® and the FCC was calculated as \( xf'(x)/f(x) \).

### 3.2.5 Others

For culture media and conditions, quantification of sugars and fermentation products, other DNA techniques and measurement of \( \beta \)-glucuronidase activity, readers are referred to section 2.2.

### 3.3 Results

#### 3.3.1 PFLs of MG1363 and IL1403 expressed in MG1363\textbackslash apfl

The PFL of IL1403 (IL-PFL) and the PFL of MG1363 (MG-PFL) were expressed in MG1363\textbackslash apfl to see if they behaved differently. This was done by integration of a copy of \( pfl \) either from MG1363 or IL1403 (MG-\textit{pfl} or IL-\textit{pfl}) following a SPL into the chromosome of MG1363 through the use of pLB85 (Figure 3.1). A control strain with the wild-type promoter region (WTP) followed by MG-\textit{pfl} integrated in the same way was also constructed, called SC119. The library of clones with different expression of MG-\textit{pfl} was called SC105 and the corresponding library of clones for IL-\textit{pfl} was SC106.

![Figure 3.1. MG1363 derivatives with modulated PFL from MG1363 or IL1403.](image)

The preliminary screening was done by quantifying the formate production in overnight culture of these strains (Figure 3.2). The preliminary screening showed that IL-PFL was indeed able to produce formate in significant amounts. The profile of formate production of the two libraries was very similar. Five strains from each library were selected for further characterization, named SC105a – e and...
SC106a–e respectively. Growth experiments of these strains on glucose and maltose in SALN media were conducted respectively. Growth rate, production formation and gusA activity during exponential growth phase were measured (see Table 3.4 for all data).

![Figure 3.2. Screening of strains with modulated pfl.](image)

Libraries of clones with modulated pfl from (a) MG1363 (library SC105) and (b) IL1403 (library SC106), respectively. Clones producing different amounts of formate in overnight cultures on maltose (in green bars) were selected for further characterization.

The gusA activity of the chosen strains ranged from 1 to 200 Miller units for growth on glucose and from around 10 to 400 Miller units for growth on maltose (Figure 3.3). The activity of all synthetic promoters appeared to have increased from 2- to 7-fold on maltose compared to glucose, with the fold change decreasing with the activity. The WTP in SC119 showed a 7-fold induction of activity on maltose compared to glucose, similar to the results in 1.7 (Figure 2.3). The fluxes of SC119 (Table 3.4) are also fairly comparable to the wild type MG1363 (Table 2.5).

![Figure 3.3. gusA activity of wild-type promoter and synthetic promoters.](image)
Table 3.4. Growth data and gusA activities of SC105, SC106, SC100 and SC119.

<table>
<thead>
<tr>
<th>Glucose</th>
<th>gusA activity (Miller unit)</th>
<th>Growth rate (h⁻¹)</th>
<th>Specific rate of consumption/production (mmol h⁻¹ gdw⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALN</td>
<td></td>
<td>Sugar</td>
<td>Lactate</td>
</tr>
<tr>
<td>SC105a</td>
<td>2 ± 0.1</td>
<td>1.01 ± 0.01</td>
<td>23.3 ± 0.4</td>
</tr>
<tr>
<td>SC105b</td>
<td>21 ± 0.3</td>
<td>1.01 ± 0.01</td>
<td>21.7 ± 0.6</td>
</tr>
<tr>
<td>SC105c</td>
<td>59 ± 1.1</td>
<td>0.98 ± 0.01</td>
<td>21.2 ± 1.1</td>
</tr>
<tr>
<td>SC105d</td>
<td>68 ± 1.2</td>
<td>0.98 ± 0.01</td>
<td>22.3 ± 0.5</td>
</tr>
<tr>
<td>SC105e</td>
<td>188 ± 3.5</td>
<td>0.94 ± 0.01</td>
<td>20.9 ± 0.5</td>
</tr>
<tr>
<td>SC106a</td>
<td>1 ± 0.1</td>
<td>1.01 ± 0.01</td>
<td>23.3 ± 0.2</td>
</tr>
<tr>
<td>SC106b</td>
<td>3 ± 0.1</td>
<td>0.98 ± 0.01</td>
<td>23.6 ± 0.1</td>
</tr>
<tr>
<td>SC106c</td>
<td>23 ± 0.4</td>
<td>1.00 ± 0.01</td>
<td>23.4 ± 0.5</td>
</tr>
<tr>
<td>SC106d</td>
<td>66 ± 2.0</td>
<td>0.99 ± 0.01</td>
<td>22.2 ± 0.7</td>
</tr>
<tr>
<td>SC106e</td>
<td>202 ± 3.7</td>
<td>0.91 ± 0.01</td>
<td>21.2 ± 0.5</td>
</tr>
<tr>
<td>SC100</td>
<td>N.D.</td>
<td>1.25 ± 0.03</td>
<td>25.0 ± 0.7</td>
</tr>
<tr>
<td>SC119</td>
<td>39 ± 2.0</td>
<td>1.02 ± 0.02</td>
<td>23.2 ± 0.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maltose</th>
<th>gusA activity (Miller unit)</th>
<th>Growth rate (h⁻¹)</th>
<th>Specific rate of consumption/production (mmol h⁻¹ gdw⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALN</td>
<td></td>
<td>Sugar</td>
<td>Lactate</td>
</tr>
<tr>
<td>SC105a</td>
<td>17 ± 0.4</td>
<td>0.43 ± 0.01</td>
<td>5.8 ± 0.8</td>
</tr>
<tr>
<td>SC105b</td>
<td>82 ± 2.8</td>
<td>0.50 ± 0.01</td>
<td>6.8 ± 0.6</td>
</tr>
<tr>
<td>SC105c</td>
<td>185 ± 1.5</td>
<td>0.54 ± 0.01</td>
<td>7.6 ± 0.5</td>
</tr>
<tr>
<td>SC105d</td>
<td>227 ± 2.8</td>
<td>0.56 ± 0.01</td>
<td>7.2 ± 0.6</td>
</tr>
<tr>
<td>SC105e</td>
<td>399 ± 15</td>
<td>0.59 ± 0.01</td>
<td>7.0 ± 0.6</td>
</tr>
<tr>
<td>SC106a</td>
<td>7 ± 0.2</td>
<td>0.39 ± 0.01</td>
<td>6.5 ± 0.6</td>
</tr>
<tr>
<td>SC106b</td>
<td>21 ± 1.6</td>
<td>0.41 ± 0.01</td>
<td>6.6 ± 0.5</td>
</tr>
<tr>
<td>SC106c</td>
<td>63 ± 1.9</td>
<td>0.49 ± 0.02</td>
<td>7.4 ± 0.9</td>
</tr>
<tr>
<td>SC106d</td>
<td>203 ± 8.9</td>
<td>0.56 ± 0.01</td>
<td>7.7 ± 0.6</td>
</tr>
<tr>
<td>SC106e</td>
<td>383 ± 17</td>
<td>0.59 ± 0.01</td>
<td>7.5 ± 0.1</td>
</tr>
<tr>
<td>SC100</td>
<td>N.D.</td>
<td>0.40 ± 0.02</td>
<td>7.5 ± 1.0</td>
</tr>
<tr>
<td>SC119</td>
<td>282 ± 30</td>
<td>0.62 ± 0.01</td>
<td>8.0 ± 0.3</td>
</tr>
</tbody>
</table>

SC100, a pfl-deleted strain; SC105, strains SPL-MG-PFL; SC106, strains with SPL-IL-PFL; SC119, a strain with WTP-MG-PFL. N.D., not detectable.

3.3.2 Flux control by pfl expression level

**Formate flux controlled by pfl expression**

How the level of pfl expression controls the growth rate and different fluxes was investigated. An obvious question is the effect of pfl expression on formate production. For growth on glucose, the formate flux appeared to increase initially with pfl expression but decrease at high expression levels (SC105e and SC106e, see Figure 3.4a). The level of pfl expression level in the wild-type MG1363 and IL1403 might be estimated from the flux-response curve by comparing the formate fluxes in the wild types under the assumption that the formate flux responds to change in pfl expression similarly in IL1403. From the similar wild-type flux levels (from Table 2.5), similar pfl expression level was
suggested flux and control coefficients (FCCs) between 0.78 and 0.84 were estimated from curve fitting (Figure 3.4a(ii)).

Figure 3.4. Dependence of formate flux on pfl expression.
SC105 (SPL-MG-PFL) (blue square), SC106 (SPL-IL-PFL) (red diamond), and SC119 (WTP-MG-PFL) (green circle) growing on (a) glucose or (b) maltose. (i) Formate flux and the derived (ii) flux control coefficient (FCC) plotted against pfl expression level. The fluxes in the wild-type IL1403 and MG1363 were provided for reference. SPL, synthetic promoter library; WTP, wild-type pfl promoter in MG1363.

For growth on maltose, a clear increasing trend of formate flux from 0 to ≈20 mmol h⁻¹gdw⁻¹ along with the increase of pfl expression level was observed (Figure 3.4b(i)). pfl expression in MG1363 and IL1403 estimated from the flux values and the flux-response curve was largely different from each other (≈340 and ≈40 Miller units respectively). The FCC decreases from 0.85 at the estimated wild-type IL1403’s level to 0.31 at the estimated wild-type MG1363’s level (Figure 3.4b(ii)).

**Acetate fluxes controlled by PFL**

Similar trends were observed for acetate production. Acetate flux increased initially but reached the apparent maximum level at ≈60 Miller units of gusA activity for growth on glucose while it kept increasing significantly for growth on maltose (Figure 3.5). The FCCs estimated for the glucose case are close to 0 and 0.19 for IL1403 and MG1363 respectively. The corresponding estimation for the maltose case is 0.52 and 0.33 respectively. It should be remarked, however, that the estimation is not consistent with the estimation from the formate flux data because different pfl expression levels reflected by gusA activity were estimated from data of different production fluxes. The issue will be discussed in later sections.
Chapter 3. Mixed-acid fermentation controlled by pyruvate-formate lyase

Figure 3.5. Dependence of acetate flux on pfl expression.
SC105 (SPL-MG-PFL), SC106 (SPL-IL-PFL), and SC119 (WTP-MG-PFL) growing on (a) glucose or (b) maltose. (i) Acetate flux and the derived (ii) FCC plotted against pfl expression level. The fluxes in the wild-type IL1403 and MG1363 were provided for reference. SPL, synthetic promoter library; WTP, wild-type pfl promoter in MG1363.

Ethanol fluxes controlled by PFL

Regarding ethanol production, for growth on glucose, no clear trend could be observed and the relative error of estimation of fluxes was large at such low values (Figure 3.6). For growth on maltose, an increasing pattern of ethanol flux along with the pfl expression level was observed. The estimated curve for FCC decreased from infinity initially because no ethanol production was observed at small positive gusA activities. The FCC tends to infinity when the value of the flux function approaches zero.

Figure 3.6. Dependence of ethanol flux on pfl expression.
SC105 (SPL-MG-PFL), SC106 (SPL-IL-PFL), and SC119 (WTP-MG-PFL) growing on (a) glucose or (b) maltose. (i) Ethanol flux and the derived (ii) FCC plotted against pfl expression level. The fluxes in the wild-type IL1403 and MG1363 were provided for reference. SPL, synthetic promoter library; WTP, wild-type pfl promoter in MG1363.
Negative control of lactate production by PFL

In contrast to the generally positive control of the mixed acids by PFL, lactate flux was found to be negatively controlled by pfl expression with a very small magnitude of negative FCC in the glucose case and a larger magnitude in the maltose case (Figure 3.7). Beside the flux control, it can also be seen that the wild-type lactate fluxes were out of the range of those of modulated strains in the glucose case. For growth on maltose, it is in particular interesting that the lactate flux of IL1403 (34 mmolh$^{-1}$gdw$^{-1}$) was almost 2-fold higher than all other strains, including strains with very low gusA activities (SC105a, SC106a, 20 mmolh$^{-1}$gdw$^{-1}$) and even the pfl-deletion strain (SC100, 21 mmolh$^{-1}$gdw$^{-1}$).

![Figure 3.7. Dependence of lactate flux on pfl expression.](image)

SC105 (SPL-MG-PFL), SC106 (SPL-IL-PFL), and SC119 (WTP-MG-PFL) growing on (a) glucose or (b) maltose. (i) Lactate flux and the derived (ii) FCC plotted against pfl expression level. The fluxes in the wild-type IL1403 and MG1363 were provided for reference. SPL, synthetic promoter library; WTP, wild-type pfl promoter in MG1363.
Control of glycolytic flux
The specific consumption rate of glucose roughly decreased with pfl expression with very low FCCs ranging from -0.04 to -0.01 (Figure 3.8). The corresponding values for IL1403 and MG1363 are apparently above those of modulated strains. For growth on maltose, the maximum glycolytic flux (with zero FCC) appeared to occur at an intermediate level of pfl expression, close to the level of SC119 which contained WTP-MG-PFL integrated in the attachment site attB. This maximum also almost coincides with the flux in MG1363. Consistent to the significantly higher lactate flux in the homolactic IL1403, the maltose flux was also substantially higher than all other strains.

Figure 3.8. Dependence of glycolytic flux on pfl expression.
SC105 (SPL-MG-PFL), SC106 (SPL-IL-PFL), and SC119 (WTP-MG-PFL) growing on (a) glucose or (b) maltose. (i) Glycolytic flux and the derived (ii) FCC plotted against pfl expression level. The fluxes in the wild-type IL1403 and MG1363 were provided for reference. SPL, synthetic promoter library; WTP, wild-type pfl promoter in MG1363.
Maximum growth rate near the wild-type pfl level

For growth on glucose, the growth rate appeared to be quite constant at a level of \( \approx 1 \text{ h}^{-1} \) for gusA activity \( \leq 70 \) Miller units and then started to drop, resulting in a curve of negative FCC (Figure 3.9). This includes the strain SC119 with WTP-MG-PFL. When the strains were grown on maltose, in contrast, a significant positive control of growth rate by pfl expression was observed initially and the maximum was then attained, also close to the level of SC119 which also had a growth rate very similar to the wild-type MG1363 (0.61 h\(^{-1}\)). The FCC therefore approached zero and became negative at higher pfl expression levels, corresponding to decreased growth rates in SC105e and SC106e. The growth rate of IL1403 is similar to that of strains with lowest pfl expression (SC105a, SC106a – b).

![Figure 3.9. Dependence of growth rate on pfl expression.](image)

SC105 (SPL-MG-PFL), SC106 (SPL-IL-PFL), and SC119 (WTP-MG-PFL) growing on (a) glucose or (b) maltose. (i) Growth rate and the derived (ii) FCC plotted against pfl expression level. The growth rates in the wild-type IL1403 and MG1363 were provided for reference. SPL, synthetic promoter library; WTP, wild-type pfl promoter in MG1363.

3.3.3 Yield on sugars

We also looked into the biomass and product yield on sugars (Figure 3.10). The patterns are in general very similar to the case of fluxes reported above. Here, however, several points not revealed by flux comparison are worth noting. First, for growth on glucose, the wild-type MG1363 and IL1403 showed higher growth rates, higher lactate fluxes than all other strains but their biomass and lactate yields on glucose indeed lied within the range of other strains (Figure 3.10a – b). Second, the biomass and lactate yield on glucose did not change significantly. Third, interestingly, SC119 containing WTP-MG-PFL did not have the highest biomass yield and the highest acetate yield which probably reflects the ATP yield although the strain had the apparently highest growth rate. MG1363 also had a slightly lower biomass and acetate yields than strains with highest pfl expression. The significance of this observation is discussed in the next section.
Figure 3.10. Biomass and product yield on sugars.
The (a) biomass, (b) lactate, (c) formate, (d) acetate and (e) ethanol yield on sugars for SC105 (SPL-MG-PFL) (blue square), SC106 (SPL-IL-PFL) (red diamond), SC119 (WTP-MG-PFL) (green circle) and SC100 (pfl deletion) (magenta triangle) growing on (i) glucose or (ii) maltose. Yields were calculated as growth rates or product fluxes divided by sugar fluxes in C_6 unit. The corresponding levels of MG1363 and IL1403 are marked by blue and red lines respectively.

3.3.4 Loss of formate-producing ability during the modulation of pfl in IL1403

To modulate pfl expression in IL1403, the reporter gene gusA following a proper leader sequence was inserted immediately after the native pfl in IL1403, resulting in the strain SC183 which produced blue spots on agar plates with X-gluc. The fragment from the pfl sequence to the region downstream of pfl used for the deletion by double crossover, including the inserted gusA in between, was confirmed to be correct by DNA sequencing. Then the plasmid pRC1-SPL-ILpfl was used to modulate pfl by integrating a SPL preceding pfl by a single-crossover event. A library of 9 strains, the SC193 series, was isolated. Preliminary tests on the formate-producing ability of SC183 and SC193 library, nonetheless, showed no formate production in all strains. The full copy of pfl in each strain and additionally the wild-type promoter region in SC183 were sequenced but no any mutation could be found (data not shown). The strains were therefore not further characterized. The possible reasons were discussed in the next section.
3.4 Discussion

3.4.1 Similar behaviour of MG-PFL and IL-PFL when expressed in MG1363

For testing the functionality of IL-PFL, both MG-PFL and IL-PFL were expressed in MG1363 where the native \( pfl \) had been deleted. The expression made use of the site-specific integration system at the attachment site using pLB85. The reporter gene \( gusA \) present on pLB85 could serve as an indicator of the expression level of \( pfl \) as an alternative to measure PFL activity directly which is not straightforward because of the extreme sensitivity of PFL toward oxygen [10]. Results of growth characterization indicated that the control of formate flux by both MG-PFL and IL-PFL was very similar. The change in growth rate and flux distribution in response to the change in MG-\( pfl \) and IL-\( pfl \) expression level were also very much alike. This suggested that IL-PFL should also be able to function properly in IL1403 provided that the PFL activating enzyme (PFL-AE) also expresses and functions properly in IL1403.

3.4.2 Flux control by PFL in MG1363

For growth on glucose, the growth rate, lactate flux and glycolytic flux were all negatively controlled by PFL, with very small negative FCCs (-0.04 ~ 0) at the estimated wild-type \( pfl \) levels for MG1363 and IL1403 respectively. Formate and acetate production fluxes meanwhile were positively controlled by PFL. The sum of FCCs for a flux by all enzymes is equal to 1 by the summation theorem of FCCs. Since previous studies found the FCC for formate flux by LDH and PYK is close to -1 and 1 respectively [11, 12]. The sum of FCCs of the rest of the reactions should be about 1, which is close to the value found for PFL. The sum of FCCs by enzymes other than LDH, PYK and PFL should thus be close to zero.

For growth on maltose, PFL appeared to have negative control on only lactate flux with FCC being close to zero (-0.06 ~ -0.04) and -0.6 ~ -0.7 at the estimated \( pfl \) levels for IL1403 and MG1363 respectively. Mixed-acid fluxes were positively controlled by PFL whereas the growth rate and glycolytic flux increased initially along \( pfl \) expression but decreased at larger levels after the maximum. The estimated FCCs were summarized in Table 3.5.

It should be noted again that the estimation of FCCs is dependent on the choice of fitted curves. In this study, general functions (exponential, power, rational, polynomial) were adopted and best fits in terms of lowest sum of squared residuals, highest coefficient of determination and fewer parameters were chosen. Some variations to the selected curves, however, can still be intuitively possible in some cases by looking at the plots. For example, for the formate and ethanol fluxes in the maltose case, good fits with more parameters registering maximum values at levels close to SC119 are also plausible. Similarly a curve fitting the lactate fluxes for growth on maltose with a minimum point attained at a level close to SC119 is another choice. This different selection of curves can lead to the conclusion of zero flux control by PFL on formate, ethanol and lactate flux at the wild-type level. More strains from
the libraries with *gusA* activity in the range of the estimated wild-type levels are required to be characterized for more accurate estimation.

| Table 3.5. Estimated flux control coefficients (FCCs) at the wild-type pfl level. |
|---------------------------------|----------------|----------------|----------------|----------------|
|                                 | Glucose        |                | Maltose        |                |
|                                 | MG136          | IL1403         | MG1363         | IL1403         |
| Estimated relative pfl expression level in wild types (gusA activity in Miller unit) | 30 ~ 40        | 30 ~ 40        | 270 ~ 290      | 30 ~ 40        |
| Growth rate                    | -0.006 ~ -0.001| 0.06 ~ 0.09    | 0.08 ~ 0.10    |                |
| FCC at the estimated wild-type level | Sugar         | -0.02         | -0.05 ~ -0.10  | 0.05 ~ 0.06    |
| Lactate                        | -0.03          | -0.71 ~ -0.63 | -0.06 ~ -0.04  |                |
| Formate                        | 0.88 ~ 1.02    | 0.42 ~ 0.48    | 0.83 ~ 0.85    |                |
| Acetate                        | 0.19           | 0.30 ~ 0.36    | 0.49 ~ 0.55    |                |
| Ethanol                        | /              | 0.47 ~ 0.53    | 1.38 ~ 1.63    |                |

### 3.4.3 Rate or yield optimization

An interesting observation from comparing flux and yield is that optimal growth rate in the wild-type MG1363 occurred at a sub-optimal growth yield and a sub-optimal ATP yield on maltose (positively correlated to acetate yield, assuming all sugar is converted into pyruvate in glycolysis). Homolactic fermentation has a lower ATP yield than mixed-acid fermentation and has been known to be out of the prediction by the well-known flux balance analysis (FBA) using genome-scale metabolic model if additional capacity constraints on flux were not added [13]. It is because FBA is in principle a method of yield optimization which, given the biomass reaction as the objective function, always selects a flux distribution with highest biomass yield and may not coincide with the flux distribution maximizing the true growth rate [14, 15]. The observation here apparently corresponds to this case, even happening in growth on maltose where ATP production rate should be the growth-limiting factor.

The data of the control strain SC119 with WTP-MG-PFL further supported this observation. SC119 was constructed in the same way as other strains with SPL-MG-PFL or SPL-IL-PFL. It behaved very similarly to the wild-type MG1363 for growth on maltose and had a *gusA* activity and biomass yield lower than SC105e and SC106e but still grew faster.

### 3.4.4 Comparison to previous results

Control of the shift from homolactic to mixed-acid fermentation in MG1363 has been demonstrated by Melchiorse et al. (2002) [1]. In their study, similar pattern of flux control by PFL was observed. There are still several points of difference in our study. First, the functionality of IL-PFL was also tested and compared to MG-PFL. Second, the construction method was different from the previous study. In the previous study, two replicating plasmids expressing *pfl* with promoters of different strength were respectively transformed into MG1363 and a *pfl*-deleted strain. Multiple copies of *pfl*
were thus present in the constructed strains. In this study, all tested strains were constructed by site-
specific chromosomal integration of pfl with synthetic promoters into SC100, ensuring a single copy
on the chromosome. This precluded possible phenotypic differences brought by the existence of
multiple copies of pfl and replicating plasmids. An obvious one is the more cellular resource spent on
their replication. Third, we attempted to look into the control of different fluxes by estimating the flux
control coefficients.

3.4.5 Unsuccessful modulation of pfl in IL1403

The absence of formate production in SC183 and the SC193 series suggested that PFL in these strains
were not functional. The obvious possibility of mutation in the coding region and promoter region of
pfl in SC183 was ruled out by sequencing. The lack of formate production might therefore be the
result of mutation in PFL-activating enzyme (PFL-AE) which is required for PFL activity or other
unknown reasons.

The results suggested that inserting gusA after pfl might be detrimental. Indeed, we also attempted to
transform pLB85-SPL-ILpfl into IL1403 but no blue colonies on agar plates with X-gluc could be
seen. A probable effect of inserting gusA into the transcription unit of pfl is increase in the
transcription of pfl due to the increased mRNA stability of the longer transcript. This is supported by
the ≈4-fold increase in gusA activity in strain SC119 compared to LB436/MGpfl (Table 2.4) when
growing on either glucose or maltose (Figure 2.3 in Chapter 2 and Table 3.4). In both strains, the
transcription of gusA is induced by the pfl promoter in MG1363. The difference lies in that in SC119
gusA is the second gene in an operon after pfl whereas the gusA in LB436/MGpfl directly follows the
promoter.

If this is true, then this implied that significant overexpression of PFL in IL1403 might be lethal. The
reason for IL1403 unfavorable of overexpressing PFL remains unknown. One possibility is the non-
functionality of alcohol dehydrogenase downstream of PFL causing redox imbalance after PFL
converts too much of the pyruvate into acetyl-CoA.

3.4.6 Possible reasons for the low formate production in IL1403

Different properties of IL-PFL and MG-PFL originating from the difference in the amino acid
sequences is one of the possible causes for the lower formate production in IL1403. It has however
been ruled out by the present study. Several other reasons for the low PFL activity in IL1403 are
possible.

