Department of Chemical Engineering and Chemical Technology

Imperial College London

Integrated multi-scale models for biologics process development

A thesis submitted for the degree of Doctor of Philosophy

in Chemical Engineering and the Diploma of Imperial College

António Carlos Martins de Jesus Lima Grilo

April 2019

- à Mafalda

- ao Pai, à Mãe, à Ana

"There are three paths to failure: not teaching what one knows, not doing as one preaches and not asking what one ignores"

-St. Bede

"If you refuse to accept anything but the best, you very often get it"

-William Somerset Maugham

Abstract

The therapeutic monoclonal antibody (mAb) market is witnessing an unprecedented growth having nearly duplicated since 2012. Faster regulatory approvals and an increasing number of biosimilars are among the most important mAb market expansion drivers. As the market grows, so does competition making timelines more aggressive. Additionally, political pressure from governments and regulatory authorities over prices of life-saving biological drugs is also increasing. As a consequence, biopharmaceutical companies need to reduce costs while being faster, particularly in having drug substance ready for clinical trials. To this aim, reducing process development times and costs is crucial.

In this thesis, mathematical models integrating several cellular activities important for biomanufacturing are developed. A previously published model-development framework is followed to ensure the development of predictive models. An integrated framework combining multivariate data analysis and biomarker identification techniques for cell culture understanding supports the development of predictive mathematical models for two industrially relevant cell lines, GS-NS0 and GS-CHO, capturing cell cycle, metabolism, energy production, mAb production and apoptosis. As the models herein suggested can capture population heterogeneity by describing the cell cycle and gene expression, their applications to bioreactor optimization is envisaged. To this aim, the models developed in this work can be combined with the description of product quality attributes and/or computational fluid dynamics descriptions of bioreactors. These are deemed important applications to accomplish significant reduction in process development time and costs.

Declaration

The work described in this thesis was carried out by the author (or in collaboration) in the Department of Chemical Engineering and Chemical Technology of Imperial College London, United Kingdom between June 2016 and March 2019. To the best of my knowledge, it does not contain work by any other person, unless otherwise acknowledged.

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND).

Under this licence, you may copy and redistribute the material in any medium or format on the condition that: you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgments

The Ph.D. is a long journey, which one walks alongside several others without whom the final result would be impossible.

The first word could only go to Professor Athanasios Mantalaris. Saki, thank you very much, first of all, for giving me the opportunity of pursuing this project in your group and with you. Thank you for your support, coaching and motivation without which I could never have completed this project. Thank you for sharing my passion for bioprocessing and for helping me focus on the goals we set together for this project while always allowing me to explore new paths. Thank you for understanding my strengths as well you understood my weaknesses. Without this, I would never have progressed as much as I have. Finally, thank you for creating opportunities for my personal development outside of the project, in particular for allowing me to teach.

I would like to thank Dr. Alexandros Kyparissidis and Dr. Karen Polizzi for examining this work and providing meaningful suggestions for improvement.

When this project was first presented to me, it was described as "creating a super-model, a combination of several years of prior knowledge and work". Most of the people who had been involved in this journey had left and only one was in BSEL when I joined: Ana. Thank you, Ana, for sharing your knowledge with me without limits. Thank you for training me, for teaching me how to culture cells, how to operate the bioreactor, how to conduct the analysis and for giving me your protocols, which were essential in streamlining my entire project. ¡Muchas gracias!

A word of acknowledgement is due to all colleagues in BSEL. Michali, thank you very much for being my doctor when I needed, for teaching me how to do PCR, for the company and fun times during the overnight experiments, for all the scientific discussions and conversations, for reading parts of this work and, of

course, for the Greek music and vocabulary: Ευχαριστώ! The Portuguese team was essential in keeping me closer to home. It was a privilege to be able to live in a foreign country, in such a diverse community, and still be able to come to work every day and speak my own language. Zé, thank you for being always present, for your patience and resilience, for taking care of my cells when I could not, and for reading parts of this work. Joana, thank you for keeping me company during the overnight work, for taking care of my cells and for everything you taught me about flow cytometry and for reviewing the corresponding section of this work. Susana, thank you for sharing your profound knowledge of flow cytometry with all of us, and for being always around and keen to help. Aos três, muitíssimo obrigado por tudo! Sophie, thank you so much for sharing many motivational afternoons and evenings, accompanied by our common Japanese friend, for always telling me I could do this, for never doubting me and for proof-reading parts of this work. Thank you very much indeed! Simo, thank you very much for teaching me GSA and sharing the algorithms with me and, Lucien Gomoescu, thank you for all your help troubleshooting GSA. David García-Münzer, thank you for introducing me to PBM modelling. Maria, thank you for teaching me how to work with gProms and how to properly build and run a model. Thank you for constantly telling me how great an achievement this Ph.D is. Finally, Asma, Tamador and Mark thank you very much for making BSEL a better place.

To each and every ESR in Modlife, congratulations on great projects you have all accomplished and thank you for a great environment in this ITN and for being a great crew I was always looking forward to meet every semester in a different place across Europe. Thank you Sasha Zubov for organizing everything so well.

Thank you very much Maximilian Lularevic, Vincent Wiegman and Dr. Alexandros Kyparissidis for granting access and all the help necessary for the bioprofile analysis. Thank you Jane and Jess for all the help in the use of the Imperial's flow cytometry facility.

Thank you all of you who were students in both modules I had the pleasure to teach – Advanced Bioprocess Engineering and Modelling of Biological Systems. Thank you for your kindness, words of appreciation, for trusting me for advice and, I could not help mentioning, for the SACA nomination! In particular, thank you Sophia Kotzamanidis, Ke Meng, Jerome Tan and Christopher Gozali for your effort and dedication to your projects.

This Ph.D will have been impossible without the presence of my friends. Thank you very much Pedro Couto, Pedro Pereira, Maria Braga, Rita Blanco and Marta Heleno, aci. Thank you very much Álvaro Gonzalez-Garcinuño for your devotion to the project with did together, for trusting me and for your constant true friendship. It was a privilege to work with you and to be your friend. A particular thank you to my dear friend João Figueiredo, for visiting, for an amazing fado evening that brought Portugal to London, for introducing me to the world of Portuguese guitar and for being always present and supportive.

Ana, muito obrigado por teres sido a irmã mais velha (mas só por quinze dias!) que me acolheu em Londres de portas abertas, pela ajuda toda quando, de repente, não sabia nada do sítio onde morava, pelos mimos todos, pela companhia. Enfim, por tudo.

Ao Pai e à Mãe, a quem tudo devo, uma palavra muito especial por terem estado sempre presentes, todos os dias, por terem sempre acreditado em mim, por me terem dado todas, mesmo todas, as oportunidades para que hoje esteja onde estou e porque, para vocês, nós estamos sempre primeiro e não há como agradecê-lo.

Mafalda... por onde começar... Obrigado por teres deixado tudo para trás para vires comigo para Londres e por estares sempre disposta a fazê-lo de novo. Obrigado por teres tratado de tudo para eu poder estar focado no meu trabalho: pelas noitadas, pelos fins-de-semana sem me veres, pelos dias que passaste ao meu lado no Imperial, por teres tratado de mim e por vires ter comigo todos os dias com um sorriso que me dá sempre força para continuar. Sem ti, nunca teria conseguido.

Contents

List of Figures **9**
List of Tables **15**
List of Publications **16**
1 Introduction **17**
1.1 Aims and Objectives 19
2 Literature Review **21**
2.1 Monoclonal Antibodies 21
2.2 Cell Culture for mAb Production 23
2.2 Mathematical Modelling of Cell Culture 47
2.3 Model Development 54
2.4 Multivariate Analysis of Cell Culture 59
3 Materials and Methods **61**
3.1 Culture of GS-NS0 and GS-CHO cells 61
3.2 Assays 63
3.3 Computational Methods 72
4 A predictive mathematical model of cell cycle, metabolism and apoptosis of GS-NS0 cells. **75**
4.1 Introduction 75
4.2 Objectives 78
4.3 Model Structure 79
4.4 Global Sensitivity Analysis 93
4.5 Parameter Estimation 98
4.6 Results and Discussion 100
4.7 Conclusions 118
5 Mapping the growth and metabolism of GS-CHO cells **121**
5.1 Introduction 121
5.2 Objectives 123
5.3 Methodology 123
5.4 Results and Discussion 125
5.5 Conclusions 152
6 An integrated mathematical model of GS-CHO cell culture **155**
6.1 Introduction 155
6.2 Objectives 157
6.3 Model Development 157
6.4 Model Analysis 171
6.5 Global Sensitivity Analysis 172
6.6 Parameter Estimation 175

6.7 Results and Discussion 177
6.8 Conclusions 188
7 Future work 191
References 195
Appendices 215

List of Figures

Figure 1: Simplified representation of the structure of monoclonal antibodies based on IgG. 22

Figure 2: Host organisms used in biologics production and brief description of their advantages and disadvantages. 25

Figure 3: The cell cycle in mammalian cells and its division in three lumped phases based on the DNA content. 28

Figure 4: Schematic representation of the tricarboxylic acid (TCA) cycle. 31

Figure 5: Schematic representation of the integration of glycolysis, lactate production, the TCA cycle and the oxidative phosphorylation chain (only 4 acceptors included for simplicity). NADH produced in the cytoplasm has to be transported into the mitochondrion which may require ATP consumption. NADH and $FADH_2$ yield electrons to a series of acceptors and ultimately to molecular oxygen forming water. The electron gradient is compensated by ATPase which produces ATP by pumping protons. 32

Figure 6: The Pentose Phosphate Pathway 34

Figure 7: TCA cycle intermediates can be originated from amino acids (green rectangles). Adapted from McAtee Pereira *et al.* (McAtee Pereira *et al.*, 2018). 35

Figure 8: Five main apoptosis pathways have been identified: (1) the extrinsic pathway activated by external stimuli (Ichim and Tait, 2016); (2) the internal pathway activated by MOMP promoted by Bax channels (Taylor, Cullen and Martin, 2008); (3) the DNA-damage-induced caspase-2-dependent pathway (Vakifahmetoglu-Norberg and Zhivotovsky, 2010); (4) the Granzyme B induced pathway where this molecule acts as a caspase-3 inducer (Taylor, Cullen and Martin, 2008); and (5) the Granzyme A induced pathway where no caspases are involved – Granzyme A induces a DNase, which induces DNA fragmentation and cell death (Orrenius, Zhivotovsky and Nicotera, 2003). 40

Figure 9: Suspension cultures across bioprocess development, commonly used reactors, goals, scales, mode of operation and relevant parameters (adapted from (Grilo and Mantalaris, 2019a) 45

Figure 10: Classification of biological systems models according to the framework suggested by Fredrickson *et al.* (Fredrickson and Tsuchiya, 1963) 49

Figure 11: Framework for mathematical model development (adapted from (Kontoravdi, Asprey, *et al.*, 2007)) 54

Figure 12: Flow cytometry work flow. Cells were gated from debris (A), singlets were gated from doublets (B) and viable cells were gated from apoptotic cells (C). FMO controls were used to gate EdU, cyclins and Ki-67 positive regions (D) and the DNA was deconvoluted using the software (E). 69

Figure 13: Use of amino acids in the TCA cycle and their consumption/production to/from other nutrients as considered in the metabolic model. 87

Figure 14: Flow of information between cell-cycle, metabolism and apoptosis sub-models. 91

Figure 15: Results of the conservation study. The number of bins in each cyclin domain was varied between 50 and 300 while keeping the DNA domain at 20 bins. The red dashed line shows the 99% threshold. 92

Figure 16: GSA simulations for viable cells (A), fraction of cells in phase G_0/G_1 (B), and the concentration of ATP (C) and mAb (D). 94

Figure 17: Global sensitivity analysis results obtained at 48, 72 and 120h of batch culture. The effect of each parameter on viable cells (X_V), fraction of cells in phase G_0/G_1 (fG), apoptotic cells (X_A), ATP, mAb, bcl2 transcripts (mRNAbcl2) and glutamate is quantified based on their Sobol sensitivity indices. A minimum threshold of 10% has been set to discriminate the significant parameters. Important parameters are shaded in light grey, dark grey and black for SIs of 0.1-0.3, 0.3-0.5 and >0.5, respectively. 96

Figure 18: Model-aided experimental development workflow. GSA is performed on a mathematical model to determine the important parameters. This gives important information on which parameters should be estimated in the laboratory (if possible), experiment duration etc. 97

Figure 19: Evolution of cell cycle distribution (G_0/G_1 in green, S in blue and G_2/M in pink) in EdU-positive (A) and EdU-negative (B) populations during 24 hours of culture in EdU-free medium after 3h exposure to 30µM EdU (*p-$value$<0.05, one-way ANOVA). 101

Figure 20: Experimental data (dots) and model predictions (lines) for viable (A), apoptotic (B) and dead (C) cell densities obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 101

Figure 21: Distributions of viable cells across phases G_0/G_1 (blue circles, solid lines), S (pink squares, dotted lines) and G_2/M (green circles, dashed lines) observed for batch cultures of GS-NS0 cells starting at 1x10^{5}cells/mL (A), 2x10^{5}cells/mL (B) and 4x10^{5}cells/mL (C) inoculation densities. 102

Figure 22: Experimental data (dots) and model predictions (lines) for mAb production in GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, full lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 103

Figure 23: Experimental data (dots) and model predictions (lines) for glucose (A) and alanine (B) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 104

Figure 24: Experimental data (dots) and model predictions (lines) for (A) glutamate (closed symbols) and arginine, (B) phenylalanine (closed symbols) and tyrosine (open symbols) and (C) proline obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines). 105

Figure 25: Experimental data (dots) and model predictions (lines) for aspartate (open symbols) and asparagine (closed symbols) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 106

Figure 26: Experimental data (dots) and model predictions (lines) for serine (closed symbols, A), glycine (open symbols, A) and cysteine (B) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 106

Figure 27: Experimental data (dots) and model predictions (lines) for threonine (closed symbols, A), tryptophan (open symbols, A), leucine (close symbol, B), histidine (open symbols, B), valine (closed symbols, C), methionine (open symbols, B), isoleucine (closed symbols, D) and lysine (open symbols, D) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 107

Figure 28: A: Experimental data (dots) and model predictions (lines) of ATP concentration in GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, full lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. B: Normalized ATP content (in fmol per viable cell) obtained during 6 days of batch cultures with inoculation densities of 1 (green), 2 (blue) and 4x10^5 cells/mL (pink) (*p-$value$<0.05, one-way ANOVA). 108

Figure 29: Experimental data (dots) and model predictions (lines) of caspase-3 (A), caspase-8 (B), *bcl-2* (C) and *bax* (D) transcripts in GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, full lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. 110

Figure 30: Hierarchical clustering obtained with profiles of 23 metabolites (20 amino acids, glucose, lactate and ATP) measured in the course of batch cultures of GS-NS0 cells at three different inoculation densities are superimposed to the time profiles of growth (left column panels - viable cells in green circles, apoptotic cells in pink diamonds and dead cells in blue squares); specific mAb production rate (central panels) and cell cycle distributions (right column panels - G_0/G_1 in blue circles, pink S in pink diamonds and G_2/M in green squares). 114

Figure 31: Hierarchical clustering (A), heat map (B) and PCA (C) of metabolic profiles obtained from all GS-NS0 cultures. A 2D PCA plot based on PC1 and PC2 only is also shown (D). Samples in hierarchical clustering are identified by culture time (t=0-144h) and density level (D1, D2 and D3 respectively for inoculation densities of 1,2 and 4x10^5 cells/mL). The last time point of the densest culture (144h, D3) is an outlier confirmed by all techniques. The two clusters identified at the maximum distance in the dendrogram (A) correspond to groups above and below the black dashed bisector line in the 2D PCA plot (D) or the hyperplane in the 3D PCA plot (E). The purple dashed line in the 2D PCA plot (D) and the hyperplane in 3D (F) separate productive and non-productive clusters although this separation cannot be identified from hierarchical clustering (A). 117

Figure 32: Viable cell density (A), viability (B) and mAb titre (C) measured in the course of 240h of GS-CHO cells batch cultures inoculated at 1 (green), 2 (blue) and 4x10^5 cells/mL (red). All cultures were conducted in triplicates (N=3 biological replicates, error bars show one standard deviation). 125

Figure 33: Multivariate metabolic profile analysis. Glucose, lactate, ammonia, 20 amino acids and ATP profiles obtained in the course of 10-day batch cultures of GS-CHO cells were analysed by hierarchical clustering (A), the corresponding heat map (B) and

Principal Component Analysis (C). A 2D representation of PCA using only PC1 and PC2 is shown in (D). Samples are identified by time (0-240h) and density level (D1, D2 and D3 for 1, 2 and 4x10^{5}cells/mL). Clusters separated at the longest distance in the dendrogram (A) can be distinguished by a hyperplane in the PCA space (E). Productive and non-productive clusters can be distinguished by drawing a hyperplane in the 3D PCA space (F) or the bisector line in 2D PCA plot (D) and confirmed by specific mAb productivities (G) obtained for inoculation levels (i) D1, (ii) D2 and (iii) D3. Colours in panels C-G refer to clusters identified in panel A. 127

Figure 34: Dynamic profiles of (A) glucose (closed circles) and lactate (open circles); (B) asparagine (closed circles), aspartate (open circles), glutamate (closed diamonds), glutamine (open diamonds) and isoleucine (close squares); (C) alanine (open squares) and ammonia (open diamonds); (D) serine (closed circles), glycine (open circles) and threonine (closed diamonds). Y-axis for plots B-D is the same. (E) Suggested metabolic network for the first 72h of culture. Alternative routes for glutamate to α-ketoglutarate conversion are shown in purple dashed lines (inoculation at 2x10^5 cells/mL; N=3 biological replicates; error bars show 1 standard deviation). 133

Figure 35: Dynamic profiles in panels (A)-(D) show the same nutrients as in Figure 34. Panel (E) shows arginine (closed triangles), histidine (open triangles) and leucine (closed squares). The transition into exponential growth is shown by a dashed green line in all dynamic profiles. A possible metabolic network and changes related to the transition to exponential growth is suggested (F). 135

Figure 36: Dynamic profiles of (A) glucose, lactate, (B) asparagine, aspartate, glutamate, glutamine, isoleucine, (C) alanine, ammonia, (D) serine, glycine, threonine, (E) arginine, histidine, leucine (key as in Figure 34 and Figure 35) and valine (open squares) whose consumption starts during stationary phase. The transition into exponential growth and stationary phase are shown by green and yellow dashed line, respectively, in all dynamic profiles. A map of metabolism changes related to the transition to the stationary phase is suggested (F). 138

Figure 37: ATP concentrations over 240h of GS-CHO cultures inoculated at 1, 2 and 4x10^5 cells/mL respectively in green, blue and red (A) and the corresponding normalised values (B) which were analysed by hierarchical clustering (C). Time points are labelled according to time (t0-t240h) and inoculation density level (D1-D3). Time points where the clustering does not correspond to the culture stage where they were collected are marked with red ovals. There are also two outliers on the right-most side of panel C. The distributions of viabilities of samples in the "Late-stage" (brown) and "Early-stage" (purple) clusters are shown in a box plot (D) where the in the "Early-stage" cluster is confirmed. (In panels A and B, N=3 biological replicates, error bars show one standard deviation). 140

Figure 38: Concentrations of ATP (filled circles, full lines) and lactate (empty circles, dashed lines) obtained in the course of 240h GS-CHO cultures inoculated at (A) 1, (B) 2 and (C) 4x10^5 cells/mL and correlation of ATP and lactate contents obtained for all cultures (D) where points relating to plots A, B and C are shown in green, blue and red, respectively. (N=3 biological replicates, error bars show one standard deviation). 142

Figure 39: Specific ATP contents obtained over 240h of GS-CHO cultures inoculated at 1, 2 and 4x10^5 cells/mL respectively in green, blue and pink columns. The red line and

the shaded region show the critical ATP content of 4.14±1.2 fmol ATP/cell estimated in this work. Columns with diagonal red stripes correspond to samples classified as "Decline phase" in the metabolic profile analysis of Figure 33A. (N=3 biological replicates, error bars show one standard deviation).....143

Figure 40: Cell cycle distributions of EdU-positive (A) and EdU-negative (B) populations over 30 hours of GS-CHO cell cultures inoculated at $2x10^5$ cells/mL (N=3 biological replicates, error bars show one standard deviation, $^*p\text{-}value<0.05$ one-way ANOVA)...144

Figure 41: Expression of cyclins E1 (blue, normalized to phase G_2/M), B1 (green, normalized to phase G_1) and D1 (orange, normalized to phase S) in GS-CHO cell cultures inoculated at $2x10^5$ cells/mL in different phases of the cell cycle after 48h of batch culture (N=3 biological replicates, error bars show one standard deviation, $^*p\text{-}value<0.05$ one-way ANOVA).....145

Figure 42: Expression of cyclins E1 (A-B) and B1 (C-D) for EdU positive (A, C) and EdU negative (B, D) populations. Lighter colours show expression in phase S. 147

Figure 43: Expressions of cyclins E1 (blue) and B1 (green) at 48 (early exponential growth), 120 (exponential growth), 144 and 168h (stationary phase) of culture. Growth stages refer to in Figure 33A (N=3 biological replicates, error bars show one standard deviation, $^*p\text{-}value<0.05$ one-way ANOVA). 148

Figure 44: Expression of Ki-67 (open diamonds) and culture viability (closed circles) over 240h of batch cultures of GS-CHO cells inoculated at $2x10^5$ cells/mL. Growth stages refer to in Figure 33A (N=3 biological replicates, error bars show one standard deviation)..150

Figure 45: (A) Apoptotic (empty circles) and dead (closed circles) cell densities and (B) expression of caspase-3 (red), caspase-8 (blue), bcl-2 (green) and bax (grey) in the course of 240h of GS-CHO batch cultures inoculated at $2x10^5$ cells/mL. Growth stages refer to Figure 33A. 152

Figure 46: Asparagine, aspartate, glutamate, and alanine metabolic interplay.....163

Figure 47: Conservation study results. In both cyclin domains the number of bins was varied from 50 to 300 while keeping 20 bins in the DNA domain. The red dashed line shows 99% conservation. 172

Figure 48: Global Sensitivity Analysis results. The significance of each parameter was assessed using the Sobol method at 60, 160 and 200h on viable cells (xV), fraction of cells in phase G1 (Fg), apoptotic cells (xA), energy production (ATP), monoclonal antibody (mAb), caspase-8 transcripts (casp-8) and aspartate concentration (Asp). A minimum threshold of 10% was set to discriminate important parameters. Significant parameters are shaded according to their SIs in light grey (0.1 – 0.3), dark grey (0.3 – 0.5) or black (>0.5).....174

Figure 49. Model predictions (lines) and experimental data (dots, N=3 biological replicates, error bars show one standard deviation) of viable (A), apoptotic (B) and dead (C) cell densities obtained in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and $4x10^8$ cells/L (red diamonds and dashed lines). 178

Figure 50. Cell cycle predictions (lines) and experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for phases G_1 (blue circles, solid lines),

S (red squares and dotted lines) and G_2/M (green diamonds and dashed line) obtained in GS-CHO batch cultures inoculated at 2 (A), 1 (A) and 4x10^8 cells/L (C)..... 178

Figure 51. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for mAb production in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 179

Figure 52. Model predictions (lines) and experimental data (dots, N=3 biological replicates, error bars show 1 standard deviation) of glucose (A), lactate (B) and alanine (C) profiles obtained in the course of GS-CHO batch cultures inoculated at 1 (green squares, dotted lines), 2 (blue circles and solid lines) and 4x108cells/mL (red diamonds and dashed lines)..... 180

Figure 53. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for ammonia production in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 180

Figure 54. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) asparagine (solid symbols, dotted lines) and aspartate (empty symbols and solid lines) and (B) glutamate metabolism in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 182

Figure 55. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) arginine (solid symbols, dotted lines) and glutamine (empty symbols and solid lines), (B) proline (solid symbols, dotted lines) and phenylalanine (empty symbols and solid lines) and (C) tyrosine (solid symbols, dotted lines) and cysteine (empty symbols and solid lines) metabolism in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 183

Figure 56. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) serine (solid symbols, dotted lines) and glycine (empty symbols and solid lines), (B) threonine metabolism in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 184

Figure 57: Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) leucine (solid symbols, dotted lines) and histidine (empty symbols and solid lines), (B) isoleucine (solid symbols, dotted lines) and tryptophan (empty symbols and solid lines), (C) lysine (solid symbols, dotted lines) and valine (empty symbols and solid lines) and (D) methionine (solid symbols, dotted lines) in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 185

Figure 58. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for ATP production in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines)..... 186

Figure 59. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for the expression of (A) caspase-8, (B) bax, (C) caspase-3 and (D) bcl-2 in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and $4x10^8$ cells/L (red diamonds and dashed lines). 187

Figure 60. Bioprocess optimization route combining multivariate analysis, steady state techniques and kinetic models. 191

Figure 61: Model based control of product quality based on apoptosis reversal. 193

Figure 62: Integration of biological models and CFD model for integrated bioreactor optimization 193

List of Tables

Table 1. Flow cytometry panels used for cell cycle, EdU, cyclins and Ki-67 analyses 68

Table 2. FMO controls and sample configurations for which they were used. 69

Table 3. Primers used for q-RT-PCR and references of their prior use. 71

Table 4: Estimated values of significant model parameters. Final values and 95% confidence intervals obtained using the MLE algorithm and experimental data obtained in this work. 99

Table 5: Parameter estimation results. Significant model parameters were re-estimated with experimental data obtained in this work and the MLE algorithm implemented in gProms. 176

List of Publications

Journal Publications

1. Grilo, A.L. & Mantalaris, A. "Apoptosis: a mammalian cell bioprocessing perspective" (2019) *Biotechnology Advances* vol. 37, 459-475

2. Grilo, A.L. & Mantalaris, A. "The increasingly human and profitable monoclonal antibody market" *Trends in Biotechnology* (2019) vol. 37, 1, 9-16.

3. Grilo, A.L. & Mantalaris, A. "A predictive mathematical model of cell-cycle, metabolism and apoptosis of mAb-producing GS-NS0 cells". *Biotechnol. J.* (submitted)[1].

4. Grilo, A.L.; Gonzalez-Garcinuño, A.; Martin Del Valle, E.; Mantalaris, A. "Mathematical modelling of fluid flow and mass transfer in a three-column downstream process for monoclonal antibodies purification" *Ind. Eng. Chem. Res.* (submitted).

5. Grilo, A.L; Mantalaris, A.; "Multivariate exo-metabolome analysis: a tool for cell-culture understanding and optimization" (manuscript in preparation)[2]

Conference Oral and Poster Communications

6. Grilo, A. L., Meng, K., Ong, T. Mantalaris, A.: "*Biological kinetic coupled with CFD simulations: towards rational bioreactor design".* **12th European Symposium in Biochemical Engineering Sciences,** Lisbon, Portugal, 2018. Oral communication

7. Grilo, A. L., Quiroga-Campano, A.; Mantalaris, A.: *"A predictive model of cell-cycle, metabolism and apoptosis of mAb-producing GS-NS0 cells"* **XVI Cell Culture Engineering Conference**, Tampa, FL, 2018. Poster communication[3].

8. Grilo, A. L., Gonzalez, A. Quiroga-Campano, A.; Mantalaris, A.: *"Model-based process development"* **Commercializing Continuous Processing in Pharma Summit,** Boston, MA, 2018. Oral communication.

[1] This manuscript is based on part of the work presented in Chapter 4 of the present dissertation
[2] This manuscript is based on work presented in Chapters 4 and 5 of the present dissertation
[3] This communication included part of the work presented in Chapter 4 of the present dissertation

Chapter 1

Introduction

The use of process systems engineering (PSE) tools for the understanding, modelling and optimization of biological processes has been discussed extensively in literature (Fussenegger, Bailey and Varner, 2000; Sidoli, Mantalaris and Asprey, 2005; Kiparissides *et al.*, 2009, 2011; Kontoravdi, Pistikopoulos and Mantalaris, 2010; Rogers and Ierapetritou, 2015; Hong, Lakshmanan, *et al.*, 2018; Kontoravdi and Jimenez del Val, 2018). Although the biochemical process industries still have not implemented these tools and technologies to the same extent as other chemical industries (Rogers and Ierapetritou, 2015), research developments achieved over the past decades have generated tools able to satisfy important needs of this industry. Among those, are process optimization (Xie and Wang, 1994; Kiparissides, Pistikopoulos and Mantalaris, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018), improved process understanding (Carinhas *et al.*, 2013; Templeton *et al.*, 2017; Haringa *et al.*, 2018; Kontoravdi and Jimenez del Val, 2018), the development of process control strategies (Steinebach *et al.*, 2016) or integrated models of entire plants for the purpose of estimating costs, optimizing the overall process or understanding the environmental impact (Pollock, Ho and Farid, 2013; Grilo *et al.*, 2017; Yang, Qadan and Ierapetritou, 2019).

Among the modelling techniques often used for bioprocess systems are steady-state models (Carinhas *et al.*, 2013; Templeton *et al.*, 2017) or kinetic models (Kontoravdi, Wong, *et al.*, 2007; Quiroga-Campano, Panoskaltsis and Mantalaris,

2018) for cell metabolism studies, detailed models of important cellular functions such as the cell cycle (Martens *et al.*, 1995; García-Münzer, Kostoglou, *et al.*, 2015), apoptosis (Fussenegger, Bailey and Varner, 2000; Naderi *et al.*, 2011), protein glycosylation (Del Val, Polizzi and Kontoravdi, 2016), detailed bioreactor models (Haringa *et al.*, 2018) or even multi-unit operation models (Canales *et al.*, 2010; Freydell *et al.*, 2011; Grilo *et al.*, 2017). Owing to the inherently dynamic nature of biological processes, mathematical models able to describe them in a dynamic manner, such as kinetic mathematical models, are generally preferred (Kyriakopoulos, Ang, Lakshmanan, Huang, Yoon, Gunawan and Lee., 2018) while steady state models are, by design, unable to predict dynamic cellular activities (Robitaille, Chen and Jolicoeur, 2015).