First, the PFL activating enzyme (PFL-AE) encoded by pflA might not be expressed sufficiently or
not function properly in IL1403. It has long been known that PFL in E. coli requires activation by
PFL-AE for catalytic activity [16] to protect PFL from irreversible oxygen cleavage [17]. Similar
mechanisms activating PFL has been observed in L. lactis [10] and demonstrated in vitro using PFL
and PFL-AE from the closely related *Streptococcus mutans* [18, 19]. A study even showed that overexpressing PFL and PFL-AE simultaneously increased the formate-to-lactate ratio significantly in *Streptococcus bovis* [20].

Second, PFL might be expressed at a lower quantity in IL1403 as suggested by the study on promoter activity (Figure 2.3, Chapter 2) and the unsuccessful modulation of PFL in IL1403 suggested a high expression of IL1403 was probably unfavourable, unlike the case of MG1363.

Third, the flux might be controlled by other enzymes or regulated at the metabolic level. For example, the bi-functional alcohol dehydrogenase (ADH) in IL1403 might be defective or not expressed, as suggested in [21] (significant transcript abundance but low activity) and the close-to-zero ethanol production in our data (Table 2.5, Chapter 2). If this is the case, then the low PFL activity could be a result of metabolic regulation following the model proposed by Garrigues et al. [22]. Since one NADH is not reduced for each pyruvate converted into acetate through PFL, phosphotransacetylase and acetate kinase in the absence of ADH activity, NADH would accumulate and inhibit Glyceraldehyde 3-phosphate dehydrogenase (GAPDH). This would cause a large triose-phosphate pool which inhibits PFL. In this case, ADH or other NADH-producing enzymes are expected to have control on the formate flux. Interestingly, in this sense, the regulation of PFL by triose-phosphate can be a vital trait to prevent glycolysis from termination due to lack of NADH regeneration during acetate production. This question can be answered by studying the evolutionary sequences of different fermentation enzymes and the regulatory mechanisms.

It should be noted that these possibilities are not mutually exclusive. Indeed, the first and second cases may be the evolutionary consequences of the third if the third is true.

### 3.5 Acknowledgement

We thank Jun Chen for help with constructing the strain MG1363Δpfl.

### 3.6 References

Chapter 3. Mixed-acid fermentation controlled by pyruvate-formate lyase


Chapter 4. Acetate kinase isozymes confer robustness in acetate metabolism


Siu Hung Joshua Chan, Lasse Nørregaard, Christian Solem, Peter Ruhdal Jensen

4.1 Abstract
Acetate kinase (ACK) (EC no: 2.7.2.1) interconverts acetyl-phosphate and acetate to either catabolize or synthesize acetyl-CoA dependent on the metabolic requirement. Among all ACK entries available in UniProt, we found that around 45% are multiple ACKs in some organisms including more than 300 species but surprisingly, little work has been done to clarify whether this has any significance. In an attempt to gain further insight we have studied the two ACKs (AckA1, AckA2) encoded by two neighboring genes conserved in Lactococcus lactis (L. lactis) by analyzing protein sequences, characterizing transcription structure, determining enzyme characteristics and effect on growth physiology. The results show that the two ACKs are most likely individually transcribed. AckA1 has a much higher turnover number and AckA2 has a much higher affinity for acetate in vitro. Consistently, growth experiments of mutant strains reveal that AckA1 has a higher capacity for acetate production which allows faster growth in an environment with high acetate concentration. Meanwhile, AckA2 is important for fast acetate-dependent growth at low concentration of acetate. The results demonstrate that the two ACKs have complementary physiological roles in L. lactis to maintain a robust acetate metabolism for fast growth at different extracellular acetate concentrations. The existence of ACK isozymes may reflect a common evolutionary strategy in bacteria in an environment with varying concentrations of acetate.

4.2 Introduction
There are many examples where products of metabolism are excreted and later re-assimilated by organisms, e.g. acetate (OAc) in E. coli [1] and ethanol in yeast [2]. The ability to switch between dissimilation and assimilation of the same metabolite is an important trait for maximizing growth in a changing environment. The acetate switch is a prominent example of this type of behavior. Acetate is one of the main metabolic products, and re-assimilation happens in order to exploit an available carbon source for further biomass formation after the primary carbon source has been depleted. A general review of the switch in acetate metabolism can be found in Wolfe [3].
*Lactococcus lactis* (*L. lactis*) is an important Gram-positive model organism which belongs to the group of Lactic Acid Bacteria (LAB) and is widely used in cheese production. In *L. lactis*, acetate may also be excreted or assimilated, depending on the environmental conditions. Figure 4.1 summarizes the relevant metabolic reactions. Under anaerobic conditions *L. lactis* produces mainly lactate as well as formate, ethanol and acetate. The amounts of formate, ethanol and acetate become very significant when growth depends on slowly fermentable sugars like maltose and galactose [4]. Acetate can also be a precursor of Ac-CoA which is vital to *L. lactis* because Ac-CoA is a precursor in fatty acid biosynthesis [5], cysteine biosynthesis [6] and peptidoglycan biosynthesis [7], etc. Acetate is required for growth when other routes to Ac-CoA are blocked. In *L. lactis* there are three known pathways leading to Ac-CoA, including the pyruvate dehydrogenase complex (PDHc), pyruvate formate lyase (PFL) and phosphotransacetylase (PTA) in conjunction with acetate kinase (ACK) (Figure 4.1). PDHc is mainly active under aerobic conditions in the presence of the cofactor lipoic acid. Under strict anaerobiosis, PDHc was shown to have low activity [8] and growth depends on PFL in the absence of acetate [9], [10]. PFL is active only anaerobically due to inactivation by oxygen [11]. The PTA-ACK pathway, which converts acetate into Ac-CoA, can support growth when PDHc and PFL are both inactive provided that acetate is added to the media. If all the three known pathways leading to Ac-CoA are blocked, *L. lactis* is unable to grow [9]. A gene predicted to encode either the AMP-forming Ac-CoA synthetase (ACS) (which is important for Ac-CoA production from acetate in *E. coli* [3]) or acyl-CoA synthetase is also present in some *L. lactis* strains but no ACS activity has been reported. There are no other annotated genes in *L. lactis* encoding enzymes known to catalyze the conversion between Ac-CoA and acetate.

![Figure 4.1. Pyruvate and Ac-CoA metabolism.](image)

Figure 4.1. Pyruvate and Ac-CoA metabolism.

Pi: inorganic phosphate. CO₂: carbon dioxide. Enzyme names are in bold. LDH: lactate dehydrogenase. PFL: pyruvate formate lyase. PDHc: pyruvate dehydrogenase complex. PTA: phosphotransacetylase. ACK: acetate kinase. ADHE: bifunctional alcohol dehydrogenase. doi:10.1371/journal.pone.0092256.g001

Interestingly, in all available sequenced genomes of *L. lactis*, two well conserved neighboring homologous genes are predicted to encode ACK. A further search on all available ACK sequences shows that in fact >300 species have multiple ACKs, including some other LAB species. The
prevalence suggests a plausible advantage conferred by multiple ACK genes to *L. lactis* and other bacteria. In the literature, the kinetics and mechanism of ACK have long been a subject of study. The kinetics of ACK among different organisms has been characterized, e.g. [12]–[15]. Crystallographic studies, site-specific mutagenesis and sequence comparisons have revealed sites important for substrate binding and catalysis, e.g. [13], [16]–[19]. Several reaction mechanisms have also been proposed [20]–[22]. Despite the massive amount of literature on ACK, surprisingly, little work has been done on ACK isozymes which exist in many species, with the exception of one kinetic study in a spirochete [23].

In the hope of revealing the significance of ACK isozymes, in this study we investigated the two ACKs in *L. lactis* by sequence analysis, characterization of transcription structure, enzyme activity and effect on growth physiology.

### 4.3 Materials and methods

#### 4.3.1 Bacterial strains and plasmids

All the *L. lactis* strains involved in this study were derived from the plasmid-free laboratory strain *L. lactis subsp. cremoris* MG1363 [24]. For overexpression of ACKs, *E. coli strain* M15 pREP4 groE5L [25] was used. The plasmid pCS1966 containing genes encoding erythromycin resistance and an orotate transporter was used for markerless gene inactivation in *L. lactis* [26]. The plasmids pLB65, harboring a gene encoding a site-specific integrase, and pLB85, containing the *gusA* reporter gene and a gene encoding erythromycin resistance, were used for constructing strains needed for *in vivo* promoter strength assessment [27]. The primers, plasmids and strains used in the study are listed in Table B.1, Table B.2 and Table B.3 respectively.

#### 4.3.2 Antibiotics

When needed erythromycin was added at 5 μg ml⁻¹ for *L. lactis*. Ampicillin and kanamycin were applied at 100 μg ml⁻¹ and 25 μg ml⁻¹ respectively for *E. coli*.

#### 4.3.3 Culture media and growth conditions

*E. coli* was grown aerobically at 37°C in Lysogeny Broth (LB). *L. lactis* strains were cultivated at 30°C without aeration in M17 broth supplemented with 2 g L⁻¹ of glucose or in chemically defined SA medium [28] devoid of acetate and supplemented with 2 g L⁻¹ of maltose (MSA). *L. lactis* growth experiments were carried out in flasks at 30°C under static conditions with slow stirring and optical density at 600 nm (OD₆₀₀) was measured regularly. As inoculum an over-night exponentially growing culture in the same medium was used and the start OD₆₀₀≈0.02. The growth rate was calculated as the average of three replications. The cell density was correlated to the cell mass of *L. lactis* to be 0.36 g (dry weight) per liter of SA medium of OD₆₀₀ = 1.
4.3.4 Quantification of maltose and fermentation products
HPLC was employed to measure the concentration of maltose, lactate, formate and acetate in the samples taken during the growth experiments as previously described [29].

4.3.5 DNA techniques
The method used to isolate the chromosomal DNA from *L. lactis* was modified from a previous method [30]. PCR amplification, restriction, ligation, transformation and plasmid purification from *E. coli* were performed following procedures described in Sambrook et al.[31] and the description from the manufacturer of the enzymes used. Electroporant cells of *L. lactis* were grown in M17 broth supplemented with 10 g L\(^{-1}\) glucose and 10 g L\(^{-1}\) glycine and transformed by electroporation as described previously [32].

4.3.6 Gene inactivation
Gene inactivation was achieved by deleting the whole gene or part of the gene containing the necessary active sites using the plasmid pCS1966 [26]. \(\approx\)800-bp regions upstream and downstream of the target to be deleted were PCR amplified and inserted into pCS1966. The resulting plasmids were used as previously described [26].

4.3.7 Construction of gusA reporter strains
The promoter containing region upstream a specific gene was PCR amplified and inserted into plasmid pLB85 and transformed into the desired *L. lactis* strain expressing phage TP901-1 integrase as described previously [27]. Transformants were selected on GM17 with erythromycin and verified by sequencing using primers CSO50 and CSO263 (Table B.1).

4.3.8 Rapid amplification of cDNA ends (RACE)
For RNA isolation, cells of MG1363 were harvested from an exponentially growing SA culture supplemented with 2 g L\(^{-1}\) glucose or maltose, with 2 μg ml\(^{-1}\) of lipoic acid and nucleosides. Cells were then resuspended in 200 μl Solution I (0.3 M sucrose and 0.01 M NaAc, pH 4.8) and 200 μl preheated Solution II (2% SDS and 0.01 M NaAc, pH 4.8). 400 μl phenol/acetate (phenol equilibrated with 100 mM NaAc, pH 4.8) was added and the mixture was disrupted by glass beads (106-μm diameter; Sigma, Prod. No. G4649) using a FastPrep (MP Biomedicals, Santa Ana, USA). The resulting lysate was centrifuged and the water phase was extracted by phenol/acetate two times and finally by phenol/acetate mixed with chloroform in a 1:1 ratio. RNA was precipitated by ethanol and dissolved in DEPC-treated water. RACE was performed using the SMARTer™ RACE cDNA Amplification Kit (Clontech) according to the instructions of the manufacturer.
4.3.9 Overproduction of *L. lactis* ACK in *E. coli*

The two ACK genes (*ackA1, ackA2*) from MG1363 were PCR amplified using primers 71f, 71r and 62f, 62r respectively. After digestion with BglII, SalI and BamHI, SalI respectively, the fragments were inserted into the vector pQE30 (Qiagen) digested with the same enzymes and subsequently introduced into the *E. coli* strain M15 pREP4groESL [25]. The strains were grown and His-tagged ACKs were produced via IPTG induction and purification on a Ni-NTA resin (Qiagen) according to the manufacturer's instruction. Purified protein was gel-filtrated on a PD-10 column (GE Healthcare) thereby transferring it to Solution A (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10% glycerol). Protein concentration was determined using Bradford Reagent (Sigma, Prod. No. B6916) and a protein standard (200 mg ml\(^{-1}\) BSA, Sigma, Prod. No. P5369), following the protocol provided by the manufacturer. The molecular weight of the protein was estimated by gel filtration using a HiPrep™ 16/60 Sephacryl™ S-300 High Resolution column (GE Healthcare) and a Gel Filtration Standard (BioRad, Cat. No. 151-1901). The mobile phase used was 0.05 M sodium phosphate, 0.15 M NaCl, pH 7.0 and the flow rate was 0.2 ml min\(^{-1}\). Proteins were detected using the Ultimate 3000 Diode Array Detector (Dionex) at 280 nm.

4.3.10 Measurement of ACK activities

ACK activities were measured on either purified proteins or in cell extracts. Cell extracts were obtained by harvesting exponentially growing cells which were then resuspended in extract buffer [29] and disrupted by glass beads (106-μm diameter; Sigma, Prod. No. G4649) using a FastPrep (MP Biomedicals, Santa Ana, USA). The master buffer used for the assay was adapted from Goel et al. [33]: 100 mM HEPES, 50 mM NaCl, 400 mM potassium glutamate, 1 mM potassium phosphate and 10× diluted metal ions present in SA medium, adjusted to pH 7.5 with potassium hydroxide. For the production of acetate from Ac-P, the same assay mix was used as in Goel et al. [33]: master buffer, 5 mM MgSO\(_4\), 2 mM D-glucose, 0.4 mM NAD\(^+\), 8.5 U ml\(^{-1}\) hexokinase, 12.7 U ml\(^{-1}\) D-glucose 6-phosphate dehydrogenase, with varying amounts of ADP and Ac-P. For measurements of \(V_{max}\) in cell extracts, 3 mM ADP and 2 mM Ac-P were used. For the reverse direction, the assay mix was modified from a previous article [34]: master buffer, 4.2 mM MgCl\(_2\), 1.7 mM phosphoenolpyruvate, 0.24 mM NADH, 9 U ml\(^{-1}\) pyruvate kinase, 12 U ml\(^{-1}\) lactate dehydrogenase, with varying amounts of ATP and potassium acetate. For measurement of \(V_{max}\) in cell extracts, 4 mM ATP and 200 mM acetate were used. The enzyme activities were determined by monitoring OD\(_{340}\) corresponding to the concentration of NADH using the Infinite® M1000 PRO microplate reader (TECAN) and the accompanying software Magellan. The 96-well microplates used were purchased from Greiner Bio-one (Cat. No. 655901).
4.3.11 Measurement of β-glucuronidase activity
The procedure used for measuring β-glucuronidase activities was modified from Miller [35] and Israelsen et al. [36].

4.3.12 Sequence analysis
Protein sequences were obtained from UniProt (http://www.uniprot.org/). Nucleotide sequences were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank/). Multiple alignments and phylogenetic construction were performed using MUSCLE [37] in CLC Main Workbench (http://www.clcbio.com/products/clc-main-workbench/). Phylogenies were visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree/). RNA secondary structure was predicted using Vienna RNA Web Services [38].

4.4 Results

4.4.1 Homologous sequences of AckA1 and AckA2 in L. lactis MG1363
Protein sequences of the two ACKs in MG1363, AckA1 and AckA2 encoded by ackA1 and ackA2 respectively, are homologous with an identity of 68%. Alignment to ACKs from Salmonella typhimurium and Methanosarcina thermophila (M. thermophila) whose structures are known [16], [17] (Figure B.1) showed conservation of active site, substrate, nucleotide triphosphate and metal binding sites except one residue in an ATP binding site (V331 of AckA1 and I331 of AckA2 respectively) which is also not conserved among other organisms. It is thus difficult to predict differences in enzymatic properties based on sequences alone.

4.4.2 Multiple ackA genes existing in L. lactis and many other species
To understand the evolutionary relationship between the ACKs from Lactococcus and other closely related LAB, a search for species with multiple acetate kinases was initiated. Under the genus Lactococcus, there are a total of 15 strains of three species, ten L. lactis, four L. garvieae and one L. raffinolactis. Interestingly, all Lactococcus strains except L. raffinolactis harbor two homologous ACK genes. For Streptococcus which is closest to Lactococcus, in contrast, among more than 500 strains with available ACK sequences, only 16 of them have two or more ACKs. It is however not a distinctive feature for Lactococcus but in fact a very general phenomenon in bacteria. Among the 11,100 entries predicted to encode ACKs in Uniprot, around 5,000 of them are not the unique gene for ACK in an organism. These multiple ACK genes exist in 2,242 strains from 320 species under 135 genera. Interested readers are referred to Appendix AB.2 for the complete list and the criteria for distinguishing multiple ACKs.
4.4.3 Two types of ACK conserved in *Lactococcus*

A multiple alignment including sequences from all *Lactococcus*, several *Streptococcus* and *Lactobacillus* representatives with the experimentally studied species as outgroup was performed to construct a phylogeny (Figure 4.2). From the phylogeny, it became clear that, except for *L. raffinolactis*, one ACK in each *Lactococcus* strain forms a monophyly and the other ACK in each strain forms another (filled triangles in Figure 4.2). A multiple alignment indicating the conserved differences of amino acid sequences between the two types of ACK in all *Lactococcus* strains is also shown in Fig. S2. For *Streptococcus* and *Lactobacillus*, similar but more complex relationships could be observed. For example, a *S. urinalis* strain has two ACKs more similar to ACKs in other species than to each other. One of the three ACKs of *Lb. sakei* also has a larger divergence with the other two ACKs than with the ACKs from bacteria in other phyla.

From the alignment of the two types of ACK in *L. lactis* with ACKs of known structure [16],[17], [39] (Figure B.2), it was observed that some important residues conserved within each type of ACK were different between the two types, e.g. position 331 (relative to MG1363’s AckA1) in an ATP binding site and position 287–291 including a deletion on a helix containing an ATP binding site.

**Figure 4.2. Phylogeny of acetate kinases from Lactococcus, Streptococcus and other species.**

One ACK from each *L. lactis* and *L. garvieae* strain forms a monophyletic group and the other ACK forms another (filled triangles). A triangle represents a cluster of sequences lumped in one line and the number in the bracket is the size of the cluster. doi:10.1371/journal.pone.0092256.g002
4.4.4 Predicted transcription terminator between the two ACK genes in *L. lactis*

In all *L. lactis* strains, the genes for the two ACKs are neighbors of each other. In MG1363, the gene upstream was annotated as *ackA1* and the other as *ackA2*. To see if they form an operon, the intergenic RNA secondary structure was predicted using ViennaRNA Web Services [38]. A stem-loop structure followed by a poly-U sequence which is a potential transcription terminator located 8 bp downstream of the stop codon of *ackA1* was predicted (Figure 4.3). The prediction is conserved for all sequenced *L. lactis*.

![Figure 4.3. Nucleotide sequence upstream of *ackA1* and *ackA2*.
TSS: putative transcription start site. Putative −35 and extended −10 element are underlined. Shaded nucleotides in −10 element: T Gn motif. cre site responsible for carbon catabolite repression is dotted underscored. Bracket pairs represent the base pairing in the predicted stem-loop structure which is conserved among all *L. lactis* strains. doi:10.1371/journal.pone.0092256.g003]

4.4.5 Distinct transcription start sites for *ackA1* and *ackA2*

A 5′-end RACE was conducted on RNA samples from MG1363 growing on glucose and maltose respectively to locate the transcription start site (TSS) of the two *ackA* genes. For each gene, an individual TSS was identified and putative −35 element and extended −10 element containing a TGn motif [40] were proposed (Figure 4.3). We were unable to demonstrate the existence of an additional transcript containing both *ackA2* and *ackA1* although this should have been possible for the RACE approach used.

4.4.6 Distinct transcription units and activities

Reporter fusions were constructed as a quantitative approach to examine the transcription activity of *ackA1* and *ackA2*. Five resulting strains with fragments A–E (Figure 4.4a) respectively fused transcriptionally to *gusA* were grown on MSA medium and the β-glucuronidase activities were determined (Figure 4.4b).
Chapter 4. Acetate kinase isozymes confer robustness in acetate metabolism

4.4.7  **Huge differences in $k_{cat}$ and $K_m$ for acetate**

The molecular weights of the His-tagged ACKs were estimated to be 100 kDa for AckA1 and 84 kDa for AckA2 using gel filtration, close to a double of the monomer (43 kDa). Both proteins are concluded to be homodimeric. ACK activities of the enzymes were measured (Figure B.3). Both AckA1 and AckA2 were active in both directions. $k_{cat}$ and $K_m$ for all four substrates were estimated (Table 4.1). Two exceptional differences between AckA1 and AckA2 were first the much higher $k_{cat}$ of AckA1 in both directions (8-fold higher for acetate production and 4-fold higher for the reverse) and second the much lower apparent $K_m$ for acetate of AckA2 (1.87 mM) compared to that of AckA1 (22.07 mM).

---

**Figure 4.4. Transcription activities of $ackA$ genes.**

(A) Fragments used for transcriptional fusion. Numbers on arrows refer to primers in Table B.1. PP: putative promoter. UR$_{1s}$: ≈500 bp upstream of $ackA1$. UR$_{1e}$: ≈50 bp after the start codon of $ackA1$. UR$_{2s}$: ≈500 bp upstream of $ackA2$. UR$_{2e}$: ≈50 bp after the start codon of $ackA2$. T: position just after the predicted terminator. (B) The β-glucuronidase activities induced by the corresponding fragments. bg: background activity from the control strain without any promoter upstream of $gusA$. Error bars are equal to standard deviations of measurements on three replications. doi:10.1371/journal.pone.0092256.g004

Fragment A including the putative promoter (PP) of $ackA1$ resulted in the highest activity of 5 Miller units. Fragment B includes both the PPs of $ackA1$ and $ackA2$ whereas fragment C includes only the PP of $ackA2$. They resulted in very similar activities (≈2 Miller units). This shows that the PP of $ackA1$ had negligible effect on the transcription of $ackA2$. Fragment D and E, starting at the same 5′ end of fragment B and C respectively and both ending just after the predicted terminator but before the PP of $ackA2$, resulted in activities indistinguishable from the background activity (‘bg’ in Figure 4.4b). This demonstrates that the predicted terminator is effective and that the transcription of $ackA2$ is governed by its own promoter.
Table 4.1. Estimated $K_m$ and $k_{cat}$ for AckA1 and AckA2.

<table>
<thead>
<tr>
<th></th>
<th>AckA1</th>
<th>AckA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_m$ (mM)</td>
<td>0.35 (0.02)</td>
<td>0.086 (0.02)</td>
</tr>
<tr>
<td>Ac-P</td>
<td>0.94 (0.09)</td>
<td>1.15 (0.12)</td>
</tr>
<tr>
<td>ADP</td>
<td>22.07 (1.05)</td>
<td>1.87 (0.29)</td>
</tr>
<tr>
<td>OAc</td>
<td>0.086 (0.0078)</td>
<td>0.21 (0.12)</td>
</tr>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>OAc $\rightarrow$ Ac-P</td>
<td>3234 (221)</td>
</tr>
<tr>
<td>Ac-P $\rightarrow$ OAc</td>
<td>1033 (98)</td>
<td>282 (28)</td>
</tr>
</tbody>
</table>

Ac-P: acetyl-phosphate. OAc: acetate. Values in bracket represent the standard error of estimation from ≥20 data points. doi:10.1371/journal.pone.0092256.t001

4.4.8 Additive cell extract activities of mutant strains

To study the physiological roles of the two ACKs in *L. lactis*, three mutant strains, MG1363$\Delta$ackA1, MG1363$\Delta$ackA2 and MG1363$\Delta$ackA1$\Delta$ackA2, were constructed by inactivating ackA1, ackA2 and both respectively. $V_{max}$ of ACK in MG1363 and the three mutant strains growing on MSA media were measured (Figure 4.5a). MG1363 showed the highest activity, followed by MG1363$\Delta$ackA2 and then MG1363$\Delta$ackA1. MG1363$\Delta$ackA1$\Delta$ackA2 had the lowest activity. An interesting observation was the additivity of the activities. When subtracting the activity in MG1363$\Delta$ackA1$\Delta$ackA2 (which represents the background activity) from the activities in the other three strains, the sum of the activities in MG1363$\Delta$ackA1 and MG1363$\Delta$ackA2 was approximately equal to the activity in MG1363. The implication is discussed below.

To verify the much lower $K_m$ for acetate of AckA2, $V_{max}$ was also determined in the presence of 1 mM acetate (Figure 4.5b). MG1363$\Delta$ackA1 did show a higher activity than MG1363$\Delta$ackA2. This is opposite to what was observed under normal assay conditions with 200 mM acetate and agrees with the differences in the $K_m$.

4.4.9 Slower acetate production by MG1363$\Delta$ackA1 at a high extracellular acetate concentration

To test whether the two ACKs performed differently *in vivo*, growth experiments of MG1363, MG1363$\Delta$ackA1, MG1363$\Delta$ackA2 and MG1363$\Delta$ackA1$\Delta$ackA2 were conducted on MSA media with or without 50 mM acetate. Figure 4.6 shows representative growth curves of the four strains and the average growth rates. In all experiments the wild type and single-deletion strains were able to grow up to an OD$_{600}$=1 where the HPLC analysis showed that the sugar had been consumed (data not shown). The double deletion strain MG1363$\Delta$ackA1$\Delta$ackA2 stopped growing at a much lower cell density of OD$_{600}$=0.1. When acetate was absent, MG1363 and the two single deletion strains grew with similar growth rates about 0.45 h$^{-1}$. With 50 mM acetate present in the media, MG1363$\Delta$ackA1 had a significantly reduced growth rate of 0.38 h$^{-1}$, equal to a 20% reduction compared to MG1363 and MG1363$\Delta$ackA2 (0.49 h$^{-1}$).
Chapter 4. Acetate kinase isozymes confer robustness in acetate metabolism

Figure 4.5. ACK activities in crude extracts of MG1363 and derived ackA deletion strains.

(A) Activities under normal assay conditions with 200 mM acetate. When subtracting the activity in MG1363ΔackA1ΔackA2 (which represents background activity) from the activities in the other three strains, the sum of the activities in MG1363ΔackA1 and MG1363ΔackA2 was approximately equal to the activity in MG1363. (B) Activities of MG1363ΔackA1 and MG1363ΔackA2 in the presence of 1 mM or 200 mM of acetate (OAc). Error bars are equal to standard deviations of measurements on three replications. doi:10.1371/journal.pone.0092256.g005

Figure 4.6. Representative growth curves and growth rates of MG1363 and derived ackA deletion strains.

Growth with (A) 0 mM and (B) 50 mM acetate. Only data obtained 5 hours after the start of the experiments are plotted for better visualization. (C) Table showing the legend and average growth rates. Values in brackets represent standard deviations of three replications. doi:10.1371/journal.pone.0092256.g006

The product formation from the three growing strains was also measured (Table 4.2). In the absence of acetate, the distribution of fermentation products was very similar for all three strains. Since MG1363ΔackA1ΔackA2 was unable to grow and produce acetate after OD$_{600}$=0.1, the acetate
production in MG1363ΔackA1 and MG1363ΔackA2 can be attributed to AckA2 and AckA1 respectively. This implies that both individual ACKs were able to sustain a flux equal to that in the wild type where both ACKs were present. In the presence of 50 mM acetate, however, the reduced acetate production rate concomitant with the reduced growth rate (Figure 4.6c), formate production rate and maltose uptake rate in MG1363ΔackA1 reveals that the flux entering the mixed acid branch decreased while the lactate flux remained unchanged (Table 4.2). This indicates that AckA2 alone in MG1363ΔackA1 was unable to maintain the same flux as in MG1363 and MG1363ΔackA2 where AckA1 was present. It can thus be concluded that AckA1 performed better than AckA2 in acetate production in the presence of a high concentration of acetate (50 mM) in the media.

Table 4.2. Average specific rates of consumption of maltose, production of lactate, formate and acetate of MG1363, MG1363ΔackA1 and MG1363ΔackA2 at 0 or 50 mM of extracellular acetate.

<table>
<thead>
<tr>
<th>Specific rate of consumption/production (mmol h⁻¹ gdw⁻¹)</th>
<th>MSA, 0 mM OAc</th>
<th>MSA, 50 mM OAc</th>
</tr>
</thead>
<tbody>
<tr>
<td>maltose, lactate, formate, acetate</td>
<td>MSA, 0 mM OAc</td>
<td>MSA, 50 mM OAc</td>
</tr>
<tr>
<td>MG1363</td>
<td>7.9 (0.5)</td>
<td>20.6 (1.5)</td>
</tr>
<tr>
<td>MG1363ΔackA1</td>
<td>8.0 (0.5)</td>
<td>19.9 (1.4)</td>
</tr>
<tr>
<td>MG1363ΔackA2</td>
<td>7.3 (0.6)</td>
<td>18.8 (1.1)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2</td>
<td>12.5 (0.9)</td>
<td>20.1 (1.6)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.5 (0.8)</td>
<td>14.0 (1.4)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>18.0 (1.1)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>20.1 (1.6)</td>
<td>14.0 (1.4)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.7 (1.0)</td>
<td>19.0 (1.8)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>18.0 (1.1)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.5 (1.2)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>18.0 (1.1)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.5 (1.2)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>18.0 (1.1)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.5 (1.2)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>18.0 (1.1)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.5 (1.2)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>18.0 (1.1)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.5 (1.2)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>8.0 (0.6)</td>
<td>7.7 (0.7)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>18.0 (1.1)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>12.5 (1.2)</td>
<td>14.3 (1.2)</td>
</tr>
<tr>
<td>MG1363ΔackA1ΔackA2Δpfl</td>
<td>7.4 (0.4)</td>
<td>8.9 (1.0)</td>
</tr>
</tbody>
</table>
| MG1363ΔackA1ΔackA2Δpfl                                 | 8.0 (0.6)    | 7.7 (0.7)      | OAc: acetate. Values in brackets represent standard deviations of three replications. doi:10.1371/journal.pone.0092256.t002

4.4.10 Slower acetate uptake by MG1363ΔackA2Δpfl at low acetate concentrations

An acetate-assimilating growth condition was created by excluding lipoic acid from the medium and knocking out PFL in the ackA deletion strains. The PFL-deleted strains, MG1363Δpfl, MG1363ΔackA1Δpfl, MG1363ΔackA2Δpfl and MG1363ΔackA1ΔackA2Δpfl, were grown in MSA media supplemented with acetate. Figure 4.7 shows the growth in media supplemented with no acetate, 8 mM, 12 mM and 50 mM acetate respectively. When acetate was absent, all four strains stopped growing at OD₆₀₀ ≈0.1 (Figure 4.7a), showing that the cells depended on acetate for growth beyond this point. When acetate was added to the media, further growth could be seen for all strains except MG1363ΔackA1ΔackA2Δpfl. A clear transition occurred between OD₆₀₀ = 0.2 and 0.3, after which growth depends on acetate. The final OD₆₀₀ was around 0.6 where HPLC analysis indicated that all maltose had been consumed.

Very slow acetate-dependent growth was observed for MG1363ΔackA2Δpfl at 8 or 12 mM of extracellular acetate, >10-fold slower than MG1363ΔackA1-Δpfl and MG1363Δpfl (Figure 4.7b, c). MG1363ΔackA1Δpfl also grew even faster than MG1363Δpfl at these acetate concentrations. When the acetate concentration increased to 50 mM, however, the growth rates were similar for all three strains (0.22–0.23 h⁻¹, Figure 4.7d).
The consumption of acetate was also quantified (Table 4.3). The acetate uptake rate of MG1363ΔackA1Δpfl was significantly higher than that of MG1363Δpfl (2–3 fold) and MG1363ΔackA2Δpfl (>6 fold) at 8 or 12 mM extracellular acetate. They were nonetheless indistinguishable at 50 mM extracellular acetate. Since MG1363ΔackA1ΔackA2Δpfl did not show acetate-dependent growth, the acetate uptake in MG1363ΔackA1Δpfl and MG1363ΔackA2Δpfl could be attributed to the presence of AckA2 and AckA1 respectively. It can thus be concluded that at a high acetate concentration (50 mM), both individual ACKs were able to take up acetate as fast as in MG1363Δpfl whereas at low acetate concentrations (≤12 mM), AckA2 has a significant higher capability for acetate uptake than AckA1.

Table 4.3. Average specific rates of consumption of maltose, acetate and production of lactate of MG1363Δpfl, MG1363ΔackA1Δpfl and MG1363ΔackA2Δpfl.

<table>
<thead>
<tr>
<th></th>
<th>Specific consumption/production rate (mmol h⁻¹ gdw⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8 mM OAc</td>
</tr>
<tr>
<td>Maltose consumption</td>
<td>lactose</td>
</tr>
<tr>
<td>MG1363Δpfl</td>
<td>1.3 (0.1)</td>
</tr>
<tr>
<td>MG1363ΔackA1Δpfl</td>
<td>2.1 (0.2)</td>
</tr>
<tr>
<td>MG1363ΔackA2Δpfl</td>
<td>0.2 (0.03)</td>
</tr>
</tbody>
</table>

Formate production was not detectable in all cases. Values in brackets represent standard deviations of three replications.

N.D.: not detectable. doi:10.1371/journal.pone.0092256.t003
4.5 Discussion

4.5.1 Prevalence of multiple ACKs in bacteria

ACK sequence data from Uniprot revealed that the existence of multiple ACKs is a very common phenomenon in bacteria which is interesting and emphasizes its possible importance. Unfortunately, there are few studies on ACK isozymes in the literature except for a spirochete [23]. We believed that the importance of ACK isozymes on acetate metabolism is being neglected while the acetate metabolism in bacteria is still a subject under active research, for instances, current generation of Geobacter sulfurreducens's growth on acetate [41] and acetate dependency of the probiotic Lactobacillus johnsonii [42]. Interestingly, the strains used in the two studies also have multiple ACKs. Knowledge on ACK isozymes in these organisms may provide insights into these studies and their applications. The current study attempts to fill the gap by studying the two ACKs in L. lactis at different levels.

4.5.2 ackA1 and ackA2 in L. lactis probably resulted from gene duplication

Protein sequence analysis revealed the conserved differences between the two ACKs found in Lactococcus. This brought an insight into the potential evolutionary advantage of having ACK isozymes. From the simple phylogenetic analysis, the two ackA genes may have resulted from a duplication event in a common ancestor of Lactococcus which had already been differentiated from Streptococcus. To prove this point, however, a more formal phylogenetic analysis is required.

4.5.3 AckA1 and AckA2 in MG1363 being isozymes rather than subunits

With respect to the expression of ackA1 and ackA2, our results suggest that the two genes are transcribed individually rather than in an operon. This is actually consistent with the Northern Blot results in Lopez de Felipe and Gaudu (2009) [43] where a transcript of around 1 kb was found for ackA1. Interestingly, in their study, ackA1 was assumed to encode one subunit of ACK in L. lactis. This question may be worth asking because the two ACKs are in fact homologous to each other.

In the current study, nevertheless, the state of being isozymes rather than subunits of one ACK for the two ackA gene products was assumed for several reasons. First, we found that the individual His-tagged enzymes could catalyze the reaction in both directions. Second, the ackA mutant strains could produce as well as utilize acetate. Third, the crude extracts from ackA mutant strains showed additive activities (\( V_{\text{MG1363 max}} = V_{\text{MG1363 ackA1 max}} + V_{\text{MG1363 ackA2 max}} - V_{\text{MG1363 ackA1ACKA2 max}} \)). If a hetero-oligomeric form of ACK with different kinetic properties exists, the additivity is less likely to hold. Fourth, most of the ACKs reported in the literature appear to be homodimeric [16],[17],[22],[44]–[47]. Finally, we have data analogous to those in Fig. 4b showing that when growing on glucose, the promoter activity of ackA2 was 10-fold lower than that of ackA1 and was close to the value of background activity (unpublished results). This is consistent with the cre site identified 6 bp downstream of the ackA2's TSS (Fig. 3) which is subject to carbon catabolite repression [48]. In light of this huge difference
between the transcriptional activities, it is unlikely that the two gene products from \textit{ackA1} and \textit{ackA2} form one protein complex. However, the possibility that hetero-oligomeric ACK exists cannot be entirely ruled out.

### 4.5.4 Possible different roles suggested by enzyme kinetics

From the His-tagged purified enzymes, AckA1 was shown to have much higher turnover number $k_{\text{cat}}$ than AckA2 whereas AckA2 has a higher affinity towards acetate. The apparent $K_m$ is the lowest of all the reported ACKs where $K_m$ for acetate usually is notoriously high (the highest being 300 mM in \textit{E. coli} [13]). The kinetic properties of the two ACKs thus suggest a possible complementary role in metabolism. For the effect of His-tagging on enzymes, we have looked into the 3D structure of the ACK from \textit{M. thermophile} (PDB ID: 1TUY) [22] which is homologous to AckA1 and AckA2 in \textit{L. lactis}. The first few residues from the N-terminus are outside the catalytic core. Together with the consistency between the assays on purified enzymes and crude extracts, we believed that the His-tag is unlikely to interfere with the reaction.

### 4.5.5 Physiological roles of ACK reported in literature

Among the ACKs reported previously, some were found to have ATP production as their primary physiological role while some are more likely to be responsible for acetate activation. For instance, the ACK in \textit{Lactobacillus sanfranciscensis} was suggested to take the role of ATP formation [13]. The ACKs in \textit{Bacillus subtilis} were shown to be non-essential for growth on acetate and meanwhile important for excretion of excess carbohydrate by producing acetate [49].

With respect to examples of acetate activation, in \textit{Corynebacterium glutamicum}, ACK activities were proven to be necessary for growth on acetate [50]. Another interesting case is \textit{M. thermophile} which is acetotrophic. The $K_m$ for acetate of the ACK from \textit{M. thermophile} was found to be 22 mM [18]. Site-directed mutagenesis in the same study revealed that only a single-residue mutation could cause a 10-fold lower $K_m$ for acetate concomitant with a 6-fold reduction in $k_{\text{cat}}$. This striking similarity between the pair of ACKs in \textit{M. thermophile} (wild-type and mutated) and the pair of ACKs in \textit{L. lactis} (AckA1 and AckA2) leaves a possible hint for how AckA1 and AckA2 differentiated and specialized. The author suggested that the sacrifice of a low $K_m$ in return for a high $k_{\text{cat}}$ conferred the advantage of more rapid acetate uptake to \textit{Methanosarcina} species in an environment with a high acetate concentration [18].

These are only a few examples among many different studies. It must be noted that despite the particular functions of ACK demonstrated in the mentioned studies, one should not exclude other possibilities. The physiological role might be dependent on the nutrients available and complementary to other enzymes like the AMP-forming ACS in bacteria. For example, in \textit{E. coli}, a number of studies on ACK-deficient mutants showed that PTA-ACK is the primary pathway for acetate production, e.g. [51], [52]. Other studies found that it is important for growth on high acetate concentration ($\geq$25
Acetate kinase isozymes confer robustness in acetate metabolism

90 mM) whereas growth on low acetate concentration (≤2.5 mM) depends on ACS [1], [53] (reviewed in [3]). This example of PTA-ACK complementary to ACS in E. coli also resembles AckA1 and AckA2 in the sense that ACS has a much lower $K_m$ for acetate (0.2 mM) and lower $V_{max}$ [53]. The difference lies in the irreversibility of ACS in E. coli and the dependence of AckA2 on PTA to produce Ac-CoA in L. lactis. A final example is a spirochete with two ACKs [23]. They had a lower $K_m$ for Ac-P and acetate respectively. The authors mentioned the possibility of the two ACKs being specialized in different directions respectively.

4.5.6 Growth on maltose as a test of the physiological roles in L. lactis

To find out whether the two ACKs have different physiological roles, mutant strains were constructed and their response to acetate during growth was examined. In our growth experiments, maltose was chosen as the carbon source because for MG1363 growing on maltose, a more significant amount of formate, acetate and ethanol is produced than on glucose [54]. Via growth on maltose the capacity of the two ACKs to bear a high flux from glycolysis can be tested. Another reason is the much lower ATP/ADP ratio in L. lactis when growing on maltose than on glucose. The ATP/ADP ratio in MG1363 was around 9 when growing on glucose [55] and was around 4 when growing on maltose [56]. The lower ATP/ADP ratio provided a more stringent condition for acetate uptake in the mutant strains.

4.5.7 Complementary roles in acetate metabolism

Our results show that under favorable conditions either one of the ACKs is sufficient for the dual function of acetate production and uptake. In an environment where the concentrations of acetate, lipoic acid (activating PDHc) and oxygen (inactivating PFL) are varying, nonetheless, AckA1 and AckA2 have their own advantages and complement each other to allow fast growth at different extracellular acetate concentrations. In an environment with high acetate concentration (50 mM), AckA1 showed its superior capability of acetate production. This is consistent with the kinetic properties in vitro. The much higher affinity for acetate of AckA2 probably led to a larger effect of product inhibition by extracellular acetate diffused into the cells. In contrast, in a dynamic environment where PFL and PDHc are inactive, e.g. containing oxygen and without lipoic acid, our results from the growth experiments of PFL and ACK defective strains show that AckA2 is important for acetate uptake when the acetate source is scarce (≤12 mM). The lower growth yield compared to PFL-effective strains (OD$_{600}$ = 0.6 vs 1) was probably a result of loss of the ATP generated from acetate production combined with additional ATP consumed for acetate uptake.

Another possible function of the two ACKs that should not be overlooked is the emergent properties of combining the isozymes. A possibility is thus a switch to fine-tune the direction and rate of the reaction in response to the cellular requirement by altering the expression of the two ackA genes. It would be interesting to look into the regulation of the expression of $ackA1$ and $ackA2$ to examine this hypothesis.
4.5.8 PTA-ACK as the only pathway interconverting Ac-CoA and acetate in *L. lactis*

The inability of MG1363ΔackA1ΔackA2 and MG1363ΔackA1ΔackA2Δpfl to grow on MSA media suggests the absence of other pathways involved in the interconversion between Ac-CoA and acetate. For MG1363ΔackA1ΔackA2, the only known pathway left for catabolizing Ac-CoA is the NAD$^+$-generating bi-functional alcohol dehydrogenase (ADHE) because of the lack of ACKs. The redox imbalance could lead to the accumulation of toxic intermediate metabolites such as acetaldehyde which prevents growth. For MG1363ΔackA1ΔackA2Δpfl, no acetate was assimilated to form Ac-CoA to satisfy anabolic requirements. If other pathways for either direction exist, one of the strains should be able to grow. Thus, these results further emphasize the importance of AckA1 and AckA2 in the acetate metabolism of *L. lactis*. For the initial growth up to OD$_{600}$≈0.1 of the two strains, it was found to be caused by the small amounts of lipoic acid present as impurity in the amino acids composing the media which could activate PDHc. Adding lipoic acid to these cultures indeed allowed growth beyond OD$_{600} = 0.1$ and growth experiments in media with reduced amounts of amino acids showed that these strains stopped growing at a lower OD$_{600}$ whereas the wild-type MG1363 was unaffected (data not shown). This indicated the presence of small amounts of lipoic acid in the amino acid stock which caused the initial growth of MG1363ΔackA1ΔackA2 and MG1363ΔackA1ΔackA2Δpfl.

4.6 Conclusions

In conclusion, the present study demonstrated the different and yet complementary roles of the two acetate kinases in *L. lactis* MG1363 with one being specialized in acetate production and the other in acetate uptake. It was observed from the sequence and phylogenetic analysis, supported with transcriptional analysis and the enzyme kinetics, and finally confirmed by the different growth behavior of mutant strains harboring only ackA1 and ackA2 respectively. The findings can be of great significance in bacterial metabolism in light of the fact that more than 300 species of organisms actually have multiple ACKs. Evolution to multiple ACKs specialized in complementary functions may be a common strategy in bacteria in response to the dual nature of acetate which can be an essential substrate but also an inhibitor for growth depending on environmental conditions.

4.7 Acknowledgments

We thank Jun Chen for help with constructing the strain MG1363Δpfl. We also thank the transnational SysMOLAB2 consortium for a fruitful discussion.

4.8 Author contributions

Conceived and designed the experiments: SHJC CS PRJ. Performed the experiments: SHJC LN. Analyzed the data: SHJC LN. Contributed reagents/materials/analysis tools: SHJC LN CS RPJ. Wrote the paper: SHJC LN CS PRJ.
References

Chapter 5. Influence of NADH/NAD\(^+\) on mixed-acid fermentation

5.1 Introduction

The cofactor ratio NADH/NAD\(^+\) has been proposed as a key determinant of the switch between homolactic fermentation and mixed-acid fermentation in *L. lactis* [1, 2]. In these two early studies, a higher NADH/NAD\(^+\) ratio and higher triosephosphate pool, especially glyceraldehyde 3-phosphate (GAP), were observed at higher glycolytic fluxes. Together with the findings by *in vitro* enzyme assay that pyruvate formate-lyase (PFL) was inhibited by GAP and that glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was very sensitive to high NADH/NAD\(^+\) ratio, the authors of the two studies established the model in which a high NADH/NAD\(^+\) ratio accompanying a high glycolytic flux activates lactate dehydrogenase (LDH), inhibits GAPDH and thus cause a high GAP pool which in turn inhibits PFL, resulting in homolactic fermentation. Mixed-acid fermentation was proposed to take place conversely when the glycolytic flux is low.

This mechanism, however, was only concluded from the difference in measured metabolite concentrations and enzyme activities. The question of whether change in the NADH/NAD\(^+\) ratio would in fact cause any change in the fermentation mode of *L. lactis* remains to be tested *in vivo*. There have been attempts to perturb the redox state by introducing NADH oxidase (NOX) activities [3–6]. Growth in these cases, nonetheless, had to be aerobic for NOX to be active while the pyruvate metabolism of *L. lactis* changes significantly under aerobic conditions compared with anaerobic conditions, e.g. no formate production due to PFL inactivated by oxygen [7], flavour compound production [8, 9].

In this study, we tested the proposed roles of NADH/NAD\(^+\) on the shift of fermentation modes in *L. lactis* growing on maltose by introducing enzyme activities to perturb the ratio anaerobically. Formate dehydrogenase (FDH) which produces CO\(_2\) and NADH from formate and NAD\(^+\) has recently been introduced into *L. lactis* MG1363 [10]. It was found that flux towards mixed-acid fermentation decreased and the ethanol-to-acetate ratio increased. Here the enzyme, 2,3-butanediol dehydrogenase (23BDH), was introduced for NAD\(^+\) regeneration by converting acetoin supplemented in the media into 2,3-butanediol (23BD) (Figure 5.1).

In the ideal case, a system for studying the effect of cofactors perturbation should be isolated from the rest of the metabolism except the exchange of cofactors. Here, under anaerobic conditions, overexpressing 23BDH can be a good approximation of an ideal system based on the following observations: (i) the flavour compounds diacetyl, acetoin and 23BD are generally not produced or produced at very low amounts under anaerobic or microaerobic conditions [8, 9, 11]; (ii) acetalactate synthase (ALS) has very low activity in non-aerated conditions [8, 11]; acetalactate decarboxylase (ALDC) favors the acetoin-forming direction (\(\Delta_r G^*\) estimated to be -1.7 or -1 kcal/mol in MetaCyc [12]) and requires the fixation of CO\(_2\) for the acetoin-consuming direction; (iii) diacetyl formation
from α-acetolactate requires oxygen; and (iv) diacetyl reductase (DAR) is irreversible. Under anaerobic conditions, therefore, since the flux going into the flavour-forming pathway is absent or negligible and meanwhile acetoin from the medium is unlikely to induce fluxes in reverse directions affecting the upstream metabolism, the overexpression of 23BDH with provided extracellular acetoin can be considered to reasonably approximate an ideal system to perturb NADH/NAD+ under anaerobic conditions.

Figure 5.1. Introducing NAD\(^+\)-regenerating enzyme activities.
Reactions in color refer to introduced activities to regenerate NAD\(^+\). Reactions in grey are thought to have low or no activity under anaerobic conditions. Enzyme or pathway names are in round rectangles. The stoichiometry from glucose to pyruvate, and from pyruvate to α-acetolactate is 1 to 2 and 2 to 1 respectively, with all other reactions being 1 to 1. LDH, lactate dehydrogenase; PDHc, pyruvate dehydrogenase complex; PFL, pyruvate formate-lyase; ADH, bi-functional alcohol dehydrogenase; PTA, phosphotransacetylase; ACK, acetate kinase; ALS, acetolactate synthase; ALDC, acetolactate decarboxylase; DAR, diacetyl reductase; 23BDH, 2,3-butanediol dehydrogenase. The reaction producing diacetyl from α-acetolactate is non-enzymatic.