In most studies mentioned above, the different biological processes and/or the different tools available are, in most cases, used isolated from each other. However, future applications will profit from integrating different biological processes and different tools available to optimize processes (Carinhas *et al.*, 2013; Galbraith *et al.*, 2018; Hong, Lakshmanan, *et al.*, 2018). In the multi-billion dollar biologics industry, any advancement made in improving process understanding or optimizing yield and product quality means significant reductions in development costs and increased earnings from shortening time-to-clinic and, ultimately, time-to-market.

In this work, a contribution towards model-based process development applications supported by integrating knowledge pertaining to different biological functions and/or different experimental/modelling tools is provided, using the monoclonal antibodies industry as a case study. The present thesis has four sections comprising Introduction, Literature Review and Materials/Methods (Chapters 1-3); development of an integrated predictive mathematical model of GS-NS0 cell culture and exploration of unsupervised methodologies for cell culture analysis (Chapter 4); mapping the growth and metabolism of a different cell line (GS-CHO) using multivariate analysis of metabolic data and identifying suitable biomarkers

(Chapter 5); use the information acquired in previous chapters to develop an integrated mathematical model of GS-CHO (Chapter 6) and, lastly, identification of possible future research lines (Chapter 7).

1.1 AIMS AND OBJECTIVES

The following aims and objectives have been set for the present work:

1. **Reviewing relevant scientific contributions (Chapter 2).**

   State-of-the art literature on upstream process development for monoclonal antibody processes; host organisms, specially mammalian cells and their biology; important optimization goals and dynamic modelling techniques available is critically reviewed;

2. **Description of the methodologies used (Chapter 3).**

   The methodology used to achieve the goals set in this work is described comprising experimental methods and materials and computational approaches and frameworks used for model development.

3. **Establishing a methodology for the development of mathematical models describing several cellular functions (Chapter 4).**

   Three important biological functions – cell cycle, apoptosis, and metabolism – are integrated in one kinetic mathematical model of mAb producing GS-NS0 cells. The model structure is analysed using global sensitivity analysis, allowing to identify the relevant parameters, which are re-estimated with experimental data. Two independent data-sets are used for model validation. Finally, multivariate analysis of metabolic profiles is attempted as an unsupervised methodology to identify the time horizons of different growth stages, distinguish productive from

non-productive clusters and identify important metabolic shifts related to different growth stages.

4. **Mapping the metabolism and growth of a different cell line, GS-CHO (Chapter 5).**

   Multivariate data analysis methodologies developed in chapter 5 are used in conjunction with other experimental techniques, namely flow cytometry and q-RT-PCR to obtain an integrated view of metabolism and growth of GS-CHO cell cultures. Metabolism is used to establish time horizons of growth stages in culture and identify important metabolic shifts and how they affect growth; minimum energy requirements are established; cell cycle biomarkers are identified, and their expression related to metabolism; expression of apoptosis-related genes is measured.

5. **Developing an integrated mathematical model of GS-CHO cells. (Chapter 6).**

   The model structure developed in Chapter 4 and the information acquired in Chapter 5 are used to develop a kinetic integrated mathematical model of a different cell line (GS-CHO).

6. **Identification and discussion of future directions (Chapter 7).**

   Future directions for the present research considering improvements and applications of this work are identified and discussed.

Chapter 2

Literature Review

The state-of-the art of research in mAb bioprocessing from both experimental and modelling standpoints is reviewed in this section. Relevant aspects of upstream bioprocessing are discussed before explaining modelling approaches and methodologies.

2.1 MONOCLONAL ANTIBODIES

Monoclonal antibodies (mAbs) are blockbuster drugs used in several clinical indications including treatment of liquid and solid tumours, dermatology conditions and many others. The production of therapeutic monoclonal antibodies has witnessed an unprecedented growth having reached a market size of $98 billion in 2017 (Grilo and Mantalaris, 2019b), which excludes mAbs used in diagnostic and research applications. Structurally, mAbs are Y-shaped glycoproteins used by the immune system to mark antigens for destruction. They consist of two antigen-binding sites called fragment antigen binding (Fab) and one fragment crystallisable (Fc) regions. Each antibody is composed of two light chains and two heavy chains. Heavy chains have a molecular weight of 50-70 kDa and each may comprise of 4 to 5 domains in total, one of which being a variable domain (V_H) and the others being conserved (C_H). The light chains (25 kDa) are composed of one variable (V_L) and one constant (C_L) domain (Schroeder and Cavacini, 2010). The Fab is made of two light-chain and two heavy-chain domains (one variable and one

constant in each chain). The remaining constant domains of the heavy chains make up the Fc region. There are 5 types of heavy chains (α, ε, γ, δ, and μ) and hence immunoglobulin (Ig) classes A, E, G, D and M, respectively. Additionally, there are two types of light chains (κ and λ). The light chains are disulphide-bonded to the heavy chains and the two heavy chains are also disulphide-bonded to each other (Schroeder and Cavacini, 2010). The integrity of these bonds is a product quality criteria in industrial processes (Rogers *et al.*, 2015). These structural aspects are depicted in Figure 1.

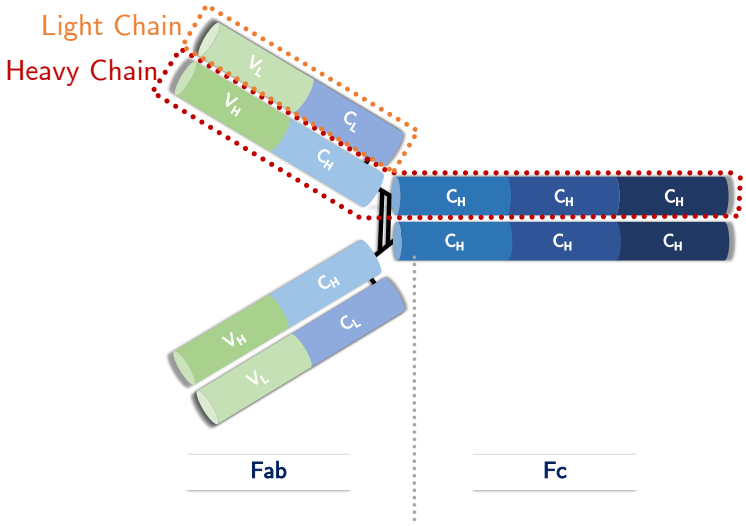


Figure 1: Simplified representation of the structure of monoclonal antibodies based on IgG.

Most antibodies found in animal sera target different receptors (epitopes) of the same antigen, hence being called polyclonal mixtures. This is undesirable for a therapeutic application where the structure and mode of action of the molecules used must be well established. These requirements are fulfilled by monoclonal antibodies which target one specific epitope of the antigen. This requires collecting only the antibodies produced by one specific B-cell. In order to satisfy market demand, culturing mAb-producing cells in large bioreactors is required.

2.2 CELL CULTURE FOR MAB PRODUCTION

Culturing B-cells renders difficult due to their short lifetimes (Fulcher and Basten, 1997). However, a major breakthrough was made when the hybridoma technology was developed providing hybrid cells obtained by fusing a B-cell with an immortal myeloma cell. The first commercial monoclonal antibody produced for therapeutic applications using the hybridoma technology was described in a scientific article in 1979 (Kung *et al.*, 1979) and approved for human therapy by the Food and Drug Administration (FDA) of the United States of America in 1986 (Waldmann, 2003). This technology used murine cells but murine mAbs posed immunogenicity concerns, namely related to the Human Anti Mouse Antibody response (HAMA) and alternative technologies were sought. Fully human monoclonal antibodies are now predominant in the market (Grilo and Mantalaris, 2019b) but before this could be possible, chimeric (34% murine content) and humanized (5-10% murine content) mAbs (Lynch and Yang, 2002; Kunert and Reinhart, 2016) were developed. Essential for the desired therapeutic action of mAbs and their residence time in the human body is ensuring the appropriate glycosylation profile (Higel *et al.*, 2016). The unique ability of mammalian cell lines to deliver mAbs with human-like glycosylation patterns has made them preferential host systems for the heterologous expression of mAbs and indeed the only class of hosts used in manufacturing of full length therapeutic mAbs (Grilo and Mantalaris, 2019b).

2.2.1 HOST ORGANISMS FOR MAB PRODUCTION

There are a number of host systems available which have been used for the production of recombinant biopharmaceuticals including mammalian cells, bacteria, yeasts, insect cells, plant cells and transgenic animals (Lalonde and Durocher, 2017). *Escherichia coli* was the preferred host organism until mid-1990s when it was first overtaken by mammalian cells (Ferrer-Miralles *et al.*, 2009). The latter have been used since mid- to late 1980s and from mid-1990s to mid-2000s competed

fiercely with *E.coli* in the market. Since then, mammalian cells have been preferred (Ferrer-Miralles *et al.*, 2009; Lalonde and Durocher, 2017). Insect and plant cells have also been attempted but different glycosylation machinery, and hence production of different glycoforms, have hindered their use in mAbs bioprocessing, although they are used to produce vaccines and recombinant proteins (Lalonde and Durocher, 2017). At present, only three therapeutic proteins are produced in transgenic animals (Lalonde and Durocher, 2017). For non-mammalian hosts, modifications to wild-type strains may enable their use in mAb bioprocessing. Bacterial hosts have been used to produce mAb fragments (Fab) and other non-glycosylated proteins, at commercial scale (Lalonde and Durocher, 2017; Grilo and Mantalaris, 2019b), yeast hosts, chiefly *Pischia pastoris,* modified to perform glycosylation, have been shown able to produce full-length fully functional mAbs (Potgieter *et al.*, 2009) and insect cells with human glycosylation machinery have also been studied (Dumont *et al.*, 2016).

2.2.2 MAMMALIAN CELLS

Therapeutic full length mAbs are produced in Chinese Hamster Ovary cells (CHO) or murine myelomas (NS0 and Sp2/0) (Grilo and Mantalaris, 2019b) but baby hamster kidney (BHK), human embryonic kidney (HEK293), and human retina (PER.C6) cell lines are used to produce other biologics (Dumont *et al.*, 2016).

CHO cells are generally preferred for mAb manufacturing (Kunert and Reinhart, 2016; Grilo and Mantalaris, 2019b) due to their ability to grow robustly and stably in serum-free suspension cultures and are regarded as safe in terms of human pathogenic virus replication (Lalonde and Durocher, 2017). Further, they are able to express human-like patterns of post-translational modifications (Ha and Lee, 2015). Unlike NS0 cell lines, for example, CHO are known to incorporate low levels of NGNA (N-glycolyl neuraminic acid), a sialic acid not synthesized by humans, in the glycosylated proteins they produce (Li *et al.*, 2010). This has tempered immunogenicity concerns that rose previously regarding NS0 cultures

which produce the alpha-Gal-alpha(1,3)Gal linkage (Jacobi *et al.*, 2014) and Neu5Gc (Lalonde and Durocher, 2017), immunogenic for humans. CHO cells are also less susceptive to viral infections, as they lack viral entry genes (Dumont *et al.*, 2016). These concerns have limited the application of NS0 cells in favour of CHO cell lines. Since the CHO genome was sequenced (for CHO-K1 cells) in 2011, and -omics data have been made public between 2011 and 2014, biotechnology applications development has further accelerated (Fischer, Handrick and Otte, 2015). The information availability and the capabilities of synthetizing mAbs with the right glycosylation patterns are among the main arguments in favour of mammalian cells, and particularly of CHO (Figure 2).

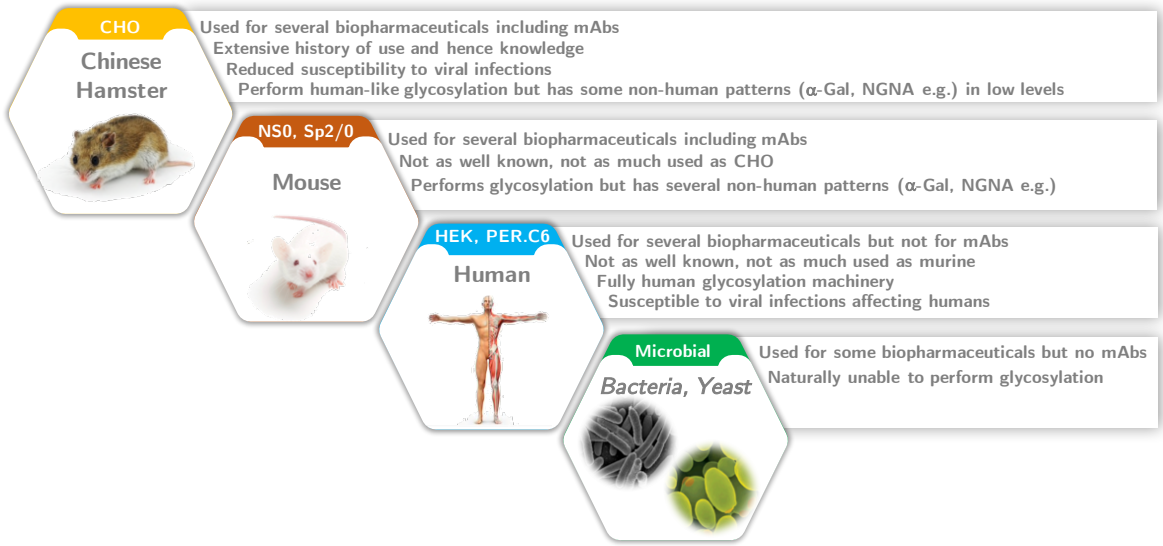


Figure 2: Host organisms used in biologics production and brief description of their advantages and disadvantages.

Glycosylation has been widely recognized as one of the most difficult critical quality attributes (CQAs) to control in therapeutic glycoprotein production (Kontoravdi and Jimenez del Val, 2018) and an essential feature to ensure the desired therapeutic action and serum half-life while avoiding undesirable immune responses (Lalonde and Durocher, 2017; Kontoravdi and Jimenez del Val, 2018).

Several expression systems to select high-producing clones have been developed. The two most widely used ones are the DHFR (DiHydroFolate

Reductase) system and the GS (Glutamine Synthetase) system both developed in the early 1990s (Butler and Jenkins, 1989; Bebbington *et al.*, 1992; Wurm, 2004; Lalonde and Durocher, 2017). The DHFR system is based on inhibiting endogenous dihydrofolate reductase activity with MTX (methotrexate) preventing nucleotide synthesis and selecting clones in nucleotide-free medium. The GS system is based on inhibiting endogenous glutamine synthetase activity with glutamate analogue methionine sulfoximine (MSX) and selecting clones in glutamine-free medium (Lalonde and Durocher, 2017). The DHFR system has been reported to lead to longer development times (Barnes, Bentley and Dickson, 2000) because it requires a larger number of amplification rounds before reaching levels of expression similar to those obtained using the GS system, which has been reported to require a single amplification round (Bebbington *et al.*, 1992). However, the application of this system to CHO cells requires higher inhibitor concentrations due to their higher native levels of the GS enzyme compared to NS0 or Sp2/0 cells (Bebbington *et al.*, 1992; Barnes, Bentley and Dickson, 2000). Nevertheless, a GS-CHO-K1 clone with an 18% increase in recombinant protein production and only one fourth of the ammonia production rate has been developed (Ha and Lee, 2015). Difficulties in producing GS-CHO cell lines can be overcome using GS-knockout CHO cells, an approach attempted by pharmaceutical company Eli Lilly (Lalonde and Durocher, 2017). This approach has enabled the use of significantly lower MSX concentrations for the development of high producing GS-CHO clones (Noh, Shin and Lee, 2018). This is consistent with prior reports of concentrations of up to 500µM being required for successful clone development in CHO cells (Bebbington *et al.*, 1992).

2.1.1.1 The Cell Cycle

The cell cycle is the process of cell duplication. In eukaryotic cells, the cycle includes growth, DNA replication, distribution of the chromosomes and cell division, giving birth to two genetically identical daughter cells. These steps occur in five highly-regulated discrete phases (Cooper and Hausman, 2007) where cell

machinery will ensure that all events occur smoothly and evaluate the external signals that promote or inhibit cell growth. It is not surprising, therefore, that one cycle may take several hours to complete in mammalian cells. After birth, cells may enter a non-proliferative (quiescent) stage of the cycle, the G_0 phase, resuming proliferation if and when the conditions are favourable. The process starts with the first gap phase (G_1 phase) where the cell grows under normal physiological activity. During this phase, RNA and proteins are synthetized. If, at any point, the conditions become adverse, cells will revert to the G_0 phase where they can remain until proliferation resumes again or the cell dies (Nelson and Cox, 2013; García-Münzer, 2014). The following phase is characterized by intense synthetic activity, hence the name Synthesis (S phase). DNA is duplicated during this phase and RNA/protein synthesis continues. This phase of the cycle can last as long as 8 hours in mammalian cells (Nelson and Cox, 2013). When DNA replication is complete, cells enter another gap phase, G_2, in preparation for cell division. This phase is generally shorter taking 3 to 4 hours (Nelson and Cox, 2013) to ensure cells are ready for mitosis (division). Mitosis (M) is the last phase of the cycle, just before division, whereby each cell gives birth to two identical daughter cells. The nucleus divides and cell energy is directed towards ensuring a successful division. A reference time of 1 hour is usually considered (Nelson and Cox, 2013; García-Münzer, 2014). The transition between phases is regulated by a wide range of proteins including cyclins and cyclin dependent kinases (cdk) (Siu, Rosner and Minella, 2012). Showcasing the important role of cell cycle proteins, accumulation of cyclin E has been shown to trigger proliferation in a variety of cell lines (Darzynkiewicz *et al.*, 1996) including CHO (Renner *et al.*, 1995), NS0 (García-Münzer *et al.*, 2013) and cancer cells (Fuentes-Garí *et al.*, 2015) suggesting a conserved role for this protein. Likewise, cyclin B1 has been shown to be expressed at late stages of the S-phase and throughout the G2/M phase (Darzynkiewicz *et al.*, 1996; García-Münzer *et al.*, 2013; Fuentes-Garí *et al.*, 2015).

Based on the DNA content, it is possible to distinguish three cell cycle stages with either 1 copy of the DNA (G_0/G_1), undergoing DNA duplication (S) or with 2 DNA copies (G_2/M) (Cooper and Hausman, 2007) as shown schematically in Figure 3. This DNA-based division of the cell cycle sets the fundamentals for its flow cytometry analysis using DNA stains, such as propidium iodide (PI) (Al-Rubeai and Emery, 1993). Online cell cycle monitoring techniques are being developed and have been tested in CHO cells (Fuge *et al.*, 2017).

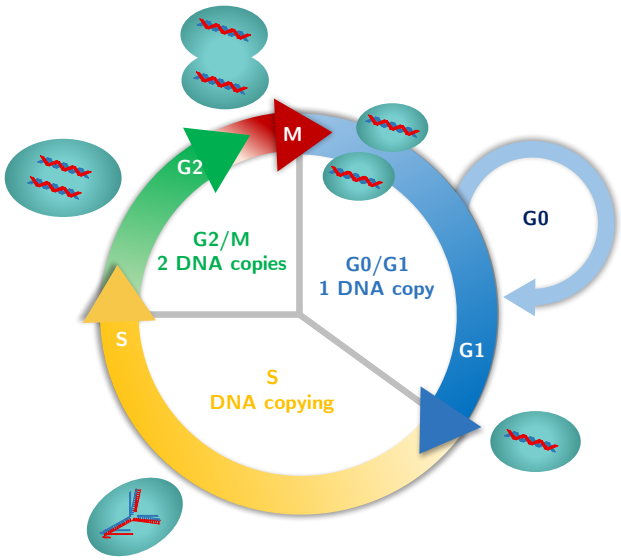


Figure 3: The cell cycle in mammalian cells and its division in three lumped phases based on the DNA content.

The cell cycle is indeed directly influenced by the environment in which cells are grown. Some nutritional stresses have been used to induce growth arrest leading to synchronised populations (cells arrested in one phase). Deprivation of some amino acids including leucine (Everhart, 1972; Ley, 1973; Xiao *et al.*, 2016) or arginine (Lamb and Wheatley, 2000) has been reported as a possible arrest strategy. At the genetic level, the overexpression of p27, a cyclin-dependent kinase inhibitor, in CHO cells, leads to G_1-phase arrest with significant productivity increases of up to 15-fold (Kaufmann *et al.*, 2001). The overexpression of p21, a cyclin E inhibitor, in NS0 cells synchronised the population in phase G_1 without changing mAb production (Khoo and Al-Rubeai, 2009). Other agents such as sodium butyrate (Hendrick *et al.*, 2001), valeric acid (Park *et al.*, 2015), thymidine (Al-Rubeai *et*

al., 1992; García Münzer *et al.*, 2014; Park *et al.*, 2015), DMSO (Ling *et al.,* 2003; García Münzer *et al.*, 2014) and several others (Park *et al.*, 2015) have also been used to arrest CHO, hybridoma and NS0 cells in phase G_1. Culturing at lower temperatures has also been attempted as a means to increase productivity by synchronising the cell population. Cultures of hybridomas at 32°C showed a 5% increase in specific productivity (Chong *et al.*, 2008) while mAb production was lower in a different hybridoma cell line cultured at 33°C (Sureshkumar and Mutharasan, 1991) and a third hybridoma cell line showed no changes in productivity at 34°C (Bloemkolk *et al.*, 1992). Shi *et al.* have reported a G_2 arrest and a positive increase in recombinant protein production in CHO cells when temperature was decreased to 30°C (Shi *et al.*, 2005) similar to what was observed by Swiderek *et al.* for NS0 cells at 22°C (Swiderek and Al-Rubeai, 2007) while Seo *et al.* working with F2N78 human cells at 33°C have not observed any increase in productivity (Seo *et al.*, 2013). Overall, these observations suggest a cell-line specific effect of temperature on cell-cycle arrest.

The cell cycle is an important biological feature for process development as it has been suggested that recombinant protein production may occur preferentially in one stage of the cycle (Suzuki and Ollis, 1989; Hendrick *et al.*, 2001; Dutton, Scharer and Moo-Young, 2006; Du *et al.*, 2015). However, the most productive phase of the cell cycle appears to be cell-line specific. While hybridomas and CHO cells have shown increased titres upon G_1 arrest (Al-Rubeai *et al.*, 1991; Park *et al.*, 2015), this has not been observed in NS0 cells (Khoo and Al-Rubeai, 2009; García-Münzer, Kostoglou, *et al.*, 2015) where mAb production has been correlated with phase G_2 (Tey and Al-Rubeai, 2004).

2.1.1.2 Mammalian Cell Metabolism

Mammalian cells require a number of nutrients including sugars, amino acids (AA), vitamins and trace elements (Galbraith *et al.*, 2018) in order to maintain viability, grow, produce recombinant proteins and the required energy for

all these processes. Energy is mainly stored in the form of Adenosine Tri-Phosphate (ATP) (Yuan, Xiong and Guan, 2013) and consumed whenever required. In a mammalian cell, there are several metabolic pathways that lead to energy (ATP) production including fermentative routes, glycolysis and the tri-carboxylic acid (TCA) cycle also known as the Krebs cycle.

Analysis of multiple process runs at Genentech has shown that glucose metabolism is linked to productivity in CHO cell cultures (Le *et al.*, 2012). In the presence of glucose, glycolysis (from Greek *glucos*=sugar and *lusis*=break) is the first metabolic pathway to be activated, comprising of a series of enzymatic reactions which eventually lead to the breakage of glucose into two 3-carbon molecules of pyruvate. The pathway includes a preparation phase and a payoff phase. During the former, glucose is degraded to glyceraldehyde 3-phosphate, co-existing in equilibrium with its isomer, dihydroxyacetone-phosphate, and consuming two molecules of ATP per molecule of glucose degraded. The payoff phase encompasses the conversion of glyceraldehyde 3-phosphate to pyruvate with the production of 4 molecules of ATP and 2 molecules of NADH (protonated form of nicotinamide adenine dinucleotide). Given that two molecules of ATP had been consumed in the first step, the overall energy yield of glycolysis is 2 molecules of ATP per molecule of glucose consumed. Pyruvate molecules formed during glycolysis can have several fates: under aerobic conditions, pyruvate is more likely to be oxidised to acetyl-CoA which can enter the TCA cycle whereas under anaerobic conditions it is more likely to be reduced to lactate in a fermentative route without any further production (or consumption) of ATP (Nelson and Cox, 2013). The conversion of pyruvate to lactate produces NAD^+, a cofactor required in glycolysis and several other metabolic reactions.

Under aerobic physiological conditions, the TCA cycle is the energy-production pathway of choice (Heiden, Cantley and Thompson, 2009; Zou and Al-Rubeai, 2016) as it yields more energy per carbon than any other. However, in cell lines used for mAb manufacturing, the glycolytic flux is higher than that of

the TCA cycle, particularly during the early stages of culture (Carinhas *et al.*, 2013; Templeton *et al.*, 2013).The TCA cycle runs inside the mitochondrion and it comprises of a series of reactions which occur in a cyclic manner to ensure storage of electrons into two redox co-factors, which will act as carriers, namely NADH and $FADH_2$ (protonated form of flavin adenine dinucleotide) (Nelson and Cox, 2013). Initially, pyruvate is transported into the mitochondrion by a transporter in the inner membrane of this organelle (Nelson and Cox, 2013). Once inside, pyruvate is decarboxylated to acetyl-CoA producing one NADH molecule. Acetyl-CoA undergoes condensation with oxaloacetate to produce citrate which can be decarboxylated to α-ketoglutarate in a series of three reactions producing one further molecule of NADH. α-Ketoglutarate can then be further decarboxylated to succinyl-CoA producing one more NADH molecule.

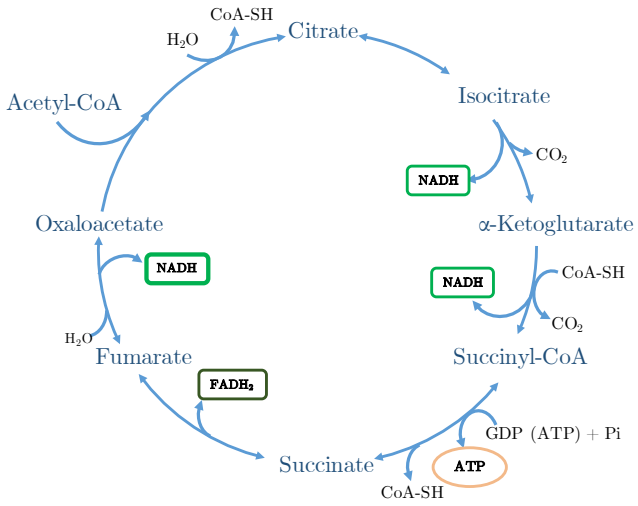


Figure 4: Schematic representation of the tricarboxylic acid (TCA) cycle.

The following reaction is succinyl-CoA's substrate-level phosphorylation to succinate, which is subsequently dehydrogenated to fumarate producing one molecule of $FADH_2$. Fumarate is hydrated to Malate which is, in its turn, dehydrogenated to oxaloacetate with production of one NADH thus regenerating the cycle. The cycle is schematically represented in figure 4.

As mentioned above, the redox co-factors produced in glycolysis and the TCA cycle (namely NADH and $FADH_2$) play the role of electron carriers. Most co-factors will be inside the mitochondrion already, apart from those formed during glycolysis. Those must be transported into the organelle if they are to be used to produce energy and the transport system used may be energy-dependent. The co-factors undergo a series of redox reactions through which electrons will be transferred in stages to a series of intermediate acceptors/donors which will, ultimately, yield the electrons to oxygen, the final acceptor. This is called the oxidative phosphorylation chain (OxPhos).

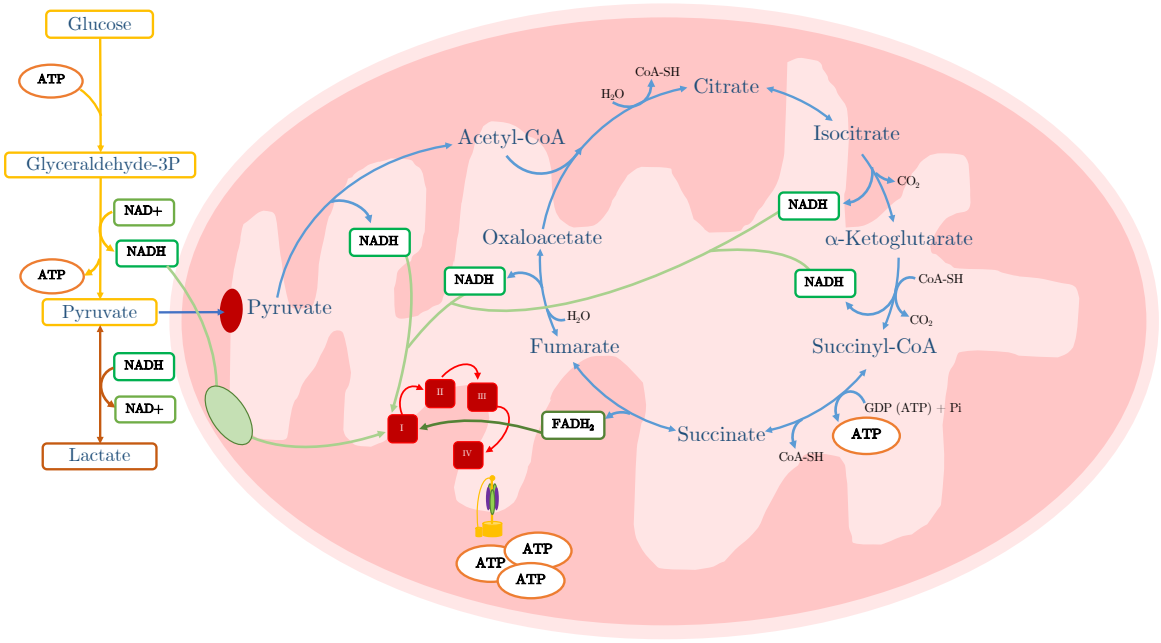


Figure 5: Schematic representation of the integration of glycolysis, lactate production, the TCA cycle and the oxidative phosphorylation chain (only 4 acceptors included for simplicity). NADH produced in the cytoplasm has to be transported into the mitochondrion which may require ATP consumption. NADH and $FADH_2$ yield electrons to a series of acceptors and ultimately to molecular oxygen forming water. The electron gradient is compensated by ATPase which produces ATP by pumping protons.

These redox reactions create a potential difference across the mitochondrion inner membrane which the organelle will compensate by pumping protons using ATPase, producing ATP (Nelson and Cox, 2013). The integration of glycolysis, lactate production, the TCA cycle and oxidative phosphorylation is shown, schematically, in Figure 5.

In summary, the glycolysis ATP balance is positive with a net production of 2 ATP (and 2 NADH). The TCA cycle runs twice per molecule of glucose (each yielding two molecules of pyruvate) and hence the overall balance is 2 molecules of GTP, 6 molecules of NADH and 2 molecules of $FADH_2$ per molecule of glucose. When the co-factors are run through the oxidative phosphorylation chain, each $FADH_2$ molecule yields an average of 1.5 ATP molecules, each NADH formed in the cycle yields about 2.5 ATP molecules and each NADH formed in the cytosol can yield 1.5 or 2.5 molecules of ATP (depending on the transport system used). All co-factors are then regenerated. Hence, the overall ATP produced per molecule of glucose consumed amounts to 30 to 32 molecules. Note that the 2 molecules of ATP produced during glycolysis can be readily used in the cytoplasm and can be considered negligible for the energy yield of glucose (Nelson and Cox, 2013; Quiroga-Campano, 2016).

Although the majority of the glucose oxidation occurs in the glycolytic pathway (80-90%), some glucose will flow into the pentose phosphate pathway (PPP) (Wamelink, Struys and Jakobs, 2008) schematically shown in Figure 6, where glucose is initially irreversibly oxidised to ribulose-5-phosphate, an essential precursor in nucleotide synthesis. This first section of the pathway further produces nicotinamide adenine dinucleotide phosphate (NADPH) required for lipid synthesis. Ribulose-5-phosphate can then be reversibly converted into several metabolic intermediates of the glycolytic pathway through a network of rearrangements. Some studies in CHO cells suggest that the fraction of glucose flowing into this pathway is small (Ahn and Antoniewicz, 2011; Templeton *et al.*, 2014, 2017; McAtee Pereira *et al.*, 2018) and one study has quantified it to be about 1% during exponential growth, decreasing during the stationary phase (Carinhas *et al.*, 2013). In GS-CHO cells, the activity of glucose-6-phosphate dehydrogenase (G6PDH), the first enzyme in the PPP, has been quantified to be only 20 to 28% of that of glucose-6-phosphate isomerase (PGI) (Zou and Al-Rubeai, 2016). However, Sengupta *et al.* have shown that during the non-growth phase this may depend on

process variables such as oxygen and carbon dioxide concentrations and temperature (Sengupta, Rose and Morgan, 2011).

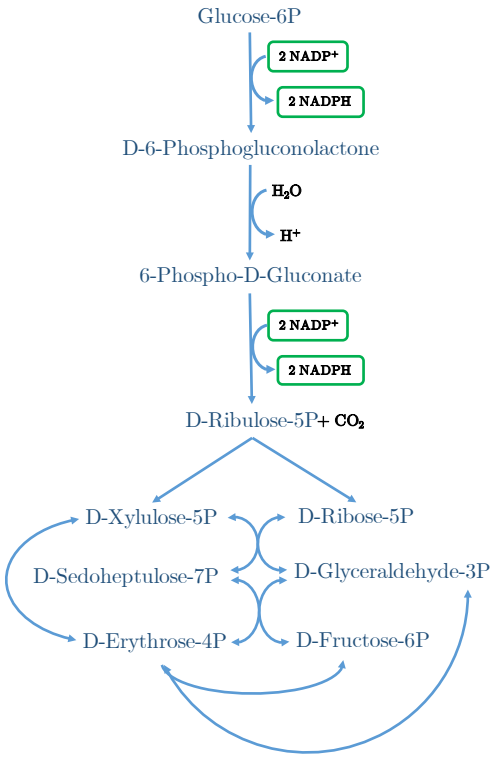


Figure 6: The Pentose Phosphate Pathway

Other essential nutrients for mammalian cells are amino acids which are the building blocks of proteins being essential for cell growth and viability but also for recombinant protein production (Carinhas *et al.*, 2013). In particular, glutamine has been identified as a key nutrient for mammalian cells in culture (Heiden, Cantley and Thompson, 2009) due to its essential role as a main nitrogen and energy source (Yuan, Xiong and Guan, 2013; Zagari *et al.*, 2013; Ha and Lee, 2015; Hartley *et al.*, 2018). Glutamine can be converted to glutamate and then to α-ketoglutarate. This TCA cycle intermediate and has also been identified a major source of carbon for lipid synthesis through feeding the TCA cycle to produce oxaloacetate in the mitochondrion which maintains citrate production (Heiden, Cantley and Thompson, 2009). Glutamine metabolism to produce TCA cycle intermediates, known as glutaminolysis, produces ammonia which is detrimental to cell growth and recombinant protein production and quality (Ha and Lee, 2015)

for a wide range of mammalian cells (Schneider, Marison and Von Stockar, 1996; Pereira, Kildegaard and Andersen, 2018). Figure 7 illustrates some of the possible conversions of AAs to TCA cycle intermediates.

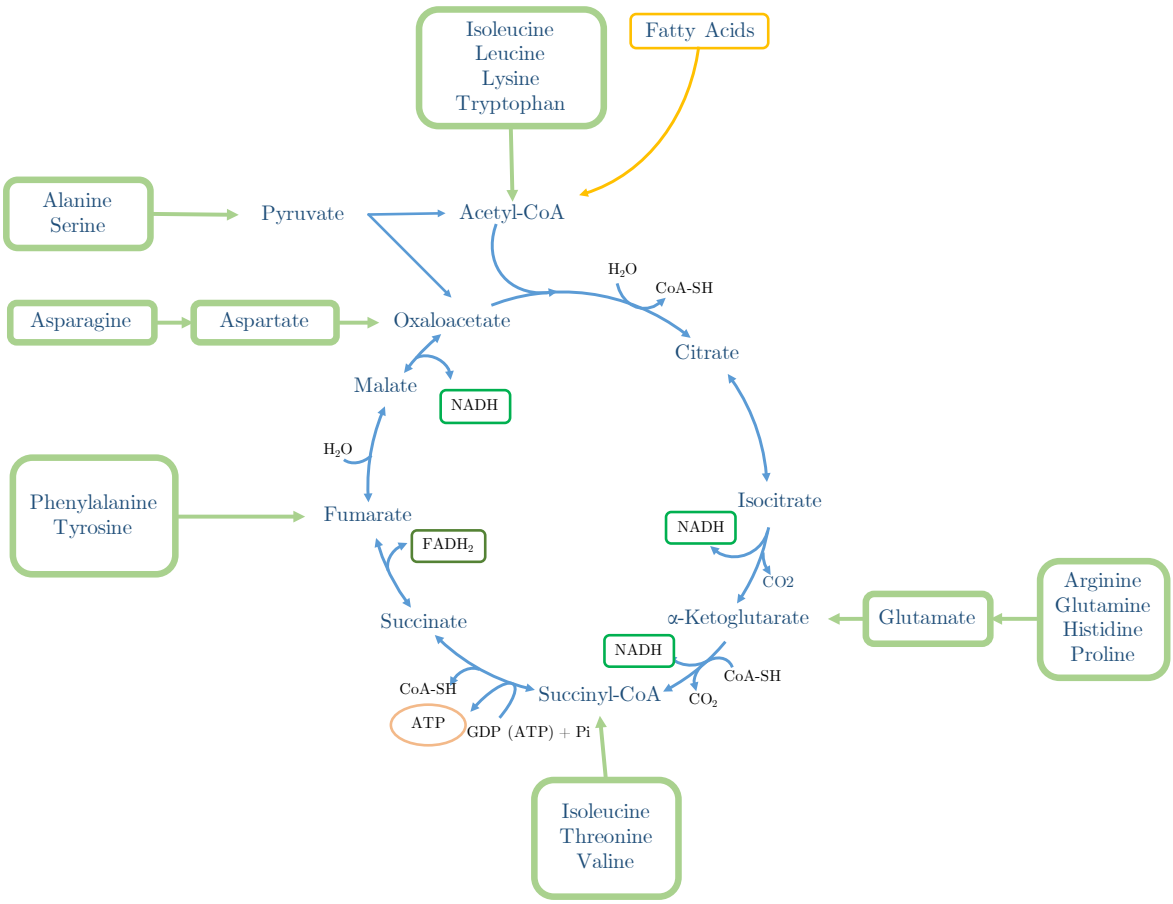


Figure 7: TCA cycle intermediates can be originated from amino acids (green rectangles). Adapted from McAtee Pereira *et al.* (McAtee Pereira *et al.*, 2018).

It has been hypothesized that production of TCA cycle intermediates, particularly malate, from glutamine via glutaminolysis may also fuel lactate production. This suggestion has been made upon observations that a lactate metabolic shift, observed in several mammalian cell cultures, coincides with glutamine depletion (Zagari *et al.*, 2013; Hartley *et al.*, 2018; Hong, Nargund, *et al.*, 2018). In order to provide enough co-factors for rapid proliferation, supported by glucose consumption, mammalian cells convert some pyruvate to lactate in order to regenerate the NAD^+ co-factor (Hong, Nargund, *et al.*, 2018) and re-equilibrate

the tightly controlled NAD^+/NADH ratio reduced by the overflow of NADH produced in glycolysis (Hartley *et al.*, 2018). Before this metabolic shift, cell metabolism is inefficient as carbon sources are only partially oxidised (Hong, Nargund, *et al.*, 2018). Although the exact mechanism of this metabolic shift remains unclear (Hartley *et al.*, 2018), Zagari *et al.* suggested a lactate-induced mitochondrion impairment as they observed increased OxPhos activity beyond the lactate metabolic shift (Zagari *et al.*, 2013). However, as reviewed by Hartley *et al.*, various experimental observations point towards a NAD^+/NADH ratio imbalance-driven process (Hartley *et al.*, 2018). Although further clarification is required, this hypothesis is consistent with increased OxPhos activity after the lactate metabolic shift (Hong, Nargund, *et al.*, 2018) as the regeneration of NADH in this pathway may replace lactate production in replenishing NAD^+ restoring the NAD^+/NADH equilibrium. More evidence of the role of OxPhos in replacing lactate formation as a source of NAD^+ has been provided by the works of Kang *et al.* showing that cooper deficiency inhibits the lactate metabolic shift. Cooper is an essential element in complex IV of the electron transport chain and its depletion means that NADH cannot be regenerated requiring lactate production to remain active (Kang *et al.*, 2014; Hartley *et al.*, 2018).

The production of toxic by-products, namely lactate and ammonia, as a consequence of the metabolism of glucose and glutamine, has prompted research on alternative medium compositions to avoid the formation of undesirable by-products. Galactose and glutamate have been tested as glucose and glutamine replacements, respectively, leading to significantly lower lactate and ammonia production (Altamirano *et al.*, 2000). Replacing asparagine with aspartate, glutamine with glutamate and reducing the serine content has significantly reduced the ammonia production rate while not impacting cell growth, viability or mAb specific productivity (McAtee Pereira *et al.*, 2018). Understanding cell metabolism is essential to reach optimal cell culture performance. In the past, before current status of metabolic understanding and analytical technologies, serum was added to

cell culture media as a source of amino acids, proteins and other elements regarded as essential. Serum stimulates cell growth but brings a source of uncertainty and uncontrolled variability into the process (Kunert and Reinhart, 2016) as it contains unknown elements, it foams under bioreactor aeration conditions, it can pose problems downstream due to its protein content and it is deemed the main source of viral contamination (Butler, 2005; Heinzle, Biwer and Cooney, 2006). The technical arguments listed above combined with others including serum supply chain reliability and price fluctuations (Jayme, 2007) undermine its use. At present, serum is regarded as undesirable, driving industrial production towards chemically defined media compositions (Heinzle, Biwer and Cooney, 2006; Jayme, 2007; Cote, 2013; Fischer, Handrick and Otte, 2015; Kunert and Reinhart, 2016) where the most widely used cell lines for mAb production can proliferate indefinitely (Li *et al.*, 2010). Salazar *et al.* reviewed how cellular, formulation and process parameters are related to the composition of the cell culture media (namely in terms of amino acids) directly influencing product quality and stability, recommending that the culture media formulation be cell line- and clone-specific (Salazar, Keusgen and Von Hagen, 2016). Evidence of the importance of media formulations has been shown by Reinhart *et al.* who have cultured CHO cells in eight commercially available cell culture media and witnessed different metabolic profiles and hence variable cell densities, viabilities, monoclonal antibody production and product glycosylation in different formulations. With the exceptions of glutamine – always reaching, at least, 20% of its initial concentration – and glycine – consistently above 120% of its initial value – no other AA had the same profile across all media formulations analysed (Reinhart *et al.*, 2015). In batch mode, while most AAs were either consumed or produced irrespective of the media formulation, aspartate and glutamate were produced in some media formulations and consumed in others (Reinhart *et al.*, 2015). In fed-batch mode, only two media were tested, and the supplements used have further been shown to have a direct impact on metabolism of some AAs. Glycosylation has shown differences in fed-batch mode as a function

of the additives used in the basal medium (Reinhart *et al.*, 2015) and a subsequent study identified product quality differences between batch and fed-batch cultures (Reinhart *et al.*, 2018). In addition, Hefzi *et al.* have analysed the auxotrophies reported in literature for CHO cells and compared conclusions of prior studies to their genome-scale models of three different CHO cells lines (S, DG44 and K1) to conclude that some auxotrophies identified in earlier studies are, in fact, clone-specific (Hefzi *et al.*, 2016). These observations support that knowledge of clone-specific metabolism is essential for process optimization.

2.1.1.3 Apoptosis and cell death

Cell death is one of the major issues in mammalian cell bioprocessing. Apoptosis, programmed cell death, has been reported to be the most important death pathway for industrially relevant cell lines (Zhang, 2010) responsible for 80% of the CHO cell deaths (Goswami *et al.*, 1999) and 90% of hybridoma deaths (Singh *et al.*, 1994).

In apoptosis, internal and/or external stimuli initiate a series of highly controlled reactions, which ultimately lead to cell death. The two most important groups of proteins involved in apoptosis, caspases and proteins of the Bcl-2 (B-cell lymphoma 2) family, participate in all pathways of apoptotic cell death (Grilo and Mantalaris, 2019a). Caspases are a group of cysteine-dependent endoproteases that catalyse the breaking of the peptide bond. Most caspases are produced as monomeric species with little or no activity, generally known as pro-caspases or zymogens (Fussenegger, Bailey and Varner, 2000; Taylor, Cullen and Martin, 2008; McIlwain, Berger and Mak, 2015). Upon activation, they homodimerize forming biologically active proteins. Fourteen caspases (numbered 1 to 14) have been reported in murine or human cells, thus being of importance in bioprocessing. In addition, four other caspases have been identified (numbered 15 to 18) in the genomes of mammalian cells, other than mice or humans (Eckhart *et al.*, 2008), with caspase 16 found to be a pseudogene (not translated). Most caspases can be

classified, according to their role, in three main groups: a) initiator caspases (caspases 8 and 9); b) executioner caspases (caspases 3, 6 and 7); and c) caspases involved in inflammation (caspases 1, 4, 5, 11 and 12L) (McIlwain, Berger and Mak, 2015). Caspase-2, which is not classified as it shares structural elements with both initiator and executioner caspases (Vakifahmetoglu-Norberg and Zhivotovsky, 2010; Fava and Bock, 2012), is known to be a substrate for caspases 3 and 8 (Fava and Bock, 2012) and to trigger cytochrome-C release leading to the formation of the apoptosome (Orrenius, Zhivotovsky and Nicotera, 2003). Caspase-2 can also be activated by downstream targets of p53 (Vakifahmetoglu-Norberg and Zhivotovsky, 2010). The role of caspase-10, which has been reported to be activated by caspase-6 (Taylor, Cullen and Martin, 2008), has not been fully elucidated yet alongside caspases 12S and 14 that also remain unclassified (McIlwain, Berger and Mak, 2015). The caspase cascade is self-catalysed once it has been triggered (Taylor, Cullen and Martin, 2008), as depicted in Figure 8.

Proteins of the Bcl-2 family are heavily involved in the regulation of apoptosis both with pro- and anti-apoptotic activities and can be divided into three sub-groups. The anti-apoptotic group comprises of Bcl-2, Bcl-xL (B-cell lymphoma extra-large), Bcl-W (Bcl-2-like protein 2), Mcl-1 (induced myeloid leukaemia cell differentiation protein), Bcl2A1 (Bcl-2 related protein A1) and Bcl-B (a Bcl-2-like protein similar in amino acid sequence to the Boo protein(Ke, Godzik and Reed, 2001)). These proteins have four Bcl-2 homology domains (BH1, BH2, BH3 and BH4). Apart from Bcl2A1, all other members have transmembrane domains (Taylor, Cullen and Martin, 2008) and are typically associated with either cellular, nuclear or mitochondrial membranes(Reynolds *et al.*, 1996). The second sub-group composed of Bax (Bcl-2-associated protein X), Bak (Bcl-2 homologous antagonist killer) and Bok (Bcl-2-related ovarian killer) lack the BH4 domain. Finally, the third sub-group, has only got the BH3 domain (Taylor, Cullen and Martin, 2008) and is composed of Bik (Bcl-2-interacting killer), Hrk (Harakiri), Bim (Bcl-2-like protein 11), Bad (Bcl-2-associated death promoter), Bid (BH3 interacting-domain

death agonist), PUMA (p53 upregulated modulator of apoptosis), NOXA (Phorbol-12-myristate-13-acetate-induced protein 1) and Bmf (Bcl-2 modifying factor). The BH4-defficient and BH3-only proteins are pro-apoptotic.

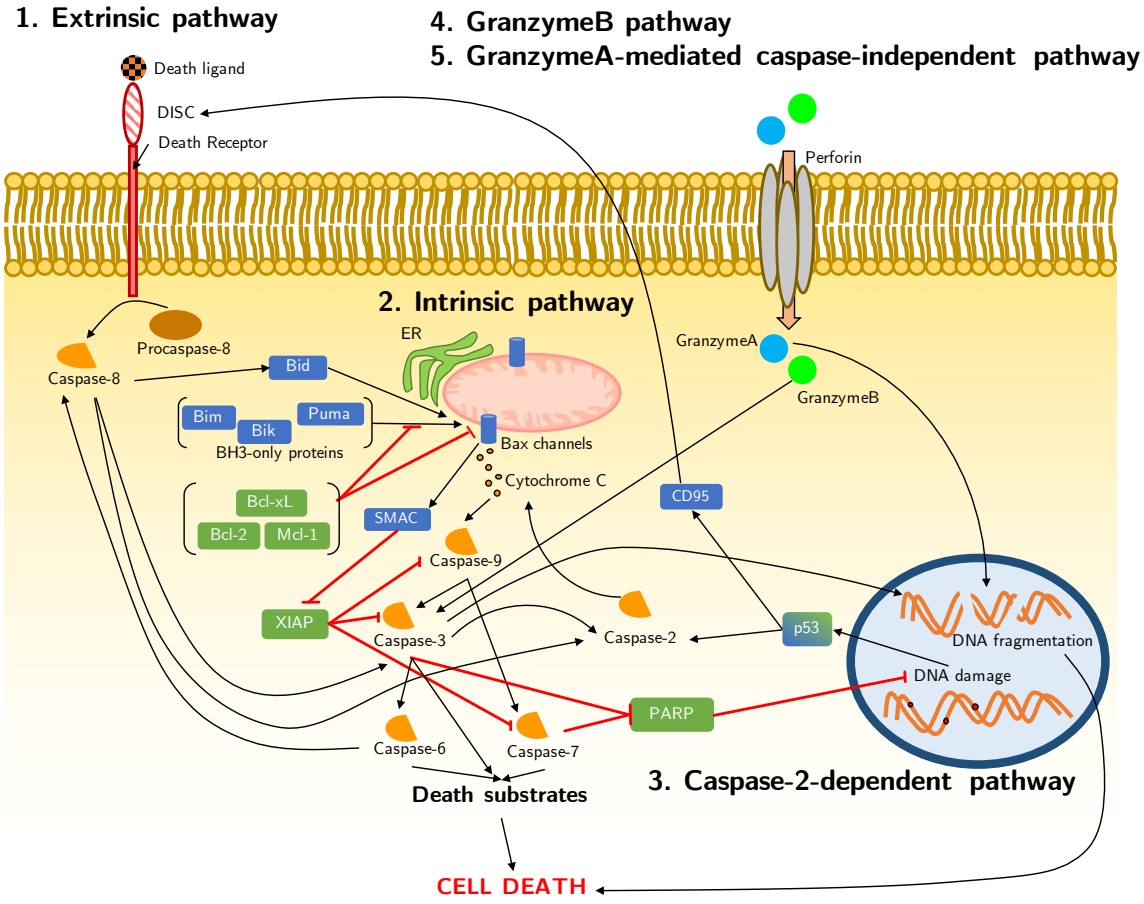


Figure 8: Five main apoptosis pathways have been identified: (1) the extrinsic pathway activated by external stimuli (Ichim and Tait, 2016); (2) the internal pathway activated by MOMP promoted by Bax channels (Taylor, Cullen and Martin, 2008); (3) the DNA-damage-induced caspase-2-dependent pathway (Vakifahmetoglu-Norberg and Zhivotovsky, 2010); (4) the Granzyme B induced pathway where this molecule acts as a caspase-3 inducer (Taylor, Cullen and Martin, 2008); and (5) the Granzyme A induced pathway where no caspases are involved – Granzyme A induces a DNase, which induces DNA fragmentation and cell death (Orrenius, Zhivotovsky and Nicotera, 2003).

Five main apoptosis pathways have been identified among several proposed. The *extrinsic pathway* (often referred as Type I) is triggered by stimuli external to the cells. A death ligand (such as FasL or Tumour Necrosis Factor α, TNFα) binds to receptors on the cell membrane and triggers the activation of caspase-8, which initiates apoptosis (Figure 8). The *intrinsic pathway* (often referred to as Type II

or mitochondrial pathway) is activated by internal stimuli (metabolic or hypoxic stresses) and can also be triggered by endoplasmic reticulum stress (Ichim and Tait, 2016). Apoptosis type II-triggering internal stimuli promote the activation of caspases and pro-apoptotic members of the Bcl-2 family (Figure 8) leading to mitochondrion outer membrane permeabilisation (MOMP) via Bax channels. Bax activates SMAC (second mitochondria-derived activator of caspases), which inhibits the anti-apoptotic activity of XIAP (X-linked inhibitor of apoptosis protein) (Ashkenazi *et al.*, 2017). In its turn, MOMP leads to the release of cytochrome-C (Taylor, Cullen and Martin, 2008), which induces the formation of the apoptosome, a complex of seven caspase-9 units (Taylor, Cullen and Martin, 2008). Caspase-9 will trigger an executioner caspase (3 or 7) ensuring the continuity of the process (Taylor, Cullen and Martin, 2008). The anti-apoptotic members of the Bcl-2 family (Bcl-2, Bcl-xL, Mcl-1 for instance) are able to inhibit apoptosis by quenching the activity of their pro-apoptotic counterparts. In this process, the activation of the *extrinsic pathway* can also occur through the activation of Bid by caspase-8. Conversely, activation of the *extrinsic pathway* as a downstream effect of the *intrinsic pathway* may occur when CD95 is activated by p53. A third pathway is dependent on the presence of caspase-2, which can be activated by tumour suppressor gene p53 upon persistent DNA damage (Vakifahmetoglu-Norberg and Zhivotovsky, 2010). Apoptosis can also be triggered via direct activation of caspase-3 by *Granzyme B* (Taylor, Cullen and Martin, 2008). Finally, DNA fragmentation promoted by *Granzyme A* can also lead to apoptosis (Orrenius, Zhivotovsky and Nicotera, 2003). Proteomic studies on CHO cells indicate that the *intrinsic pathway* is the most relevant apoptotic driver in culture (Wei *et al.*, 2011). It can be directly linked to the operating conditions of a bioreactor rendering it important in bioprocessing as it can be triggered by lack of nutrients, excess protein production, temperature or pH shifts, and high osmolality, among others.

Regulation of apoptosis is achieved by protein-protein interactions, which may involve a third protein as a mediator. Examples of protein-protein interactions

in the absence of a mediator include the BH3 domains of Bax proteins yielding a tetramer that forms a channel in the mitochondrial membrane or the inhibition of Bax channel formation by Bcl-2 (Zha *et al.*, 1996; Antonsson *et al.*, 1997). In contrast, the interaction between Bcl-2 or Bcl-xL with caspase-8 requires mediation by BAP31 (Bcl-2 associated protein 31) (Orrenius, Zhivotovsky and Nicotera, 2003). The *intrinsic pathway* results in MOMP caused by either Bax channels on the mitochondrial membrane (Taylor, Cullen and Martin, 2008) or calcium ions (Orrenius, Zhivotovsky and Nicotera, 2003). MOMP will in turn cause the release of a number of pro-apoptotic molecules such as cytochrome-C. As depicted in Figure 8, caspase-3 is instrumental in apoptosis and can be activated by initiator caspases 8 and 9 or by Granzyme-B. Once activated, it will back-activate upstream caspases to ensure the continuity of the process (Taylor, Cullen and Martin, 2008). Caspase-3 is also an activator of CAD (caspase-activated DNase) which, in proliferative cells, is complexed with its inhibitor. When caspase-3 is overexpressed, this complex is cleaved and CAD degrades chromosomal DNA (Nagata *et al.*, 2003), which also occurs when Poly(ADP-ribose) polymerase (PARP) is inhibited by caspase-3 (or 7) (Cohen, 1997). During late apoptosis, caspase-3 acts as a cytoskeletal reorganization and disintegration inducer leading to the formation of apoptosis bodies (Elmore, 2007).

Several structural features change in the cell during apoptosis namely at the cytoskeleton level where several elements are degraded leading to membrane blebbing. Other physical changes include destruction of the nuclear lamina leading to the collapse of the nuclear envelope, DNA fragmentation and chromatin condensation (Goswami *et al.*, 1999; Elmore, 2007); formation of apoptotic bodies containing intact organelles(Elmore, 2007); and mitochondria swelling and elongation (Häcker, 2000; Orrenius, Zhivotovsky and Nicotera, 2003; Blackstone and Chang, 2011). Under certain conditions, apoptosis can be reversed as reported in several cell lines upon removal of the apoptotic stimuli (Tang *et al.*, 2009) or by p53 activation inducing DNA repair, cell cycle arrest and apoptosis arrest

(Batchelor *et al.*, 2008). In addition, late stage apoptosis may also be reversible even after DNA damage has been induced. Anastasis (Greek for "Rising to Life") has been proposed as a name for the process of "resurrection" of cells. The suggested reversal mechanism requires a cohort of several proteins and a number of genetic changes have been observed following reversal of apoptosis, in both chromosomal aberrations and number (Tang *et al.*, 2012).

Controlling apoptosis is important in therapeutic protein production as significant differences in aggregation, charge variants and glycoforms have been observed throughout batch and fed-batch cultures of GS-CHO cells (Park *et al.*, 2017). Decreased quality has been observed at early stationary phased compared to exponential growth, further worsening during the decline phase. As a consequence of cell death proteins released to the culture medium are mostly of intracellular origin (Park *et al.*, 2017) and in increased amounts (Albrecht *et al.*, 2018). While it is already common practice that cultures be halted when viability drops below a given threshold, 60% for example (Khoo and Al-Rubeai, 2009; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018), this may already compromise product quality.

Detailed mathematical descriptions of all proteins and interactions involved in the different apoptosis pathways pose significant challenges in (1) capturing the biology, (2) obtaining the required experimental data to validate model predictions, and (3) producing computationally trackable, predictive models, which can be used in model-based process optimization. Consequently, there are two *a-priori* choices in apoptosis modelling. On the biology side, a careful evidence-based decision on the intervenors and interactions to be modelled, and on the modelling side deciding on a model structure. While some researchers have chosen formulations based on chain reactions (initiation, amplification and execution (Zhang, Brazhnik and Tyson, 2009)) or modelled the transitions of cell populations between states (viable, apoptotic and dead (Meshram *et al.*, 2012)), most have opted for kinetic/mechanistic formulations. Generally, apoptosis mathematical models either

include few interactions in great depth capturing detailed kinetic associations or capture entire pathways considering few intervenors. Examples of models describing few species in great detail include (1) models of DNA damage (Talemi *et al.*, 2015), (2) models of caspase-3 activation by various pathways (Stucki and Simon, 2005), (3) models of Bax/Bcl-2 interactions (Chen *et al.*, 2007), (4) models of caspase 3 and 8 interaction including bistability studies (Eissing *et al.*, 2004) which were later compared to experimental data (Pace *et al.*, 2010), and (5) models for the interplay between autophagy and apoptosis (Tyson *et al.*, 2011). Models of complex networks include a Fas signalling model where certain simulations were compared to experimental data (Hua *et al.*, 2005; Okazaki *et al.*, 2008) and a model of the CD95-induced apoptosis pathway (Bentele *et al.*, 2004) that described the minimum apoptosis-inducing ligand/receptor ratio using sensitivity analysis and experimentally validated. The work of Bentele and co-workers provides a suitable example of model features that would be of interest to the bioprocess community especially in relation to bioprocess optimization. Integrating metabolism and apoptosis has been attempted in CHO cells using metabolic flux analysis (MFA) for glucose, lactate, five amino acids and ammonia; alas, metabolite predictions did not agree with experimental data beyond the exponential growth failing to capture metabolic shifts (Naderi *et al.*, 2011). The effect of nutritional stresses on gene expression of GS-NS0 cells has been analysed using a combined modelling/experimental approach whereby a simplified kinetic model captured nutrient and metabolite profiles and described the expression of six apoptosis-related genes (caspases 3 and 8, atf5, bcl-2, bax and trp53bp2 (p53 binding protein 2)) in batch mode (Usaku, 2015). The main drawback of this work was the lack of predictive capability in fed-batch mode.