5.2 Materials and methods

5.2.1 Bacteria strains and plasmids
All L. lactis strains were derived from MG1363 and IL1403. For introducing 23BDH activities, pCS4701 and pCS4639, each containing a synthetic codon-optimized operon consisting of budC/dar (encoding diacetyl/acetoin reductase, accession no. AF098800) from Klebsiella pneumonia and budC/bdh (encoding L-23BDH, accession no. AB009078) from Brevibacterium saccharolyticum ordered from GenScript following a synthetic promoter (kindly provided by Christian Solem), were used. Plasmids and strains used or constructed in this study are listed in Table 5.1 and Table 5.2 respectively.
Table 5.1. Plasmid used in the study.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCI372</td>
<td>Shuttle vector between <em>E. coli</em> and <em>L. lactis</em>; chloramphenicol resistance</td>
</tr>
<tr>
<td>pCS4639</td>
<td>pCI372 derivatives, each carrying a synthetic promoter followed by a synthetic codon-optimized operon consisting of <em>budC/dar</em> (encoding diacetyl/acetoin reductase, accession no. AF098800) from <em>Klebsiella pneumonia</em> and <em>budC/bdh</em> (encoding L-23BDH, accession no. AB009078) from <em>Brevibacterium saccharolyticum</em> ordered from GenScript; chloramphenicol resistance</td>
</tr>
<tr>
<td>pCS4701</td>
<td><em>Klebsiella pneumonia</em> and <em>budC/bdh</em> (encoding L-23BDH, accession no. AB009078) from <em>Brevibacterium saccharolyticum</em> ordered from GenScript; chloramphenicol resistance</td>
</tr>
</tbody>
</table>

Table 5.2. Strains used or constructed in the study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent strain</td>
<td></td>
</tr>
<tr>
<td>MG1363 [14]</td>
<td>A plasmid-free strain derived from <em>L. lactis</em> subsp. cremoris NCDO 712</td>
</tr>
<tr>
<td>IL1403 [15]</td>
<td>A plasmid-free strain derived from <em>L. lactis</em> subsp. lactis biovar diacetylactis CNRZ157</td>
</tr>
<tr>
<td>MG/SP1-<em>bdh</em></td>
<td>MG1363 transformed with pCS4701; Cam(^r)</td>
</tr>
<tr>
<td>MG/SP2-<em>bdh</em></td>
<td>MG1363 transformed with pCS4639; Cam(^r)</td>
</tr>
<tr>
<td>IL/SP1-<em>bdh</em></td>
<td>IL1403 transformed with pCS4701; Cam(^r)</td>
</tr>
</tbody>
</table>

Cam\(^r\), chloramphenicol resistance

5.2.2 Culture media and conditions

*L. lactis* was cultivated in M17 media or chemically defined SA media [16] at 30 °C without aeration, supplemented with 0.2% (w/v) glucose (GM17). When needed chloramphenicol were added at 5 μg ml\(^{-1}\) for *L. lactis* respectively. For growth experiments, defined SA media devoid of sodium acetate and supplemented with nucleosides (adenosine, cytidine, guanosine, thymidine, uridine and inosine, 20 mg/L), α-lipoic acid (2 mg/L) and either glucose or maltose (0.2%) were used, abbreviated as GluSALN or MalSALN respectively. To enable the detection of acetoin consumption and 23BD production by HPLC, in some cases, the buffer MOPS used in normal SALN media was replaced by 1.9% β-glycerophosphate (BPG), which is applied in the M17 media. The resultant medium is denoted by SALN(BPG) medium. Growth experiments were conducted in flasks at 30 °C under static conditions with slow stirring. Optical density at 600 nm (OD\(_{600}\)) was regularly measured. To prepare a preculture, a single colony from an agar plate was inoculated into GluSALN or MalSALN media in dilution series. Growth experiments were started by inoculating an exponentially growing overnight culture selected from the dilution series into a flask containing the same medium, up to OD\(_{600}\) = 0.001 or 0.02 for growth on glucose or maltose respectively. Biological triplicates for each experiment were performed. The cell density equal to 0.36 g (dry weight) (gdw) per liter of SALN medium at OD\(_{600}\) = 1 was used for the calculation of specific rates.

5.2.3 Quantification of sugar and fermentation products

High-performance liquid chromatography (HPLC) was employed to measure the concentration of glucose, maltose, lactate, formate, acetate, ethanol, acetoin and 23BD in the samples taken during the
growth experiments as previously described [17]. Specific rates of consumption/production for mid-exponential growth phase (OD$_{600}$ ≈ 0.4 – 0.6) were calculated. Carbon recovery was calculated as the sum of acetate, ethanol and lactate fluxes divided by sugar uptake flux in C$_3$-mole.

### 5.2.4 Extraction and quantification of NADH and NAD$^+$
Growing culture of _L. lactis_ in the mid-exponential phase (OD$_{600}$=0.4) was quenched by liquid nitrogen and stored at -20 °C. Extraction and quantification of nicotinamide adenine dinucleotides were performed using the kit NAD/NADH-Glo™ Assay (Promega) following the instruction of the provided manual. NAD$^+$ and NADH were extracted by heating in acidic and basic solutions respectively. They were then individually quantified by a luciferase assay perform on 96-well white microplates (Greiner Bio-one, cat. no. 655904) using the Infinite® M1000 PRO microplate reader (TECAN). Intracellular concentrations were estimated by assuming that 1 g (dry weight) of cells corresponded to 1.67 ml of intracellular volume [18].

### 5.2.5 DNA techniques
Electrocompeotent cells of _L. lactis_ were grown in M17 broth supplemented with 10 g L$^{-1}$ glucose and 10 g L$^{-1}$ glycine and transformed by electroporation as described previously [19].

### 5.2.6 Measurement of 2,3-butanediol dehydrogenase activity
Exponentially growing culture at OD$_{600}$=0.6 was quenched on ice and harvested, then resuspended in extract buffer [17] and disrupted by glass beads (106-μm diameter; Sigma, Prod. No. G4649) using a FastPrep (MP Biomedicals, Santa Ana, USA). The master buffer used for the assay was adapted from Goel et al. [20]: 100 mM HEPES, 50 mM NaCl, 400 mM potassium glutamate, 1 mM potassium phosphate and 10× diluted metal ions present in SA medium, adjusted to pH 7.5 with potassium hydroxide. The reaction mix included the master buffer, 5 mM acetoin and 0.25 mM NADH. The enzyme activities were determined by monitoring OD$_{340}$ corresponding to the concentration of NADH using the Infinite® M1000 PRO microplate reader (TECAN) and the accompanying software Magellan. The 96-well microplates used were purchased from Greiner Bio-one (Cat. No. 655901). Protein concentration was determined using Bradford Reagent (Sigma, Prod. No. B6916) and a protein standard (200 mg ml$^{-1}$ BSA, Sigma, Prod. No. P5369), following the protocol provided by the manufacturer.

### 5.3 Results

#### 5.3.1 Introduction of 2,3-butanediol dehydrogenase activity
To introduce 23BDH activity, the vector pCS4701 containing a gene encoding 23BDH following a synthetic promoter was transformed into MG1363 and IL1403 respectively. The resultant strains derived from MG1363 and IL1403 were called MG/SP1-<i>bdh</i> and IL/SP1-<i>bdh</i> respectively.
As a preliminary test, the fermentation products of MG1363, MG/SP1-bdh, IL1403 and IL/SP1-bdh in overnight cultures in MalSALN media were quantified (Figure 5.2). For MG/SP1-bdh, the fermentation pattern was very similar to that of the wild-type MG1363. When the substrate of 23BDH, acetoin, was added into the medium, however, substantial increase in acetate (from 5.3 to 8.3 mM) and decrease in ethanol (from 4.6 to 1.9 mM) were observed. The formate produced was also reduced (from 11.5 to 10 mM). For IL/SP1-bdh, while in the absence of acetoin, it behaved similarly, interestingly, the formate production increased substantially (from 1.4 to 3.8 mM) and the effect for acetate was even more pronounced (from 0.8 to 4 mM).

5.3.2 Faster growth of MG/SP1-bdh in the presence of acetoin in the media

To further investigate the effect of 23BDH activity and the presence of acetoin in media on the fermentation pattern, growth experiments of MG1363 and MG/SP1-bdh were conducted in MalSALN media with 0, 5 mM (MG/SP1-bdh only) or 20 mM acetoin. Interestingly, the growth rate of MG/SP1-bdh increased from 0.61 h⁻¹, a value close to the wild-type MG1363, to 0.67 h⁻¹ and 0.72 h⁻¹ as extracellular acetoin increased from 0 to 5 mM and 20 mM respectively (Figure 5.3a). Meanwhile, 20 mM acetoin present in the medium appeared to inhibit the growth of MG1363 slightly.

The specific maltose consumption rates of MG/SP1-bdh were in all cases lower than MG1363 growing in the medium with no acetoin (Figure 5.3b). Since the growth rates of MG/SP1-bdh are higher, the biomass yields (= growth rate / maltose flux) on maltose of MG/SP1-bdh were higher than those of MG1363.

5.3.3 Flux directed towards acetate production

When growing on media with acetoain, MG/SP1-bdh exhibited significant redirection of pyruvate flux compared to MG1363. For MG1363, the fermentation shifted to a more homolactic one in the presence of 20 mM action. For MG/SP1-bdh, no substantial change in lactate and formate production was observed for growth with zero or 5 mM acetoin (Figure 5.3c, d). At 20 mM acetoin, while the formate flux dropped similarly, the lactate production in MG/SP1-bdh also decreased in contrast to
the increase in MG1363. The most pronounced effect was observed in acetate production, which significantly increased with acetoin concentration in MG/SP1-\textit{bdh} (Figure 5.3e). The case of 20 mM is in particular interesting because the acetate flux (15 mmol h\(^{-1}\) gw\(^{-1}\)) was much larger than the formate flux (6 mmol h\(^{-1}\) gw\(^{-1}\)). This situation seldom happened in anaerobic conditions. Meanwhile, ethanol production was reduced significantly to virtually zero production at 20 mM acetoin (Figure 5.3f). The complete data are shown in Table C.1 and Table C.2 in Appendix C.

![Figure 5.3](image)

**Figure 5.3.** Growth rates, fluxes of MG1363 and MG/SP1-\textit{bdh} growing in MalSALN.
(a) Growth rates, specific rates of (b) maltose consumption, (c) lactate, (d) formate, (e) acetate and (f) ethanol production of MG1363 and MG/SP1-\textit{bdh} growing in MalSALN with no acetoin, 5 mM (MG/SP1-\textit{bdh}) or 20 mM acetoin were shown.

5.3.4  **No change in growth rate in media devoid of lipoic acid**

MG1363 and MG/SP1-\textit{bdh} were also cultivated in MalSAN media which was devoid of lipoic acid, the essential cofactor of pyruvate dehydrogenase complex (PDHc) at acetoin concentrations of 0, 10 and 20 mM (Table 5.3). Interestingly, the growth rate of MG/SP1-\textit{bdh} was constantly about 0.6 h\(^{-1}\) along the increase of acetoin concentrations. While the change in fermentation pattern in the presence of acetoin was in general similar to the growth in MalSALN, the decreases in formate production (from 16 to 11 or 12 mmol h\(^{-1}\) gw\(^{-1}\)) and increases in acetate production (from 9 to 10 or 11 mmol h\(^{-1}\) gw\(^{-1}\)) at higher acetoin concentrations were less pronounced. An obvious difference is the acetate
flux being lower than the formate flux in all cases contrary to the case of MG/SP1-\textit{bdh} growing in MalSALN with 20 mM acetoin.

**Table 5.3.** MG1363 and MG/SP1-\textit{bdh} growing in MalSAN with or without acetoin.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Acetoin added</th>
<th>Growth rate (h$^{-1}$)</th>
<th>Specific rate of sugar consumption and product formation (mmol h$^{-1}$ gdw$^{-1}$)</th>
<th>Maltose</th>
<th>Lactate</th>
<th>Formate</th>
<th>Acetate</th>
<th>Ethanol</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG1363</td>
<td>0</td>
<td>0.60</td>
<td>8.2 13 19 10 10.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 mM</td>
<td>0.47</td>
<td>7.0 17 6 4 1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MG/SP1-\textit{bdh}</td>
<td>0 10 mM 20 mM</td>
<td>0.58 0.56 ± 0.01 7.2 ± 0.4 10 ± 0.4 11 ± 0.3 10 ± 0.1 0.3 ± 0.1 7.4 ± 0.1 10 ± 0.4 12 ± 0.3 11 ± 0.1 0.2 ± 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.3.5 **Unexpected lower carbon recovery in MG/SP1-\textit{bdh} at ≥10mM acetoin**

When checking the carbon balance in the growth experiments, it was found that only 80% of carbon was recovered from the product for MG/SP1-\textit{bdh} growing on MalSALN with 20 mM acetoin. For growth on MalSAN with 10 or 20 mM acetoin, the recovery even reduced to 70%. The recovery was significantly lower than that of other conditions (≥90%), which was generally observed in previous experiments.

5.3.6 **Growth of MG1363, MG/SP2-\textit{bdh} and MG/SP1-\textit{bdh} in MalSALN(BGP)**

In light of the uncharacterized effect of ≥10mM acetoin which deviated the carbon balance very significantly, in the subsequent experiments, only 5 mM acetoin was applied to the media in which the carbon recovery remained at a high level (Table C.2). To more clearly quantify the effect of the NAD$^+$-regenerating activity brought by 23BDH, the strain MG/SP2-\textit{bdh}, which is MG1363 transformed with pCS4639 containing a weaker promoter than that in pCS4701, was also characterized. Finally, to allow the quantification of acetoin and 2,3-butanediol by HPLC, the buffer MOPS used in normal SALN media was replaced by 1.9% β-glycerophosphate which acts as buffer in the M17 medium.

The results of growth in MalSALN(BPG) were in general similar to those in MalSALN except the overall lower growth rates (Figure 5.4). MG/SP2-\textit{bdh} and MG/SP1-\textit{bdh}, when supplemented with 5 mM acetoin, grew slightly faster than the wild-type MG1363 albeit they had lower maltose uptake fluxes (Table 5.4).
Influence of NADH/NAD+ on mixed-acid fermentation

Chapter 5

Figure 5.4. Growth rates in MalSALN or MalSALN(BGP) media.

Growth of MG/SP1-bdh was not tested in MalSALN

Table 5.4. Specific rates of MG1363, MG/SP2–bdh and MG/SP1-bdh growing in MalSALN(BGP).

<table>
<thead>
<tr>
<th>Acetoin added</th>
<th>Strain</th>
<th>Growth rate (h⁻¹)</th>
<th>Specific rate of sugar consumption/product formation (mmol h⁻¹ gdw⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Maltose</td>
</tr>
<tr>
<td>0</td>
<td>MG1363</td>
<td>0.46</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>MG/SP2–bdh</td>
<td>0.45</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>MG/SP1-bdh</td>
<td>0.44</td>
<td>7.0</td>
</tr>
<tr>
<td>5 mM</td>
<td>MG1363</td>
<td>0.39 ± 0.004</td>
<td>7.8 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>MG/SP2–bdh</td>
<td>0.47 ± 0.006</td>
<td>6.8 ± 0.1</td>
</tr>
<tr>
<td></td>
<td>MG/SP1-bdh</td>
<td>0.49 ± 0.002</td>
<td>7.0 ± 0.1</td>
</tr>
</tbody>
</table>

Acetate and ethanol production significantly increased and decreased respectively while the formate flux was reduced slightly and no significant change in lactate flux was observed when comparing MG/SP2–bdh and MG/SP1-bdh growing at 5 mM acetoin to MG1363 growing without acetoin (Table 5.4). The changes in growth rate, acetate and ethanol production were found to be correlated with the in vitro activity of 23BDH in the strains and so were the increases in acetoin consumption and 23BD production (Figure 5.5). No production of diacetyl was observed.

Figure 5.5. Fluxes of MG1361, MG/SP2–bdh and MG/SP1-bdh plotted against 23BDH activity.

(a) Growth rate, (b) acetate, ethanol, (c) acetoin (uptake) and 2,3-butanediol (23BD) fluxes plotted against the specific activity of 23BDH measured in MG/SP2–bdh and MG/SP1-bdh. MG1363 growing without acetoin was assumed to have zero activity.

101
Here we compared MG1363 growing without acetoin and assigned a zero specific 23BDH activity to this case because no acetoin was added to the medium and no acetoin was detectable while 23BD was produced at a very low rate of 0.3 mmol h\(^{-1}\)gdw\(^{-1}\), compared to the pyruvate flux of 29 mmol h\(^{-1}\)gdw\(^{-1}\). It is also because in the ideal case, 23BDH acts as an isolated system exchanging only NADH and NAD\(^{+}\) with the rest of the metabolism in the presence of acetoin or 23BD, so there would be no activity in the absence of any substrates. Furthermore, although for MG1363 growing at 5 mM acetoin, a very low 23BDH activity concomitant with acetoin consumption and 23BD production at rates lower than MG/SP2\(-bdh\) and MG/SP1\(-bdh\) were measured (see Table C.5 in Appendix C for the data), the case appeared to be not a good reference that was comparable with others because it grew substantially slower and deviated to a more homolactic mode of fermentation in the presence of 5 mM extracellular acetoin. More discussion can be found below in ‘Unexpected effect of extracellular acetoin’.

### 5.3.7 Slower growth of MG/SP2\(-bdh\) and MG/SP1\(-bdh\) on glucose

The three strains were also cultivated on glucose under the same conditions. Similar to the case of maltose, when supplied with 5 mM acetoin, MG/SP2\(-bdh\) and MG/SP1\(-bdh\) had slightly lower glycolytic fluxes and significantly more acetate production compared to MG1363 which similarly appeared to be more homolactic in the presence of 5 mM acetoin (Table 5.5). Moreover, MG/SP1\(-bdh\) consumed acetoin and produced 23BD at highest rates to which MG/SP2\(-bdh\) was second (Table C.5 in Appendix C).

<table>
<thead>
<tr>
<th>Acetoin added</th>
<th>Strain</th>
<th>Growth rate (h(^{-1}))</th>
<th>Specific rate of sugar consumption/product formation (mmol h(^{-1})gdw(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Glucose</td>
</tr>
<tr>
<td>0</td>
<td>MG1363</td>
<td>0.99</td>
<td>24.8</td>
</tr>
<tr>
<td></td>
<td>MG/SP2(-bdh)</td>
<td>0.98</td>
<td>24.0</td>
</tr>
<tr>
<td></td>
<td>MG/SP1(-bdh)</td>
<td>0.97</td>
<td>25.4</td>
</tr>
<tr>
<td>5 mM</td>
<td>MG1363</td>
<td>0.98 ± 0.01</td>
<td>24.5 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>MG/SP2(-bdh)</td>
<td>0.95 ± 0.00</td>
<td>23.5 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>MG/SP1(-bdh)</td>
<td>0.95 ± 0.00</td>
<td>23.3 ± 0.3</td>
</tr>
</tbody>
</table>

Despite the resemblance, several distinctive features of interest could also been observed. First, MG/SP2\(-bdh\) and MG/SP1\(-bdh\) grew slower than MG1363 when extracellular acetoin was present (Table 5.5). Second, lactate flux decreased whereas ethanol production remained indefinite. For each replicate of the same experimental conditions, either production or consumption of ethanol at a very low rate might be measured. This indeed appeared to be a feature commonly observed for growth on glucose in our experimental setup (e.g. Table 2.5 in Chapter 2). Third, the production of formate by MG/SP2\(-bdh\) and MG/SP1\(-bdh\) was lower than that of MG1363 growing at 5 mM acetoin. More interestingly, it was much less than the production of acetate.
### 5.3.8 Increased NADH/NAD⁺ ratio

Concentrations of NADH and NAD⁺ of cultures in MalSALN(BGP) and GluSALN(BGP) were quantified. The NADH/NAD⁺ ratio measured in MG1363 growing on glucose or maltose was comparable to previous studies (e.g. [2]) and was higher on glucose than on maltose (Table 5.6). For MG/SP2--bdh and MG/SP1-bdh, surprisingly, their measured NADH/NAD⁺ ratios were higher than that of MG1363. The increase was most significant for growth on maltose, from 0.06 in MG1363 to 0.26 in MG/SP2--bdh and 0.37 in MG/SP1-bdh, whereas the increase in the ratio of MG/SP2--bdh and MG/SP1-bdh compared to MG1363 growing on maltose without acetoin was much smaller. For growth on maltose, the concentrations of both NADH and NAD⁺ increased in MG/SP1-bdh and MG/SP2-bdh whereas for growth on glucose, NAD⁺ concentration showed no significant change.

### 5.4 Discussion

#### 5.4.1 Strategy to perturb NADH/NAD⁺

In this study, 23BDH together with acetoin added to media was proposed as an approximation for an ideal system which perturbs the NADH/NAD⁺ ratio by converting NADH into NAD⁺ and is isolated from the rest of the metabolism. While under aerobic conditions NOX is a good choice, under anaerobic conditions the construction of such a system is not straightforward because of the complexity of the metabolic network. The design of the present system was based on the knowledge on *L. lactis* to try to find a reaction with low activity whose substrate and product can both be easily transported into and out of the cells.

| Table 5.6. Measurements of NADH/NAD⁺ ratio and their intracellular concentrations. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | 0 mM acetoin    | 5 mM acetoin    |                 |                 |                 |                 |
|                 | MG1363          | MG/SP2--bdh     | MG/SP1-bdh      | MG1363          | MG/SP2--bdh     | MG/SP1-bdh      |
| MalSALN (BGP)   |                 |                 |                 |                 |                 |                 |
| NADH (mM)       | 0.08            | 0.09            | 0.13            | 0.02 ± 0.01     | 0.13 ± 0.03     | 0.20 ± 0.04     |
| NAD⁺ (mM)       | 1.01            | 0.88            | 0.92            | 0.33 ± 0.06     | 0.51 ± 0.07     | 0.55 ± 0.03     |
| NAD(H) pool (mM)| 1.09            | 0.96            | 1.04            | 0.35 ± 0.07     | 0.64 ± 0.10     | 0.76 ± 0.07     |
| NADH/NAD⁺       | 0.08 ± 0.10     | 0.14            | 0.06 ± 0.02     | 0.26 ± 0.05     | 0.37 ± 0.06     |
| GluSALN (BGP)   |                 |                 |                 |                 |                 |                 |
| NADH (mM)       | 0.19 ± 0.03     | 0.19 ± 0.01     | 0.28 ± 0.05     |
| NAD⁺ (mM)       | 0.91 ± 0.07     | 0.80 ± 0.03     | 0.88 ± 0.09     |
| NAD(H) pool (mM)| Not determined  | 1.11 ± 0.10     | 0.99 ± 0.03     | 1.16 ± 0.14     |
| NADH/NAD⁺       | 0.21 ± 0.02     | 0.24 ± 0.01     | 0.31 ± 0.03     |

NAD(H) pool is the sum of the concentrations of NADH and NAD⁺.

Characterization of growth and fermentation pattern and quantification of NADH/NAD⁺ ratio showed that the growth of MG/SP2--bdh and MG/SP1-bdh which contain plasmids expressing 23BDH was
similar to MG1363 in the absence of acetoin but significantly different in the presence of acetoin. This indicated that the difference originated from the conversion of acetoin into 23BD by the 23BDH activity expressed in MG/SP2--bdh and MG/SP1--bdh. This is strongly supported by the similar measured rates of acetoin consumption and 23BD production. In this sense, the introduced system can answer the question of the effect of NADH/NAD\(^+\) ratio on fermentation modes to a certain extent except a disadvantage described below.

5.4.2 Unexpected effect of extracellular acetoin on MG1363

From experimental data regarding MG1363, the presence of acetoin in media appeared to turn the fermentation mode into more homolactic and even with carbon lost probably to other pathways at a higher acetoin concentration of 20 mM. This to a certain extent confused the interpretation of the results. A hypothesis that can explain the phenomenon and fit the data is that acetoin inhibits the activity of PFL. The formate production of MG1363 in either MalSALN(BGP) or GluSALN(BGP) was 2-fold lower in the presence of 5 mM acetoin present. In fact, we observed that for MG/SP2--bdh and MG/SP1--bdh growing in MalSALN(BGP) with 5 mM acetoin, the formate production rate increased in the early exponential phase (OD\(600 = 0.15 – 0.3\)) and became fairly constant in the reported mid-exponential phase (OD\(600 = 0.35 – 0.6\)) along with the decrease of acetoin concentration in the media (data not shown). Meanwhile, the consumption rate of sugar, production rate of lactate and acetate remained constant over the entire range.

This hypothesis can explain the slower growth of MG1363 in the presence of acetoin because the acetate production and thus ATP production rate were limited by the supply of acetyl-CoA by PFL. The exact effect of acetoin requires further investigation. For the system to work in a more elegant way, for example, acetoin can be added at a lower amount (say 1 mM) and growth is characterized at a lower cell density or weaker synthetic promoters can be used so that acetoin is consumed slow but steadily. Chemostat can also be used in which the acetoin concentration can be finetuned to a low level enough for full activity.