2.1.2 CELL CULTURE

Cell culture is the process by which cells are allowed proliferation under a controlled environment aiming at the production of a target number of cells or a pre-specified product for which the conditions would have been optimized. There are a number of methods to perform cell culture in suspension including batch, fed-batch and different types of continuous methods but also T-flask cultures and other monolayer methods (not reviewed here).

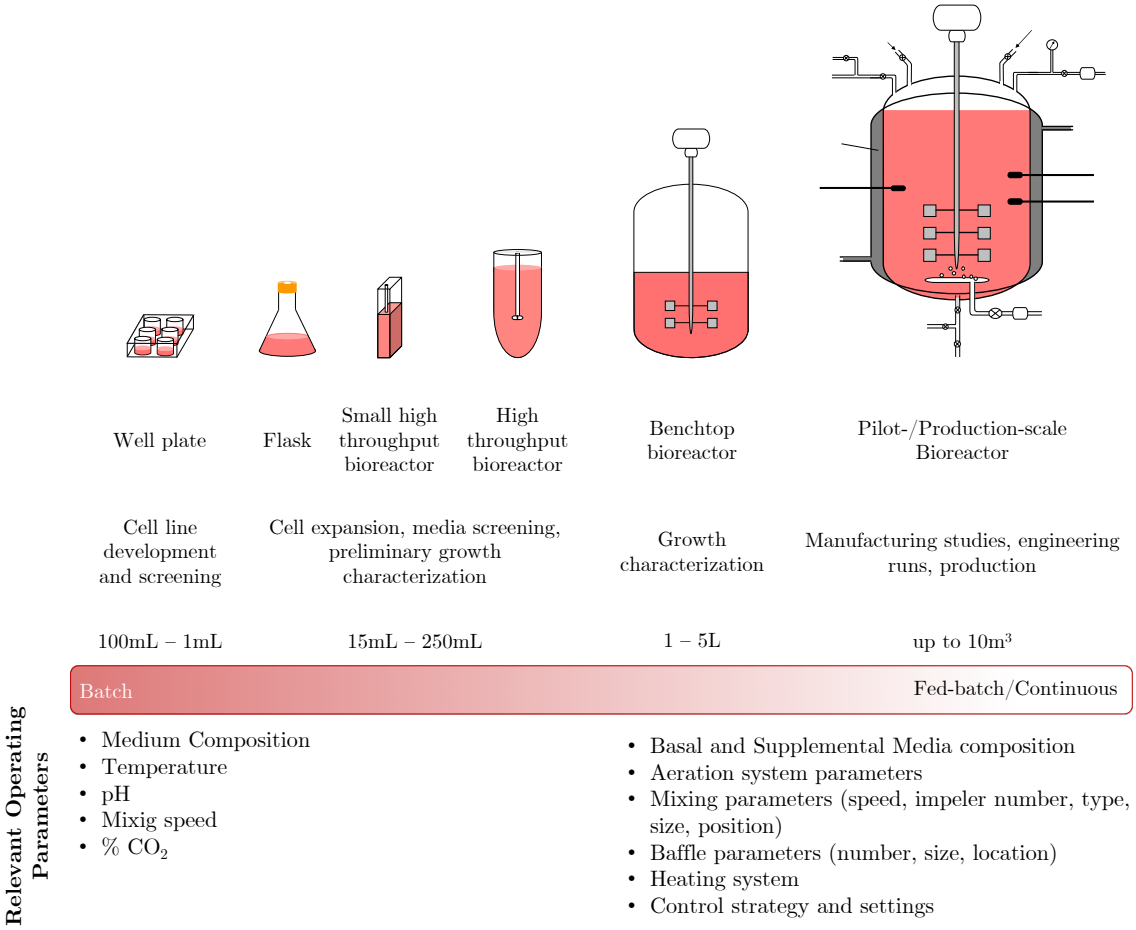


Figure 9: Suspension cultures across bioprocess development, commonly used reactors, goals, scales, mode of operation and relevant parameters (adapted from (Grilo and Mantalaris, 2019a).

Production of mAbs is conducted in suspension cultures and hence, throughout process development, suspension cultures are conducted at different scales, as depicted in Figure 9. As scale varies, there are a number of parameters to be controlled which vary in nature and importance and can be physical

(temperature, gas flow, impeller speed), chemical (pH, osmolality, dissolved O_2 and CO_2, metabolite levels) and biological (cell density, cell cycle, mitochondrial activity) (Kunert and Reinhart, 2016). In this section, the most relevant cell culture strategies for the production of monoclonal antibodies will be outlined.

Batch cell culture is the simplest mode of operation, allowing cells to proliferate for a defined time in a closed system after an initial charge of medium and inoculum. There are no additions of nutrients nor removals of exhaust media or toxic metabolites throughout the culture timespan. Cells will initially spend some time adapting to the new environment (lag-phase) requiring most nutrients for energy and protein production before they start exponential growth. A stationary phase follows, characterized by an equilibrium between growth and death, and finally, when the growth rate is offset by the death rate, the culture will begin the decline phase. Culture viability tends to remain high until mid-stationary phase, when a decrease will be observed, and a pronounced cut will be seen towards the end of the culture. Product accumulation occurs throughout culture time but this culture mode is disadvantageous in terms of culture duration, which will be limited by the accumulation of toxic metabolites (Cacciuttolo, 2007) and the depletion of key nutrients.

Issues encountered in batch cultures can be partially overcome by fed-batch cultures where fresh media (possibly with a different formulation) is added continuously or at pre-defined times without product, cells or exhaust media removal. Controlling the feeding allows for the control of the growth rate thus providing a slightly higher degree of control (Gaughan, 2016). Ideally, operating in fed-batch mode requires prior knowledge of the process, namely in terms of nutritional requirements of the cell lines used. This will allow for stringent control of process variables and optimal feeding scheduling. This strategy requires relatively simple process equipment and provides relatively easy scale-up, validation, process characterization and reasonable costs making it the most popular culture mode employed in industry and academia for growing cells in suspension. Other relevant

variables one can control in fed-batch operation include the composition of the medium initially in the bioreactor (basal medium), the medium fed to the bioreactor (feed medium, obtained generally by concentration of the basal medium (Abu-Absi *et al.*, 2013), but not necessarily), the feeding strategy or any combination of the above making optimality a challenging target in practice (Jiang *et al.*, 2012).

2.2 MATHEMATICAL MODELLING OF CELL CULTURE

Mathematical modelling offers a cheaper and rational way of optimizing systems compared to traditional experimentation-intensive methods. Mathematical descriptions of the cellular and biochemical processes as well as fluid flow within a bioreactor, can be used to create a model that describes the required operation(s). Such models can be used as a tool to design the required experiments in order to obtain key model parameters. When the model is refined and experimentally validated, in can be used to optimize the system (model-based optimization) as well as to perform model-based control. Examples of mathematical modelling applications are optimal media design (Xie and Wang, 1994; Galbraith *et al.*, 2018; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018), identification of previously missed variables (deZengotita *et al.*, 2000), product quality monitoring (Kontoravdi, Asprey, *et al.*, 2007; Kontoravdi and Jimenez del Val, 2018), feeding strategy optimization (Kiparissides, Pistikopoulos and Mantalaris, 2015), understanding of physiological pathways (Fussenegger, Bailey and Varner, 2000; Huber *et al.*, 2011), oxygen uptake rates prediction (Rigual-González *et al.*, 2016) or hydrodynamic stress studies (Sieck *et al.*, 2013).

Key advantages of using mathematical modelling are rational design of experiments, determination of the feasible design space, fast screening of several hundreds of scenarios and the possibility of achieving global optimality. Hence, laborious and expensive experimental work is reduced to a minimum. However, using modelling requires a significant accumulated knowledge of the system/process under study (cellular metabolism, reactor performance etc.). Even if knowledge is

available, model predictions may carry errors, especially if the models are too large or if they require more experimental data than that available/collectable. Right opposite, sit simplistic models which may not be fully able to capture the reality under study. Therefore, the ability to select the phenomena to be modelled must be combined with significant model development skills in such a way that the model can capture the desired reality and has only the required complexity (Hand, 2014).

2.2.1 MODEL CLASSIFICATION

Modelling biological systems is a complex task due to the dynamic nature of the processes to be described and their multiple time-scales. A systematic classification framework of mathematical models of biological systems has been suggested by Fredrickson *et al.* (Fredrickson and Tsuchiya, 1963), and is shown in Figure 10. According to this framework, models are classified according to their structure and segregation. Segregation refers to whether heterogeneity of a cell population is considered, and structure refers to compartmentalisation, i.e., consideration of different sub-compartments within a cell. Therefore, unsegregated models treat cells as a homogeneous population and unstructured models give little or no detail in terms of cellular compartmentalization.

At any level of structure/segregation, models can be steady-state or dynamic according to their ability to capture system variations in time and they can be deterministic or probabilistic, depending on how system variations are treated. A steady-state model will capture one time point of the system but its validity can be compromised if used to predict other time points (Robitaille, Chen and Jolicoeur, 2015). Time changes can be computed in a deterministic manner, where the model would use experimental observations to compute changes assuming the most common system behaviour holds (Kyparissidis, 2012), or in a probabilistic fashion where uncertainty is accounted for through stochastic variations of the system state.

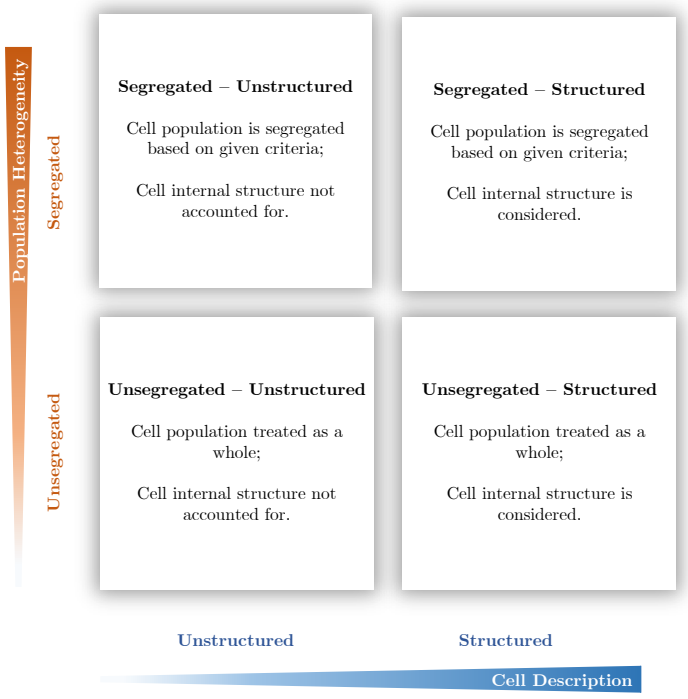


Figure 10: Classification of biological systems models according to the framework suggested by Fredrickson *et al.* (Fredrickson and Tsuchiya, 1963)

A number of unsegregated, unstructured dynamic models of mAb production cells have been published. Kontoravdi *et al.* have developed an unstructured, unsegregated, dynamic model of cell growth and AA metabolism in two mammalian cell lines (Kontoravdi, Wong, *et al.*, 2007) and Kyparissidis *et al.* have developed a model of similar structure where more attention is given to the role of glutamate, the importance of ammonia, and the lag-phase (Kiparissides, Pistikopoulos and Mantalaris, 2015). Robitaille and co-workers (Robitaille, Chen and Jolicoeur, 2015) have attempted to predict the metabolism of CHO cells using a steady-state model. While the model can predict some of the authors' data, it fails to capture metabolic shifts. Usaku has also developed a model of the same category, where a significant amount of experimental and modelling work was combined in order to include gene expression so as to be able to capture apoptosis in mammalian cell culture (Usaku, 2015).

A structured-unsegregated dynamic model has been derived to describe the inter-organelle monoclonal antibody transport and secretion of mAbs in hybridoma

cells by Bibila and Flickinger (Bibila and Flickinger, 1991). The authors have considered three compartments – ER, Golgi apparatus and medium – and previously published times of transfer to establish their model. This model was tested at steady state (Bibila and Flickinger, 1992a) and transient conditions (Bibila and Flickinger, 1992b). Another example is the work of Kontoravdi *et al.* who have developed a model to describe monoclonal antibody production and glycosylation in a pioneering attempt to include product quality in a cell culture model to be able to monitor it in culture. The monoclonal antibody production and glycosylation occur in sub-compartments (structure) (Kontoravdi, Asprey, *et al.*, 2007).

The importance of segregating a cell population has been identified since the 1960s (Fredrickson and Tsuchiya, 1963). One possibility of doing so would be to evaluate a single cell behaviour in time with a great level of compartmentalisation using single cell models (SCM) (García-Münzer, 2014). However, this approach is very time consuming, computationally demanding and work intensive, given the need to obtain data for each of the millions of cells in the system. Sanderson *et al.* have suggested a structured SCM model. The model shows a very considerable degree of complexity comprising 50 compartments, detailed metabolism, membrane transport and product formation, in an extraordinary attempt to describe the complexity of a cell. However, this comes at a high computational cost (Sanderson, 1997). A simpler version of Sanderson's model has been used by Ho *et al.* who have used the SCM part coupled with a simpler metabolic model (Ho, Varley and Mantalaris, 2006). A more efficient alternative to SCMs would be the use of a population balance model (PBM). Several models have used the PBM formalism, often using age, mass or volume as segregation criteria (Martens *et al.*, 1995), which are difficult to assess and may have little biological meaning. More recently, proteins and DNA have been used as distribution variables to develop a cell-cycle model of GS-NS0 cells (García-Münzer, Kostoglou, *et al.*, 2015) and cancer cells (Fuentes-Garí *et al.*, 2015).

2.2.2 MODEL EQUATIONS

Mathematical models of cell cultures in bioreactors require kinetic equations and mass balance equations. The most general mass balance is that for the continuous bioreactor where possible inflow and outflow streams are included. If a fed-batch bioreactor is considered, terms related to outflow streams are removed, and if the inlet streams are also discarded the model would describe a batch bioreactor. For biological systems, mass balances to sugars, AAs, by-products, recombinant proteins etc. are usually required and take the form of equation (1)

$$\frac{dP}{dt} = F_{in}P_{in} + X_V(Q_{prod,P^-} Q_{cons,P})V^- F_{out}P \tag{1}$$

where P is a given metabolite/nutrient/product, F_{in} and F_{out} are the inlet and outlet flow rates, P_{in} is the concentration of P at the inlet, xV is the concentration of viable cells, $Q_{prod,}$P and $Q_{cons,}$P are the specific rates of production and consumption of P, respectively, and V is the volume of the bioreactor (Kontoravdi, Asprey, et al., 2007; Kontoravdi, Wong, et al., 2007; Kiparissides, Pistikopoulos and Mantalaris, 2010; García-Münzer, Kostoglou, et al., 2015; Usaku, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

The derivation of the equations for the rates of production and consumption is also important. In most literature, metabolic rates are related to the organism's growth rate using yield coefficients (proportionality constants between biomass growth rate and substrate consumption/production rates). There are a number of kinetic equations available in literature which can be used for bioprocess modelling purposes (Kyparissidis, 2012). These include for example the Hill equation (equation (2)), of which the widely used Monod equation is a particular solution (for n=1).

$$\mu = \mu_{max}\left(\frac{S}{K_S + S}\right)^n \tag{2}$$

Notably, a similar form of the Monod equation (shown in equation (3) for an apoptotic rate) has been used in processes related to inhibition or, more generally, when increasing S leads to reducing μ. It has been used to describe inhibition of cell growth by lactate and ammonia (Kontoravdi, Asprey, *et al.*, 2007) or to model cell death as a function of glutamate concentration (García-Münzer, Kostoglou, *et al.*, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

$$\mu_a = \mu_{a,max}\left(\frac{K_S}{K_S + S}\right) \tag{3}$$

2.2.2.1 Population balance equations for biological systems

Population Balance Equations (PBEs) offer a good computational time/biological meaning trade-off. PBEs use an array of variables to distribute (i.e. segregate) a population into sub-populations. The general structure of a PBE is shown in equation (4). The case herein described is that where a continuous, one-variable PBE is used. A population is split into several subpopulations (stages) of states k=*i, j*, etc. Each N_k subpopulation (stage) is described by an internal coordinate, x_k, which accumulates at a rate r_k, where N_i is called the number distribution on stage k=i.

$$\frac{\partial N_i(x_i,t)}{\partial t} + \frac{\partial}{\partial x}\left[r_i(x_i,S)N_i(x_i,t)\right] =$$

$$= -\Gamma_i(x_i,S)N_i(x_i,t) + \Gamma_{i+1}(x_{i+1},S)N_{i+1}(x_{i+1},t) - DN_i(x_i,t) - \tag{4}$$

$$-k_d(S)N_i(x_i,t) + h\int_{x_{i-1,min}}^{x_{i-1,max}} \Gamma_{i-1}(x_{i-1}S)N_{i-1}(x_{i-1},t)\,p(x_{i-1},S,t)dy$$

The left-hand side of equation (4) describes the variation of the number of cells with a given value of the internal coordinate (x_k) in time (first term) and the second term describes the progression of the population in space (second term). Space is, in this context, the domain of the internal variable used for segregation

(x). The right-hand side of equation (4) accounts for the means by which population variations can happen. The first and second terms account, respectively, for outward and inward flow of entities at rates Γ_i and Γ_j, respectively; the third and fourth terms account for entities which leave the system by dilution or death, respectively; the fifth term accounts for the generation of particles (2 in the case of cell growth), which can enter state i at whichever level of the internal variable of the previous state (x_{i-1}) and at a probability p, which is in turn dependent on the internal coordinate used, time and substrate. Rates of accumulation of internal variables (r) are considered dependent on time (t) and substrate (S), and transition between states (Γ) are considered dependent on the internal coordinate (x) and substrate.

Examples of the use of population balance equations (PBEs) include age-based models (A. Hjortso and Nielsen, 1995), although difficulties with experimental measurements and limited biological significance of cell age, especially in indefinitely replicating cell lines, as those used in biologics production, have limited further application (Mantzaris, Daoutidis and Srienc, 2001). Volume-based PBEs have also been widely used, often combined with DNA content, to segregate the cell population, differentiating between gap phases (G_0/G_1 and G_2) and the S-phase of the cycle (Liu *et al.*, 2007), which has been supported by recent evidence that mAb productivity and cell size are related (Pan *et al.*, 2017). Mass-based PBMs have been coupled with SCMs, an approach that has proven to be too computationally expensive (Sidoli, Asprey and Mantalaris, 2006). In summary, difficulties in identifying suitable biologically meaningful segregation criteria and the complexity of the partial-integro-differential equations involved have limited the application of PBEs in the design, control and optimization of biological processes (Lencastre Fernandes *et al.*, 2011).

A suitable discretization strategy, and extensive experimental work, enabled the use of PBEs to describe the cell cycle in the industrially relevant GS-NS0 cell line (García-Münzer *et al.*, 2013; García-Münzer, Kostoglou, *et al.*, 2015). The

authors used cyclin and DNA contents as internal variables to distribute the population of cells across the cell cycle, using three PBEs, one for each phase. Cyclin E1, DNA and cyclin B1 were used as internal coordinates for phases G_1, S and G_2/M, respectively, using experimentally determined profiles and transition thresholds (García-Münzer *et al.,* 2013). Each phase of the cell cycle was described using one PBE of the kind of equation (4), considering only the relevant terms. One of the most evident practical motivations for the use of segregated models is the possibility of varying recombinant protein production rates across the different phases of the cell cycle.

2.3 MODEL DEVELOPMENT

2.3.1 MODEL DEVELOPMENT FRAMEWORK

As modelling gained an increased interest, model development has also become more rational. One of the most systematic and fruitful model development frameworks is that suggested by Kontoravdi *et al.* (Kontoravdi, Pistikopoulos and Mantalaris, 2010), which has been further expanded in order to enable its use in bioprocess development linking *in-silico* model development, evaluation and refinement with *in-vitro* experiments (Kiparissides *et al.*, 2011). This framework is schematically represented in figure 11.

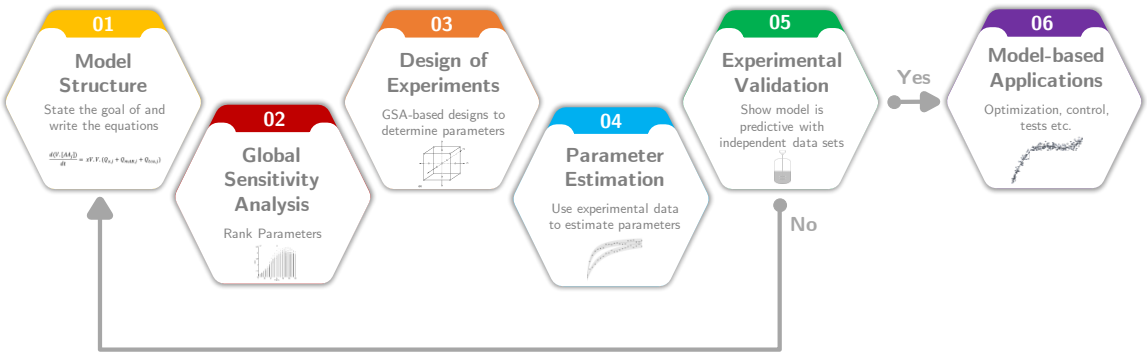


Figure 11: Framework for mathematical model development (adapted from (Kontoravdi, Asprey, *et al.*, 2007))

2.3.1.1 Model structure

As depicted in figure 11, the first step is designing the model structure, whereby the modeller will gain knowledge of the processes to be modelled so as to be able to choose the type of model that better serves the purpose and to compile the required equations. Initial model simulations will be run using nominal values for all model parameters taken ideally from prior experiments or literature data from similar processes.

2.3.1.2 Global Sensitivity Analysis

An important step in reducing the subsequent workload is the identification of the most significant model parameters. Global Sensitivity Analysis (GSA) is the process by which the model is evaluated so as to identify which model parameters most greatly affect the model outputs. Spencer *et al.* explain sensitivity analysis as the modelling equivalent to the experimental task of knocking down or overexpressing genes and evaluating the impact on phenotype (Spencer and Sorger, 2011). In this analogy, the phenotype is the model output and the genes are the model parameters. Given that, in the general case, several sources will have been used to obtain nominal parameter values, it is important to conduct GSA at an early stage so as to immediately determine which model parameters need to be re-estimated. This allows to reduce the experimental burden in model development.

A number of mathematical algorithms exist to perform GSA. The choice of the algorithm must be made upon the type of model chosen. Currently, three main types of sensitivity analysis methods exist: screening methods, local methods and global methods. While screening methods pose problems with non-linear systems and local methods evaluate only a vicinity of the nominal values of the parameters, global methods, in particular variance-based methods, are suitable for complex, highly non-linear models (Kiparissides *et al.*, 2009).

In global methods, all parameters are varied simultaneously, and the analysis is performed over a wide space for each parameter. Furthermore, these methods

are quantitative thus allowing for full parameter ranking. The main drawback of these methods is the computational time required for running the algorithms. Currently, the two most widely used methods are Sobol's sensitivity indices and Fourier Amplitude Sensitivity Test (FAST). Both methods provide identical results and there is no consensus as to which method prevails, although the Sobol method is easier to implement (Kiparissides *et al.*, 2009). More recently, Kyparissidis *et al.* have shown that the Derivative-based Global Sensitivity Measures (DGSM) method is able to provide similar results to the Sobol's indices algorithm while running about 10 to 20 times quicker (Kiparissides *et al.*, 2009). This method is, in fact, a one-step-at-a-time approach although with superior accuracy (Kyparissidis, 2012). The Sobol method has been successfully used for several large-scale highly complex models of biological systems with proven ability of ranking model parameters and providing biologically significant information (Kontoravdi *et al.*, 2008; Fuentes-Garí *et al.*, 2015; García-Münzer, Kostoglou, *et al.*, 2015; Usaku, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

It should be observed that sensitivity analysis is not the only possible technique for model reduction. Alternatives include time-scale separation – where a time scale is fixed and anything faster is approximated by steady-state/equilibrium approaches – or lumping, where a subset of variables is lumped thus reducing the dimensionality of the model (Almquist *et al.*, 2014). These approaches may compromise the dynamic nature of the model and/or reduce its biological significance, especially if significant processes occurring at different time-scales are modelled together, and hence sensitivity analysis is a preferred technique.

2.3.1.3 Model-aided Experimental Development

As GSA is a dynamic analysis allowing for ranking model parameters at different time points, it also allows for the design of suitable experiments to measure the required parameters. These experiments are not only essential for model

development as they significantly improve parameter confidence. Furthermore, design of experiments aids the development of predictive models whose applicability holds for a wide range of operating conditions (Kontoravdi, Pistikopoulos and Mantalaris, 2010; García-Münzer, Kostoglou, *et al.*, 2015).

An integrated framework for experimental analysis, model development and experimental design has also been suggested (Marquardt, 2005). This framework, called MEXA (Model-based Experimental Analysis) starts with an experimental design, uses it to develop a model and improves the experimental set-up using model results and the needs for estimating model parameters. This approach is analogous to that suggested recently for model-based design-of-experiments in bioprocess development (Abt *et al.,* 2018).

2.3.1.4 Parameter estimation

The procedure of fine-tuning model parameters in order to better describe experimental data is usually called parameter estimation. Two possible approaches can be used: the parameters are either fixed one at a time or jointly estimated given a set of experimental data. Once model reduction is achieved, only a certain subset of the parameters need to be estimated, while others may remain at previously set nominal values (Almquist *et al.*, 2014). Two strategies commonly used for parameter estimation are the Maximum Likelihood Estimation (MLE) algorithm or Bayesian algorithms.

The MLE algorithm implemented in gProms follows equation (5), where N is the total number of experimental measurements, q is the parameter vector, NE, NV and NM are the numbers of experiments, variables and measurements performed, respectively, $\sigma^2_{i,j,k}$ is the standard deviation of the k^{th} measurement, of the j^{th} variable of the i^{th} experiment, $\bar{z}_{\text{i,j,k}}$ is the k^{th} value of the j^{th} variable in the i^{th} experiment and z_{ijk} is the corresponding model predicted value.

$$\Phi = \frac{N}{2}\ln(2\pi) + \frac{1}{2}\min_{\theta}\left\{\sum_{i=1}^{NE}\sum_{j=1}^{NV_i}\sum_{k=1}^{NM_{i,j}}\left[\ln\sigma_{i,j,k}^2 + \frac{\left(\overline{z_{i,j,k}} - z_{i,j,k}\right)^2}{\sigma_{i,j,k}^2}\right]\right\} \quad (5)$$

One of the caveats of using the MLE method is that results depend on the initial guess provided by the user. On the contrary, Baysian parameter estimation methods use the information obtained at each run to improve the algorithm and hence do not require as accurate an initial guess as the deterministic methods, such as the MLE. In this approach, model parameters are treated as random variables and, using Bayes's rule, the posterior, i.e. the probability density function of the parameters given the data is given in equation (6) where θ is the parameter vector and Z is the experimental data (Almquist *et al.*, 2014).

$$p(\theta|Z) = \frac{p(Z|\theta).p(\theta)}{p(Z)} \quad (6)$$

2.3.1.5 Validation and use

Validation of the model must be accomplished using experimental data from an independent experiment which has not been used for model construction (Almquist *et al.*, 2014), namely for parameter estimation. Once such an orthogonal experiment has been performed so as to validate the model (i.e. to show that the model can in fact describe the system and not just one experiment) and the model has been shown able to describe the data, it can be used for optimization or control.

One example of the use of model-based control has been described by Lu *et al.* in 2013. The authors describe an algorithm for automated dynamic feeding strategies. Every 3 days, the nutrient concentrations are analysed, and the consumption rates are calculated. The authors choose a single nutrient as an

indicator of cell metabolism and normalize the feed composition with respect to this nutrient. This iterative process is repeated until the end of the culture (Lu *et al.*, 2013). This method requires little *a-priori* knowledge of the system to be optimized but optimization is limited by the use of one single component to calculate the feeding. According to the authors this is an advantage because several cell lines are to be tested and under different conditions making it impossible to determine the feeding beforehand (Lu *et al.*, 2013). In fact, several components can be limiting at the same time and, as seen above, metabolic shifts can occur. Models based on significant metabolic understanding avoid this by predicting the whole culture duration ahead of time and pre-determining the culture behaviour. If accurately validated, such models are therefore advantageous. The work of Quiroga-Campano *et al.* has recently shown that the determination of optimal feeding strategies can be done beforehand using clone-, culture media- and molecule-specific modelling with great improvements on performance (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

2.4 MULTIVARIATE ANALYSIS OF CELL CULTURE

MVA of metabolites in samples collected during cell culture has been used to increase process understanding and information available for process development.

The main techniques used are hierarchical clustering (HCL) and principal component analysis (PCA). HCL is a classification technique widely used for data analysis and pattern recognition which evaluates the closeness of k observations based on their N coordinates using a given metric. After calculating the distance between any two observations in the N-dimensional space, the algorithm usually produces a dendrogram. The traditional way to compute the distance (L) between 2 observations (x_a and x_b) in an N-dimensional space is the Minkowski distance (equation (7)), which is, in fact, a family of metrics (one for each value of p). The

widely used Manhattan and Euclidean metrics are the particular solutions for *p=1* and *p=2*, respectively (Murtagh and Contreras, 2012).

$$L_p(x_a;x_b)=\left(\sum_{i=1}^{N}|x_{i,a}-x_{i,b}|^p\right)^{1/p} \quad ; \forall p\geq 1;\ p\in\mathbb{Z}^+ \tag{7}$$

Principal Component Analysis (PCA) is a multivariate analysis technique which looks at the data in the space defined by the (orthogonal) directions of largest variance. The first principal component (PC1) is therefore the direction of maximum data variance, the second component is the second direction of maximal variance subject to being orthogonal to PC1 etc. The number of principal components to use in a PCA is usually dictated by a pre-set threshold on the fraction of the total variance to be described (80 or 90% e.g.).