5.4.3 PDHc activity in anaerobic conditions

Another side observation of the study is the suspected activity of pyruvate dehydrogenase complex (PDHc) which oxidases pyruvate into acetyl-CoA. In the literature, PDHc was proposed to be inactive anaerobically due to the high NADH/NAD\(^+\) ratio \[21\] and was found to have low or no expression anaerobically \[2, 8, 11\]. In our results, PDHc activity was suggested by the larger acetate flux than formate flux when MG/SP1--bdh grew in MalSALN with 20 mM acetoin and GluSALN(BPG) with 5 mM acetoin. In contrast, the acetate flux was lower and very close to the formate flux for the slower growth of MG/SP1--bdh in MalSAN without \(\alpha\)-lipoic acid, an essential cofactor for PDHc activity that \textit{L. lactis} is auxotrophic for. This further supported that PDHc was indeed active under such conditions. These cases with acetate production probably supported by PDHc activity can be interpreted as an
alternative acetate-producing capacity brought by PDHc when PFL was not available, very much alike the case of aerobic growth where PFL was inactivated by oxygen, PDHc was found to be active and acetate was produced [11]. The magnitude of the PDHc activity observed in this study has rarely been reported in the literature.

5.4.4 Shift towards acetate production in MG1363 and additionally formate production in IL1403

From different sets of experiments, it is clear that in MG1363, additional NAD⁺-regenerating activities brought by 23BDH had a positive control on acetate production flux. For growth on maltose, the flux was redirected from ethanol production and allowed faster growth. This corresponds well to the recent results which found an increased ethanol-to-acetate ratio and meanwhile a slower growth rate when introducing FDH activity into MG1363 to produce NADH from NAD⁺ by oxidizing formate into CO₂ [10]. The effect on formate production, however, could not be concluded because it coincided with the apparent inhibitory effect of acetoin on formate production.

For growth on glucose, the increase in acetate production was redirected from lactate production. The decrement in formate flux in MG/SP2--bdh and MG/SP1-bdh can be savely considered as an effect of the additional NAD⁺-regenerating activity because the formate flux of MG1363 growing with 5 mM acetoin which did not consume acetoin was even higher than that of MG/SP2--bdh and MG/SP1-bdh which consumed acetoin down to 4 mM and 3 mM respectively (data not shown). Consequently, the inhibitory effect of acetoin on formate production had to be less significant in MG/SP2--bdh and MG/SP1-bdh than MG1363. The effect should thus originate from the additional NAD⁺-regenerating activity.

The mechanism of the redirection of flux from ethanol production in the maltose case or from lactate in the glucose case is not clear. It might be simply the competition of NADH between lactate dehydrogenase (LDH), alcohol dehydrogenase (ADH) and 23BDH. For the inhibition of formate production on glucose, one possible explanation is the higher NADH/NAD⁺ ratio. By the model proposed in [1, 2], this would inhibit GAPDH and cause a higher GAP pool, which inhibits PFL, to accumulate for maintaining a high glycolytic flux.

From the test of IL1403 and IL/SP1-bdh in overnight culture, however, a more than two-fold increase of both formate and acetate were observed in the presence of extracellular acetoin. This suggested that acetoin may not be inhibitory to the mixed-acid fermentation pathway in IL1403 as in MG1363. Moreover, the more mixed-acid fermentation pattern in IL1403 caused by the additional NAD⁺-regenerating activity to a certain extent confirm one of the conjecture made in Chapter 3 of why IL1403 remains homolactic even at low glycolytic flux. However, more efforts are required to characterize IL1403 and IL/SP1-bdh for more concrete conclusions.
5.4.5 Counterintuitive NADH/NAD\(^+\) ratio

Among all the results, the most confusing one is the higher NADH/NAD\(^+\) ratio measured. Intuitively, introducing a reaction that acts in the direction of NAD\(^+\) production should lower the NADH/NAD\(^+\) ratio. From the flux data, it seems that the reason is the additional NAD\(^+\)-producing flux by 23BDH could not compensate the decrease in the other NAD\(^+\)-producing flux (summarized in Table 5.7). This might involve some very interesting kinetic behavior which can be studied through kinetic modeling, for example, oscillation of the system. This might also help to explain why PDHc appeared to be active under a high NADH/NAD\(^+\) ratio. The worst possibility is that the quantification procedure was problematic and thus caused large systematic errors in all experiments.

| Table 5.7. Estimated net production of NADH from glycolysis and fermentation. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Flux through pathway/enzyme (mmolh\(^{-1}\)gdw\(^{-1}\)) | Net NADH production* (mmolh\(^{-1}\)gdw\(^{-1}\)) |
| Glycolysis (C\(_6\)-mol) | LDH | PDHc | ADH | 23BDH |
| GluSALN(BGP) | MG1363 | 24.8 | 45.0 | 0 | -0.22 | 0 | 5.0 |
| MG/SP1-bdh | 23.3 | 37.3 | 4.7 | 0.1 | 6.7 | 7.1 |
| MalSALN(BGP) | MG1363 | 14.6 | 10 | 0 | 6.6 | 0.3 | 5.7 |
| MG/SP1-bdh | 14.0 | 10 | 0 | 3.6 | 4.0 | 6.8 |

*Production = 2 × glycolysis + PDHc; consumption = LDH + 2 × ADH + 23BDH. 23BDH, 2,3-butanediol dehydrogenase; ADH, alcohol dehydrogenase; PDHc, pyruvate dehydrogenase complex.

5.5 References


Chapter 6. Amino acid metabolism and fermentation modes

6.1 Introduction

In addition to the gene expression level and cofactor level, the nutrient availability has once been shown to be one of the factors influencing the fermentation modes. Garrigues et al. observed that for *L. lactis* growing on the same carbon source galactose, the richer the medium was, the faster the growth rate was and unexpectedly the more mixed-acid the fermentation became [1]. This disproved the hypothesis that the extent of mixed-acid fermentation is simply negatively correlated to the growth rate. Moreover from these experimental results, the authors proposed the theory that the mixed-acid switch depends on the balance between catabolism and anabolism. When the catabolic flux is the growth-limiting factor for reasons such as growth on slowly fermentable sugars like galactose, mixed acids are produced for ATP maximization by acetate production whereas when other anabolic requirements such as amino acid (AA) biosynthesis are limiting for growth, homolactic fermentation prevails. In our previous study, the same pattern was observed for *L. lactis* MG1363 growing on maltose in the define SA medium omitting certain AAs [2]. Some of the results are summerized in Table 6.1. Different AAs could have very different effects on both growth rate and mixed-acid fermentation.

In this study, we attempted to explore this aspect from the computational perspective by using the recently published genome-scale metabolic network for MG1363 and a set of experimental data from that publication [3]. Elemenary flux modes (EFMs) were chosen as the tool for metabolic pathway analysis [4] because they represent all the minimal functional units in a metabolic network which may reveal unexpected connections between different metabolic activities.

<table>
<thead>
<tr>
<th>Medium ²</th>
<th>Growth rate (h⁻¹)</th>
<th>Formate % ³</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALN</td>
<td>0.65</td>
<td>56%</td>
</tr>
<tr>
<td>SALN ÷ cysteine</td>
<td>0.64</td>
<td>45%</td>
</tr>
<tr>
<td>SALN ÷ phenylalanine</td>
<td>0.59</td>
<td>36%</td>
</tr>
<tr>
<td>SALN ÷ alanine</td>
<td>0.58</td>
<td>52%</td>
</tr>
<tr>
<td>SALN ÷ lysine</td>
<td>0.55</td>
<td>45%</td>
</tr>
<tr>
<td>SALN ÷ threonine</td>
<td>0.43</td>
<td>30%</td>
</tr>
</tbody>
</table>

¹ Adapted from [2].
² 0.2% maltose, SA medium supplemented with α-lipoic acid and nucleosides (SALN).
³ Formate percentage was calculated as: \( \frac{\text{flux}_{\text{formate}}}{\text{flux}_{\text{formate}} + \text{flux}_{\text{lactate}}} \)
6.2 Results and discussion

6.2.1 Computational method developed
A decomposition method termed ‘minimal branching decomposition’ (MBD) was developed to interpret a flux distribution in terms of a set of EFMs with an overall minimum number of branching pathways with respect to a given set of metabolites. By choosing amino acids as the set of metabolites with minimum branching pathways involving them, the computational method was applied to a flux distribution estimated from the experimental data of MG1363 exhibiting significant mixed-acid fermentation [3]. This may reveal the distribution of the amino acid fluxes into different metabolic activities including catabolism as well as biomass formation.

The computational method, as a method of representing flux distributions by EFMs that is applicable to general metabolic networks, has been published. Please see Appendix A for the full paper which includes the derivation, validation and application of the method to the amino acid metabolism of *L. lactis* MG1363 and other metabolic models.

6.2.2 Amino acid utilization and mixed-acid fermentation
Here more discussion is presented regarding the relationship between amino acid metabolism and mixed-acid fermentation in addition to the results presented in the article in Appendix A. Table 6.2 shows the rates of production of fermentation products and consumption of glucose and amino acids of the estimated flux distribution and the EFMs decomposing it. The redox balance and ATP production from the module of glycolysis and fermentation pathways were also calculated.

Lactate production was only involved in two of the EFMs which were redox-balanced in the module of glycolysis and fermentation pathways. In all other modes acetate and formate were produced. An interesting observation is that the module of glycolysis and fermentation in the overall flux distribution was indeed redox-unbalanced with net production in NADH. Flahaut et al. has ascribed this to the production of flavour compounds from amino acids which oxidized NADH and shared the role of alcohol dehydrogenase to allow more acetate production [3].

In our resolution of the flux distribution by EFMs, the same could be observed. EFM 5, 6 and 13 in Table 6.2 together accounted for 80% of the net NADH produced in glycolysis. In these modes, 50% or less carbon from glucose uptake ended up in acetate production. All or most of the rest of carbon entered into the amino acid metabolism and participated in flavour compound production together with the amino acids assimilated in these EFMs, glycine, threonine and methionine. The flavour compounds produced in these three modes accounted for 60% of all flavour compound production. These NAD⁺-regenerating pathways in additional to fermentation contributed to around 10% of the ATP and acetate production. In the sugar-uptake-limited case being analysed (chemostat at dilution rate of 0.05h⁻¹), this 10% contribution is expected to be quite important for growth. This finding also
supports the finding in Chapter 5 which suggested additional NAD⁺-regenerating activities switched the metabolism towards acetate production.

Another interesting observation is that even though the overall carbon recovery was lower than 100% indicating some glucose ended up in products or biomass other than lactate, formate, acetate and ethanol, EFMs with recovery higher than 100% might actually be active. In EFM 7 and 14, though at small fluxes, formate and acetate were directly produced from cysteine and serine.

The above analysis provided hints for the possible connection of amino acid metabolism and the mixed-acid fermentation, in particular acetate production. The flux distribution being analyzed, however, is only one of the infinitely many possible solutions of the flux space constrained by exchange fluxes. A more comprehensive framework considering the whole feasible flux space and EFMs simultaneously should allow more systematic assessment of the significance of the decomposition by EFMs.

### 6.3 References


### Table 6.2: Production and uptake fluxes of the estimated flux distribution and its decomposition by EFMs.

<table>
<thead>
<tr>
<th>Substrate/Product</th>
<th>Flux (mmolh⁻¹gdw⁻¹)</th>
<th>Flux distribution</th>
<th>EFM¹</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactate</td>
<td>0.606</td>
<td>0.365</td>
<td>0.2</td>
</tr>
<tr>
<td>Formate</td>
<td>2.305</td>
<td>0.0065</td>
<td>0.863</td>
</tr>
<tr>
<td>Acetate</td>
<td>1.213</td>
<td>0.0043</td>
<td>0.449</td>
</tr>
<tr>
<td>Ethanol</td>
<td>0.994</td>
<td>0.0022</td>
<td>0.414</td>
</tr>
<tr>
<td><strong>Uptake</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td>1.523</td>
<td>0.0029</td>
<td>0.445</td>
</tr>
<tr>
<td>Ala</td>
<td>0.039</td>
<td>0.0006</td>
<td>0.002</td>
</tr>
<tr>
<td>Arg</td>
<td>0.009</td>
<td>0.0003</td>
<td>0.0001</td>
</tr>
<tr>
<td>Asp</td>
<td>0.008</td>
<td>0.0002</td>
<td>0.008</td>
</tr>
<tr>
<td>Cys</td>
<td>0.031</td>
<td>0.004</td>
<td>0.0001</td>
</tr>
<tr>
<td>Glu</td>
<td>0.046</td>
<td>0.0005</td>
<td>0.026</td>
</tr>
<tr>
<td>Gly</td>
<td>0.046</td>
<td>0.0004</td>
<td>0.002</td>
</tr>
<tr>
<td>His</td>
<td>0.004</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Ile</td>
<td>0.018</td>
<td>0.0006</td>
<td>0.0002</td>
</tr>
<tr>
<td>Leu</td>
<td>0.049</td>
<td>0.018</td>
<td>0.0004</td>
</tr>
<tr>
<td>Lys</td>
<td>0.018</td>
<td>0.004</td>
<td>0.0001</td>
</tr>
<tr>
<td>Met</td>
<td>0.005</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Phe</td>
<td>0.008</td>
<td>0.0003</td>
<td>0.0001</td>
</tr>
<tr>
<td>Ser</td>
<td>0.085</td>
<td>0.010</td>
<td>0.0022</td>
</tr>
<tr>
<td>Thr</td>
<td>0.062</td>
<td>0.0005</td>
<td>0.0002</td>
</tr>
<tr>
<td>Val</td>
<td>0.016</td>
<td>0.0005</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Glycolysis + fermentation pathways:

- **NADH production**: 2.9 0.2 0.0047 0.87 0.89 0.08 0.10 0 0.003 0.37 0.0009 0.2 0.005 0.11 0 0.03
- **NADH consumption**: 2.6 0.2 0.0045 0.83 0.87 0.03 0 0.003 0.37 0.0005 0.2 0.005 0.02 0 0.03
- **NADH net production**: 0.3 0 0.0002 0.04 0.02 0.08 0.07 0 0.000 0.00 0.0004 0 0 0.09 0 0.00
- **ATP production**: 4.1 0.3 0.0082 1.29 1.32 0.12 0.13 0.01 0.006 0.39 0.0015 0.2 0.007 0.15 0.01 0.03
- **ATP production %**: 100% 8% 0% 32% 33% 3% 3% 3% 0% 0% 0% 6% 0% 4% 0% 1%
- **Carbon recovery**: 92% 87% 113% 97% 96% 50% 33% 124% 104% 78% 100% 100% 51% 74%

### Notes:

- Four EFMs with ≤ 0.1% of the sum of fluxes of the flux distribution were not shown.
- NADH production from glycolysis was equal to the flux of GAPDH.
- NADH consumption by fermentation modes was calculated as the sum of flux through LDH and 2 × flux through ADH.
- ATP production was calculated as the sum of fluxes through PGK, PYK and ACK subtracted by the flux through PFK.
- Carbon recovery was calculated as the sum of production of lactate, acetate and ethanol divided by the consumption of glucose.
- Carbon recovery equal to infinity implies fermentation products produced from carbon sources other than glucose.
Chapter 7. Conclusion and future directions

7.1 Conclusion

In the present study, efforts have been made to investigate the switch between homolactic fermentation and mixed-acid fermentation influenced by three aspects of metabolism, namely the level of gene expression in the pathway for mixed-acid fermentation, the level of the cofactor NADH/NAD$^+$ and the availability of amino acids as nutrients.

7.1.1 Different promoter activities of mixed-acid genes

At the level of gene expression, the promoter activities of the genes in the pathway branch that produces formate, acetate and ethanol were compared between the two important laboratory strains MG1363 and IL1403 which differ from each other with respect to the mixed-acid switch. The results suggested that different types of regulatory elements governed the transcriptional regulation of different genes in the two strains, including the obvious indication of trans-regulation on pyruvate formate-lyase (PFL) and cis-regulation on phosphotransacetylase (PTA), as well as the more unobvious regulation pattern for the two genes encoding for acetate kinase (ACK).

7.1.2 Control on mixed-acid fermentation by pyruvate formate-lyase

The difference observed in the promoter activity of PFL between the two strains led to the attempt to modulating the expression of PFL in both MG1363 and IL1403. For MG1363, consistent with results from the literature, PFL was found to have control on the mixed-acid switch. In particular, complementary to the previous control analysis which implied the existence of one or more reactions positively controlling the formate flux in the wild-type MG1363 growing on glucose, we estimated a control coefficient close to one by PFL on the formate flux at the wild-type PFL level for growth on glucose. For IL1403, however, the modulation was not successful and this suggested that a high expression of PFL might be lethal. Overexpressing PFL is thus probably not the first step to take to restore a mixed-acid-fermenting phenotype of IL1403.

7.1.3 Complementary roles of acetate kinase isozymes

Meanwhile, the interesting observation that one of the genes for ACK was the only gene having a higher promoter activity in IL1403 than in MG1363 gave rise to the conjecture that the two genes for ACK might play different roles in acetate metabolism. We then characterized the two genes in MG1363. We found different kinetic properties of the two ACKs encoded by the two genes and their complementary roles for acetate production and uptake under different conditions. It was also observed that the existence of acetate kinase isozymes was very common among bacteria.
7.1.4 Acetate production induced by additional NAD$^{+}$-regenerating activities

At the level of metabolic regulation, we tested the proposed roles of the cofactors NADH and NAD$^{+}$ on the mixed-acid switch by introducing an enzyme activity to approximate an ideal isolated system that only oxidized NADH from the metabolism to NAD$^{+}$ anaerobically. The construction partially restored the mixed-acid fermenting ability in IL1403 as reflected by a substantial increase in formate and acetate production. For MG1363, no any positive effect on formate production was observed but significant increase in acetate and decrease in ethanol were observed. An interesting observation was the apparent high flux through pyruvate dehydrogenase complex which has not been reported in the literature.

7.1.5 Connection between amino acid metabolism and fermentation modes

The extent of the mixed-acid fermentation of MG1363 was found to depend on amino acid availability and individual amino acids could have largely different impact. Meanwhile, a computational method for combining metabolic flux analysis and elementary mode analysis was developed and applied to a flux distribution estimated from experimental data. Elementary modes connecting amino acid metabolism and fermentation modes were observed.

7.2 Future directions

Based on the results from the current studies, many interesting questions are worth further research. They are described below.

7.2.1 Identify transcriptional regulators and cis-regulatory elements

In Chapter 2, trans-regulatory elements for the transcriptional regulation of pfl, probably activators for growth on maltose or repressors for growth on glucose were suggested to exist. The exact genetic element responsible for this observation can be identified through random mutagenesis using existing tools such as insertional sequence ISS1 for L. lactis and subsequent screening for the loss of reporter activity on maltose or gain on glucose.

Moreover, cis-regulatory elements regulating the transcription of eutD, the gene for phosphotransacetylase, were also expected to exist and the mutated cre site upstream of eutD in IL1403 was proposed as a candidate. It can be tested by restoring the cre site in IL1403.

7.2.2 Restore the phenotype of mixed-acid fermentation for IL1403

In Chapter 3, the modulation of PFL in IL1403 turned out to be unsuccessful while in Chapter 5 the formate- and acetate-producing ability of IL1403 was increased. Further efforts can be made to restore the mixed-acid-fermenting phenotype of IL1403.
7.2.3 Characterize acetate-dependent growth of IL1403

In Chapter 4 only the ACK isozymes of MG1363 were studied though the study was originally motivated by a stronger induction of the promoter of one of the ACKs in IL1403 whose ortholog in MG1363 was later found to be specialized in acetate uptake. It would be interesting to compare the acetate-dependent growth of IL1403 to that of MG1363 and see whether the effect in IL1403 is even more pronounced.

7.2.4 Model the perturbation of NADH/NAD$^+$ ratio

In Chapter 5, introducing an NADH-consuming enzyme unexpectedly resulted in increase in the measured NADH/NAD$^+$ ratio. If the trend of the measurement is correct, very interesting kinetic behaviour may be involved in this situation. Kinetic modeling would be a powerful tool to make sense of the observation.

7.2.5 Study the effect of another important cofactors, ATP/ADP ratio

Beside NADH/NAD$^+$, in the literature, the ATP/ADP ratio has also been proposed to be an important factor of the mixed-acid switch. It would be interesting to perturb the ATP/ADP ratio by introducing ATPase activities as in previous studies.

7.2.6 Comprehensive computational analysis on the connection between amino acid metabolism and fermentation modes

For a more in-depth understanding of the relationship between amino acid metabolism and fermentation modes, a more comprehensive computational analysis using genome-scale models and elementary modes can be performed. The experimental data of growth in media with different amino acid compositions, together with other available -omics data such as gene-expression data can be integrated for identifying active subnetworks. Such analysis is expected to shed light on some unobvious pathways or hidden regulation between the two subsystems.
Appendix A Estimating biological elementary flux modes that decompose a flux distribution by the minimal branching property


Siu Hung Joshua Chan1, Christian Solem1, Peter Ruhdal Jensen1 and Ping Ji2
1Systems Biotechnology and Biorefining, National Food Institute, Technical University of Denmark
2Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University

A.1 Abstract

Motivation: Elementary flux mode (EFM) is a useful tool in constraint-based modeling of metabolic networks. The property that every flux distribution can be decomposed as a weighted sum of EFMs allows certain applications of EFMs to studying flux distributions. The existence of biologically infeasible EFMs and the non-uniqueness of the decomposition, however, undermine the applicability of such methods. Efforts have been made to find biologically feasible EFMs by incorporating information from transcriptional regulation and thermodynamics. Yet, no attempt has been made to distinguish biologically feasible EFMs by considering their graphical properties. A previous study on the transcriptional regulation of metabolic genes found that distinct branches at a branch point metabolite usually belong to distinct metabolic pathways. This suggests an intuitive property of biologically feasible EFMs, i.e. minimal branching.

Results: We developed the concept of minimal branching EFM and derived the minimal branching decomposition (MBD) to decompose flux distributions. Testing in the core Escherichia coli metabolic network indicated that MBD can distinguish branches at branch points and greatly reduced the solution space in which the decomposition is often unique. An experimental flux distribution from a previous study on mouse cardiomyocyte was decomposed using MBD. Comparison with decomposition by a minimum number of EFMs showed that MBD found EFMs more consistent with established biological knowledge, which facilitates interpretation. Comparison of the methods applied to a complex flux distribution in Lactococcus lactis similarly showed the advantages of MBD. The minimal branching EFM concept underlying MBD should be useful in other applications.

A.2 Introduction

Metabolic networks consisting of biochemical reactions and compounds in a cell have proven to be an extremely useful in silico tool to study cellular metabolism (Edwards and Palsson, 1999; Oberhardt et al., 2009; Price et al., 2004). The structure of a metabolic network is reflected in the stoichiometric matrix S, and the state of the network is embodied in the flux distribution (FD) v, which is a vector
containing all reaction rates in the network. Based on the well-established quasi-steady-state assumption, \( \mathbf{v} \) can be constrained in a polyhedral cone \( (S\mathbf{v} = 0, \mathbf{v}_{\text{ub}} \geq \mathbf{v} \geq \mathbf{v}_{\text{lb}}) \) and studied using a series of so-called constraint-based reconstruction and analysis (COBRA) methods (Lewis et al., 2012).

One of the approaches of theoretical importance and application value is the elementary flux mode (EFM) analysis (Schuster and Hilgetag, 1994). An EFM is a FD with a minimal number of active reactions that together satisfy the steady-state condition. It is thus a minimal metabolic pathway in a metabolic network. EFMs have been successfully applied to examine network structures, explore new metabolic pathways, suggest rational strain design, etc. (reviewed in Trinh et al., 2009; Zanghellini et al., 2013).

Another property of EFMs useful for studying FDs is that every FD \( \mathbf{v} \) is a non-negative sum of EFMs:

\[
\mathbf{v} = \sum_{k=1}^{K} \alpha_k \mathbf{e}_k \tag{1}
\]

where \( \mathbf{e}_1 \cdots \mathbf{e}_K \) are all EFMs of the network, and \( \alpha_k \) is the corresponding weight coefficient. This decomposition of FDs from EFMs or extreme pathways (another type of pathways with a similar definition) has been investigated and applied to certain biological studies by choosing particular decompositions (reviewed in Chan and Ji, 2011]. The usefulness of these methods, however, is limited for two main reasons. First, the practical computation of the decomposition requires the set of all EFMs \textit{a priori}, which is usually not available in large-scale networks owing to combinatorial explosion. This computational issue can be overcome in two ways, either by sampling EFMs (Machado et al., 2012; Tabe-Bordbar and Marashi, 2013) or by finding EFMs with particular properties to solve for the decomposition without requiring the entire set of EFMs (Chan and Ji, 2011; de Figueiredo et al., 2009; Ip et al., 2011; Pey and Planes, 2014; Rezola et al., 2011). The second limitation is the non-uniqueness of decomposition and its dependence on the choice of algorithms or optimization objective (Zanghellini et al., 2013). The ambiguity severely weakens the biological significance of decomposition. Poolman et al. (2004) has proposed a method to obtain unique decompositions by combining EFMs with variable weights into compound modes with unique weights. Another method for unique decompositions is the minimization of the sum of squares of weights (Schwartz and Kanehisa, 2005, 2006), but the resulting decomposition can be dependent on how EFMs are scaled.

The usefulness of EFMs can be enhanced by finding biologically feasible EFMs. Attempts have been made to incorporate transcriptional regulatory information to enumerate biologically feasible EFMs (Jungreuthmayer et al., 2013) and to rule out thermodynamically infeasible EFMs (Jol et al., 2012).
These approaches integrating information other than stoichiometry did give promising results. The information, for instance, metabolomics data and transcriptomics data, however, is not always available. Surprisingly, all the proposed methods to find EFMs or to choose decomposition have not made full use of their graphical properties, which should be intuitively important for interpreting EFMs. An earlier study on the co-regulation of metabolic pathways found that transcriptional regulation tends to favor linear metabolic flow by co-expressing only distinct branches at metabolic branch points (Ihmels et al., 2004). This means, in Figure A.1 where ‘A’ is a hypothetical branch point metabolite, the genes for the linear pair of reactions ‘R1’ and ‘R3’, or the pair ‘R2’ and ‘R4’, are more likely to be co-expressed than ‘R3’, ‘R4’ and ‘R5’, which are distinct branches. When decomposing a FD containing ‘R1’ to ‘R4’, two EFMs containing pair ‘R1’ and ‘R3’ and pair ‘R2’ and ‘R4’ are thus a more preferable choice than one EFM containing all ‘R1’ to ‘R4’, as there are probably two distinct pathways. This suggests the rationale of minimizing branching pathways to choose EFMs that are more likely to be biologically feasible. This rationale is also expected to allow clearer interpretation of the roles of the chosen EFMs because they are expected to contain fewer distinct pathways. We believed that such EFMs can have wide applicability. As the first attempt based on this rationale, in this article, a new optimization objective called minimal branching decomposition (MBD) is proposed to identify more biologically relevant EFMs to decompose FDs using stoichiometric information alone.

Figure A.1. The graphical meaning of NP\(_{A1}\).

A is a metabolite in the network. If EFM 1 contains all five reactions, then NP\(_{A1}\) = \(2 \times 3 = 6\), equal to the number of all locally different paths passing through A.

### A.3 Methods

#### A.3.1 Optimization objective

The goal is to derive an optimization objective for finding \(a_1 \cdots a_K\) to compute an MBD composed of EFMs with as few branching pathways as possible. Given the set of all EFM contained in the matrix \(E=[e_1 \cdots e_K] \in \mathbb{R}^{m \times K}\) and the stoichiometric matrix \(S=[S_{ij}] \in \mathbb{R}^{m \times n}\) in a metabolic network with \(m\) metabolites and \(n\) reactions, assume that the FD to be decomposed is \(v=[v_j] \in \mathbb{R}^n\).

First, define the consumption matrix \(C=[C_{ij}]\) and the production matrix \(P=[P_{ij}]\) for metabolites:
In other words, if $C_{ij} = 1$, then metabolite $i$ is consumed by reaction $j$ and if $P_{ij} = 1$, then metabolite $i$ is produced by reaction $j$. Based on $C$ and $P$, compute the number of paths $NP = [NP_{ik}] \in \mathbb{R}^{n \times K}$ passing through metabolite $i$ in EFM $k$:

$$NP_{ik} = \left( \sum_{j=1}^{n} C_{ij} \delta_{jk} \right) \left( \sum_{j=1}^{n} P_{ij} \delta_{jk} \right)$$

(3)

where $\delta_{jk} = 1$ if $e_{jk} \neq 0$ and $\delta_{jk} = 0$ if $e_{jk} = 0$. $NP_{ik}$ is in fact the number of reactions consuming metabolite $i$ multiplied by the number of reactions producing metabolite $i$ in EFM $k$. It is thus the number of locally different paths passing through metabolite $i$. Figure A.1 illustrates an example. $NP$ is used for distinguishing EFMs which deviate from linear flow through branch point metabolites. Perfectly linear EFMs would have $NP_{ik} = 1$ for all metabolites and EFMs with more branching flow would have higher $NP_{ik}$. The minimum use of EFMs with high $NP_{ik}$ therefore defines the optimization objective to obtain an MBD. Hence, the objective coefficient $c_k$ for EFM $k$ is expressed as the following sum:

$$c_k = \sum_{i=1}^{m} w_i NP_{ik}$$

(4)

where $w_i$ is the weight for metabolite $i$. It is manually assigned to account for the relative importance of a metabolite as a branch point metabolite. The simplest way for assigning weights to metabolites is to just assign 1 for all weights meaning that there is no bias toward the branching of any metabolite in an EFM. In practice, biological knowledge is required to choose suitable weights to obtain better decomposition. For example, for pyruvate which is a true branch point metabolite involved in different metabolic pathways, a decomposition by EFMs with different fates of pyruvate, for instance, one going into the anaerobic fermentation and the other going into the TCA cycle, is preferred to a decomposition by all EFMs involving both the fermentation and the TCA cycle. In this case, branching of pyruvate within an EFM is strongly undesired so the weight can be set to be much higher than 1. In contrast, cofactors like ATP/ADP are global cellular currencies and their distinct linear flow may not be biologically relevant. Low or zero weights can be assigned to them. The weighting also depends on the particular purpose of application. For example, if one focuses on comparing the glycolysis and the pentose phosphate pathway, a high weight can then be assigned to glucose 6-phosphate, which is a branch point of the two pathways.
A.3.2 Minimal branching decomposition

Having derived the objective, an MBD can be obtained by solving the following mixed integer linear programming problem (MILP):

\[
\min \sum_{k=1}^{K} \sum_{i=1}^{m} w_i NP_{ik} \beta_k
\]  

(5)

subject to \( \sum_{k=1}^{K} e_{jk} \alpha_k = v_j \) for \( j = 1, \ldots, N \)

(6)

\[ 0 \leq \alpha_k \leq \beta_k \text{ for } k = 1, \ldots, K \]  

(7)

\[ \beta_k \in \{0,1\} \text{ for } k = 1, \ldots, K \]  

(8)

Objective (5) minimizes the use of EFMs with higher numbers of non-linear pathways being active at branch point metabolites. Equation (6) is the constraint for decomposition. Constraint (7) is the on/off switch for an EFM. If \( \beta_k = 0 \), EFM \( k \) is switched off. If \( \beta_k = 1 \), EFM \( k \) is switched on. In the problem stated here, two assumptions without loss of generality are made on the EFM matrix \( E = [e_{jk}] \) for simpler presentation. First, \( E \) is calculated from a stoichiometric matrix consisting of irreversible reactions only. It can be achieved by dividing each reversible reaction into two irreversible reactions with stoichiometry negative to each other. This avoids negative fluxes and additional constraints on reversible EFMs. Second, all EFMs are assumed to be scaled to have maximum weights not greater than 1 (\( \alpha_k \leq 1 \)). One way for scaling is to multiply each \( e_k \) by \( n_k = \min_j (v_j / e_{jk} \mid e_{jk} > 0) \). Another method used in Wiback et al. (2003) is scaling to an uptake reaction by multiplying each \( e_k \) by 

\[ n_k = v_{uptake} / e_{uptake,k} \cdot \]

It is noted that the ‘no cancellation’ property of decomposition by EFMs (Llaneras and Picó, 2010) ensures that if for a reaction \( j, v_j = 0 \), then all EFMs with \( e_{jk} \neq 0 \) must have \( \alpha_k = 0 \). It facilitates optimization by first selecting EFMs able to have non-zero weights for a FD, called ‘contributable EFMs’ and defined by \( E_{CONTRI} = \{ e_k \mid e_{jk} = 0 \text{ if } v_j = 0 \} \), out of \( E \) which is usually intractable for larger networks (Klamt and Stelling, 2002; Trinh et al., 2009).

A.3.3 Enumeration of alternative optima

One of the major drawbacks of interpreting decomposition by EFMs is the non-uniqueness. Despite the use of certain optimization objectives, for instance, a minimum number of EFMs as in Wiback et al. (2003), the number of alternative optimal solutions can still be large. The performance of MBD regarding the uniqueness of decomposition was examined by enumerating alternative optima. The \( R \)-th alternative optimal MBD was calculated by adding the following constraint to the model:
\[
\sum_{k=1}^{K} \beta^r_k \beta_k \leq \left( \sum_{k=1}^{K} \beta^r_k \right)^{-1} \quad \text{for } r = 1, \ldots, R-1
\]  

(10)

where \( b_r = [\beta^r_1, \ldots, \beta^r_K] \) is the binary integer solution of the \( r \)-th alternative optimal MBD found. Suboptimal solutions can also be found.

### A.3.4 Other decompositions for comparison

MBD was compared with decomposition by a minimum number of EFMs (MinEFMD) and random decomposition (RD). MinEFMD was calculated by replacing eq. (5) with \( \sum_{j \in \text{contr}} e_{jk} = 0 \) for all \( k \). RD was sampled as follows:

1. **Initialization:** \( D = \emptyset \).
2. **Step 1:** find contributable EFMs \( E_{\text{CONTRI}} = \{ e_k \mid e_{jk} = 0 \text{ if } v_j = 0 \} \).
3. **Step 2:** randomly choose an \( e_k \) from \( D \leftarrow D \cup \{ e_k \} \).
4. **Step 3:** \( v \leftarrow v - r e_k \) where \( r = \min_j \{ v_j/e_{jk} \mid e_{jk} > 0 \} \).
   - If \( v = 0 \), terminate and \( D \) is a decomposition. Else, go to step 1.

### A.3.5 Implementations

EFMs were calculated using efmtool (Terzer and Stelling, 2008) and optimization models were solved by Gurobi 5.5® via the COBRA toolbox in MATLAB® in which all other calculations were performed. The MATLAB script file is available in the Supplementary Material (section A.8). The desktop computer used had a 3.1 GHz quad-core CPU and 16 GB of RAM.

### A.4 Results

MBD was tested on randomly sampled FDs in the core *Escherichia coli* metabolic network as a validation of the method (Orth et al., 2010). It was then applied to a FD determined by isotopomer distributions reported previously for the cardiomyocyte in perfused mouse heart (Vo and Palsson, 2006) and to a FD estimated from a dataset of metabolite assimilation and excretion in the recently published genome-scale metabolic network of *Lactococcus lactis* (Flahaut et al., 2013). In the latter two cases, FDs were not sampled because they were estimated from experimental data.

### A.4.1 Core *E. coli* metabolic network

The core *E. coli* metabolic network (Orth et al., 2010) with 95 reactions and 72 metabolites simplified from the iAF1260 genome-scale network of *E. coli* (Feist et al., 2007) includes glycolysis, TCA cycle, oxidative phosphorylation, pentose phosphate pathway and a biomass reaction, alongside with a set of transcriptional regulatory rules that can be readily implemented. Three different tests were performed to examine (i) the ability to separate fluxes at branch point metabolites; (ii) the performance in a
simulated case consistent with regulatory rules and (iii) the issue of multiple optimal decompositions. As the set of all possible FDs has infinitely many solutions, random sampling based on the hit-and-run method (Almaas et al., 2004) was applied to sample FDs for the tests. Table A.1 summarizes the results. More details can be found in Supplementary Material (section A.8).

Table A.1. Summary of test results in the core *E. coli* metabolic network.

<table>
<thead>
<tr>
<th>Test</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Separation of combined fluxes at branch points.</td>
<td>Correct recovery rate</td>
</tr>
<tr>
<td>83 cases in scenario 1 (Figure A.2a), 94 cases in scenario 2 (Figure A.2b)</td>
<td>Scenario 1: MBD: 100% MBD: 94% RD: 7.5% RD: 1%</td>
</tr>
<tr>
<td>2. Coupling of glyoxylate cycle and acetate uptake implied by the regulatory rules in the presence of fructose and acetate (Figure A.2c).</td>
<td>No EFMs with simultaneous fructose uptake and glyoxylate cycle activity MBD: 100% MinEFMD: 0% RD: 1%</td>
</tr>
<tr>
<td>3. Uniqueness of decompositions. Enumeration of alternative MBDs and MinEFMDs</td>
<td>MBD: 2,000 FDs tested. 45% unique; 95% ≤ 8 optima. MinEFMD: 20 FDs tested. 100% ≥ 1,000 optima</td>
</tr>
</tbody>
</table>

RD: random decomposition.

*Separation of fluxes at branch points*

To check that the proposed method has the desired ability to choose EFMs of distinct branches at branch points, two scenarios were tested (Figure A.2a). First, two of the consuming reactions \((C_1, C_2)\) of a branch point were chosen. Two sets of FDs, one with \(v_{C_1} \neq 0, v_{C_2} = 0\) and vice versa, were sampled and summed up as the set of FDs being tested. The second scenario is more difficult in which metabolites have two producing reactions \((P_1, P_2)\) individually and coupled to \(C_1\) and \(C_2\) respectively (Figure A.2b). FDs were sampled and summed up as complex FDs for testing. In both scenarios, FDs were decomposed into EFMs to check whether the distinct branches were correctly associated in the chosen EFMs. In this test, the available carbon sources were set to be fructose and acetate to allow more branch points while their simultaneous uptake does not conflict the given regulatory rules. As this test aimed to assess the basic ability of MBD, the weight \(w_i\) was set to 1 for the branch point being tested and 0 for the rest of metabolites. All metabolites were balanced during the calculation of EFMs. 19 500 EFMs were calculated.
Appendix A. Minimal branching decomposition

Figure A.2. Testing the ability of MBD to separate fluxes at branch points.
(a) Scenario 1 in test 1: separation of two individual consuming reactions. (b) Scenario 2 in test 1: correct separation of two pairs of coupled producing and consuming reactions. (c) Test 2: a simulated case of glyoxylate cycle coupled to acetate uptake implied by regulatory rules.

There are 11 branch points in the first scenario. In each case, 10 FDs were randomly sampled and then subjected to MBD. In all, 1000 RDs were obtained for each FD as comparison. In all cases, no EFMs in an MBD use two consuming reactions simultaneously whereas it is true for only 7.5% of the RDs overall. In scenario 2, six branch points with two or more producing and consuming reactions were identified and 94 combinations were tested. For each combination, 10 FDs for each of the four producing and consuming pairs were sampled and combined as shown in Figure A.2b. MBDs for the resultant FDs were then calculated. In 94% of the MBDs for near 2000 FDs, all EFMs involve the correct reaction pairs whereas only 1% of the RDs is correct. These results confirmed that the proposed MBD is capable of separating fluxes at branch point metabolites.

**Coupling of glyoxylate cycle and acetate uptake**

The given regulatory rules of the core *E. coli* network imply that the glyoxylate cycle is active only if acetate is being consumed. Thus, in the presence of fructose and acetate, the system can consume both substrates with the glyoxylate cycle used primarily for acetate metabolism. Similar experiments and results have been reported in yeast (dos Santos et al., 2003). To test whether MBD can distinguish the acetate-induced glyoxylate cycle from the TCA cycle, two sets of FDs were sampled, one without acetate uptake and glyoxylate cycle activity and the other without fructose uptake (Figure A.3c). One FD from each set was summed to form a complex FD which was subject to MBD. \( w_i = 1 \) for all metabolites except the obvious branch point Acetyl-CoA with \( w_{Acetyl-CoA} = 1000 \). MinEFMDs and RDs were also computed for comparison.

In all 200 FDs tested, no EFMs in any MBDs have simultaneous fructose uptake and glyoxylate cycle activity whereas all MinEFMDs do. Only 1.5% of MBDs use one EFM cometabolizing fructose and
acetate whereas all MinEFMDs do. For RDs, >99% have neither properties. This example demonstrated the unique capability of MBD to distinguish individually regulated pathways in a realistic system setting with underlying regulatory rules.

**Uniqueness of MBDs**

To investigate the uniqueness of MBDs, 2000 FDs with glucose as carbon source were sampled. Alternative optimal MBDs and MinEFMDs were enumerated. To obtain a smaller set of EFMs for easier optimization, several cofactors and small metabolites were unbalanced when calculating EFMs. It is noted that though EFMs resulting from unbalancing cofactors can still properly decompose any FDs, they cannot reflect the subtle interdependence between different metabolic pathways. More discussion can be found in the Supplementary Material (section A.8).

MBD has a satisfactory performance regarding uniqueness. There are 45% of FDs that have a unique MBD and 95% of the FDs have no more than eight multiple MBDs. To compare the results with MinEFMDs, we randomly picked 20 FDs and enumerated the alternative MinEFMDs. All FDs turned out to have ≥1000 different MinEFMDs. As the minimum number of EFMs required for decomposition is much smaller than the number of contributable EFMs, e.g. dozens out of hundreds or thousands, there are virtually countless MinEFMDs. The number of multiple MBDs is extremely small in this sense. The results suggested that MBD is effective in identifying a small subset of decompositions by EFMs with fewer branching pathways which are suitable for understanding the FDs.

**A.4.2 Mouse myocardial metabolic network**

To demonstrate that the proposed method indeed yield a more intuitive and biologically relevant decomposition, MBD was applied to a FD of mouse cardiomyocyte estimated from isotopomer labeling experiments performed in perfused mouse hearts (Vo and Palsson, 2006). The model and the FD are available as the supplementary information in that publication. The mouse myocardial metabolic network consists of 240 metabolites and 257 reactions compartmentalized into cytoplasm, mitochondrion and extracellular space. There are 99 active reactions in the FD (Figure A.3). Glucose and oleate (a fatty acid) were available as carbon sources. Lactate, pyruvate and the ketone bodies acetoacetate and 3-hydroxybutanoate were excreted. A small amount of citrate and succinate were also produced.
Figure A.3. Reactions with non-zero fluxes in the mouse myocardial metabolic network.
Intermediate steps of glycolysis are omitted. Water involved in reactions is not shown. All abbreviations of reactions and metabolites are the same as in the original network. Colored version is available in Supplementary Material.

The primary function of cardiomyocyte is energy production for heart contraction. It is represented by the reaction ‘DMatp’ hydrolyzing ATP in the network. By decomposing the FD into a set of EFMs, different pathways for energy production and other cellular activities can be revealed. All metabolites were balanced for EFM calculation. In all, 1236 EFMs were computed. Optimality could not be proved after 24 h in the initial trial to find an MBD using a uniform weight. We then attempted to put emphasis on the obvious branch point of glucose and fatty acid oxidation, acetyl-CoA, by giving a weight of 1000 to it and 1 to other metabolites. An optimal MBD comprising 14 EFMs was obtained and was proved to be unique. EFMs in the MBD can be divided into four types. The first is three EFMs of fatty acid oxidation that produce CO$_2$ and ketone bodies. The second is nine EFMs producing CO$_2$ and different organic acids from glucose. The third is an EFM consuming glucose and oleate and producing citrate, β-hydroxybutyrate (BHB) and CO$_2$. In this EFM, it is unambiguous that BHB is produced from oleate because pyruvate dehydrogenase, the only reaction in the FD converting pyruvate into acetyl-CoA, is not included in this EFM and acetyl-CoA is the precursor of BHB. The fourth type is an isolated cycle between citrulline and ornithine. The difference between EFMs lies in the products and the different cellular activities or pathways, including the malate-aspartate shuttle, pseudoketogenesis and an apparent glutamate-glutamine futile cycle.
We examined whether finding decomposition after removing EFMs with simultaneous oleate and glucose uptake is possible. No solution could be found, showing that at least one such EFM is required in any decomposition. To see whether this necessary interdependence of glucose and fatty acid oxidation originated from the balance of cofactors, we tried to find MBDs using sets of EFMs calculated from unbalancing some cofactors and small metabolites. When unbalancing NAD$^+$ and NADH, an MBD without any EFM using both carbon sources simultaneously was found, whereas unbalancing other metabolites such as ATP, ADP, H$^+$ or Pi did not give the desired result. This suggested that glucose and fatty acid oxidation interacted to have a balanced use of NADH/NAD$^+$. This kind of comparison between MBDs obtained from balanced and partially balanced EFMs can provide insights into the interdependence of pathways.

For comparison, MinEFMDs and RDs were computed. There were 35 optimal MinEFMDs, each with 13 EFMs. To have a fair comparison, nine suboptimal MBDs were enumerated. Comparing the three types of decomposition can distinguish several properties of an MBD (Table A.2). First, only one EFM in each MBD consumes glucose and oleate simultaneously. As one is hexose and the other is fatty acid, their uses lie in different pathways and thus a good decomposition should avoid using EFMs consuming both. In all MinEFMDs, however, at least five such EFMs were used.

Table A.2. Comparison of MBD, MinEFMD and random decomposition for the mouse myocardial flux distribution.

<table>
<thead>
<tr>
<th>Property</th>
<th>MBD</th>
<th>MinEFMD</th>
<th>RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number of EFMs</td>
<td>14</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>2. Number of multiple optima</td>
<td>1 (unique)</td>
<td>35</td>
<td>--</td>
</tr>
<tr>
<td>3. Maximum number of products in EFMs</td>
<td>2</td>
<td>3 or 4</td>
<td>≥ 3 (88.8%)</td>
</tr>
<tr>
<td>4. Number of EFMs with simultaneous glucose and oleate uptake</td>
<td>1*</td>
<td>≥ 5</td>
<td>≥ 5 (99.5%)</td>
</tr>
<tr>
<td>5. No EFMs producing ketone bodies from glucose</td>
<td>True</td>
<td>True for 8.6%</td>
<td>True for 6.4%</td>
</tr>
<tr>
<td>6. No EFMs involving oleate uptake and malonyl-CoA simultaneously</td>
<td>True</td>
<td>None of the solutions</td>
<td>True for 0.1%</td>
</tr>
</tbody>
</table>

*Our results showed that at least one EFM consuming both carbon sources is required for decomposition. RD: random decomposition.

Second, most EFMs in the MBDs excrete at most one organic acid or ketone body except two EFMs that excrete two whereas all MinEFMDs use several EFMs producing three or four products simultaneously. Such decompositions are more difficult to interpret because the sole effect of producing a certain compound alone (e.g. role in ATP production) might not be observed by looking at individual EFMs. EFMs excreting more metabolites are also expectably longer which might be more difficult to visualize. Another relevant advantage of MBD is that because an MBD would in general use EFMs that produce as few compounds as possible, if coproduction is observed in an EFM in an MBD solution, it may suggest the possibility that the production of the compounds is coupled under the particular condition where the FD is measured.
Third, the EFMs used in the MBDs appear to be more consistent with the regulation of metabolic pathways. For instance, ketogenesis by which ketone bodies are produced is known to be associated with fatty acid oxidation rather than glycolysis. In MBDs, no EFMs producing ketone bodies from glucose are present but 19 of 35 MinEFMDs contain at least one of these EFMs. Another example is malonyl-CoA (‘malcoa’ in Fig. 3). Production of malonyl-CoA is a known mechanism leading to the repression of fatty acid oxidation by glucose by inhibiting the transfer of fatty acyl moieties into mitochondria in cardiac muscle (Depre et al., 1999). Use of EFMs that do not couple fatty acid oxidation and the cycle between malonyl-CoA and acetyl-CoA is thus a more rational choice. In the MBD, no EFMs simultaneously use oleate and involve malonyl-CoA whereas all MinEFMDs do. The same comparison was also performed between MBDs and MinEFMDs obtained from EFMs unbalancing NADH and NADH. All properties still hold (see Supplementary Material in section A.8).

From the comparison, MBD is shown to have several properties of biological relevance that are possessed by only a tiny portion of MinEFMDs. We argue that the more biologically relevant decomposition by MBD does not occur by chance but rather by the minimization of the use of EFMs with branching pathways. The reasons are 2-fold. First, from the consideration of the network structure alone, naturally occurring metabolic pathways tend to prefer linearity to branching at branch point metabolites as suggested in Ihmels et al. (2004) based on their observation on data from ≥1000 experimental conditions. Second, the flux of a certain reaction is the sum of the fluxes of different pathways containing that particular reaction. As different pathways have different fluxes, minimizing the use of those ‘hybrid’ EFMs would force the use of EFMs representing the true pathways to fit the flux values rather than a combination of ‘hybrid’ EFMs which is also able to fit the values of fluxes with a similar number of EFMs.

A.4.3 Lactococcus lactis genome-scale metabolic network

*Lactococcus lactis* (*L. lactis*) is a model organism in the group of lactic acid bacteria and has an important role in cheese production. *L. lactis* is in general auxotrophic for several amino acids (AAs) (Jensen and Hammer, 1993) and grows on complex media. It also produces compounds involved in flavour formation (Ayad et al., 1999). From the recently published genome-scale metabolic network of *L. lactis* subsp. *cremoris* MG1363 which specifies many flavour-forming pathways (Flahaut et al., 2013), a FD with a complex profile of substrate consumption and production can be obtained and can thus test the applicability of MBD. The FD was simulated by FBA using the measured consumption/production rates of glucose, lactate, formate, acetate, ethanol and different AAs in a chemostat experiment at dilution rate 0.05 h⁻¹ as constraints. The subnetwork spanned by the simulated FD contains 294 metabolites and 280 reactions, including 26 uptake and 13 excretion reactions. In all, 518 EFMs were computed when balancing all the metabolites.
A unique MBD of 11 EFMs was computed. Nine other suboptimal MBDs, 10 MinEFMDs and 10^4 RDs were obtained for comparison. First, we focused on two properties, the number of exchange reactions and the length of the EFMs. They can reflect the interpretability and visualizability of the EFMs. Figure A.4 shows their distribution in the three types of decompositions. MBD in general used shorter EFMs with lower numbers of uptake and excretion reactions whereas MinEFMD and RD performed similarly. This indicates that MBD is easier for interpretation and visualization.