PCA is a useful technique to reduce the dimensionality of data but, at the same, it also looses information as the variance is not fully captured. Hence, its combined use with HCL (where all information is kept) is important to ensure a consistent interpretation of results (Klontzas *et al.*, 2017; Vernardis *et al.*, 2017). Both techniques are unsupervised since no information on sampling times, system conditions, number of clusters or any other is provided apart from the coordinates of each observation in the *N*-dimensional system.

Goals achieved in prior studies with MVA of cell culture data include identification of metabolic features related to different culture media (Dietmair *et al.*, 2012), different productivity levels (Khoo and Al-Rubeai, 2007; Popp *et al.,* 2016), cell cycle arrest (Khoo and Al-Rubeai, 2007), different time points or different culture phases even if not in chronological order (Khoo and Al-Rubeai, 2007; Chrysanthopoulos, Goudar and Klapa, 2010; Dietmair *et al.,* 2012; Vernardis, Goudar and Klapa, 2013), identification of good process indicators (Le *et al.*, 2012) and development of feeding strategies (Kyriakopoulos and Kontoravdi, 2014).

Chapter 3

Materials and Methods

This chapter describes the materials and methods employed in the course of the work herein presented, comprising experimental (3.1 and 3.2) and computational (3.3) methods. All medium supplements and chemicals used were purchased from Sigma-Aldrich, unless stated otherwise.

3.1 CULTURE OF GS-NS0 AND GS-CHO CELLS

Batch cultures of GS-NS0 and GS-CHO cells (Lonza Biologics) were conducted at 20% working volume in 250mL or 1L disposable Erlenmeyer flasks (Corning). Culture flasks were incubated at 37°C, 5% CO_2 and 100% humidity in an AutoFlow 5500E incubator (Nuaire) and orbitally shaken at 130 rpm on a Shaker-07 (SciQuip). Cells were harvested by centrifugation in a Rotina 46 centrifuge (Hettich). Both cell lines produce the cB72.3 IgG4 monoclonal antibody.

3.1.1 MEDIA PREPARATION AND CULTURE CONDITIONS

All GS-NS0 and GS-CHO cultures were performed in serum-free conditions using, respectively, the CD Hybridoma or the CD CHO chemically defined media (both from Life Technologies) supplemented with 1% penicillin/streptomycin and 25µM methylsulphoximine (MSX) to maintain the selective pressure on the GS system. Both media were sterilized by filtration using sterile filter units (Corning) under sterile conditions, aliquoted and stored at 4-8°C for up to 12 days before use.

The CD Hybridoma medium used for GS-NS0 cell cultures was further supplemented with 2mL of cholesterol 250x (Life Technologies), 20mL of GS supplement 50x, 1mL of serum replacement 50x and 1mL of a 1M sodium hydroxide (NaOH) solution to a final pH of 7.2.

3.1.2 INOCULUM PREPARATION

One cryovial containing cells preserved in liquid nitrogen was warmed up to 37°C. When thawed, the whole content of the vial was mixed with 10mL of fresh medium previously warmed up to the same temperature. This suspension was centrifuged at 800 rpm for 5 minutes and the supernatant was removed. The pellet was re-suspended in 25mL of fresh medium (at 37°C) and cultured for 2 days (GS-NS0) or 4 days (GS-CHO) in a 250mL flask. Cell density and viability were assessed immediately after inoculation by haemocytometry using the erythrosine B (ATCC) dye-exclusion method and counting cells under a reversed phase optical microscope (Leica) at 100x magnification.

3.1.3 PREPARATION OF CELL BANKS

GS-NS0 and GS-CHO cell banks were prepared in 5 passages. The GS-NS0 cell bank was obtained from a serum-free adapted clone developed in our group (Quiroga-Campano, 2016) and cells were sub-cultured every 48 hours. The GS-CHO cell bank was prepared from a vial obtained from Lonza Biologics and cells were sub-cultured every four days, following manufacturer's instructions.

The target viable cell density (VCD) to start each passage was $2.5x10^5$ cells/mL for both cell lines. The required volume of culture was calculated based on the VCDs assessed by haemocytometry. At each passage, a sample of the appropriate volume was taken from the culture obtained in the previous passage and centrifuged at 800rpm for 3 min. The supernatant was removed, and cells were re-suspended in pre-warmed fresh medium. Cells were cultured in 50mL for two passages, in 2x50mL for the third passage, in 3x50mL for the fourth passage and

finally in 3x200mL for the last passage. Cells were cryopreserved for future use in the same medium used for subculture but supplemented with dimethylsulfoxide (DMSO) at 10% (GS-NS0 cells) or 7.5% (GS-CHO cells) by volume. The cryopreservation medium for the GS-CHO cells did not contain MSX, following manufacturer's instructions.

At the last passage, cells in all flasks were counted by haemocytometry to assess the VCD. The volume of culture required to prepare the desired number of vials was taken and centrifuged for 5 minutes at 800 rpm. The supernatant was removed, and cells in each tube were re-suspended in the appropriate volume of cryopreservation media to a final cell density of $1x10^7$ cells/mL (GS-NS0) or $1.5x10^7$ cells/mL (GS-CHO). Cells were aliquoted in 2mL cryovials (VWR) containing 1mL of suspension. The vials were placed inside Mister Frosty cryopreservation boxes at –20°C and transferred to –80°C after 2h. After 24h, the tubes were transferred to the cell bank in liquid nitrogen.

3.1.4 BATCH CULTURES

An inoculum culture was prepared as described above and expanded in three passages starting at $2.5x10^5$ cells/mL. From the last passage, three cultures were prepared at 1, 2 and $4x10^5$ cells/mL initial cell densities in 1L Erlenmeyer flasks and grown in batch mode for 144h (GS-NS0) or 240h (GS-CHO).

3.2 ASSAYS

Every 24h, a 3mL sample was taken from each culture. From this sample, 500µL of cell suspension (not centrifuged) were used for at-line analysis of viable, apoptotic and dead cell densities and ATP content. The remaining 2.5mL were centrifuged at 800rpm for 5 minutes. The supernatant was used for metabolite and mAb concentration measurements and the pellet was washed with the same volume of PBS before cells were preserved adequately for cell cycle and gene expression analysis, as described below.

3.2.1 CELL DENSITY AND VIABILITY

Viable, apoptotic and dead cell densities were analysed by flow cytometry. Every 24h, a culture sample was taken, appropriately stained with ViaCount reagent (Merck) as per manufacturer's instructions, incubated at room temperature for 10 minutes and run in a Guava 8HT benchtop flow cytometer (Merck). Data was analysed using GuavaSoft 3.1 (Merck).

3.2.2 ATP MEASUREMENTS

The concentration of ATP was assessed using the CellTiter Glo luminescence assay (Promega). The CellTiter Glo reagent was added 1:1 to 100µL of cell suspension in a 96-well plate. The luciferase enzyme present in the reagent catalyses the reaction of luciferin with ATP, producing a luminescent signal proportional to the ATP concentration in the sample. A calibration curve of ATP concentration vs. luminescence was prepared using serial dilutions of an ATP standard solution (Promega) to obtain ATP concentrations ranging from $1x10^{-4}$ to $1x10^{-7}$ M. A blank was also prepared using cell culture media and subtracted from all signals. Samples were orbitally shaken for 2 minutes, incubated at 25°C for 10 minutes and the luminescent signal was read in a plate reader (Promega). For each sample, the corresponding filtered supernatant (cell-free) was used as a blank and its value was discounted from all measurements, which were conducted in triplicates.

3.2.3 GLUCOSE, LACTATE AND AMMONIA QUANTIFICATION

For each time point, a 1mL supernatant sample was collected from culture and preserved at –20°C. At the time of analysis, samples were thawed at room temperature and analysed using a BioProfile Flex (NovaBiomedical).

3.2.4 AMINO ACID QUANTIFICATION

A culture supernatant sample was filtered using 10kDa cut-off ultrafiltration centrifugal filter units (Ultracell YM 10, Merck) for 1h at 4°C in a Heraeus Fresco

17 micro centrifuge (ThermoScientific) and stored at –20°C for up to 1 month prior to analysis. At the time of analysis, samples were thawed at room temperature and analysed in an Infinity 1260 HPLC (Agilent Technologies). According to the manufacturer's protocol, samples were derivatised in-line with ortho-phthaldehyde (OPA) for primary amino-acids and 9-fluorenylmethyl chloroformate (FMOC) for secondary amino acids, both from Agilent Technologies. Samples were then mixed with injection diluent (50mL mobile phase A with 200µL of H_3PO_4) and 1.0µL was injected in a reversed-phase Zorbax Eclipse Plus 4.6 x 100mm C18, 1.8µm column with a guard column (4.6 x 0.2 mm), all from Agilent Technologies.

The mobile phases were (A) 10mM Na_2HPO_4, 10mM $Na_2B_4O_7.10H_2O$, 0.5mM NaN_3 pH 8.4 in DI water and (B) acetonitrile, methanol and DI water (45:45:10 % v/v). The gradient programme was 98% A from 0 to 0.35 minutes followed by a linear gradient to 43% A at 13.4 minutes; a step decrease to 0% A at 13.5 minutes and finally a step increase to 98% A at 15.8 minutes. The flow rate was kept constant at 1.5mL/min and all samples were processed at 40±0.8°C. Primary amino acids were detected at 338nm and secondary amino acids at 262nm. Results were analysed in the OpenLab CDS software (Agilent Technologies).

An amino acid standard (Sigma) was supplemented with 30mM Asparagine, Tryptophan and Glutamine solutions (in 0.1M HCl) to obtain a stock solution with all required amino acids. The stock solution was serially diluted to prepare calibration curves of concentration vs. peak area for all amino acids.

3.2.5 MONOCLONAL ANTIBODY QUANTIFICATION

Monoclonal antibody titre was measured using an analytical protein-A monolithic chromatography column in an Affinity 1260 HPLC (both from Agilent Technologies). A culture supernatant sample was collected at each time point, filtered through a 0.2µm syringe filter (VWR) and stored at –20°C for up to 4 months prior to analysis. After thawing at room temperature, samples (50µL) were injected into the column at 1.0mL/min under 50mM phosphate buffer pH 7.4

(binding buffer) at 20±0.8°C. Elution was promoted for 3.5 minutes by a step change to 100% 0.1M acetic acid, pH 2.6. The mobile phase composition was then reversed to 100% binding buffer. The monoclonal antibody is detected by UV at 280nm. The column cleaning-in-place protocol was run once every seven samples. Results were analysed in OpenLab CDS software (Agilent Technologies) and compared to a calibration curve of peak area vs. mAb concentration obtained by serially diluting a cB72.3 IgG4 standard solution (Lonza Biologics) in binding buffer.

3.2.6 CELL CYCLE ASSAYS

Flow cytometry was used to measure the expression of cyclins E1, B1 and D1, the proliferation marker Ki-67, thymidine analogue EdU (5-ethynyl-2-deoxyuridine) and to assess the cell cycle distributions. A volume of the PBS suspension obtained at each time-point containing approximately $5x10^5$ cells was centrifuged and the cells were resuspended in 200µL of ice-cold 70% ethanol and stored at –20°C, for up to 1 month prior to analysis. At the time of analysis, the samples were centrifuged for 5 minutes at 5000rpm and the ethanol was removed. The pellet was washed with 200µL of PBS (cell cycle distribution assay only) or PBS 1%BSA (EdU, Ki-67 and cyclins assays) and centrifuged again under the same conditions. The protocol continued with the EdU click chemistry reaction, cyclin and Ki-67 staining or DNA staining.

3.2.6.1 EdU click chemistry reaction

For cultures using the EdU proliferation marker, cells in mid-exponential growth were exposed to 30µM EdU for 3 hours (García Münzer *et al.*, 2014) by supplementing the culture with the appropriate volume of a 10mM EdU solution in PBS (Invitrogen). Cells were then washed twice with PBS and finally cultured in fresh medium and sampled every 2 hours for 24h(NS0) or 30h (CHO). At the time of analysis, a click chemistry reaction cocktail containing Alexa Fluor 647

(Invitrogen) was prepared according to the supplier's instructions and added to the cells. After a 30-minute incubation in the dark at room temperature, cells were washed with a saponin-based washing reagent (part of the analysis kit). The protocol continued with cyclins, Ki-67 or DNA staining.

3.2.6.2 Cyclins and Ki-67 Staining

Cyclin E1 was stained using a FITC-conjugated rabbit anti-human polyclonal antibody (Bioss Antibodies); Cyclin B1 was stained using a FITC-conjugated mouse anti-human antibody (BD Biosciences); Cyclin D1 was stained using a FITC-conjugated mouse anti-human antibody (BD Biosciences) and Ki-67 was stained using an AlexaFluor 647-conjugated mouse anti-human antibody (Biolegend). After incubation in the dark, at room temperature, for 20 minutes, cells were washed with cell staining buffer (Biolegend) before DNA staining.

3.2.6.3 DNA staining

DNA was stained with propidium iodide (PI) using the Guava CellCycle reagent (Merck). Cells were resuspended in 200µL reagent and incubated for 20 minutes in the dark. Finally, samples were filtered using 40µm cell strainers (Falcon) into a transparent round-bottom 96-well plate before analysis.

3.2.6.4 Flow cytometry configurations and workflow

Samples stained for DNA only or DNA and EdU only were analysed using the Guava cytometer. The PI signal (DNA) was read from the 583/26nm emission filter using a 488nm excitation wavelength while the AlexaFluor 647 signal (EdU) was measured from the 664/20nm emission filter with a 640nm excitation wavelength.

All other samples were analysed using a Fortessa II cytometer (BD Biosciences). The 610/20nm emission filter and the 561nm excitation laser were used for PI, a 488nm wavelength and the 530/30nm emission filter were used for

the FITC-conjugated cyclins and an excitation wavelength of 640nm was used with an emission filter of 670/14 nm for the AlexaFluor 647-stained EdU or Ki-67. Single colour controls were used for compensation. The panels used are described in Table 1. After acquiring 10000 events for each sample, data was analysed in FlowJo v10.0.

Table 1. Flow cytometry panels used for cell cycle, EdU, cyclins and Ki-67 analyses

Cytometer	**Stains (target)**	**Excitation Laser (nm)**	**Emission Filter (nm)**
Guava 8HT	PI (DNA)	488	583/26
	AlexaFluor 647 (EdU)	642	664/20
FortessaII	PI (DNA)	561	610/20
	FITC (Cyclins)	488	530/30
	AlexaFluor 647 (Ki67 or EdU)	640	670/14

The flow cytometry workflow started with gating cells from debris using a forward scatter vs. side scatter plot (Figure 12A) and then gating singlets from doublets using a forward scatter height vs. area plot (Figure 12B). Viable cells were gated from apoptotic cells in a PI histogram (Figure 12C). Using fluorescence minus one (FMO) controls, the EdU-, cyclins- and Ki-67 negative gates were identified. The positive gates were obtained by reverting the negative gates, which were then used in the corresponding DNA bivariate plots (Figure 12D). The FMO controls used are listed in Table 2. The cell cycle distributions were assessed using the software tool for DNA deconvolution (Figure 12E). For cyclin analysis, the cyclin-positive gates were applied to each cell cycle phase to obtain the fraction of positive cells. Finally, cyclin expressions were normalized to the cell cycle phase where their expression is assumed to be minimal, specifically phase G_1 for cyclin B1, phase S for cyclin D1 and phase G_2 for cyclin E1 (García-Münzer, Kostoglou, *et al.*, 2015).

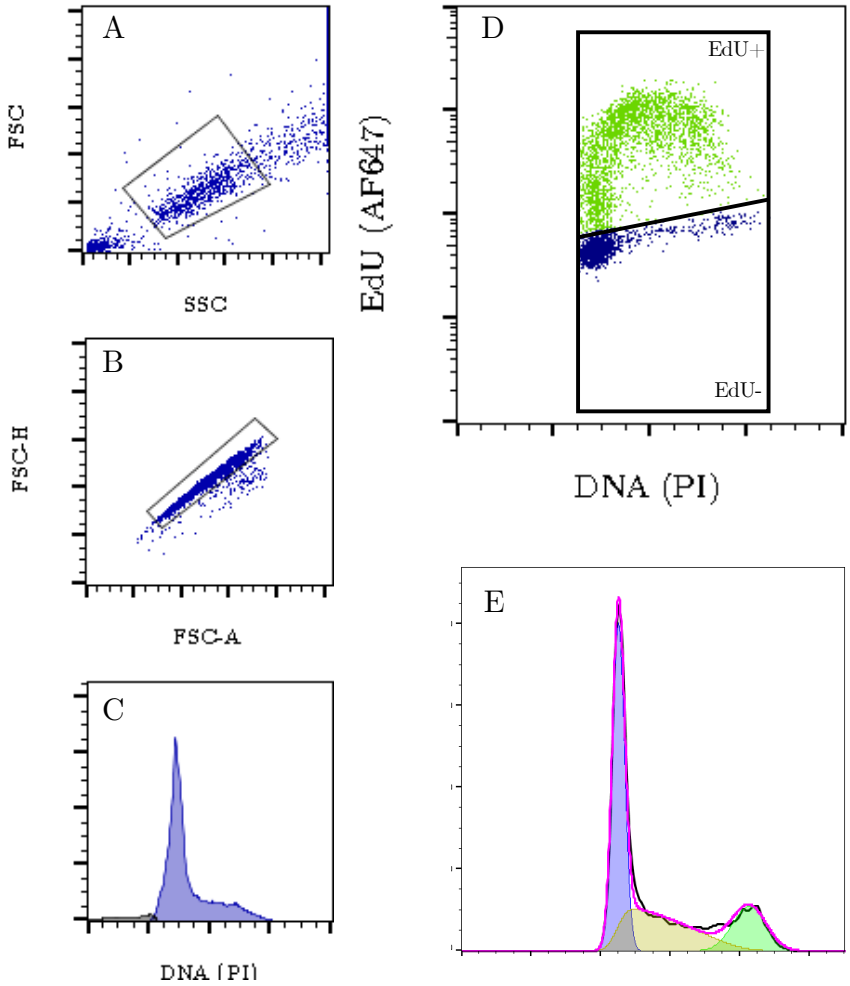


Figure 12: Flow cytometry work flow. Cells were gated from debris (A), singlets were gated from doublets (B) and viable cells were gated from apoptotic cells (C). FMO controls were used to gate EdU, cyclins and Ki-67 positive regions (D) and the DNA was deconvoluted using the software (E).

Table 2. FMO controls and sample configurations for which they were used.

Gate obtained	**FMO**	**Sample staining**
Cyclin negative	DNA + Ki67	DNA + Ki67 + Cyclin
	DNA + EdU	DNA + EdU + Cyclin
Ki-67 negative	DNA	DNA + Ki67
	DNA + Cyclin	DNA + Cyclin + Ki67
EdU negative	DNA	DNA + EdU
	DNA + Cyclin	DNA + Cyclin + EdU

3.2.7 GENE EXPRESSION ANALYSES

Expression of four genes involved in apoptosis, namely *bax*, *bcl-2*, *casp-8* and *casp-3* was analysed by q-RT-PCR. Every 24h, a sample containing approximately $5x10^5$ cells was taken, the supernatant was removed, and the cells were washed with PBS, resuspended in RNAprotect reagent (Qiagen) and stored at –20°C for up to 6 months prior to RNA extraction.

RNA was extracted using the RNeasy RNA extraction kit (Qiagen) according the manufacturer's instructions. Briefly, cell lysis was induced chemically with kit RLT buffer supplemented with 1% β-mercaptoethanol and the lysate was homogenised with a shredder (Qiagen). The material was loaded onto the purification column, washed twice with the appropriate kit buffers and eluted with 2 x 40µL RNAse-free water (Qiagen). RNA concentrations were measured by the absorbance of the purified RNA solution at 260nm, using a nanodrop spectrophotometer (ThermoScientific). Purified RNA solutions were frozen at -80°C for up to 6 months prior to analysis.

Quantitative, real-time Reverse Transcriptase Polymerase Chain Reaction (q-RT-PCR) was conducted using the Kappa Sybrfast PCR kit (Sigma-Aldrich) and forward and reverse primers (shown in Table 3, purchased from Invitrogen).

The reaction mixture contained the Sybrgreen kit master mix (nucleotides, buffer, polymerase), 0.04 nmol of each primer (forward and reverse) for the gene of interest and the reverse transcriptase. The total reaction volume was 20µL. The amplification reaction was run in a StepOnePlus thermocycler (Applied Biosystems) with a 2-stage holding programme: 5 minutes at 42°C and 3 minutes at 95°C followed by the cycling programme comprising of 3 seconds at 95°C and finally 30 seconds at 62°C where data was acquired. All samples were run in triplicates and negative controls were included. The $\Delta\Delta C_T$ method was used for data analysis normalizing the expressions of the four genes of interest to the expression of the house keeping gene, glyceraldehyde 3-phosphate

dehydrogenase (*gapdh)* (Figueroa *et al.*, 2007; Browne and Al-Rubeai, 2011), at each time point. All time points were normalized to the beginning of the culture.

Table 3. Primers used for q-RT-PCR and references of their prior use.

Cell Line	Gene	Direction	Sequence	Reference
NS0/CHO	bcl-2	Forward	GGGAGAACAGGGTATGATAACCG	(Li *et al.*, 2013)
		Reverse	TAGCCCCTCTGTGACAGCTTA	
NS0	bax	Forward	TGAAGACAGGGGCCTTTTTG	(Li *et al.*, 2016)
		Reverse	AATTCGCCGGAGACACTCG	
CHO	bax	Forward	AGGCGGCGGGCCCACCAGCTC	(Lim *et al.*, 2006)
		Reverse	CATCAGCAAACATGTCAGCTG	
NS0	casp8	Forward	CAACTTCCTAGACTGCAACCG	(Wang *et al.*, 2017)
		Reverse	TCCAACTCGCTCACTTCTTCT	
CHO	casp8	Forward	AACCCAAAATCTTTTTTATTCAGGCTTG	(Yun *et al.*, 2007a)
		Reverse	GTCAACCATGTAGAGCACACATC	
NS0/CHO	casp3	Forward	ATGGAGAACAACAAAACCTCAGT	(Wang *et al.*, 2017)
		Reverse	TTGCTCCCATGTATGGTCTTTAC	
NS0/CHO	gapdh	Forward	AGGTCGGTGTGAACGGATTTG	(Figueroa *et al.*, 2007)
		Reverse	TGTAGACCATGTAGTTGAGGTCA	

3.2.8 STATISTICAL ANALYSES

Mean and standard deviations of experimental data were computed in Microsoft Excel, using the AVERAGE and STDEV functions, respectively. All other statistical analyses were conducted in SPSS (IBM) using one- or two-way ANOVA, as appropriate, with a level of significance of 0.05.

Hierarchical clustering with Euclidean distance metric and Principal Component Analysis (PCA) were used as unsupervised methods for multivariate data analysis of metabolite profiles (Vernardis, Goudar and Klapa, 2013). Concentrations of metabolites were autoscaled to a normal distribution with zero mean and unit standard deviation using equation (8), whereby, for each metabolite, *m,* the normalized values *($\hat{m}$)* were calculated by subtracting the average value across all samples ($\bar{m}$) and dividing by the standard deviation (σ_m) *(van den Berg et al.,* 2006). SPSS was used for these analyses.

$$\widehat{m} = \frac{m - \overline{m}}{\sigma_m} \tag{8}$$

3.3 COMPUTATIONAL METHODS

3.3.1 MODEL IMPLEMENTATION

Mathematical models were implemented in gProms 4.2.0 (Process Systems Enterprise) running on a desktop personal computer with a 3.4 GHz Intel Core i7 processor, 16 GB of RAM memory running Windows 7, 64-bit.

3.3.2 GLOBAL SENSITIVITY ANALYSIS

Global Sensitivity Analysis (GSA) was run using the gO:Matlab interface tool (Process Systems Enterprise), which enables data exchange between gProms and Matlab (R2014a version, Mathworks). For all model parameters, nominal values were taken from previous works (García-Münzer, Kostoglou, *et al.*, 2015; Usaku, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018) and lower and upper bounds were defined as ±50% of the nominal value. For each simulation, values for all parameters were obtained using the Sobol quasi-random method (*sobolset* function in Matlab) and sent to gProms. At the end of each simulation, the model outputs were sent from gProms and stored in Matlab.

The Sobol sensitivity indices (SI) were used to rank the parameters. SIs were estimated using the computationally efficient High Dimensional Model Representation (HDMR) method (Ziehn, 2008), implemented into a publicly available toolbox (Ziehn and Tomlin, 2009). For each model parameter, the corresponding SI is calculated as the ratio of output variance caused by varying this parameter to the total output variance (Sobol, 2001). Therefore, SIs vary between 0 and 1 and a threshold of 0.1 has been set do discriminate important parameters (Sidoli, Mantalaris and Asprey, 2005; García-Münzer, Kostoglou, *et al.*, 2015). The criteria for setting the number of model simulations used for GSA was the convergence of SI indices, namely, that the SIs calculated for all parameters sum to unity. The values of important parameters were re-estimated while those of non-important parameters were left at their nominal values.

3.3.3 PARAMETER ESTIMATION AND VALIDATION

Significant parameters were estimated using the maximum likelihood algorithm implemented in gProms 4.2.0 using initial values from previous works or from experimental data obtained in this work. The variance method was set as constant relative variance with a value equal to the coefficient of variation obtained experimentally for each variable.

Model simulations for the data-set used for parameter estimation were compared to experimental data. In order to assess model's predictive capabilities, model predictions of two independent data-sets were also compared to the corresponding experimental data, without changing the parameter values obtained from parameter estimation.

Chapter 4

A predictive mathematical model of cell cycle, metabolism and apoptosis of GS-NS0 cells.

In the present chapter, an integrated mathematical model describing cell cycle, metabolism and apoptosis of mAb-producing GS-NS0 cells is presented. Following the model development framework presented earlier (Figure 11) global sensitivity analysis (GSA) is used to identify the significant parameters which are subsequently estimated using a maximum likelihood algorithm. Finally, independent experimental data sets are used to show the model's predictive capabilities. The model structure developed in this chapter will be combined with experimental data presented in Chapter 5 and adapted to develop a mathematical model for a different organism, presented in Chapter 6.

4.1 INTRODUCTION

The increasing competition in the mAbs market substantiates the need for tools enabling accelerated process development thereby reducing time-to-market and maximizing profit, while capable of supporting product- and host organism-specific decisions. Process development for biologics has traditionally relied on

intensive experimental work to determine optimal operating conditions. Recently, design of experiments (DoE) methodologies have become established as a means to identify optimal conditions and reduce time and resources spent (Sokolov *et al.*, 2017; Galbraith *et al.*, 2018). However, DoE remains an experimentation-based procedure with no guaranteed global optimality. For upstream processing, common optimization goals include identifying optimal media formulations (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018), feeding strategies (Kiparissides, Pistikopoulos and Mantalaris, 2015) and bioreactor designs (Haringa *et al.*, 2018) all of which could be more efficiently achieved using mathematical modelling and model-based optimization. This approach allows for achieving an optimized cell culture system and for obtaining results at a reduced cost and in shorter times. The development of predictive mathematical models capturing the biology of the host organism is central to this goal as is an appropriate choice of mathematical formulations able to capture population heterogeneity and dynamics (Kyriakopoulos, Ang, Lakshmanan, Huang, Yoon, Gunawan and Lee, 2018).

Metabolism, cell growth, cell cycle, cell death, and recombinant protein production are among the most relevant biological processes to be captured by a mathematical model of mammalian cell culture for mAb production. Different approaches have been taken to develop mathematical models of the cell cycle including the use of empirical probabilistic equations without biological significance (Jandt *et al.*, 2015), ordinary differential equations (García-Münzer, Ivarsson, *et al.*, 2015), or PBEs taking into account experimental determination of cycling times and suitable biomarkers of cell cycle as well as the impact of metabolism on their expression (García-Münzer *et al.*, 2013; García-Münzer, Kostoglou, *et al.*, 2015). Likewise, metabolic models of different kinds are also available. Kinetic metabolic models describing glucose and amino acid consumption, the formation of metabolites and protein production (Kontoravdi, Wong, *et al.*, 2007; Almquist *et al.*, 2014; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018) are generally preferred due to describing metabolism in a dynamic manner and to having the

potential to be used for process optimization (Kyriakopoulos, Ang, Lakshmanan, Huang, Yoon, Gunawan and Lee., 2018). In addition, metabolic flux analysis studies describing the active metabolic pathways at a particular stage of culture and quantifying the fluxes have also been attempted (Carinhas *et al.*, 2013; Templeton *et al.*, 2013). Although using the results of metabolic flux analysis to describe metabolism throughout culture has been attempted, model predictions deviate from experimental data particularly when attempting to describe metabolic shifts such the lactate, glutamate or alanine shifts in CHO cell culture (Robitaille, Chen and Jolicoeur, 2015). Instead, metabolic flux analysis can be used to support the development of kinetic models to be used in process optimization (Galbraith *et al.*, 2018).

Finally, mathematical models of apoptosis have also been attempted focusing on either the extrinsic pathway (Pace *et al.*, 2010) or the intrinsic pathway (Zhang, Brazhnik and Tyson, 2009) or combining elements of both (Hua *et al.*, 2005; Stucki and Simon, 2005; Okazaki *et al.*, 2008). Combined models of apoptosis usually compromise on the detail of the biological information included to avoid obtaining untraceable models. For example, Usaku used 6 genes to describe apoptosis in GS-NS0 cells (Usaku, 2015), while Meshram *et al.* considered only the expression of two caspase proteins (Meshram *et al.*, 2012).

In studies published in literature so far, the cell cycle, metabolism and apoptosis of mammalian cells have been modelled separately and, hence, an integrated model linking all these biological functions is lacking. Cell cycle and metabolism of mammalian cells have been modelled together establishing a link between energy metabolism and the cell cycle, recently shown to occur through amino acid sensing (Lee and Finkel, 2013). However, this work is limited by segregation using age and volume, both shown to have little biological significance, and by lumping all amino acids (AA) into one term, whereas requirements for each nutrient are different, hence preventing its use for culture optimization (Martens *et al.*, 1995). Therefore, there is scope for the development of an integrated model

providing a combined description of these cellular functions. Although the use of NS0 cells is phasing out, there is a significant amount of knowledge, both experimental and computational, to support the work and therefore this organism is chosen for the first integrated model of cell cycle, metabolism and apoptosis (Tey and Al-Rubeai, 2005; Swiderek and Al-Rubeai, 2007; García-Münzer *et al.*, 2013; García-Münzer, Kostoglou, *et al.*, 2015; Kiparissides, Pistikopoulos and Mantalaris, 2015; Usaku, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018)

4.2 OBJECTIVES

The aim of the present chapter is to provide a predictive mathematical model of cell cycle, metabolism and apoptosis in mAb-producing GS-NS0 cells which remains computationally inexpensive and hence useful for model-based cell culture optimization. To this end, the following objectives are set:

- Formulate an integrated dynamic mathematical model of cell-cycle, metabolism and apoptosis of GS-NS0 cells with fast computational solution;
- Perform GSA to study model structure, identify the significant model parameters and design meaningful experiments;
- Conduct serum-free batch cultures of GS-NS0 cells to obtain data to estimate the important model parameters using a maximum likelihood parameter estimation algorithm;
- Use independent experimental data sets to assess model predictive capabilities.
- Identify an unsupervised multivariate data analysis tool to link metabolic profiles to growth, death, productivity and cell-cycle of GS-NS0 cells.

4.3 MODEL STRUCTURE

In this section, the mathematical formulation used in the present chapter is presented in three subsections describing cell cycle, metabolism and apoptosis. This work is based on experimental knowledge and detailed mathematical models of cell cycle, metabolism and apoptosis of GS-NS0 cells (García-Münzer, Kostoglou, *et al.*, 2015; Usaku, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

4.3.1 CELL CYCLE AND VIABLE CELLS

The mathematical model of cell cycle used in this work is adapted from a prior work with the same cell line (García-Münzer, Kostoglou, *et al.*, 2015). The original model is based on experimental data and has been experimentally validated under three different conditions (asynchronous growth, DMSO-arrested cells and Thymidine-arrested cells). This model lacks information on apoptosis, the main cause of death in bioreactors (Grilo and Mantalaris, 2019a), and has only limited information on metabolism, including only glucose, glutamate, lactate and ammonia.

$$\frac{\partial N_{G1}(cycE,t)}{\partial t}+\frac{\partial}{\partial cycE,}[r_{G1}(cycE,S)N_{G1}(cycE,t)]= -\Gamma_{G1}(cycE,S)N_{G1}(cycE,t)$$
$$-[\mu_A(g)+\mu_D(E)]N_{G1}(cycE,t)+2\delta(cycE)\int_{cycB,min}^{cycB,max}\Gamma_{G2}(cycB,S)N_{G2}(cycB,t)\,dcycB \quad (9)$$

$$\frac{\partial N_S(DNA,t)}{\partial t}+\frac{\partial}{\partial DNA,}[r_S(DNA,S)N_S(DNA,t)]= -[\mu_A(g)+\mu_D(E)]N_S(DNA,t)+$$
$$+\ \delta(DNA\text{-}1)\int_{cycE,min}^{cycE,max}\Gamma_{G1}(cycE,S)N_{G1}(cycE,t)\,dcycE \quad (10)$$

$$\frac{\partial(cycB,t)}{\partial t}+\frac{\partial}{\partial cycB,}[r_{G2}(cycB,S)N_{G2}(cycB,t)]= -\Gamma_{G2}(cycB,S)N_{G2}(cycB,t)$$
$$-[\mu_A(g)+\mu_D(E)]N_{G2}(cycB,t)+\delta(cycB\text{-}cycB_S)r_S(DNA{=}2,S)N_S(DNA{=}2,t) \quad (11)$$

The basic structure of the cell-cycle model, comprising one PBE for each phase of the cell cycle, was maintained. Equations (9) to (11) describe phase G_0/G_1, phase S and phase G_2/M, respectively. Cyclin E1, DNA and cyclin B1 are used to distribute the population at each stage of the cycle based on prior experimental evidence (García-Münzer *et al.,* 2013).

Cells begin cycling in phase G_0/G_1 (described by equation (9)) which is tracked using cyclin E1 (cycE). They progress through this phase at a rate r_{G1} as cyclin E1 accumulates in the cells. Above a certain threshold, determined experimentally, they may transition to phase S. This transition occurs at a rate, Γ_{G1}. Once cells enter phase S (equation (10)), progression, at a rate r_s is determined by DNA replication. Upon entrance into phase S, the DNA content is 1, as expressed by the Dirac delta function in front of the integral in equation (10), while the cyclin E1 content can be variable, provided it is higher than the transition threshold determined experimentally (García-Münzer *et al.,* 2013), and hence the integral term in equation (10). Once DNA replication is complete, i.e. DNA=2, cells immediately enter phase G_2/M. Progression through this phase occurs at a rate r_{G2} and depends on the accumulation of cyclin B1 (cycB). As expressed by the Dirac delta function in equation (11), production of this cyclin starts during phase S (García-Münzer *et al.,* 2013) reaching a level $cycB_S$. Above a certain cyclin B1 threshold, they will leave phase G_2 at a rate Γ_{G2} and enter phase G_0/G_1 with a minimum value of cyclin E1 (García-Münzer *et al.,* 2013). Each cell leaving phase G_2/M gives birth to 2 identical daughter cells which will start cycling in phase G_0/G_1, and hence the number 2 in front of the integral term in equation (9)).

In equations (1) to (3), $N_{i=G1,S,G2}$ represents the number of cells in each phase "i" of the cell cycle with a given content of the respective internal coordinates (cycE, DNA or cycB). Hence, the first term in the LHS of all equations accounts for the variation of the number of cells in each phase of the cycle with a given content of the internal coordinate over time, while the second term accounts for the progression within each phase of the cycle as a function of the progression rate

(r). These rates are a function of the rate of the accumulation of the corresponding distribution variable – DNA or cyclins – taken as the reciprocal of the minimum time spent in each phase of the cycle (determined experimentally in exponential growth). Nutrient limiting effects may reduce these rates, as described by equations (12) and(13) where nutrient limitations are accounted for by the $f_{limGluArg}$ function, calculated from equation (14), and $t_{i,min}$ are the minimum times spent in each phase (measured experimentally during exponential growth).

$$r_s = \frac{1}{t_{S,min}} \cdot f_{limGluArg} \tag{12}$$

$$r_{i=G1,G2} = \frac{1}{\tau_{k=t_{G1,min},t_{G2,min}}} \cdot f_{limGluArg} \tag{13}$$

$$f_{lim}GluArg = \left(\frac{[Glu]+[Arg]}{[Glu]+[Arg]+K_{GluArg}}\right) \tag{14}$$

The original equations used have been modified to include the dependency of the cell cycle progression rates on arginine, together with glutamate. In the original model, glutamate dependency was included because cyclin expression shows significant differences before and after glutamate exhaustion and because this is the main energy yielding metabolism in this cell line (García-Münzer, Kostoglou, *et al.*, 2015). Detailed analysis of amino acid metabolism in GS-NS0 serum-free cultures confirms the essential role of glutamate and suggests that arginine feeds energy production once glutamate is exhausted (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). The limiting effect of lactate in cyclin expression, suggested in the original formulation of the model, has been neglected in this work because lactate production of GS-NS0 cells in serum-free batch cultures is negligible (Bibila *et al.*, 1994), which has been confirmed in this work, reaching maximum concentrations one order of magnitude lower than the limiting/inhibiting levels reported for this and other mammalian cell lines, which range from 20mM for CHO cells (Hartley *et al.*, 2018), 50mM for GS-NS0 (García-Münzer, Kostoglou, *et al.*,

2015) or 90mM for hybridoma cells (Jang and Barford, 2000). The rates of transition from G_0/G_1 to S and from G_2/M to G_0/G_1 are calculated from equation (15), where the constant probability of transition (p_i) and the cyclin thresholds have be determined from experimental data (García-Münzer *et al.*, 2013; García-Münzer, Kostoglou, *et al.*, 2015) and the effect of glucose in prolonging stationary phase and delaying the decline phase is accounted for by the f_{limGLC} function calculated from equation (16).

$$\Gamma_{i=G_1,G_2} = \begin{cases} 0 & \textit{below cyclin threshold} \\ p_i . f_{lim}Glc & \textit{above cyclin threshold} \end{cases} \tag{15}$$

$$f_{lim}Glc = \left(\frac{[Glc]}{[Glc] + K_{Glc}}\right) \tag{16}$$

The times spent in each phase may vary according to nutrient availability being calculated from equation (17) and used in equation (18) to compute the growth rate.

$$t_{i=G1,G2} = \frac{cyclin\ threshold}{r_{i=G1,G2}} \quad ; \quad t_S = \frac{1}{r_S} \tag{17}$$

$$\mu = \frac{ln(2)}{t_{G_0/G_1} + t_S + t_{G_2/M}} \tag{18}$$

The total population in each phase of the cell cycle is calculated by the integral across all values of the internal coordinate (cyclins or DNA) as shown by equation (19) and the total viable cell density (X_V) is calculated as the sum of cells in each phase of the cell cycle divided by the culture volume, as shown in equation (20).

$$N_{i=G1,S,G2} = \int_{x_{k,min}}^{x_{k,max}} N_i(x_k)dx_k \tag{19}$$

$$X_V = \frac{N_{G_1} + N_S + N_{G_2}}{V} \tag{20}$$

As mentioned above, apoptosis was not part of the original cell cycle model where all forms of cell death were lumped into one term. In the present formulation (equations 6 to 8), the death rate is replaced by two terms, μ_A and μ_D, respectively accounting for apoptotic or necrotic death, which may occur at any phase of the cell cycle (Al-Rubeai *et al.*, 1992; Martens *et al.*, 1995). Apoptotic cells will die at a rate μ_{AD} and dead cells can be lysed at a rate k_{lys}. Consequently, the apoptotic and dead cells populations can be computed from equations (21) and (22).

$$\frac{d(X_A V)}{dt} = \mu_A X_V V - \mu_{AD} X_A V \tag{21}$$

$$\frac{d(X_D V)}{dt} = \mu_{AD} X_A V + \mu_D X_V V - k_{lys} X_D V \tag{22}$$

The apoptotic death rates, μ_A and μ_{AD}, have been linked to the expression of two key genes, namely, caspase-8 (c8) and caspase-3 (c3), respectively. In the original apoptosis model of Usaku, the rate of apoptotic death, μ_A, was linked to the expression of caspase-3 (Usaku, 2015). However, caspase-8 is involved in the early stages of programmed cell death whereas caspase-3 acts later as an executioner caspase (McIlwain, Berger and Mak, 2015) and hence, in this work, μ_A is described using caspase-8 whereas μ_{AD} is described using caspase-3. It has also been considered that, if cells cannot produce the minimum energy that they require (ATP_{crit}), i.e., they are under severe metabolic stress, death occurs through processes other than apoptosis (Eguchi, Shimizu and Tsujimoto, 1997; Grilo and Mantalaris, 2019a). The corresponding mathematical expressions are shown in equation (23).

$$\begin{aligned} \mu_A &= \mu_{A,max} \frac{mRNA_{c8}}{mRNA_{c8} + K_{c8}} \\ \mu_{AD} &= \mu_{AD,max} \frac{mRNA_{c3}}{mRNA_{c3} + K_{c3}} \\ \mu_D &= \mu_{D,max} \frac{ATP}{ATP + ATP_{crit}} \end{aligned} \tag{23}$$

4.3.2 METABOLISM

The metabolic model considered in this work describes the metabolism of glucose and 19 amino acids (glutamate, arginine, proline, histidine, alanine, phenylalanine, leucine, isoleucine, lysine, tyrosine, methionine, tryptophan, cysteine, serine, glycine, aspartate, asparagine, threonine and valine) for biomass generation, energy production in the form of ATP and mAb production. This model is based on the work of Quiroga-Campano et al. who have successfully used their kinetic metabolic model to develop cell line- and clone-specific medium formulations for cell culture optimization (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). The original model does not include cell cycle predictions and the apoptotic cell death is linked only to metabolic stresses, including nutrient depletion and lactate accumulation. As explained above, minimal lactate production is observed in serum-free cultures of this cell line and therefore, lactate predictions have not been considered. In addition, the toxic effect of ammonia is usually not considered in mathematical models of GS-NS0 cells due to ammonia concentrations in culture being lower than 1mM (Bibila *et al.*, 1994; Wu *et al.*, 2004; Ho, Varley and Mantalaris, 2006) and hence ammonia predictions were also not included in this work.

4.3.2.1 Glucose

The mass balance to glucose in a batch bioreactor is given by equation (24) where *[Glc]* is the concentration of glucose (in mM) and the Q terms are the consumption rates (mmol/cell/h) for biomass proliferation ($Q_{x,glc}$), mAb production ($Q_{mAb,glc}$), alanine production ($Q_{ala,glc}$), energy production in the TCA cycle ($Q_{tca,glc}$). Analogous subscripts will be used for all other metabolites.

$$\frac{d(V.[Glc])}{dt} = X_V.V.(Q_{x,glc} + Q_{mAb,glc} + Q_{ala,glc} + Q_{tca,glc}) \quad (24)$$

The glucose consumption rates for biomass, energy production, mAb production and alanine production, used in equation (24), are calculated by equations (25) to (28), respectively.

$$Q_{x,glc} = \frac{\mu}{Y_{x/GLC}} \quad (25) \qquad Q_{tca,glc} = \frac{\mu}{Y_{TCA/GLC}} + M_{tca/Glc} \quad (26)$$

$$Q_{mAb,glc} = MmAbx \times mAb_{Glc} \quad (27) \qquad Q_{ala,glc} = \frac{\mu}{Y_{ala/glc}} + M_{ala/glc} \quad (28)$$

In equations (25) to (28), Y terms are called yield coefficients and are usually formally defined as ratios of rates (Hong, 1989). For instance, $Y_{x/glc}$ (cell/mmol) is the ratio of the rate of consumption of glucose for proliferation over the cells growth rate. In their turn, M terms account for maintenance requirements whereby $M_{tca/glc}$ (mmol/cell/h), for example, accounts for energy produced from glucose in the TCA cycle and used to sustain cellular viability. The term $MmAbx$, in equation (27), accounts for the formation of mAb and the mAb_{Glc} term is the glucose required per milligram of mAb formed (mmol/mg mAb). As detailed by Quiroga-Campano *et al.*, the yield coefficients for biomass proliferation have been obtained from the nutrient requirements to form a new cell while those pertaining to mAb formation have been derived from the requirements to form a new mAb molecule (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

The total amount of glucose used in the TCA cycle to produce ATP is denoted C_{glc} and is calculated from equation (29).

$$C_{Glc} = Q_{TCA,Glc}\ X_V . V \quad (29)$$

4.3.2.2 Amino Acids

The mass balances to all amino acids (AA) have the same structure, given by equation (30) where "j"is any of the 19 AA considered. Since the NS0 cells used in this work have been selected using the GS system, all glutamine required is produced from glutamate using the GS enzyme. Therefore, the terms referring to

glutamate and glutamine have been summed together in one equation. The rates of AA consumption for proliferation, mAb production and TCA cycle are given, respectively, by equations (31) to (33), and $Q_{prod,AAj}$ in equation (30) denotes the production rate for amino acid " j" – alanine can be produced from glucose and glycine can be produced from serine (Carinhas *et al.*, 2013; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

$$\frac{d(V.[AA_j])}{dt} = X_V.V.(Q_{prod,AAj} - Q_{x,AAj} - Q_{mAb,AAj} - Q_{tca,AAj}) \quad (30)$$

$$Q_{x,j} = \frac{\mu}{Y_{x/AAj}} \quad (31)$$

$$Q_{mAB,j} = MmAbx \,.\, M_{mAb/AAj} \quad (32)$$

$$Q_{tca,AAj} = \gamma_j \,.\, Q_{tca,AA} \quad (33)$$

In equation (33), the terms γ_j denote the fraction of the total AA usage in the TCA cycle that refers to AA_j and $Q_{tca,AA}$ is the rate of consumption of AAj in the TCA cycle, given by equation (34). This formulation allows for reducing the number of model parameters associated with AA usage in the TCA cycle.

$$Q_{tca,AA} = \frac{\mu}{Y_{tca/AA}} + M_{tca/AA} \quad (34)$$

Figure 13 shows that glutamate, arginine, proline, isoleucine, threonine, valine, leucine, phenylalanine, tyrosine, asparagine, aspartate and alanine are considered to fuel the TCA cycle; alanine can be produced from pyruvate or can produce pyruvate, feeding the TCA cycle, and serine is used to produce glycine.

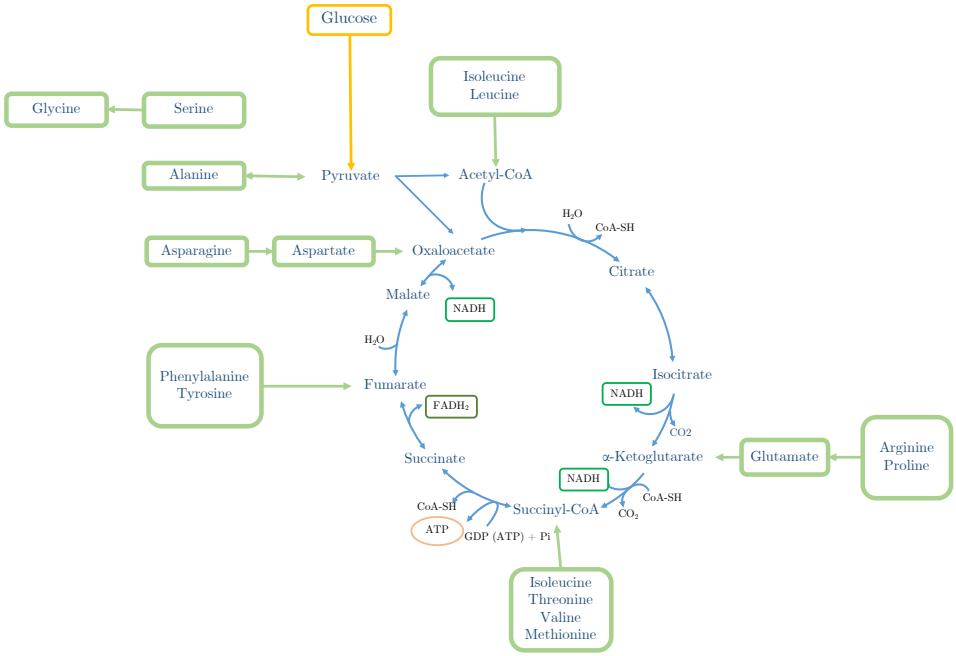


Figure 13: Use of amino acids in the TCA cycle and their consumption/production to/from other nutrients as considered in the metabolic model.

Analogously to what was described for glucose in equation (29), it is possible to calculate the amount of each AA used to produce energy in the TCA cycle given by

$$C_j = Q_{tca,AAj}\, X_V . V . \tag{35}$$

The specific consumption/production rates of some amino acids change according to metabolic shifts. For instance, it has been suggested that, once glutamate is exhausted, the required levels for biomass and glutamine production (and usage) are replenished by arginine as expressed in equation (36) (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). At the same time, proline, phenylalanine, tyrosine and alanine replace glutamate in satisfying the energetic requirements. When arginine becomes depleted, alanine covers for the energetic requirements of both glutamate and arginine. This is expressed in equation (37).

$$Q_{x,arg} = \begin{cases} \frac{\mu}{Y_{x,arg}}, & [Glu] > 0 \\ \frac{\mu}{Y_{x,arg}} + \frac{\mu}{Y_{x,glu}} + \frac{\mu}{Y_{x,gln}}, & [Glu] = 0 \end{cases} \quad (36)$$

$$Q_{tca,ala} = \\ = \begin{cases} 0, & [Glu] > 0, [Arg] > 0 \\ \frac{Q_{tca,glu}ATP_{glu} - Q_{tca,pro}ATP_{pro} - Q_{tca,phe}ATP_{phe} - Q_{tca,tyr}ATP_{tyr}}{ATP_{ala}}, & [Glu] = 0, [Arg] > 0 \\ \frac{(Q_{tca,glu}ATP_{glu} + Q_{tca,arg}ATP_{arg}) - Q_{tca,pro}ATP_{pro} - Q_{tca,phe}ATP_{phe} - Q_{tca,tyr}ATP_{tyr}}{ATP_{Ala}}, & [Glu] = 0, [Arg] = 0 \end{cases} \quad (37)$$

In equation (34), the terms ATP_j are the average number of ATP molecules generated by each metabolite, calculated based on the biochemical reactions that each metabolite undergoes. Specifically, the generation of NADH and $FADH_2$ cofactors and their subsequent use in OxPhos to generate ATP is accounted for (Quiroga-Campano *et al.*, 2016). It has been assumed that, on average, 1.5 molecules of ATP are obtained per molecule of cytosolic NADH or $FADH_2$, while mitochondrial NADH and $FADH_2$ have been assumed to produce 2.5 and 1.5 molecules of ATP, respectively (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018)

4.3.2.3 ATP

Finally, to account for all energy produced, a mass balance to ATP is derived, as shown in equation (38). ATP required for growth is accounted for by the $Y_{x/ATP}$ term while ATP required to sustain viability is accounted for using $M_{x/ATP}$. Energy required to produce mAb is accounted for using the $M_{mAb,ATP}$ (mmol ATP/mg mAb/h)

$$\frac{d(V.[ATP])}{dt} = C_{glc}ATP_{glc} + \sum_j C_j ATP_j - X_V.V.\left(\frac{\mu}{Y_{x/ATP}} + M_{x/ATP} + M_{mAb,ATP}MmAbx\right) \quad (38)$$

4.3.2.4 Monoclonal Antibody

The monoclonal antibody of interest is secreted by living cells. When cells die, their membranes become compromised releasing mAb accumulated intracellularly into the extracellular medium, contributing to the overall titre measured in the supernatant (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). This has been accounted for in the mAb mass balance (equation (39)). Since apoptotic cells in the context of this work refers to an early apoptotic population, and hence not membrane compromised, only the late apoptotic and dead cells (combined in X_D) will be deemed responsible for mAb secretion as a consequence of death. Experimentally, a titre decrease has been observed towards the end of the culture (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). One possible explanation for this is mAb degradation via proteolytic activity as internal proteases are released to the extracellular medium when cells undergo lysis (Fischer, Handrick and Otte, 2015). Under these assumptions, the mAb mass balance has the form of equation (39).

$$\frac{d(V.[mAB])}{dt} = X_V.V.mAb_{release} + (\mu_{ad} + \mu_d).xD.V(mAB_{acc} - mAB_{deg}) \tag{39}$$

The intracellular mAb accumulation term is given by the difference between mAb produced and mAb secreted, as expressed by equation (40). In order to account for different mAb productivity across the cell cycle, different mAb production rates must be used for each phase (García-Münzer, Kostoglou, *et al.*, 2015). Therefore, the $MmAb_x$ term will be given by a weighted average of the production rates at the different stages using the fraction of cells at each stage as weights (equation (41)).

$$\frac{dmAb_{acc}}{dt} = MmAb_x - mAb_{release} \tag{40}$$

$$MmAb_x = \sum_{i=G1,S,G2} q_{i,mAB}F_i \tag{41}$$

4.3.3 APOPTOSIS

Four genes involved in apoptosis have been considered in this work: caspase-8 and caspase-3, as mentioned above, together with *bcl-2* and *bax*, respectively encoding for anti- and pro-apoptotic proteins (Taylor, Cullen and Martin, 2008). Previous data obtained in the course of fed-batch cultures of GS-NS0 cells suggest that expression of caspase-8 is increased in the absence of aspartate and glutamate (Usaku, 2015). In order to account for the role of arginine in the metabolism of GS-NS0 cells the dependency on the concentration of this AA was also considered. The expression of caspase-8 has been modelled by equation (42). The fed-batch studies also revealed that the expression of caspase-3 was increased upon depletion of aspartate. Caspase-3 can also be self-activated or activated by caspase-8 (McIlwain, Berger and Mak, 2015) as described by equation (43).

$$\frac{dmRNA_{c8}}{dt} = \beta_{c8} \sum_{w=Asp,Glu,Arg} \frac{km_{w,c8}}{[w] + km_{w,c8}} - \alpha_{c8}.mRNA_{c8} \tag{42}$$

$$\frac{dmRNA_{c3}}{dt} = \beta_{c3} \left(\frac{km_{Asp,c3}}{[Asp] + km_{Asp,c3}} + \frac{mRNA_{c3}}{mRNA_{c3} + km_{c3}} + \frac{mRNA_{c8}}{mRNA_{c8} + km_{c3,c8}} \right) - \alpha_{c3}.mRNA_{c3} \tag{43}$$

The expression of *bcl-2* is enhanced in the absence of arginine and glucose (Altman and Rathmell, 2012; Bobak et al., 2016) and under overexpression of mAb, due to endoplasmic reticulum stress (Kober, Zehe and Bode, 2012), as described in equation (44). Transcription of the *bax* gene can be activated by caspase-8 and inhibited by *bcl-2* (Taylor, Cullen and Martin, 2008), as shown in equation (45).

$$\frac{dmRNA_{bcl2}}{dt} = \beta_{bcl2} \sum_{w=Glc,Arg} \frac{km_{w,bcl2}}{[w] + km_{w,bcl2}} + \left(\frac{mAB}{mAB + km_{mAB,bcl2}} \right)^2 - \alpha_{bcl2}.mRNA_{bcl2} \tag{44}$$

$$\frac{dmRNA_{bax}}{dt} = \beta_{bax} \left(\frac{mRNA_{c8}}{mRNA_{c8} + km_{bax,casp8}} + \frac{km_{bax,bcl2}}{bcl2 + km_{bax,bcl2}} \right) - \alpha_{bax}.mRNA_{bax} \tag{45}$$

4.3.4 INTEGRATION STRATEGY AND MODEL ANALYSIS

The model formulation described above includes three sub-models describing the three biological aspects considered in this work, namely cell-cycle, metabolism and apoptosis.

The cell cycle model is built on experimental data – cyclin expression, phase transition thresholds, times spent in each phase of the cycle and how metabolism affects cyclin expression. This sub-model is used to compute the cell cycle distribution and cycling times, and, hence the specific growth rate, under varying nutrient concentrations. This information is sent into the metabolic model which uses the growth rate, the cell cycle distribution and the cells and mAb compositions to compute the profiles of glucose and AA, energy generation and mAb production throughout culture. The metabolic model sends nutrient concentration information back to the cell cycle model. The nutrient, mAb and ATP profiles are sent to the apoptosis model, as these are sources of stress, in order to compute the death rates and the early- and late-stage apoptotic cell densities. The death rates are also used back in the cell cycle model to account for death. The flow of information in the model is illustrated in Figure 14.

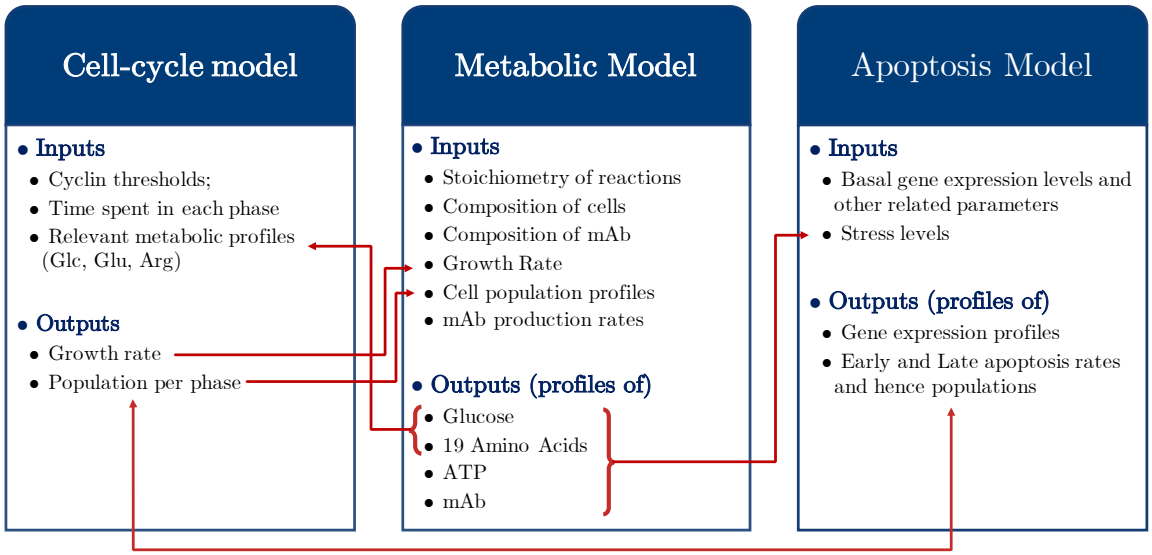


Figure 14: Flow of information between cell-cycle, metabolism and apoptosis sub-models.

With the present formulation, the model has 63 parameters. Among them, are two probabilities, namely those of transitioning from phase G_0/G_1 to phase S and from phase G_2/M to phase G_0/G_1 of the cell cycle. Due to the probabilistic nature of these transitions, some cells (entities) may transition from one phase to the next but others may remain in the same phase progressing to a higher content of the internal coordinate. Therefore, it is important to ensure that both cyclin domains (where probabilistic transitions exist) have enough bins to avoid loss of entities. To ensure appropriate domain sizes, a conservation study was conducted to determine the minimum number of bins. To this end, the birth term (h in equation (9)) and the probabilities of transition between cell cycle phases (equation (15)) were set to unity, thus maximizing transition rates (worst case scenario) and ensuring no births. Furthermore, the death rates were set to zero to ensure no deaths. The number of bins in each cyclin domain was varied between 50 and 300. As shown in Figure 15, using 200 bins in each cyclin domain a conservation of >99% of the entities (cells) is achieved at the point of maximum cell density (102h). The number of DNA bins was kept constant at 20 as these refer to a non-probabilistic transition (García-Münzer, Kostoglou, *et al.*, 2015). Under these conditions, the model has a total of 943 algebraic equations and 452 differential equations and the computational time to solve the model is 2 seconds.