Second, we looked into the organization of individual pathways. An interesting observation pertains to the fermentation mode. *L. lactis* has been well known for its shift between homolactic fermentation producing only lactate and mixed-acid fermentation that produces formate, acetate and ethanol in addition depending on the glucose uptake rate (Thomas et al., 1979). The two modes are regulated differently. The reactions responsible for lactate, formate, acetate and ethanol production are lactate dehydrogenase (LDH), pyruvate-formate lyase (PFL), acetate kinase (ACK) and alcohol dehydrogenase (ADH) respectively. No MBDs contain EFMs involving concurrent use of LDH and PFL, ACK or ADH, while 97% of RDs and all but 1 MinEFMDs do.

![Figure A.4](image)

Figure A.4. (a) Number of uptake reactions, (b) number of excretion reactions and (c) length of EFMs in MBD, MinEFMD and random decomposition.

The *L. lactis* network also models detailed AA metabolism including pathways for AA interconversion and tRNA biosynthesis. According to a recently reconstructed transcriptional regulatory network of *L. lactis* (Ravcheev et al., 2013) available in RegPrecise database (Novichkov et al., 2010), they belong to different regulons and meanwhile they coincide at many metabolites. It is relevant to check how MBD performs regarding these branch points. In the current MBD, however,
the EFM matrix used to calculate MBD includes the biomass reaction, forcing many pathways to participate in one EFM to produce biomass. This is also the reason for the small size of the EFM set.

To cope with this, in a new round of MBD calculation, the biomass reaction was split into several reactions producing individual biomass components, including DNA, RNA, proteins, lipid, peptidoglycan, etc. This then resulted in 32 962 EFMs. A weight equal to 1000 was assigned to branch points in AA metabolism (Table A.3), and 1 to other metabolites. Most branch points are AAs acting as precursors for tRNA, other biomass components or flavour compounds. Cysteine and serine as an exception contribute to the pyruvate pool. Other branch points include 2-oxobutanoate (2OBUT), a precursor for a flavour compound where the fluxes from threonine and cystathionine meet; ribose 5-phosphate (R5P), synthesized from the pentose phosphate pathway or the ribose produced from methionine and phosphoribosyl pyrophosphate (PRPP), synthesized from R5P, used for producing either tryptophan or NAD⁺.

<table>
<thead>
<tr>
<th>Branch point</th>
<th>Flow from</th>
<th>Flow towards</th>
<th>% with no EFM of &gt;1 branches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Ala[e], pyruvate</td>
<td>Ala-tRNA⁵, D-alanine⁵</td>
<td>100% 13%</td>
</tr>
<tr>
<td>Cys</td>
<td>Cys[e]</td>
<td>Cys-tRNA⁵, L-cyst, pyruvate, CoA⁴</td>
<td>80% 0</td>
</tr>
<tr>
<td>Met</td>
<td>Met[e]</td>
<td>Met-tRNA⁵, L-cyst.</td>
<td>100% 1%</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly[e]</td>
<td>Gly-tRNA⁵, Thr</td>
<td>100% 0.1%</td>
</tr>
<tr>
<td>Thr</td>
<td>Thr[e], Gly</td>
<td>Thr-tRNA⁵, 2OBUT</td>
<td>100% 0</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser[e]</td>
<td>Ser-tRNA⁵, pyruvate</td>
<td>0 0.8%</td>
</tr>
<tr>
<td>Val</td>
<td>Val[e]</td>
<td>Val-tRNA⁵, 2H3MB[e]²</td>
<td>100% 1.4%</td>
</tr>
<tr>
<td>Glu</td>
<td>Glu[e]</td>
<td>Glu-tRNA⁵, D-glutamate⁴</td>
<td>100% 14%</td>
</tr>
<tr>
<td>Leu</td>
<td>Leu[e]</td>
<td>Leu-tRNA⁵, 2HXIC[e]²²</td>
<td>0 0</td>
</tr>
<tr>
<td>Lys</td>
<td>Lys[e]</td>
<td>Lys-tRNA⁵, uAGL⁵</td>
<td>100% 14%</td>
</tr>
<tr>
<td>2OBUT</td>
<td>Thr, L-cysth</td>
<td>2H3MP[e]²²</td>
<td>100% 50%</td>
</tr>
<tr>
<td>R5P</td>
<td>ribose, RU5P</td>
<td>PRPP</td>
<td>100% 0.4%</td>
</tr>
<tr>
<td>PRPP</td>
<td>R5P</td>
<td>Trp-tRNA⁵, NAD⁺²²</td>
<td>100% 6%</td>
</tr>
</tbody>
</table>

Notes. [e], extracellular; L-cysth, L-cystathionine; 2OBUT, 2-oxobutanoate; 2H3MB, 2-hydroxy-3-methylbutanoate; 2H3MP, 2-hydroxy-3-methylpentanoate; 2HXIC, L-2-hydroxyisocaproate; uAGL, UDP-N-acetyl muramoyl-L-alanyl-D-glutamyl-L-lysine; R5P, ribose 5-phosphate; RU5P, ribulose 5-phosphate; PRPP, phosphoribosyl pyrophosphate.

Convergence to optimality seemed impractical and five suboptimal MBDs were obtained. The best one has an optimality gap of 22% after 24 h of calculation. For MinEFMD, no solution using less EFMs than RD could be found so only RD was considered.
Comparing the original MBDs with the current ones, differences can be observed. First, the present MBDs consist of eight more EFMs than the original. Second, EFMs are shorter in the present MBDs. These changes are expectable because smaller parts of metabolism are sufficient for producing individual biomass components at steady state while more EFMs are required when each EFM covers a smaller part. We also looked into the ATP turnover: in the original MBDs, ATP consumption is perfectly distributed to growth requirement in five EFMs and non-growth-associated maintenance in the other six EFMs. In the current MBDs, however, only 70% of ATP maintenance is consumed in EFMs independent of synthesis of biomass components. This may be the result of suboptimality. Alternatively, this suggests that simultaneous clear roles in both ATP consumption and AA branch points are impossible. However, as an advantage of the current MBDs, the ATP consumption by synthesizing each biomass component can be reflected. tRNA and nucleotide synthesis consume >80% of the ATP.

Regarding AA branch points, MBD successfully chose separate EFMs representing the production of biomass components and flavour compounds. Table A.3 compares the five MBDs and 10^4 RDs. All tRNAs are produced in three EFMs and other amino-acid-consuming reactions are in other EFMs, except for the cases of serine and leucine. For example, beside tRNA charging in three EFMs, cysteine participates in the reaction of phosphopantothenate-cysteine ligase for biosynthesis of coenzyme A in another EFM, goes into the pyruvate pool by cysteine desulphhydrase in three other EFMs, and produces cystathionine (for producing a flavour compound) in one EFM. Taking glycine and threonine as another example, beside tRNA charging of the two AAs in three EFMs, threonine is converted into 2OBUT by threonine deaminase for flavour compound production in four other EFMs, of which threonine is taken up from the medium in three of the EFMs and is converted from glycine by threonine aldolase in the other one. In most cases, RD failed to have such properties. More details can be found in Supplementary Material (section A.8).

For leucine, we found that in all EFMs, tRNA charging was coupled to the production of a flavour compound, thus inseparable. For serine, if EFMs with only one serine-consuming reaction are used, EFMs branching at either threonine or cysteine must be used. It was observed that all three AAs had one consuming reaction producing NH₄. We thus suspected that branches were not separated due to NH₄ balance. When finding MBD from EFMs with NH₄ unbalanced, indeed the issue was resolved.

From this and the previous examples, an interesting observation is that the FDs can be represented by most EFMs of individual pathways together with a few necessary EFMs coupling the pathways. It may suggest that while a large part of the metabolism can be interpreted as the sum of individual pathways, at least a small part must be ascribed to cooperation between pathways.
A.5 Conclusion

MBD to decompose a FD into a set of EFMs was derived. The EFMs contain as few branching pathways as possible. MBD advances the analysis of FD by decomposition into EFMs in three aspects. First, it originates from the consideration of the branching structure in metabolic network based on experimental evidence which is a novel attempt. Second, MBD results in more biologically relevant EFMs leading to clearer interpretation. Third, MBD showed good performance regarding uniqueness. It diminishes the ambiguity of decomposition.

The minimal branching EFM concept underlying MBD can be useful in further applications, e.g. identifying metabolic pathways for engineering by acting as an estimate for biologically feasible EFMs. Scaling up the method to genome-scale networks is necessary for practical applicability. In conclusion, MBD is able to bring useful insights when analyzing FDs. Combining information from e.g. transcriptional regulation and thermodynamics, wider application of EFMs can be anticipated.

A.6 Acknowledgements

The authors would like to acknowledge the helpful comments made by the three anonymous reviewers.

A.7 References


A.8 Supplementary Material

A.8.1 Contributable EFMs
The whole set of EFMs $E$ or at least the subset called ‘contributable EFMs’ $E_{\text{contri}}$ (EFMs able to have non-zero weights for a particular flux distribution) needs to be calculated before MBD or other optimization objectives can be applied to choose a particular subset of EFMs to reconstruct a flux distribution. In the article, contributable EFMs with respect to a flux distribution $v$ is defined as:

$$E_{\text{CONTRI}} = \{e_k | \forall j \text{ s.t. } v_j = 0 \Rightarrow e_{jk} = 0 \}$$

EFMs in $E$ but not in $E_{\text{contri}}$ can have only zero weight to $v$ whereas those in $E_{\text{contri}}$ are always possible to have non-zero weight to $v$ because of the ‘no cancellation’ property. Concretely, it means that because each reversible reaction is split into two irreversible reactions, fluxes in all EFMs must be positive ($e_{jk} > 0$) so from the equation of decomposition:

$$v_j = \sum_{k=1}^{K} \alpha_k e_{jk}$$

Thus $v_j = 0$ implies that $\alpha_k = 0$ whenever $e_{jk} > 0$. $E_{\text{contri}}$ can also be found given the whole set $E$ by deleting the columns with $e_{jk} \neq 0$ but $v_j = 0$ for each $j$. Or alternatively, let $s = [S_{v_j \neq 0}, S_{v_j = 0}]$ where $S_{v_j \neq 0}$ and $S_{v_j = 0}$ contain the columns of reactions with non-zero and zero fluxes respectively. Then a more practical way is to calculate $E_{\text{contri}}$ directly from the submatrix $s_{v_j \neq 0}$. In summary:

- $E$ can recover any flux distributions satisfying $Sv = 0$, $v_{ub} \geq v \geq v_{lb}$ but not always tractable
- $E_{\text{contri}}$ is defined for a given flux distribution $v$ and is more likely tractable or can be found from $E$ if $E$ is known.
- $E_{\text{contri}}$ in general cannot recover all flux distributions in the network defined by $S$ but can recover all flux distributions in the subnetwork defined by $S_{v_j \neq 0}$

A.8.2 Decompositions by EFMs with unbalanced metabolites
EFMs resulting from unbalancing cofactors cannot reflect the subtle inter-dependence between different metabolic pathways but they can still properly decompose any FDs because of the following:

Let $S = [S_{\text{bal}}, S_{\text{unbal}}] \in \mathbb{R}^{m \times n}$ be the stoichiometric matrix where $S_{\text{bal}} \in \mathbb{R}^{m_{\text{bal}} \times n}$ and $S_{\text{unbal}} \in \mathbb{R}^{(m-m_{\text{bal}}) \times n}$ are the stoichiometric matrices of balanced and unbalanced metabolites respectively.

Denote the EFM matrix derived from $S_{\text{bal}}$ by:

$$E_{\text{bal}} = [e_{\text{bal}1}, e_{\text{bal}2}, \ldots, e_{\text{bal}K}] \in \mathbb{R}^{n \times K}.$$
Note that a flux distribution $\mathbf{v}$ satisfying $S\mathbf{v} = \mathbf{0}$ also satisfies $S_{\text{bal}}\mathbf{v} = \mathbf{0}$. This implies $\mathbf{v} = \mathbf{E}_{\text{bal}} \alpha$ where $\alpha = [\alpha_1 \alpha_2 \ldots \alpha_K]$ is the weight vector of EFM.

The imbalance of unbalanced metabolites in the $k$-th EFM, mathematically equal to the vector $S_{\text{unbal}}(\alpha_k \mathbf{e}_{\text{bal}k})$ actually reflects the net production/consumption of unbalanced metabolites of that EFM in the flux distribution which is always cancelled out by other EFMs in the decomposition. Occasionally, it can bring insights into the different roles of the pathways regarding the production/consumption of the unbalanced metabolites which are overall balanced in the flux distribution.

### A.8.3 Test of uniqueness of MBDs

To obtain a smaller set of EFMs for easier optimization, several cofactors and small metabolites were not required to be balanced when calculating EFMs, including ATP and ADP, acetyl-coenzyme A and coenzyme A, NADH and NAD$^+$, NADPH and NADP$^+$, O$_2$, CO$_2$, H$_2$O, H$^+$, NH$_4$ and inorganic phosphate. In this setting, many branches balancing cofactors are excluded and not accounted by MBD. $10^5$ EFMs were computed. For each flux distribution, contributable EFMs were selected by ruling out EFMs with $e_{jk} \neq 0$ but $v_j = 0$. A summary about the tests on the two sets of flux distributions is shown in Table A.4. The number of alternative MBDs and the time required for computing the first MBD solution are plotted in Figure A.5.

<table>
<thead>
<tr>
<th>Carbon source</th>
<th>Glucose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples decomposed</td>
<td>2000</td>
</tr>
<tr>
<td>Weight $w_i$</td>
<td>1 for all metabolites</td>
</tr>
<tr>
<td>Contributable EFMs in each flux distribution</td>
<td>159 (sd = 27)</td>
</tr>
<tr>
<td>EFMs in MBD for each flux distribution</td>
<td>36</td>
</tr>
<tr>
<td>Flux distributions with unique MBD</td>
<td>898 (45%)</td>
</tr>
<tr>
<td>Alternative MinEFMDs for 20 tested FDs</td>
<td>$\geq 1000$</td>
</tr>
</tbody>
</table>

sd, standard deviation.
Appendix A. Minimal branching decomposition

Figure A.5. (a) Number of multiple MBDs and (b) time required for computing the first MBD for FDs randomly sampled from the core *E. coli* metabolic network.

A.8.4 Mouse myocardial metabolic network

![Diagram of mouse myocardial metabolic network]

Figure A.6. Reactions with non-zero fluxes in the mouse myocardial metabolic network (color).

Intermediate steps of glycolysis are omitted. Water involved in reactions is not shown. All abbreviations of reactions and metabolites are the same as in the original network.
As described in the article, we also performed analysis of MBD and MinEFMD using the set of EFMs resulted from unbalancing NADH and NAD⁺. Table A.5 summarizes the results. >20,000 alternative MinEFMDs were found whereas the MBD is unique. Figure A.7 shows the unique MBD and a randomly selected MinEFMD.

Table A.5. Comparison of the properties of MBD and MinEFMD for the mouse myocardial flux distribution calculated from the EFM matrix in which NADH and NAD⁺ are left unbalanced.

<table>
<thead>
<tr>
<th>Property</th>
<th>MBD</th>
<th>MinEFMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Number of EFMs in the solution</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>2. Number of multiple optima</td>
<td>1 (unique)</td>
<td>&gt;20,000</td>
</tr>
<tr>
<td>3. Number of products excreted in EFMs</td>
<td>All EFMs produce 1 product except one EFM producing 2</td>
<td>83% of solutions contain ≥ 2 EFMs producing 2 – 3</td>
</tr>
<tr>
<td>4. No EFMs catabolizing glucose and oleate simultaneously</td>
<td>True</td>
<td>True for 0.005% of solutions</td>
</tr>
<tr>
<td>5. No EFMs producing ketone bodies from glucose</td>
<td>True</td>
<td>True for 0.1% of the solutions</td>
</tr>
<tr>
<td>6. No EFMs catabolizing oleate and involving malonyl-CoA simultaneously</td>
<td>True</td>
<td>True for 4.3% of the solutions</td>
</tr>
<tr>
<td>7. Satisfying any two of the property 3 – 5</td>
<td>True</td>
<td>None of the solutions</td>
</tr>
</tbody>
</table>
Figure A.7. Two sets of 16 EFMs composing (a) the unique MBD and (b) a randomly selected MinEFMD calculated from EFMs in which NADH and NAD$^+$ are left unbalanced.

Each thumbnail shows the reactions of an EFM. Two EFMs containing only the transport reactions for water are not shown.

A.8.5 *Lactococcus lactis*’s metabolic network

In the second round of MBD calculation to study the amino acid metabolism in *Lactococcus lactis*, the biomass reaction was separated several reactions producing individual components. They are listed in Table A.6 put at the end of this document.

Figure A.8 shows the branch points in the amino acid metabolism. Figure A.9 shows the best MBD obtained and Figure A.10 shows a randomly selected EFM from a randomly selected random decomposition. The most obvious difference is that in the MBD, all tRNA charging reactions do not mix with reactions in other branches except for leucine and serine. For leucine, tRNA charging is necessarily coupled to a flavour compound production and for serine branches are not separated due to NH₄ balance as explained in the article. In non-tRNA-charging EFMs, amino acids are further metabolized to produce other biomass components or flavour compounds. For example, in peptidoglycan synthesis, starting from fructose 6-phosphate, alanine, glutamate and lysine are involved in turn. For cysteine, it can participate in the reaction of phosphopantothenatecysteine ligase for biosynthesis of coenzyme A from pantothenate; go into the pyruvate pool by cysteine desulhydrase; and produces cystathionine with another substrate O-acetyl-L-homoserine produced from methionine to finally produce a flavour compound. For glycine, it can be converted to threonine which is used to produce the same flavour compound. There are also three branch points which are not amino acids but are involved in the amino acid metabolism, including 2-oxobutanoate (2OBUT), a precursor for a flavour compound where the fluxes from threonine and cystathionine meet; ribose 5-phosphate (R5P), synthesized from the pentose phosphate pathway or the ribose produced from methionine; phosphoribosyl pyrophosphate (PRPP), synthesized from R5P and used for producing either tryptophan or NAD⁺. From Figure A.9, it can be seen how different parts of the amino acid metabolism and the different branches of branch points were included in different EFMs whereas in a general random decomposition, many EFMs similar to Figure A.10 containing many branches of branch points co-exist.
Figure A.8. Branch points in the amino acid metabolism in the flux distribution in the *L. lactis* metabolic network.

Extracellular metabolites are in blue. Biomass components are in green. Circled metabolites are the branch points in the amino acid metabolism which were given a high weight for the calculation of MBD. Some reactions are lumped for simplicity. All abbreviations follow the original publication of the *L. lactis* metabolic network (Flahaut, N. a L. et al. (2013) Genome-scale metabolic model for *Lactococcus lactis* MG1363 and its application to the analysis of flavour formation. *Appl. Microbiol. Biotechnol.*, 97, 8729–39).
Figure A.9. MBD with high weights on amino acid branch point.
Each thumbnail shows the active reactions in the amino acid metabolism of an EFM. Only EFMs with distinct active reactions in the amino acid metabolism are shown.
Figure A.10. A randomly selected EFM from a randomly selected random decomposition of the flux distribution in the *L. lactis* metabolic network.
### Table A.6. Splitting the biomass reaction into groups of metabolites

<table>
<thead>
<tr>
<th>metabolite</th>
<th>stoichiometry</th>
<th>group</th>
<th>metabolite</th>
<th>stoichiometry</th>
<th>group</th>
</tr>
</thead>
<tbody>
<tr>
<td>2chdeacp[c]</td>
<td>-0.00365</td>
<td>phospholipid</td>
<td>tnamet[c]</td>
<td>0.08402</td>
<td>tRNA</td>
</tr>
<tr>
<td>2coacdcp[c]</td>
<td>-0.05359</td>
<td>phospholipid</td>
<td>trnaphe[c]</td>
<td>0.16804</td>
<td>tRNA</td>
</tr>
<tr>
<td>2ctdeacp[c]</td>
<td>-0.00061</td>
<td>phospholipid</td>
<td>trnapro[c]</td>
<td>0.180643</td>
<td>tRNA</td>
</tr>
<tr>
<td>acp[c]</td>
<td>0.1218</td>
<td>phospholipid</td>
<td>trnaser[c]</td>
<td>0.247859</td>
<td>tRNA</td>
</tr>
<tr>
<td>epocdacp[c]</td>
<td>-0.01583</td>
<td>phospholipid</td>
<td>trnathr[c]</td>
<td>0.277266</td>
<td>tRNA</td>
</tr>
<tr>
<td>hdeacp[c]</td>
<td>-0.03593</td>
<td>phospholipid</td>
<td>trnattrp[c]</td>
<td>0.033608</td>
<td>tRNA</td>
</tr>
<tr>
<td>ocdacp[c]</td>
<td>-0.00122</td>
<td>phospholipid</td>
<td>trnatyr[c]</td>
<td>0.12603</td>
<td>tRNA</td>
</tr>
<tr>
<td>tdeacp[c]</td>
<td>-0.01096</td>
<td>phospholipid</td>
<td>trnaval[c]</td>
<td>0.306673</td>
<td>tRNA</td>
</tr>
<tr>
<td>glyc3p[c]</td>
<td>-0.3359</td>
<td>phospholipid</td>
<td>trptrna[c]</td>
<td>-0.03361</td>
<td>tRNA</td>
</tr>
<tr>
<td>glyc[c]</td>
<td>0.0138</td>
<td>phospholipid</td>
<td>tytrna[c]</td>
<td>-0.12603</td>
<td>tRNA</td>
</tr>
<tr>
<td>alatrna[c]</td>
<td>-0.36549</td>
<td>tRNA</td>
<td>valtrna[c]</td>
<td>-0.30667</td>
<td>tRNA</td>
</tr>
<tr>
<td>argrrna[c]</td>
<td>-0.17644</td>
<td>tRNA</td>
<td>adp[c]</td>
<td>41.22259</td>
<td>nucleotide</td>
</tr>
<tr>
<td>asnrna[c]</td>
<td>-0.19745</td>
<td>tRNA</td>
<td>atp[c]</td>
<td>-41.3081</td>
<td>nucleotide</td>
</tr>
<tr>
<td>asptrna[c]</td>
<td>-0.19745</td>
<td>tRNA</td>
<td>cmp[c]</td>
<td>0.275</td>
<td>nucleotide</td>
</tr>
<tr>
<td>cytrna[c]</td>
<td>-0.20585</td>
<td>tRNA</td>
<td>ctp[c]</td>
<td>-0.3408</td>
<td>nucleotide</td>
</tr>
<tr>
<td>gltrna[c]</td>
<td>-0.23106</td>
<td>tRNA</td>
<td>datp[c]</td>
<td>-0.02368</td>
<td>nucleotide</td>
</tr>
<tr>
<td>glutrna[c]</td>
<td>-0.23106</td>
<td>tRNA</td>
<td>dctp[c]</td>
<td>-0.01332</td>
<td>nucleotide</td>
</tr>
<tr>
<td>glytrna[c]</td>
<td>-0.24786</td>
<td>tRNA</td>
<td>dgtp[c]</td>
<td>-0.01332</td>
<td>nucleotide</td>
</tr>
<tr>
<td>histrna[c]</td>
<td>-0.07562</td>
<td>tRNA</td>
<td>dtp[c]</td>
<td>-0.02368</td>
<td>nucleotide</td>
</tr>
<tr>
<td>iletRNA[c]</td>
<td>-0.24366</td>
<td>tRNA</td>
<td>gdp[c]</td>
<td>8.402</td>
<td>nucleotide</td>
</tr>
<tr>
<td>leutrna[c]</td>
<td>-0.36969</td>
<td>tRNA</td>
<td>gtp[c]</td>
<td>-8.50728</td>
<td>nucleotide</td>
</tr>
<tr>
<td>lystrna[c]</td>
<td>-0.23656</td>
<td>tRNA</td>
<td>h2o[c]</td>
<td>-49.6253</td>
<td>nucleotide</td>
</tr>
<tr>
<td>mettrna[c]</td>
<td>-0.08402</td>
<td>tRNA</td>
<td>h[c]</td>
<td>49.76409</td>
<td>nucleotide</td>
</tr>
<tr>
<td>phetrna[c]</td>
<td>-0.16804</td>
<td>tRNA</td>
<td>pi[c]</td>
<td>49.68549</td>
<td>nucleotide</td>
</tr>
<tr>
<td>protRNA[c]</td>
<td>-0.18064</td>
<td>tRNA</td>
<td>pp[i][c]</td>
<td>0.678</td>
<td>nucleotide</td>
</tr>
<tr>
<td>sertrna[c]</td>
<td>-0.24786</td>
<td>tRNA</td>
<td>udp[c]</td>
<td>0.2103</td>
<td>nucleotide</td>
</tr>
<tr>
<td>thtrna[c]</td>
<td>-0.27727</td>
<td>tRNA</td>
<td>ump[c]</td>
<td>0.0064</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnaala[c]</td>
<td>0.365487</td>
<td>tRNA</td>
<td>utp[c]</td>
<td>-0.07238</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnaarg[c]</td>
<td>0.176442</td>
<td>tRNA</td>
<td>dtdp6dm[c]</td>
<td>-0.0064</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnaasn[c]</td>
<td>0.197447</td>
<td>tRNA</td>
<td>dtdp[c]</td>
<td>0.0064</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnaasp[c]</td>
<td>0.197447</td>
<td>tRNA</td>
<td>udc[pdp][c]</td>
<td>-0.0002</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnacy[c]</td>
<td>0.205849</td>
<td>tRNA</td>
<td>udp[g][c]</td>
<td>-0.0633</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnagln[c]</td>
<td>0.231055</td>
<td>tRNA</td>
<td>udp[g][al][c]</td>
<td>-0.1534</td>
<td>nucleotide</td>
</tr>
<tr>
<td>trnaglu[c]</td>
<td>0.231055</td>
<td>tRNA</td>
<td>PG[c]</td>
<td>-0.119</td>
<td>peptidoglycan</td>
</tr>
<tr>
<td>trnagly[c]</td>
<td>0.247859</td>
<td>tRNA</td>
<td>ala-D[c]</td>
<td>-0.09</td>
<td>D-alanine</td>
</tr>
<tr>
<td>trnahis[c]</td>
<td>0.075618</td>
<td>tRNA</td>
<td>coa[c]</td>
<td>-0.0002</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>trnaile[c]</td>
<td>0.243658</td>
<td>tRNA</td>
<td>nad[c]</td>
<td>-0.002</td>
<td>NAD⁺</td>
</tr>
<tr>
<td>trnaleu[c]</td>
<td>0.369688</td>
<td>tRNA</td>
<td>thf[c]</td>
<td>-1.00E-05</td>
<td>tetrahydrofolate</td>
</tr>
<tr>
<td>trnaly[c]</td>
<td>0.236556</td>
<td>tRNA</td>
<td>thmpp[c]</td>
<td>-1.00E-05</td>
<td>Thiamine diphosphate</td>
</tr>
</tbody>
</table>

A.8.6 Matlab script

The matlab script for MBD together with a data file containing examples can be downloaded at Bioinformatics online:

http://bioinformatics.oxfordjournals.org/

or

https://github.com/shjchan/MinBranchDecomp
### Appendix B  Supplementary Material for Chapter 4

#### B.1. Supplementary tables

<table>
<thead>
<tr>
<th>Primer</th>
<th>Restriction site</th>
<th>Amplified region/ Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>56f</td>
<td>BamHI</td>
<td>ackA1 core upstream</td>
<td>gacaggatccGGAGGATTTTAGCACAAGTG</td>
</tr>
<tr>
<td>56r</td>
<td>PstI</td>
<td></td>
<td>tactgtcagCAATGCCGCGACGATTTG</td>
</tr>
<tr>
<td>57f</td>
<td>PstI</td>
<td>ackA1 core downstream</td>
<td>tactgtcagCACCTCTTGTGAGGAGTG</td>
</tr>
<tr>
<td>57r</td>
<td>XhoI</td>
<td></td>
<td>gttacctcagCTAACGTTTCTGTTTGT</td>
</tr>
<tr>
<td>59f</td>
<td></td>
<td>Verify ackA1 core deletion with 57r</td>
<td>CAAATGCTGCGCGCATTTG</td>
</tr>
<tr>
<td>60f</td>
<td>BamHI</td>
<td>ackA2 core upstream</td>
<td>gacaggatccGATGTATGTTGACCGCATTC</td>
</tr>
<tr>
<td>60r</td>
<td>XhoI</td>
<td></td>
<td>gttacctcagAAACGTTTGCAGGATTTG</td>
</tr>
<tr>
<td>61f</td>
<td>XhoI</td>
<td>ackA2 core downstream</td>
<td>gttacctcagGCGATTTGAGGTGTAATAC</td>
</tr>
<tr>
<td>61r</td>
<td>KpnI</td>
<td></td>
<td>taggtacctCAACAGATCAATTGCTTCAT</td>
</tr>
<tr>
<td>63f</td>
<td></td>
<td>Verify ackA2 core deletion with 61r</td>
<td>CTTTAGTACTACTCCGGCAAC</td>
</tr>
<tr>
<td>CSO834</td>
<td>XbaI</td>
<td>pfl upstream</td>
<td>ctagtctagaCAAGTGATGTACCAAAATGAC</td>
</tr>
<tr>
<td>CSO835</td>
<td>BamHI</td>
<td></td>
<td>cgcggatccTTTGAATCTTCTTTGTTCT</td>
</tr>
<tr>
<td>CSO836</td>
<td>BamHI</td>
<td>pfl downstream</td>
<td>cgcggatccTTCTTAGATTAAACATATTAAG</td>
</tr>
<tr>
<td>CSO837</td>
<td>XhoI</td>
<td></td>
<td>gttacctcagTGTTGATTCACCTCATTTT</td>
</tr>
<tr>
<td>CSO852</td>
<td></td>
<td></td>
<td>CTTGGAATTCTCTTGCTATTTAC</td>
</tr>
<tr>
<td>CSO853</td>
<td></td>
<td></td>
<td>CTTTGTGACATCAAATTACTTG</td>
</tr>
<tr>
<td>71f</td>
<td>BglII</td>
<td>ackA1 gene for His-tagging</td>
<td>actgaagatccACAAACACATTGACGTAACGCTTGTTGCTAC</td>
</tr>
<tr>
<td>71r</td>
<td>SalI</td>
<td></td>
<td>actgaagatccTTATTTTTTAAGTGCTCACAAGTC</td>
</tr>
<tr>
<td>62f</td>
<td>BamHI</td>
<td>ackA2 gene for His-tagging</td>
<td>gacaggatccGAAAAACGCTCGTGGCAAT</td>
</tr>
<tr>
<td>62r</td>
<td>SalI</td>
<td></td>
<td>tacagtctgaTTATTTAGCGGAGGCCCT</td>
</tr>
</tbody>
</table>

#### Construction of gusA reporter strains

<table>
<thead>
<tr>
<th>Primer</th>
<th>Restriction site</th>
<th>Amplified region/ Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>11f</td>
<td>XbaI</td>
<td>436-bp ackA1 upstream and 36-bp</td>
<td>atcgatctagaGAGGATTTACTGACAAGTG</td>
</tr>
<tr>
<td>11r</td>
<td>PstI</td>
<td>CDS with a stop codon</td>
<td>atcgatctgaTGATGAAACCAGGCTTTAC</td>
</tr>
<tr>
<td>12f</td>
<td>XbaI</td>
<td>444-bp ackA2 upstream and 45-bp</td>
<td>atcgatctagaTGAGATGTATGTTGACCG</td>
</tr>
<tr>
<td>12r</td>
<td>PstI</td>
<td>CDS with a stop codon</td>
<td>atcgatctgaTAATGATGAGGAGGCCCT</td>
</tr>
<tr>
<td>75r</td>
<td>SalI</td>
<td>Anti-sense to ackA genes; priming</td>
<td>actgaagatccTTCTTCAACCTTTGTATCTTTGCTCAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the predicted transcription terminator of ackA1 and excluding the putative</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>promoter of ackA2</td>
<td></td>
</tr>
<tr>
<td>CSO50</td>
<td>BamHI</td>
<td>Verify chromosomal integration of</td>
<td>ggaaggatcCCCATAGTTCATCAGTTATC</td>
</tr>
<tr>
<td>CSO263</td>
<td></td>
<td>pLB85</td>
<td>CGCGATCCAGACTGAATG</td>
</tr>
</tbody>
</table>

#### RACE

<table>
<thead>
<tr>
<th>Primer</th>
<th>Restriction site</th>
<th>Amplified region/ Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>72r</td>
<td></td>
<td>5`-RACE for ackA1</td>
<td>CCAACAAGGGGTGTAAGGCCAT</td>
</tr>
<tr>
<td>73r</td>
<td></td>
<td>5`-RACE for ackA2</td>
<td>CGAAAACAGGGGACGCAAGTCAT</td>
</tr>
</tbody>
</table>

DOI: 10.1371/journal.pone.0092256.s001
Appendix B Supplementary Material for Chapter 4

Table B.2. Plasmid used in Chapter 4.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCS1966</td>
<td>Contain oroP for orotate transporter as counterselection tool for deletion; Erm' [26]</td>
</tr>
<tr>
<td>pCS1966-ackA1</td>
<td>pCS1966 containing the upstream and downstream regions of the core part of ackA1 obtained with primer pairs 56-57; Erm'</td>
</tr>
<tr>
<td>pCS1966-ackA2</td>
<td>pCS1966 containing the upstream and downstream regions of the core part of ackA2 obtained with primer pairs 60-61; Erm'</td>
</tr>
<tr>
<td>pCS1966-pfl</td>
<td>pCS1966 containing the upstream and downstream regions of pfl obtained with primers CSO834-837; Erm'</td>
</tr>
<tr>
<td>pLB65</td>
<td>Contain orf1 expressing phage TP901-1 integrase; Cam' [27]</td>
</tr>
<tr>
<td>pLB85</td>
<td>Reporter vector containing gusA reporter; Erm', Amp' [27]</td>
</tr>
<tr>
<td>pLB85-ackA1</td>
<td>pLB85 containing fragment A obtained with primer pair 11 (including 36-bp CDS and 436-bp upstream region of ackA1); Erm', Amp'</td>
</tr>
<tr>
<td>pLB85-ackA2</td>
<td>pLB85 containing fragment B obtained with primers 11f, 12r (including 36-bp CDS and the promoter region of ackA2 as well as the CDS and 436-bp upstream region of ackA1); Erm', Amp'</td>
</tr>
<tr>
<td>pLB85-pCS1966</td>
<td>pLB85 containing fragment C obtained with primer pair 12 (45-bp CDS and 444-bp upstream region of ackA2); Erm', Amp'</td>
</tr>
<tr>
<td>pQE30</td>
<td>Commercial vector (Qiagen) used for N-terminal His-tagging; Amp'</td>
</tr>
<tr>
<td>pQE30-ackA1</td>
<td>pQE30 containing the gene ackA1 amplified using primer pair 71; Amp'</td>
</tr>
<tr>
<td>pQE30-ackA2</td>
<td>pQE30 containing the gene ackA2 amplified using primer pair 62; Amp'</td>
</tr>
<tr>
<td>pCS1966-ackA1-qF</td>
<td>pCS1966 containing fragment D obtained with primers 11f, 75r (fragment B excluding the region immediately after the predicted transcription terminator of ackA1); Erm', Amp'</td>
</tr>
<tr>
<td>pCS1966-ackA2-qF</td>
<td>pCS1966 containing fragment E obtained with primers 12f, 75r (fragment C excluding the region immediately after the predicted transcription terminator of ackA1); Erm', Amp'</td>
</tr>
<tr>
<td>pQE30-ackA1-qF</td>
<td>pQE30 containing the gene ackA1 amplified using primer pair 71; Amp'</td>
</tr>
<tr>
<td>pQE30-ackA2-qF</td>
<td>pQE30 containing the gene ackA2 amplified using primer pair 62; Amp'</td>
</tr>
<tr>
<td>pCS1966-ackA1-qF</td>
<td>pCS1966 containing fragment D obtained with primers 11f, 75r (fragment B excluding the region immediately after the predicted transcription terminator of ackA1); Erm', Amp'</td>
</tr>
<tr>
<td>pCS1966-ackA2-qF</td>
<td>pCS1966 containing fragment E obtained with primers 12f, 75r (fragment C excluding the region immediately after the predicted transcription terminator of ackA1); Erm', Amp'</td>
</tr>
<tr>
<td>pQE30-ackA1-qF</td>
<td>pQE30 containing the gene ackA1 amplified using primer pair 71; Amp'</td>
</tr>
<tr>
<td>pQE30-ackA2-qF</td>
<td>pQE30 containing the gene ackA2 amplified using primer pair 62; Amp'</td>
</tr>
</tbody>
</table>

Erm', Cam' and Amp' stand for erythromycin, chloramphenicol and ampicillin resistance respectively. CDS: coding sequence. doi:10.1371/journal.pone.0092256.s002

Table B.3. Strains used in Chapter 4.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG1363</td>
<td>A plasmid-free strain derived from <em>L. lactis subsp. cremoris</em> NCDO 712 [24]</td>
</tr>
<tr>
<td>MG1363Δpfl</td>
<td>MG1363 with pfl deleted using pCS1966-pfl</td>
</tr>
<tr>
<td>MG1363ΔackA1</td>
<td>MG1363 with the core part of ackA1 deleted using pCS1966-ackA1core</td>
</tr>
<tr>
<td>MG1363ΔackA2</td>
<td>MG1363 with the core part of ackA2 deleted using pCS1966-ackA2core</td>
</tr>
<tr>
<td>MG1363ΔackA1Δpfl</td>
<td>MG1363ΔackA1 with pfl deleted using pCS1966-pfl</td>
</tr>
<tr>
<td>MG1363ΔackA2Δpfl</td>
<td>MG1363ΔackA2 with pfl deleted using pCS1966-pfl</td>
</tr>
<tr>
<td>MG1363ΔackA1Δpfl</td>
<td>MG1363ΔackA1 with pfl deleted using pCS1966-pfl</td>
</tr>
<tr>
<td>MG1363ΔackA2Δpfl</td>
<td>MG1363ΔackA2 with pfl deleted using pCS1966-pfl</td>
</tr>
<tr>
<td>LB436</td>
<td>MG1363 containing pLB65</td>
</tr>
<tr>
<td>LB436ΔqF</td>
<td>LB436 with pPB85 integrated into the TP901-1 attachment site</td>
</tr>
<tr>
<td>LB436ΔackA1</td>
<td>LB436 with pLB85-ackA1 integrated into the TP901-1 attachment site</td>
</tr>
<tr>
<td>LB436ΔackA1Δpfl</td>
<td>LB436 with pLB85-ackA1Δpfl integrated into the TP901-1 attachment site</td>
</tr>
<tr>
<td>LB436ΔackA2Δpfl</td>
<td>LB436 with pLB85-ackA2Δpfl integrated into the TP901-1 attachment site</td>
</tr>
<tr>
<td>LB436ΔackA2Δpfl</td>
<td>LB436 with pLB85-ackA2Δpfl integrated into the TP901-1 attachment site</td>
</tr>
<tr>
<td>M15 pREP4 groESL</td>
<td>An <em>E. coli</em> strain for protein overexpression [25]</td>
</tr>
<tr>
<td>SC136</td>
<td>M15 pREP4 groESL containing pQE30-ackA1</td>
</tr>
<tr>
<td>SC137</td>
<td>M15 pREP4 groESL containing pQE30-ackA2</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0092256.s004
B.2. ACK entries in Uniprot

The data for counting the number of species with multiple ACKs from the 11,122 ACK entries in Uniprot is too large to be put in the thesis. Interested readers please download the .xlsx file from PLoS ONE website:

http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0092256

B.3. Supplementary figures

Figure B.1. Multiple alignment of AckA1, AckA2 and ACKs from Salmonella typhimurium and Methanosarcina thermophila.

Structures identified in previous crystallographic studies* are annotated. doi:10.1371/journal.pone.0092256.s005


Figure B.2. Multiple alignment of all lactococcal ACKs.

Residues conserved with each type of ACK but different between the two types are annotated.

doi:10.1371/journal.pone.0092256.s006
Figure B.3. Activities of purified acetate kinases.

(a) AckA1 and (b) AckA2 converting acetate (OAc) into acetyl-phosphate (Ac-P) at different levels of (i) acetate and (ii) ATP and, converting acetyl-phosphate into acetate at different levels of (iii) Ac-P and (iv) ADP, respectively.

doi:10.1371/journal.pone.0092256.s007
Appendix C  Supplementary Material for Chapter 5

Table C.1. MG1363 and MG/SP1-bdh growing in MalSALN with or without acetoin.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Acetoin added</th>
<th>Growth rate (h⁻¹)</th>
<th>Specific rate of sugar consumption and product formation (mmol h⁻¹gdw⁻¹)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Maltose</td>
<td>Lactate</td>
<td>Formate</td>
<td>Acetate</td>
</tr>
<tr>
<td>MG1363 0</td>
<td>0.60 ± 0.01</td>
<td>7.7 ± 0.1</td>
<td>11.7 ± 0.2</td>
<td>16.6 ± 0.4</td>
<td>9.6 ± 0.7</td>
<td>6.4 ± 0.7</td>
</tr>
<tr>
<td>20 mM</td>
<td>0.54 ± 0.01</td>
<td>6.4 ± 0.4</td>
<td>17.0 ± 1.2</td>
<td>5.8 ± 0.2</td>
<td>4.9 ± 0.1</td>
<td>1.4 ± 0.1</td>
</tr>
<tr>
<td>MG/SP1-bdh 0</td>
<td>0.61 ± 0.01</td>
<td>7.0 ± 0.1</td>
<td>11.1 ± 0.2</td>
<td>15.9 ± 0.6</td>
<td>9.5 ± 0.1</td>
<td>5.9 ± 0.2</td>
</tr>
<tr>
<td>5 mM</td>
<td>0.67 ± 0.01</td>
<td>7.5 ± 0.02</td>
<td>11.8 ± 0.3</td>
<td>16.5 ± 0.5</td>
<td>12.4 ± 0.3</td>
<td>4.0 ± 0.5</td>
</tr>
<tr>
<td>20 mM</td>
<td>0.72 ± 0.01</td>
<td>6.9 ± 0.1</td>
<td>7.6 ± 0.1</td>
<td>5.5 ± 0.2</td>
<td>14.9 ± 0.1</td>
<td>-0.4 ± 0.6</td>
</tr>
</tbody>
</table>

Table C.2. Product yield, carbon recovery and biomass yield for growth in MalSALN.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Acetoin added</th>
<th>Product yield on sugar a</th>
<th>Carbon recovery b (gdw / C6-mol)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lactate</td>
<td>Formate</td>
<td>Acetate</td>
<td>Ethanol</td>
<td></td>
</tr>
<tr>
<td>MG1363 0</td>
<td>38% ± 0%</td>
<td>54% ± 1%</td>
<td>31% ± 2%</td>
<td>21% ± 2%</td>
<td>90% ± 4%</td>
<td>38.8 ± 0.7</td>
</tr>
<tr>
<td>20 mM</td>
<td>23% ± 2%</td>
<td>23% ± 2%</td>
<td>19% ± 1%</td>
<td>6% ± 0%</td>
<td>91% ± 3%</td>
<td>42.0 ± 2.8</td>
</tr>
<tr>
<td>MG/SP1-bdh 0</td>
<td>57% ± 2%</td>
<td>57% ± 2%</td>
<td>34% ± 0%</td>
<td>21% ± 1%</td>
<td>95% ± 1%</td>
<td>43.7 ± 0.9</td>
</tr>
<tr>
<td>5 mM</td>
<td>55% ± 2%</td>
<td>55% ± 2%</td>
<td>41% ± 1%</td>
<td>13% ± 2%</td>
<td>94% ± 4%</td>
<td>44.7 ± 0.5</td>
</tr>
<tr>
<td>20 mM</td>
<td>20% ± 1%</td>
<td>20% ± 1%</td>
<td>54% ± 1%</td>
<td>-1% ± 2%</td>
<td>80% ± 1%</td>
<td>51.9 ± 1.4</td>
</tr>
</tbody>
</table>

a Yield was calculated in terms of C₃-mol (= product flux / maltose flux / 4).
b Carbon recovery was the sum of lactate, ethanol and acetate yield on sugar.

Table C.3. Product yield, carbon recovery and biomass yield for growth in MalSALN(BGP).

<table>
<thead>
<tr>
<th>Acetoin added</th>
<th>Strain</th>
<th>Product yield on sugar a</th>
<th>Carbon recovery b (gdw / C₃-mol)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>MG1363 0</td>
<td>35%</td>
<td>60%</td>
<td>33%</td>
<td>23%</td>
<td>90.0%</td>
</tr>
<tr>
<td>MG/SP2--bdh</td>
<td>38%</td>
<td>57%</td>
<td>32%</td>
<td>22%</td>
<td>90.0%</td>
<td>31.4</td>
</tr>
<tr>
<td>MG/SP1-bdh</td>
<td>41%</td>
<td>58%</td>
<td>33%</td>
<td>20%</td>
<td>90.0%</td>
<td>31.4</td>
</tr>
<tr>
<td>5 mM</td>
<td>MG1363 5</td>
<td>55 ± 3%</td>
<td>30 ± 2%</td>
<td>22 ± 1%</td>
<td>8 ± 1%</td>
<td>84 ± 6%</td>
</tr>
<tr>
<td>MG/SP2--bdh</td>
<td>47 ± 1%</td>
<td>53 ± 3%</td>
<td>37 ± 0.1%</td>
<td>16 ± 0.3%</td>
<td>99 ± 1%</td>
<td>34.1 ± 0.8</td>
</tr>
<tr>
<td>MG/SP1-bdh</td>
<td>36 ± 1%</td>
<td>58 ± 2%</td>
<td>42 ± 0.2%</td>
<td>13 ± 1%</td>
<td>90 ± 1%</td>
<td>34.7 ± 0.7</td>
</tr>
</tbody>
</table>

a Yield was calculated in terms of C₃-mol (= product flux / maltose flux / 4).
b Carbon recovery was the sum of lactate, ethanol and acetate yield on sugar.
### Table C.4. Product yield, carbon recovery and biomass yield for growth in GluSALN(BGP).

<table>
<thead>
<tr>
<th>Acetoin added</th>
<th>Strain</th>
<th>Product yield on sugar</th>
<th>Carbon recovery</th>
<th>Biomass yield (gdw / C₆-mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lactate</td>
<td>Formate</td>
<td>Acetate</td>
</tr>
<tr>
<td>0</td>
<td>MG1363</td>
<td>90.7%</td>
<td>4.2%</td>
<td>3.3%</td>
</tr>
<tr>
<td>MG/SP2–bdh</td>
<td>98.2%</td>
<td>4.7%</td>
<td>4.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>MG/SP1–bdh</td>
<td>89.7%</td>
<td>4.2%</td>
<td>3.8%</td>
<td>0.0%</td>
</tr>
<tr>
<td>5 mM</td>
<td>MG1363</td>
<td>95 ± 4%</td>
<td>2.2 ± 0.3%</td>
<td>2.1 ± 0.4%</td>
</tr>
<tr>
<td>MG/SP2–bdh</td>
<td>85 ± 0%</td>
<td>1.2 ± 0.2%</td>
<td>6.5 ± 0.4%</td>
<td>-0.6 ± 1.1%</td>
</tr>
<tr>
<td>MG/SP1–bdh</td>
<td>80 ± 2%</td>
<td>1.0 ± 0.4%</td>
<td>10.9 ± 0.4%</td>
<td>0.3 ± 0.5%</td>
</tr>
</tbody>
</table>

*a* Yield was calculated in terms of C₃-mol (product flux / glucose flux / 2).

*b* Carbon recovery was the sum of lactate, ethanol and acetate yield on sugar.

### Table C.5. 23BDH activities and rates of acetoin consumption and 23BD production in SALN(BGP).

<table>
<thead>
<tr>
<th>Acetoin in medium</th>
<th>Strain</th>
<th>23BDH activity in MalSALN(BGP) (µmolmin⁻¹mg protein⁻¹)</th>
<th>Specific rate (mmolh⁻¹gdw⁻¹)*</th>
<th>MalSALN(BGP)</th>
<th>GluSALN(BGP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>MG1363</td>
<td>N.D.</td>
<td>0.0</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>MG/SP2–bdh</td>
<td>N.D.</td>
<td>0.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>MG/SP1–bdh</td>
<td>N.D.</td>
<td>0.3</td>
<td>0.0</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>5 mM</td>
<td>MG1363</td>
<td>0.07 ± 0.02</td>
<td>1.8 ± 0.2</td>
<td>1.9 ± 0.2</td>
<td>-0.7 ± 0.9</td>
</tr>
<tr>
<td>MG/SP2–bdh</td>
<td>0.50 ± 0.05</td>
<td>3.0 ± 0.6</td>
<td>2.8 ± 0.3</td>
<td>4.8 ± 0.1</td>
<td>3.2 ± 0.1</td>
</tr>
<tr>
<td>MG/SP1–bdh</td>
<td>2.67 ± 0.08</td>
<td>4.1 ± 0.5</td>
<td>4.0 ± 0.3</td>
<td>8.8 ± 0.4</td>
<td>6.7 ± 0.9</td>
</tr>
</tbody>
</table>

N.D., not determined. 23BDH, 2,3-butanediol dehydrogenase; 23BD, 2,3-butanediol.

*a* positive values for acetoin and 23BD for consumption and production respectively.