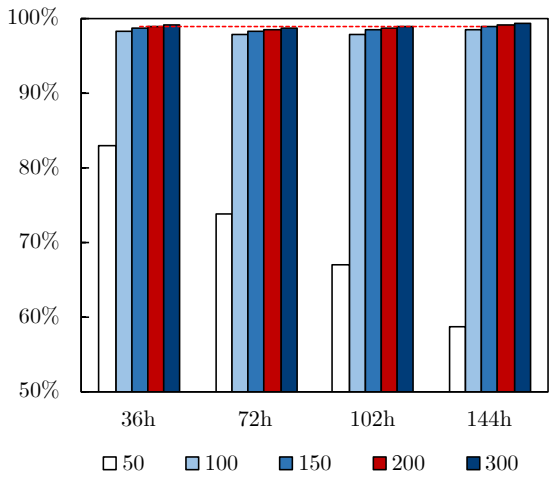


Figure 15: Results of the conservation study. The number of bins in each cyclin domain was varied between 50 and 300 while keeping the DNA domain at 20 bins. The red dashed line shows the 99% threshold.

4.4 GLOBAL SENSITIVITY ANALYSIS

GSA was employed to identify the significant model parameters and to design meaningful experiments to estimate them, thus also contributing to the understanding of the biological system. The sensitivity indices were calculated based on 5000 model simulations, varying all model parameters, simultaneously, ±50% from their nominal values. The criteria used were (1) no variation on the SIs on the viable cells (X_V) output and (2) that all SIs summed up to unity. The 5000 simulations used are between 2^{12} and 2^{13}. A number of model runs that is not an exact power of two and hence may compromise the uniformity of the Sobol generator (Sobol' *et al.*, 2011). However, in this work, it was found that several parameter combinations generated using the Sobol generator (`sobolset` function) led to unfeasible scenarios and hence, because convergence was achieved, this number was considered appropriate

The SIs were calculated at three biologically relevant time points – 48h (end of lag-phase), 72h (mid-exponential phase) and 120h (decline phase) – for 32 model outputs (viable, apoptotic and dead cell populations, cell cycle phases, mAb, ATP, 19 amino acids, glucose and four gene transcripts). Figure 16 shows GSA simulations for viable cells, ATP, cells in G_0/G_1 and mAb titre.

Given that SIs are defined as fractions of the total variance of each output, they vary between 0 (non-significant) and 1. A minimum threshold of 0.1 was set to discriminate the important parameters (Sidoli, Mantalaris and Asprey, 2005; García-Münzer, Kostoglou, *et al.*, 2015; Kiparissides, Pistikopoulos and Mantalaris, 2015; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). Parameters with SI indices above 0.1 are considered important and need to be re-estimated, whereas non-important parameters can be left at their nominal values. Figure 17 shows the results of GSA for viable cell density, fraction of cells in phase G_0/G_1, apoptotic cell density, ATP concentration, mAb concentration, *bcl2* transcripts and glutamate concentration (full GSA results shown in Appendix I). From the original 63 model parameters, only 30 were found significant allowing to significantly reduce

the dimensionality of the parameter estimation problem and the experimental work required.

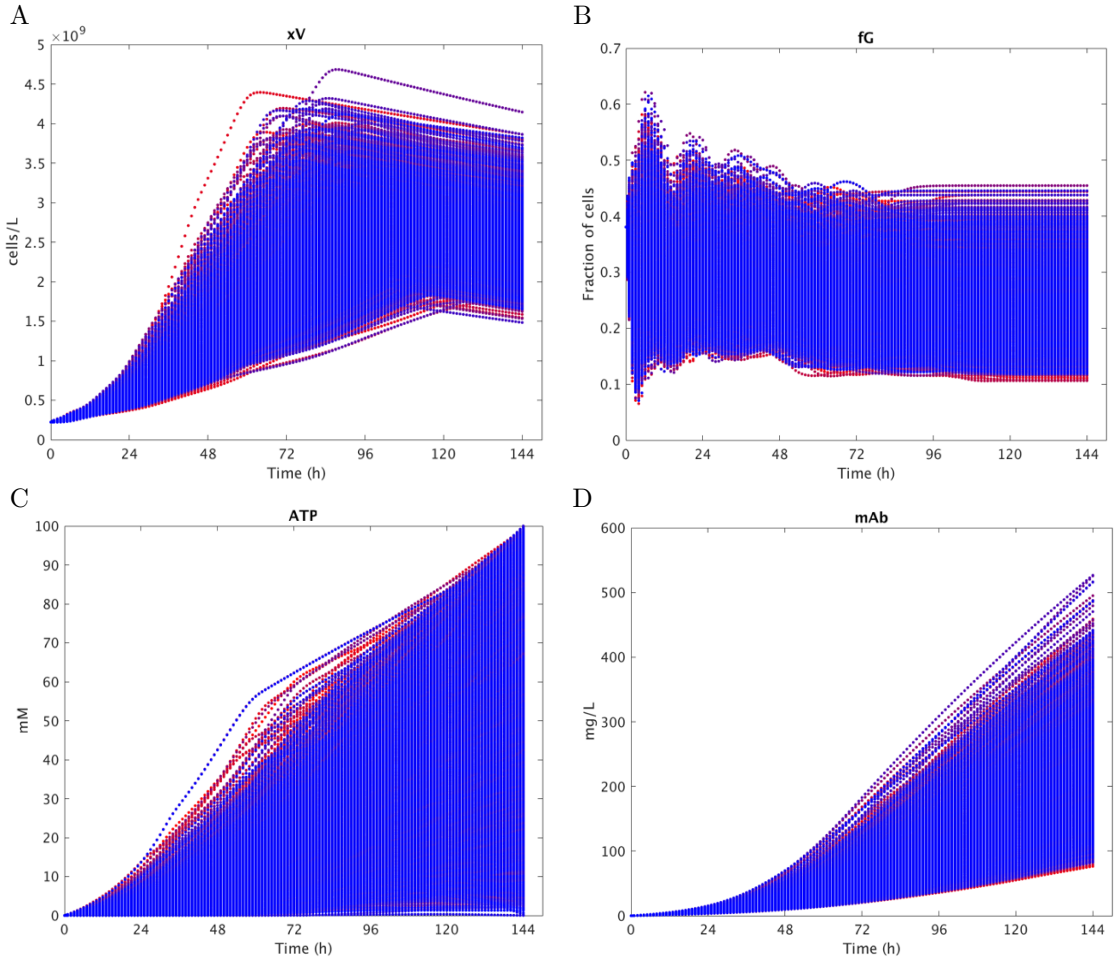


Figure 16: GSA simulations for viable cells (A), fraction of cells in phase G_0/G_1 (B), and the concentration of ATP (C) and mAb (D).

GSA provided useful biological information. For example, the results obtained for the viable cell density (Figure 17) indicate that, according to model predictions, this output is dominated by growth-related parameters during early stages of culture, while parameters related to metabolic limitations and death increase in significance as the culture reaches the stationary and decline phases. Interestingly, during the decline phase, glutamate and arginine fuelling the TCA cycle remain important for viable cells. The apoptotic death rate, unlike that of death through other processes, is an important parameter in late culture stages. The glutamate profile is most dependent on its requirements to fuel the TCA cycle, producing energy for growth, than on the maintenance requirements and the

predictions of this AA are directly affected by the cycling times. Gene expression, for apoptosis-related genes (Figure 17, Appendix I) was dominated by the corresponding basal expression and degradation rates.

It is also interesting to note that, while previously published GSA results suggested that phase G_1 was the most important for mAb production (García-Münzer, Kostoglou, *et al.*, 2015), those obtained in this work (Figure 17) suggest that mAb production is more strongly affected by G_2/M and S phases of the cell cycle. Tey and Al-Rubeai reported a correlation between cells in the G_2 phase and mAb production for this cell line (Tey and Al-Rubeai, 2005) and prior experimental evidence shows that G_0/G_1 arrests caused by p21 overexpression (Khoo and Al-Rubeai, 2009) or the addition of DMSO or Thymidine (García-Münzer, Kostoglou, *et al.*, 2015) to the culture medium did not lead to higher mAb titre in GS-NS0 cultures.

Finally, mAb degradation, as a consequence of cell death, is not predicted to play an important role in mAb titre, consistent with experimental observations reported in this work (section 4.6.3) and in literature where it was only observed for denser inoculations (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

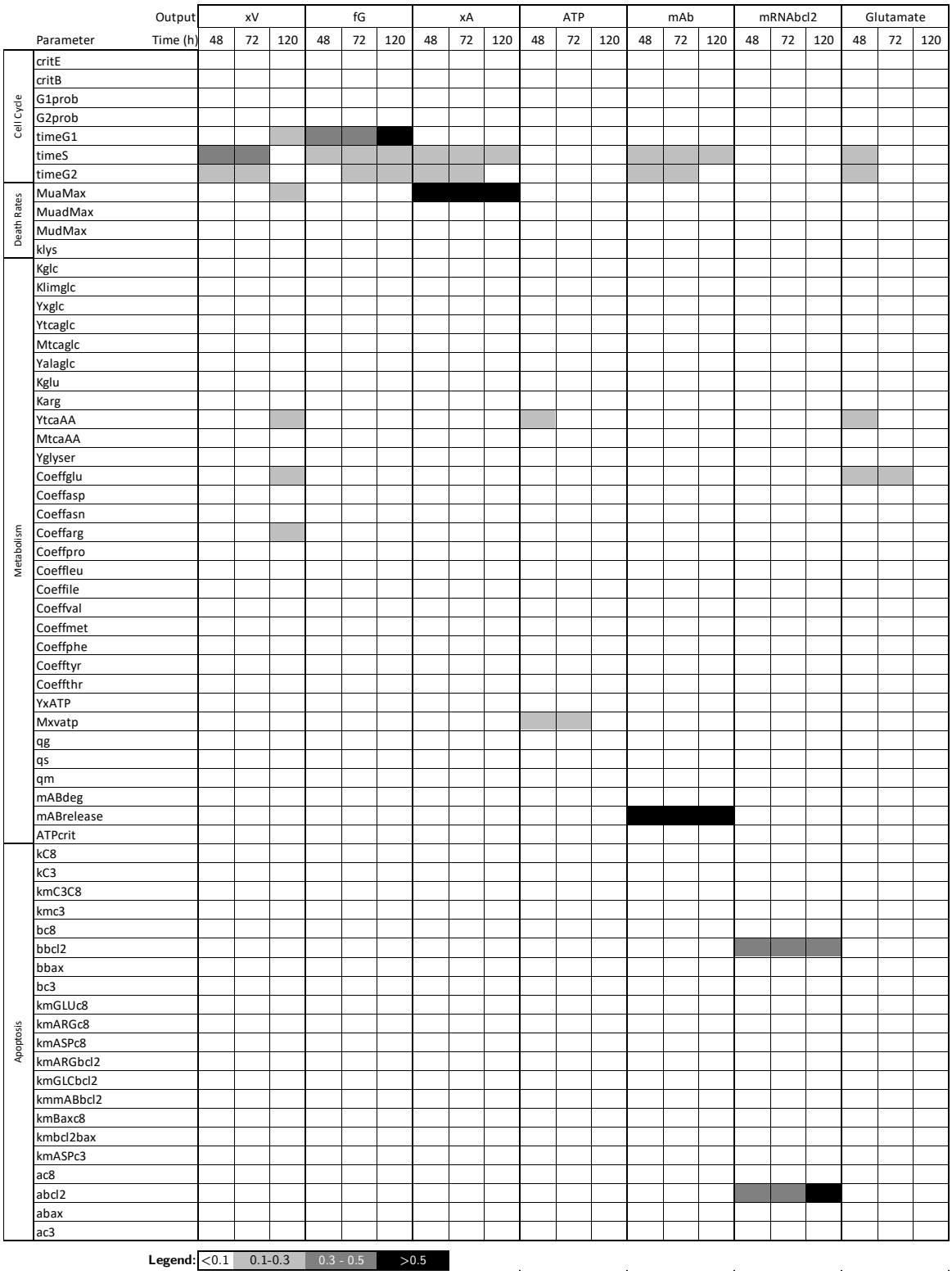


Figure 17: Global sensitivity analysis results obtained at 48, 72 and 120h of batch culture. The effect of each parameter on viable cells (X_V), fraction of cells in phase G_0/G_1 (fG), apoptotic cells (X_A), ATP, mAb, bcl2 transcripts (mRNAbcl2) and glutamate is quantified based on their Sobol sensitivity indices. A minimum threshold of 10% has been set to discriminate the significant parameters. Important parameters are shaded in light grey, dark grey and black for SIs of 0.1-0.3, 0.3-0.5 and >0.5, respectively.

4.4.1 MODEL-AIDED EXPERIMENTAL DEVELOPMENT; METHODOLOGY

In this section, the use of GSA results to support experimental development is shown. In the case of the integrated mathematical model of GS-NS0 cells, GSA has confirmed prior experimental observations of different metabolic regimes in the course of batch cultures of GS-NS0 cells, by identifying dynamically varying SIs of different parameters. This corroborates the need to conduct, at least, 120h of batch cultures in order to obtain meaningful data to estimate the relevant model parameters. GS-NS0 cells were cultured for 6 days (144h) in batch mode and sampled every 24h for *at-line* analysis of viable, apoptotic and dead cell densities and ATP content within minutes of sampling. Culture supernatant samples were used for *a-posteriori* (offline) analysis of glucose, lactate, AAs and mAb and cells were preserved for cell-cycle and gene expression analysis.

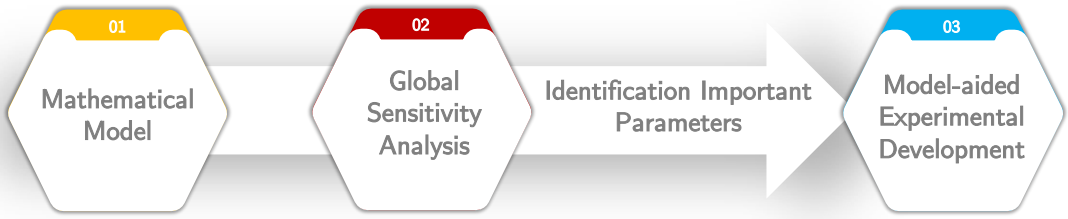


Figure 18: Model-aided experimental development workflow. GSA is performed on a mathematical model to determine the important parameters. This gives important information on which parameters should be estimated in the laboratory (if possible), experiment duration etc.

From GSA, it could also be identified that cell cycle times are important parameters and especially so at the earlier stages of the culture. In order to determine the cycling times under non-disturbed growth conditions, the EdU method was used. In brief, this method relies on exposing cells in mid-exponential growth to EdU (a Thymidine analogue) which will be uptaken by cells undergoing phase S during this period. After a click chemistry reaction, EdU can be labelled with a suitable fluorescent dye allowing for tracking of the S-phase population by flow cytometry.

4.5 PARAMETER ESTIMATION

Identification of the significant model parameters at different culture times has been achieved by GSA. The dynamic significance of parameters seen by the variation of their SIs over time supports the need to conduct batch experiments in order to collect the required data to estimate the relevant model parameters at different culture stages. Parameter estimation results are presented in Table 4 and show that, for 24 out of the 30 significant parameters, the 95% confidence intervals are narrow, being an order of magnitude lower than the value of the corresponding parameters, in most cases. In particular, cycling times were estimated in very good agreement with experimental data and with narrow confidence intervals.

Since only non-negative values are allowed for parameters accounting for the conversion of phenylalanine and tyrosine to TCA cycle intermediates, zero was set as the lower bound for these parameters in the parameter estimation algorithm. Reaching the lower bound on estimating values for these parameters has prevented the calculation of the respective confidence intervals. A value of zero corresponds to no usage of these AA to fuel the TCA cycle, which will be an underlying assumption of the present model. When estimating values for the parameters related to *bax* gene expression, the confidence intervals obtained were wider than the predicted parameter values. This means that the algorithm is unable to estimate these parameters from the experimental data provided and suggests that more experimental data would be required. Problems arising from the experimental determination of transcripts include the difficulty of such measurements and possible discrepancies between *in-vitro* transcript production/degradation rates and those observed in culture. In spite of the large size of the parameter estimation problem, it was successfully solved using the MLE algorithm to re-estimate most of the significant parameters of the model. With the exception of two parameters related to the *bax* gene expression, it was possible to obtain values in agreement with experimental observations and with a narrow 95% confidence interval, showing the quality of the solution encountered. It is possible that the quality of the gene

expression data and also the large nature of the problem may have compromised the quality of gene-expression related parameters.

Table 4: Estimated values of significant model parameters. Final values and 95% confidence intervals obtained using the MLE algorithm and experimental data obtained in this work.

Parameter (units)	**Biological Significance**	**Final Value**	**95% CI**
$timeG1_{min}$ (h)	Time spent in phase G_0/G_1	8.27	1.4
$timeS_{min}$ (h)	Time spent in phase S	5.54	0.65
$timeG2_{min}$ (h)	Time spent in phase G_2/M	8.66	0.97
$\mu_{a,max}$ (h^{-1})	Maximum rate of early apoptosis	$3.27x10^{-3}$	$5.9x10^{-4}$
$\mu_{ad,max}$ (h^{-1})	Maximum rate of late apoptosis	$1.03\ x10^{-1}$	$2.7x10^{-2}$
coeffArg (--)	Relation between total rate of AA consumption for energy production and the rate of consumption of each AA in the TCA cycle	$1.35\ x10^{-1}$	--
coeffGlu (--)		$8.96x10^{-1}$	$1.3x10^{-1}$
coeffIle (--)		$3.60x10^{-1}$	$3.5x10^{-2}$
coeffLeu (--)		$5.11x10^{-1}$	$7.6x10^{-2}$
coeffMet (--)		$3.10x10^{-2}$	$3.3x10^{-3}$
coeffPhe (--)		0	--
coeffPro (--)		$5.10x10^{-1}$	$7.6x10^{-3}$
coeffThr (--)		$2.73x10^{-2}$	$1.3x10^{-3}$
coeffTyr (--)		0	--
coeffVal (--)		$1.93x10^{-1}$	$2.0x10^{-2}$
$m_{tca,AA}$(mmol/cell)	TCA usage of AAs for maintenance	$9.99x10^{-12}$	$1.9x10^{-12}$
$m_{tca,glc}$(mmol/cell)	TCA usage of glucose for maintenance	$3.58x10^{-11}$	$1.6x10^{-12}$
$m_{xV,atp}$(mmol/cell)	ATP used for maintenance	$1.12x10^{-9}$	$5.0x10^{-11}$
$Y_{ala,glc}$ (mol/mol)	Yield of alanine on glucose	$4.94x10^{9}$	$4.1x10^{8}$
$Y_{gly,ser}$ (mol/mol)	Yield of glycine on serine	$1.10x10^{9}$	$1.6x10^{8}$
$Y_{tca,AA}$ (mol/mol)	Growth-related AA use in TCA cycle	$2.07x10^{9}$	$2.4x10^{8}$
$mAb_{release}$	Rate of mAb release from viable cells	$5.00x10^{-10}$	--
α_{bax} (h^{-1})	Bax mRNA degradation rate	$1.21x10^{-1}$	$4.3x10^{-1}$
$\alpha_{bcl\text{-}2}$ (h^{-1})	Bcl-2 mRNA degradation rate	$1.02x10^{-1}$	$3.7x10^{-2}$
α_{casp3} (h^{-1})	Casp-3 mRNA degradation rate	$6.00x10^{-3}$	$1.9x10^{-3}$
α_{casp8} (h^{-1})	Casp-8 mRNA degradation rate	$3.53x10^{-3}$	$9.9x10^{-4}$
β_{bax} (h^{-1})	Bax mRNA base production rate	$8.52x10^{-2}$	$3.1x10^{-1}$
$\beta_{bcl\text{-}2}$ (h^{-1})	Bcl-2 mRNA base production rate	$8.33x10^{-2}$	$2.9x10^{-2}$
β_{casp3} (h^{-1})	Casp-3 mRNA base production rate	$1.44x10^{-3}$	$9.0x10^{-4}$
β_{casp8} (h^{-1})	Casp-8 mRNA base production rate	$8.00x10^{-4}$	$4.6x10^{-4}$

4.6 RESULTS AND DISCUSSION

In the present section, the determination of cycling times of GS-NS0 cells in serum-free cultures and experimental measurements conducted in the course of batch cultures of GS-NS0 cells are presented and shown together with the corresponding model predictions. Cultures inoculated at $2x10^5$cells/mL were used for parameter estimation and, in order to verify whether predictive capabilities were retained for a wider range of conditions, model predictions of two cultures starting at different inoculation densities ($1x10^5$ and $4x10^5$ cells/mL) have been compared to the corresponding experimental data. All experimental data is reported as mean of N=3 biological replicates ± one standard deviation (error bars).

4.6.1 DETERMINATION OF CYCLING TIMES

As explained above, cycling times of GS-NS0 cells were found to be important parameters for the model and their accurate determination was conducted using the EdU method. The EdU-positive population shows a significant increase in the G_2/M content between 0 and 2h of culture (Figure 19A), suggesting that cells entered this phase. This transition will be completed when a significant fraction of the cells moves to phase G_0/G_1, which is seen at 6-8h, suggesting a G_2/M transit time of 6h. A significant increase in the S-phase fraction at 16-18h implies that transit through G_0/G_1 has taken around 10h. The S-phase time is determined from the EdU negative population (Figure 19B), where a significant increase in the corresponding fraction occurs between 4-6h. A significant increase in the G_2/M fraction at 10-12h indicates that cells required 6h to complete phase S. The total time is approximately 22 hours corresponding to a growth rate of 0.0315 h^{-1}. Compared to the 18h doubling time previously reported for cultures with serum, a 22% increase is reported in this work. The observation of slower growth rates in the absence of serum is consistent with previous reports for this and other cell lines (González Hernández and Fischer, 2007; Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

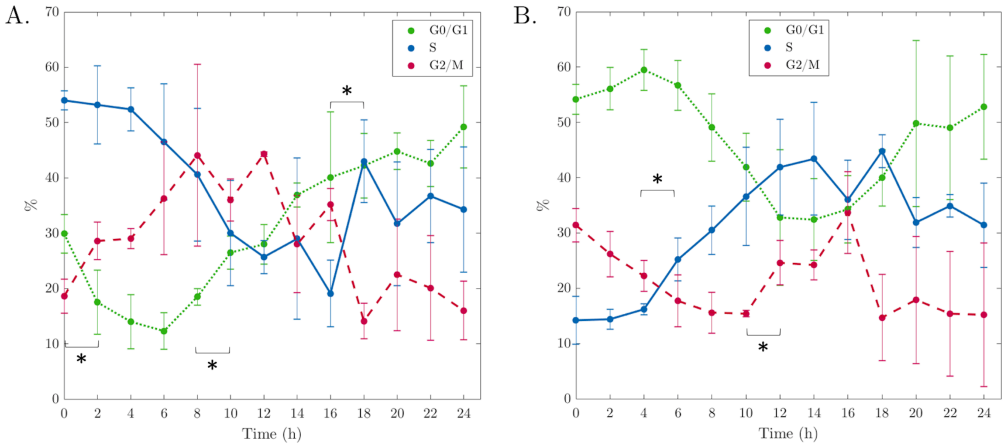


Figure 19: Evolution of cell cycle distribution (G_0/G_1 in green, S in blue and G_2/M in pink) in EdU-positive (A) and EdU-negative (B) populations during 24 hours of culture in EdU-free medium after 3h exposure to 30µM EdU (*p-$value$<0.05, one-way ANOVA).

4.6.2 CELL CYCLE, GROWTH, APOPTOSIS AND DEATH PREDICTIONS

Viable, apoptotic and dead cell densities obtained in batch cultures of GS-NS0 cells, and the corresponding model predictions, are presented in Figure 20. It can be observed that the model is able to capture these three profiles, under all three conditions run, showing that the model predictive capabilities are retained for a minimum interval of 1–4x10^{5}cells/mL in batch mode.

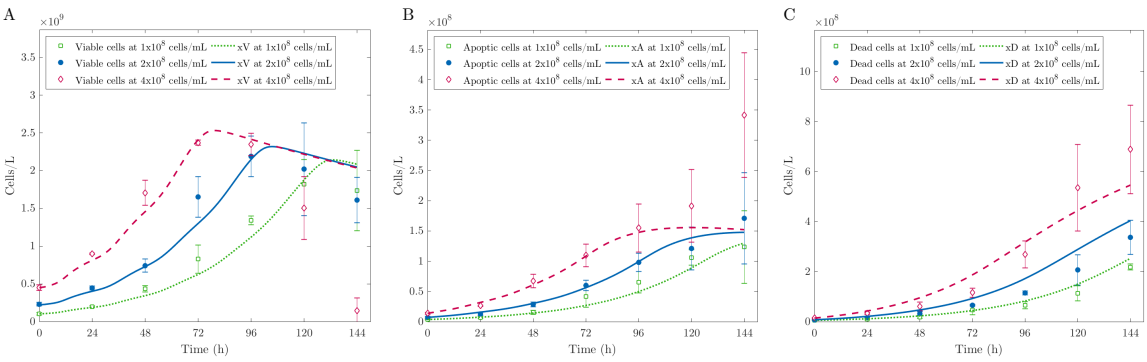


Figure 20: Experimental data (dots) and model predictions (lines) for viable (A), apoptotic (B) and dead (C) cell densities obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

The model also predicts the cell cycle distributions during most of the duration of the cultures (Figure 21), not only predicting the distribution of viable cells across the cycle but also capturing its oscillatory behaviour, seen especially

during earlier culture times as observed in this work (Figure 19) and consistent with prior observations (García-Münzer, Kostoglou, *et al.*, 2015). As the sampling time is very similar to the duration of the cell cycle it is not surprising that experimental cell cycle distribution data obtained during early culture stages and exponential growth shows relatively stable populations.

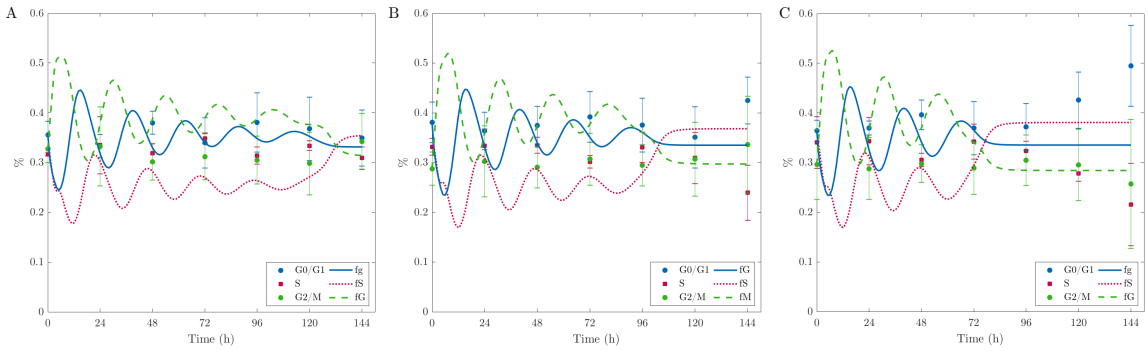


Figure 21: Distributions of viable cells across phases G_0/G_1 (blue circles, solid lines), S (pink squares, dotted lines) and G_2/M (green circles, dashed lines) observed for batch cultures of GS-NS0 cells starting at $1x10^5$cells/mL (A), $2x10^5$cells/mL (B) and $4x10^5$cells/mL (C) inoculation densities.

4.6.3 MONOCLONAL ANTIBODY PREDICTIONS

Monoclonal antibody production is well captured by the model for all batch cultures conducted in this work, as shown in Figure 22. Trends and magnitudes observed the two lower inoculation densities are well captured by the model. Predictions for the denser culture overestimate production during the decline phase, and specifically for the last 24h. At this time, sampling the culture is problematic because cell aggregation is very significant, which may compromise representative sampling (Quiroga-Campano, 2016). Additionally, the cellular activity changes towards the end of culture, compromising viability and affecting metabolism, as corroborated in section 4.6.7.

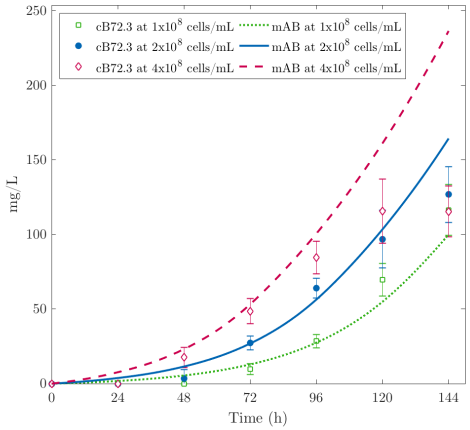


Figure 22: Experimental data (dots) and model predictions (lines) for mAb production in GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, full lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

4.6.4 GLUCOSE AND AMINO ACIDS METABOLISM PREDICTIONS

4.6.4.1 Glucose and Alanine

Under the culture conditions employed, glucose was the main carbon and energy source. Glucose is primarily metabolized via glycolysis producing pyruvate which can have several metabolic fates. It has been suggested that rapidly proliferating cells (such as cancer cells) need to satisfy their energy requirements quickly in order to continue proliferating and hence may prefer to use aerobic glycolysis producing lactate from pyruvate. Although inefficient, this metabolic option, known as the Warburg effect, offers the possibility of fast energy production (Heiden, Cantley and Thompson, 2009). Industrial cell lines, often selected based on proliferation, also rely, to some extent, on aerobic glycolysis to obtain energy (Buchsteiner *et al.*, 2018). The consequent lactate accumulation is generally considered negative for culture, having been shown to be a good process indicator of (poor) culture performance in terms of recombinant protein titres and viable cell densities achieved (Le *et al.*, 2012). In GS-NS0 cell cultures, it has been shown that lactate production in serum-free batch cultures is very low (<5mM) not inhibiting cell culture (Bibila *et al.*, 1994), unlike what has been seen in cultures with serum where concentrations reached 25mM (Campano *et al.*, 2016) or in serum-free fed-batch cultures where 40mM where reached (Mulukutla, Gramer and Hu, 2012). Pyruvate is therefore either used for alanine production or to fuel the TCA cycle

(Mulukutla, Gramer and Hu, 2012). Glucose and alanine profiles are well captured by model predictions for all three data sets, as shown in Figure 23.

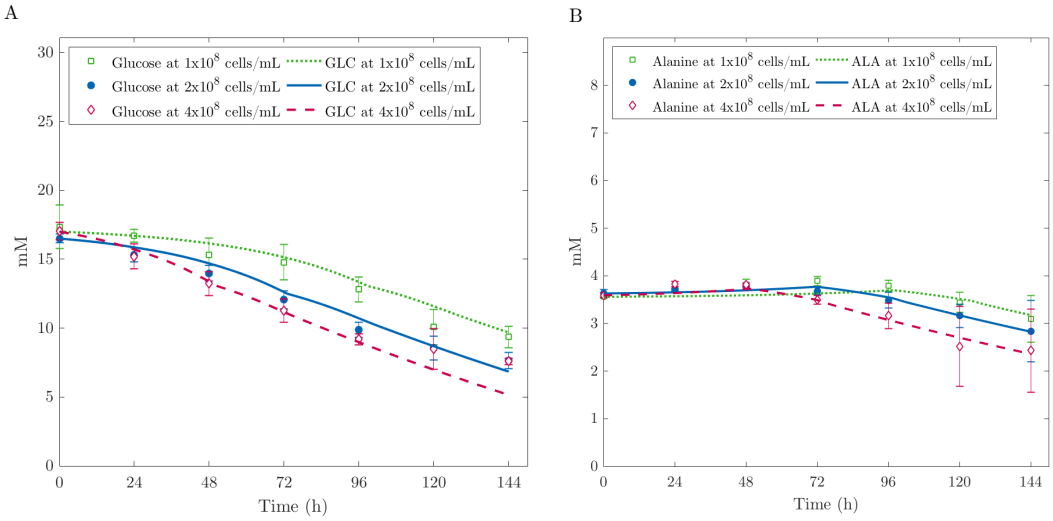


Figure 23: Experimental data (dots) and model predictions (lines) for glucose (A) and alanine (B) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

4.6.4.2 Non-Essential Amino Acids

Previous studies conducted with GS-NS0 cells in batch and fed-batch modes have reported that peak density is reached when glutamate is exhausted in the medium, or very shortly after (García-Münzer, Kostoglou, *et al.*, 2015; Kiparissides, Pistikopoulos and Mantalaris, 2015). Nevertheless, due to the importance of glutamate for these cells as source of glutamine and precursor of TCA intermediate α-ketoglutarate, glutamate consumption has been considered preferential to that of glucose (Kiparissides, Pistikopoulos and Mantalaris, 2015). Based on detailed analyses of amino acid consumption in GS-NS0 cultures with serum (Campano *et al.*, 2016) and serum-free (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018), it has been suggested that arginine sustains growth when glutamate is exhausted replenishing glutamate and glutamine requirements.

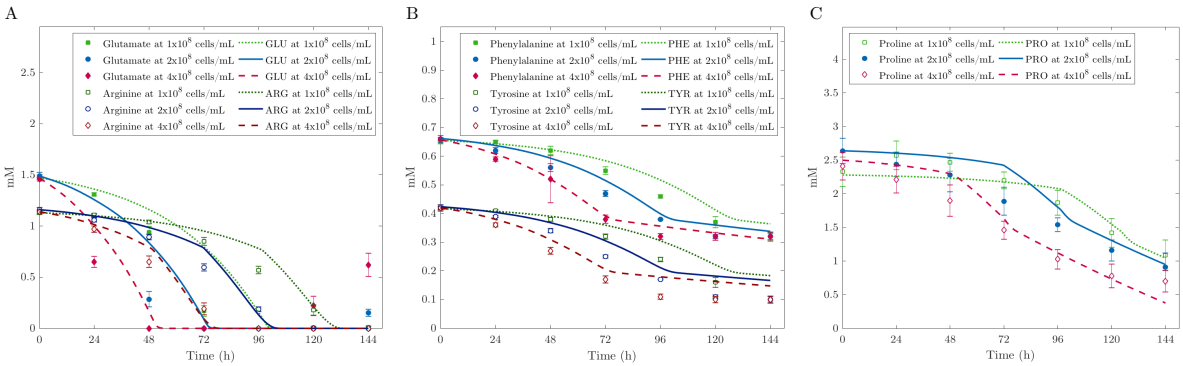


Figure 24: Experimental data (dots) and model predictions (lines) for (A) glutamate (closed symbols) and arginine, (B) phenylalanine (closed symbols) and tyrosine (open symbols) and (C) proline obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines).

As mentioned above, it has also been suggested that other amino acids play a role in satisfying metabolic needs post glutamate depletion, namely non-essential amino acids alanine, proline and tyrosine as well as essential amino acid phenylalanine (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). All these amino acid profiles are well predicted by the model, as shown in Figure 24. As seen above, alanine can be produced from pyruvate, but this reaction is reversible allowing for alanine to fuel acetyl-CoA production and hence the TCA cycle acting as energy source.

Another non-essential amino acid, and the first to be depleted in culture, is aspartate, triggering the consumption of asparagine. Figure 25 shows that the model can capture both profiles.

Serine is consumed from the beginning of the culture producing glycine during the first 96h of culture (Figure 26A). During the last 2 days of culture glycine production rate slows down and the model overestimates the serine consumption rate. This suggests that serine composition of cells and/or mAb considered are different from those of the clone used or mAb produced and the use of more accurate values could improve model predictions. Alternatively, a serine production term (from pyruvate) could be introduced in the model. Finally, cysteine is considered to be used solely for biomass and mAb production and its consumption profile is well captured by the model (Figure 26B).

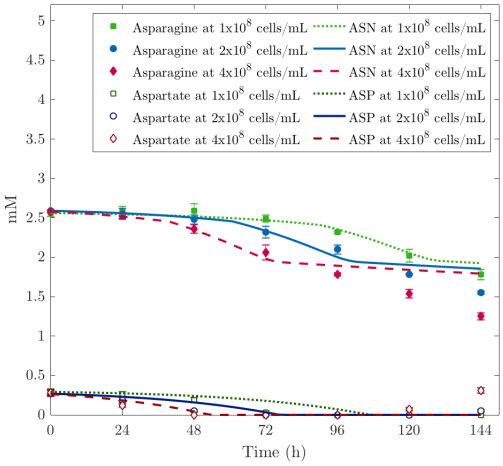


Figure 25: Experimental data (dots) and model predictions (lines) for aspartate (open symbols) and asparagine (closed symbols) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

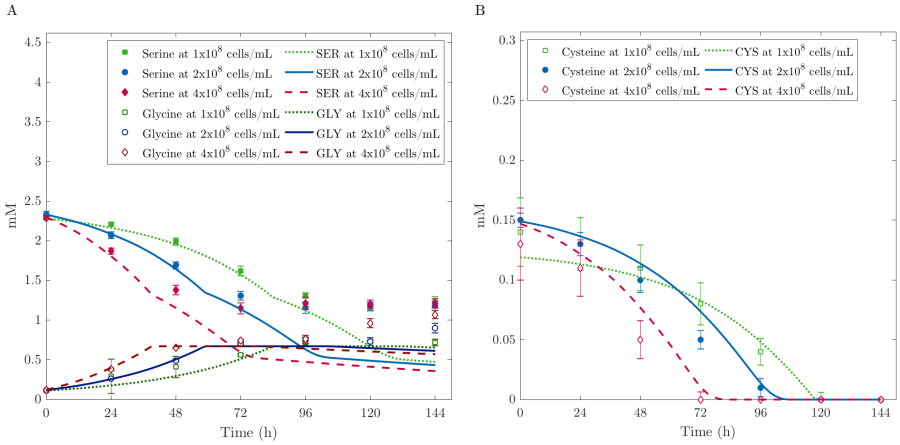


Figure 26: Experimental data (dots) and model predictions (lines) for serine (closed symbols, A), glycine (open symbols, A) and cysteine (B) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

4.6.4.3 Essential Amino Acids

The profiles of essential amino acids threonine, tryptophan, histidine, leucine, valine, methionine, lysine and isoleucine is shown in Figure 27 (the profile for essential amino acid phenylalanine is shown in Figure 24B). The consumption of tyrosine and valine is slightly underestimated throughout the cultures while the consumption of histidine is slightly overestimated. Nevertheless, metabolic profiles are closely followed by the model.

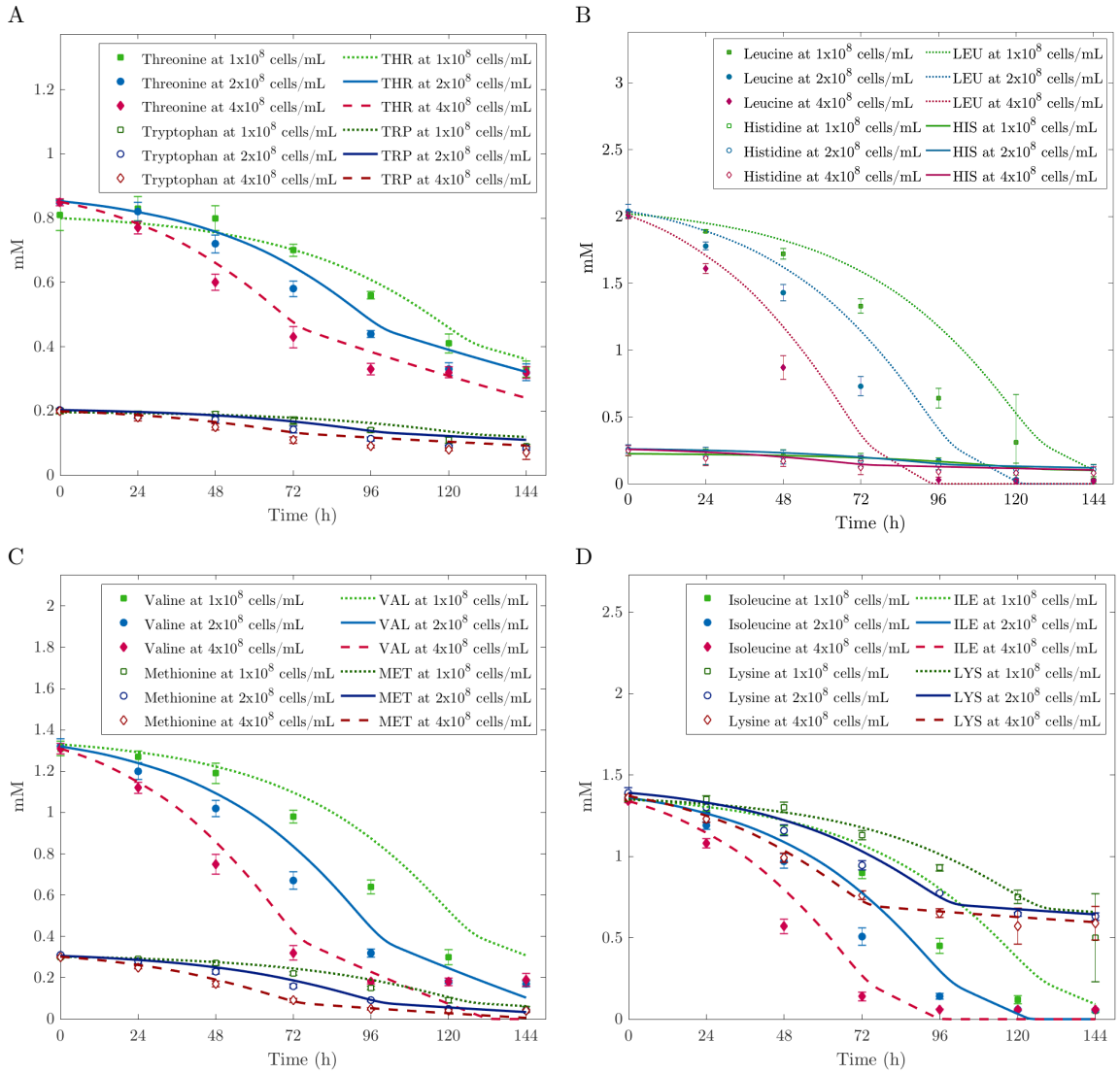


Figure 27: Experimental data (dots) and model predictions (lines) for threonine (closed symbols, A), tryptophan (open symbols, A), leucine (close symbol, B), histidine (open symbols, B), valine (closed symbols, C), methionine (open symbols, B), isoleucine (closed symbols, D) and lysine (open symbols, D) obtained for GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, solid lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

4.6.5 ATP PREDICTIONS AND MINIMUM REQUIREMENTS

As seen in the previous sections, cell growth, consumption and production of nutrients and monoclonal antibody titres are well predicted by the model. The energy produced in the cellular metabolism is also well predicted by the model for all inoculation densities, with the exception of the last time point for the densest culture as shown in Figure 28A.

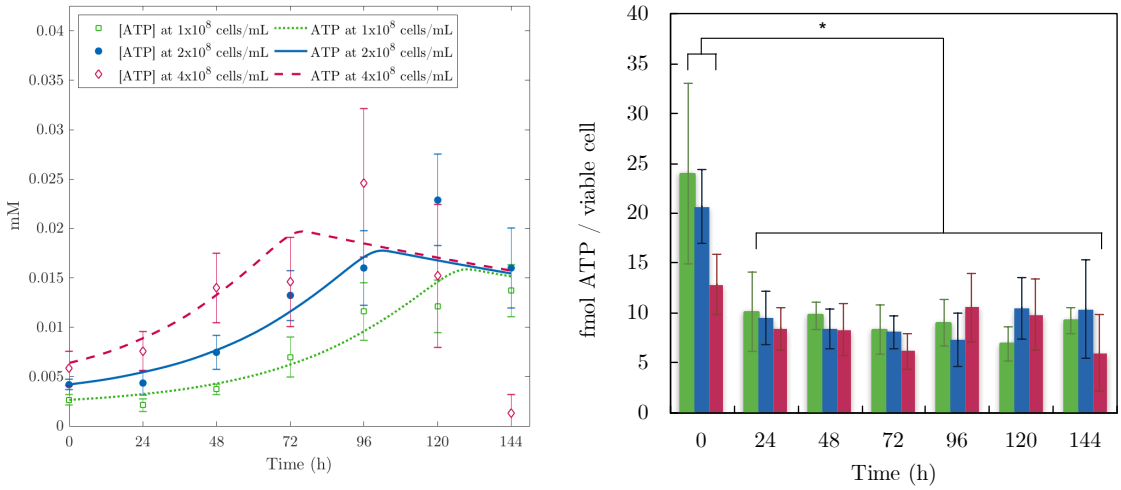


Figure 28: A: Experimental data (dots) and model predictions (lines) of ATP concentration in GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, full lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities. B: Normalized ATP content (in fmol per viable cell) obtained during 6 days of batch cultures with inoculation densities of 1 (green), 2 (blue) and 4x10^5 cells/mL (pink) (*p-$value$<0.05, one-way ANOVA).

The model assumes that, when metabolic stress is extreme, cell death may occur through processes other than apoptosis such as necrosis, which has been shown to occur upon ATP shortage (Eguchi, Shimizu and Tsujimoto, 1997). In order to estimate the critical ATP concentrations for GS-NS0 cells, the ATP concentrations measured in this work were normalized to the viable cell population (Figure 28B). In order to identify possible sources of variance in the normalized ATP contents, a two-way ANOVA was conducted with *initial inoculation density* and *time* tested as possible sources of variance and a significance of 0.05. From this analysis, only time was identified as a source of significant variance (p-$value$<0.001), while the inoculation density did not cause significant differences in the specific ATP content (p-$value$>0.05). Subsequently, one-way ANOVA, with the same significance level, was used to assess which time points were significantly different from the others. As shown in Figure 28B, only the initial time-point is different from the others (p-$value$<0.05). The average cellular ATP content can be calculated from the average of the normalized ATP contents for 24-144h across all densities to be 8.70±2.7 fmol ATP per viable cell. This value is in agreement with those reported for cultures of the same cell line with 10% serum (9.22±5.02

fmol/cell) (Campano *et al.*, 2016) and serum-free (4.82±1.09 fmol/cell) (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018) media.

4.6.6 APOPTOSIS-RELATED GENE EXPRESSION

As described above, apoptosis-related gene expression has been used to describe the transition between viable to early apoptosis and from early to late apoptosis using caspase-8 and caspase-3, respectively. In addition, one pro-apoptotic and one anti-apoptotic gene (*bax* and *bcl-2*) were included. Capturing the biology of apoptosis is important as a means to link the impact of process parameters (agitation, oxygen, shear stress e.g.) in physiology. As works are being conducted towards on-line measurement of relevant gene expressions (Pioch *et al.*, 2007), gene expression may be used in model-based control.

The oscillatory behaviour of gene expression observed was not fully tracked by the model (Figure 29). However, those observed for the expression of *bcl-2* were predicted to some extent for all densities. The expression of this gene appears to show a density-related behaviour observed for the densest culture which is not observed for the other two conditions (Figure 29C).

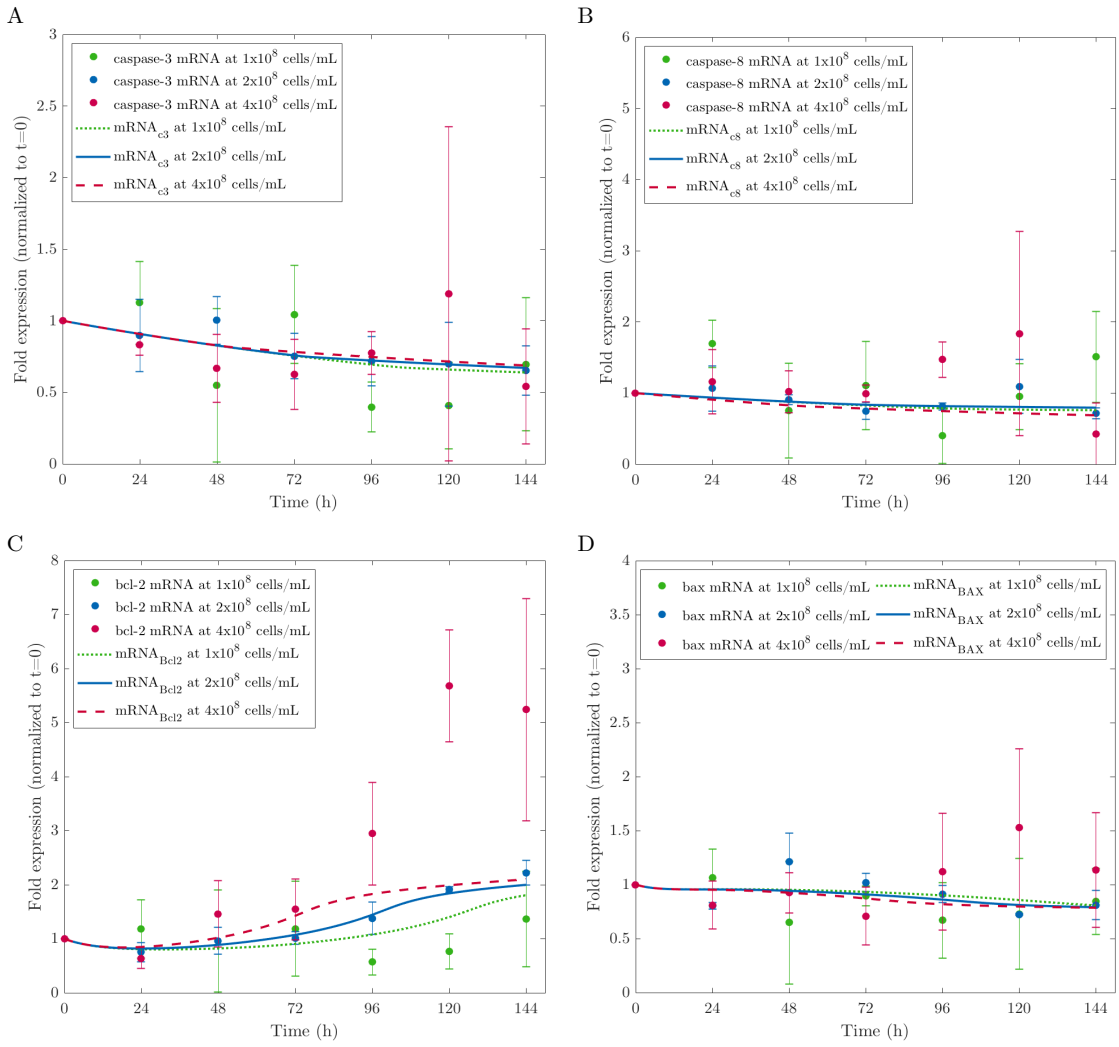


Figure 29: Experimental data (dots) and model predictions (lines) of caspase-3 (A), caspase-8 (B), *bcl-2* (C) and *bax* (D) transcripts in GS-NS0 batch cultures starting at 1x10^{5}cells/mL (green squares, dotted lines), 2x10^{5}cells/mL (blue circles, full lines) and 4x10^{5}cells/mL (pink diamonds, dashed lines) inoculation densities.

4.6.7 DYNAMIC INTERPLAY BETWEEN METABOLISM, GROWTH, AND PRODUCTIVITY

The development of an integrated mathematical model, such as that proposed in this chapter, is important because it captures population heterogeneity through cell cycle and gene expression but also because it captures the dynamic interplay between metabolism, growth, death and mAb production. It has been seen above that for all inoculation densities, maximum viable cell density (Figure 20A) correlates with the depletion of arginine, which starts to be consumed when glutamate is exhausted (Figure 24A), corroborating the important roles of these

amino acids for GS-NS0 cells. Additionally, until complete glutamate depletion (Figure 24), apoptotic (Figure 20B) and dead (Figure 20C) cell densities are low and the concentration of apoptotic cells is comparable to that of dead cells. Beyond this point, the dead cell population grows faster than the apoptotic population, and the viable cells decline. The fast transition from viable to dead cells, follows closely the ATP levels (Figure 28), suggesting that a more severe nutritional stress is imposed possibly triggering other forms of cell death (Grilo and Mantalaris, 2019a). It has also been shown that cell cycle distributions are stable for all cultures until the depletion of arginine (Figure 24), suggesting an important role of nutritional stress on cell cycle regulation which has been described in previous studies where expression of cyclins E1 and B1 has shown significant differences before and after glutamate depletion (García-Münzer *et al.*, 2013).

The central role of metabolism has motivated multivariate data analysis of metabolic profiles in an attempt to derive an unsupervised methodology to identify patterns and, potentially, correlate them with cell growth, death and productivity. Similar multivariate analyses of metabolic profiles have been conducted to discriminate osteoblasts differentiation stages using metabolomics (50 metabolites) (Klontzas *et al.*, 2017) or to track metabolic changes in human embryonic and human pluripotent stem cells (Vernardis *et al.*, 2017), using principal component analysis and hierarchical clustering (HCL) as unsupervised methods. These methods are also explored herein.

HCL is used to evaluates the closeness of k observations based on their coordinates using a given metric. In the case of the present work, the observations are the time points of each culture and the coordinates are the concentrations of the metabolites measured at each time points. The metric used to evaluate how close two observations are is the Euclidean distance in the 23-dimensional space, defined by the concentrations of the 23 metabolites analysed. Additionally, PCA is used to identify clusters and provide a visualization tool of the data. Both methodologies are unsupervised.

4.6.7.1 Inoculation density-specific metabolic profiles

For the analysis of GS-NS0 cell cultures, 23 metabolic profiles obtained for each of the three inoculation densities used in this work – 20 amino acids, glucose, lactate and ATP – were analysed using hierarchical clustering (HCL) with Euclidean distance as measure is shown in Figure 30, where it has been superimposed to the viable, apoptotic and dead cell profiles, the specific mAb production rate and the cell cycle distribution.

At the shortest distance in the dendrogram obtained from the HCL algorithm, three clusters can be identified for the lowest inoculation density and four clusters are identified for the other two inoculation densities (Figure 30). Based on the time-points included in each cluster, for each density, it can be seen that the clusters correspond to lag-phase, exponential growth, stationary phase and decline phase superimposing perfectly with the growth kinetics of each culture. The different growth stages of culture are, therefore, predicted by HCL of metabolic profiles. Termination of exponential growth correlates with glutamate depletion while stationary growth will last until arginine is depleted, suggesting that the glutamate metabolic shift should be avoided. The analysis of profiles obtained for the lowest inoculation density is not conclusive as to whether cells are in stationary phase or decline phase at the last time point. Likewise, it is important to note that the last time point in the denser culture is an outlier in terms of growth and mAb productivity.

At the furthest distance in all dendrograms obtained, two clusters can be identified, which appear to correlate with productive/non-productive stages of the culture, with the 72h time-point for the $2x10^5$ cell/mL inoculation density being the only exception. This time-point is classified as non-productive, although production is observed. Finally, cell cycle distributions remain stable throughout culture for all densities.

Although an increasing G_0/G_1 percentage is observed for the two denser cultures, suggesting a possible arrest as cultures progress through the decline phase,

the distributions obtained are not statistically significantly different from those obtained during stationary phase and hence no such conclusion can be drawn.

The relationships identified above between metabolism and growth kinetics, productivity and cell cycle are identical for all inoculation densities. Therefore, it is worth trying to analyse the entire dataset at the same time. An integrated analysis would bring more data, and hence more information, so as to make a stronger tool to analyse cultures based on metabolic profiles. This is shown in the next section.

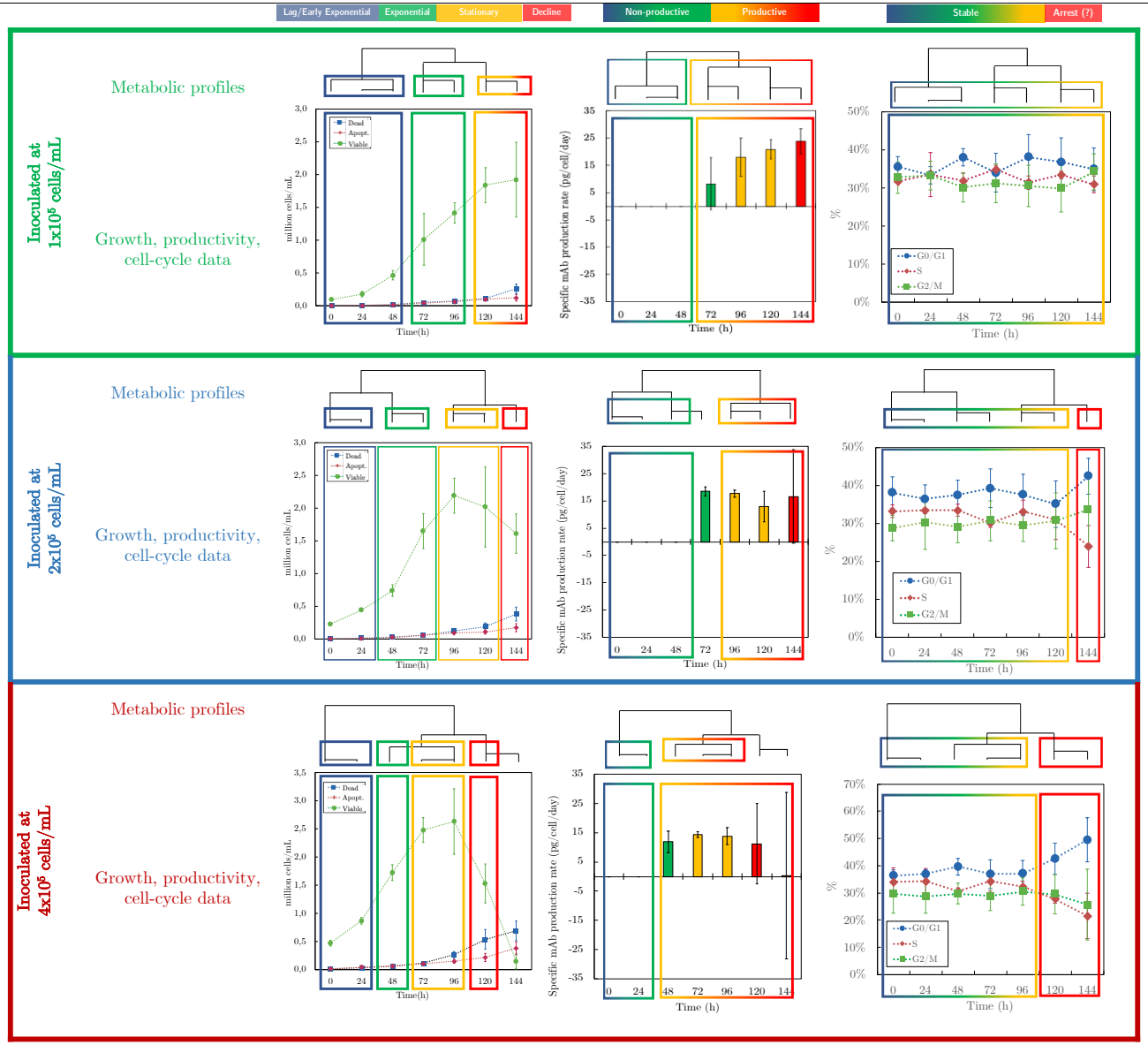


Figure 30: Hierarchical clustering obtained with profiles of 23 metabolites (20 amino acids, glucose, lactate and ATP) measured in the course of batch cultures of GS-NS0 cells at three different inoculation densities are superimposed to the time profiles of growth (left column panels - viable cells in green circles, apoptotic cells in pink diamonds and dead cells in blue squares); specific mAb production rate (central panels) and cell cycle distributions (right column panels - G_0/G_1 in blue circles, pink S in pink diamonds and G_2/M in green squares).

4.6.7.2 Integrated analysis of metabolic profiles for batch cultures

The conclusions drawn from hierarchical clustering, in the previous section, appear to be valid across the three inoculation densities tested. Therefore, an integrated analysis of all metabolic profiles obtained for all thee inoculation densities was conducted in order to assess whether metabolic features related to important culture information could be reconciled for all culture conditions, making it a feature of GS-NS0 cultures.

HCL was performed on all data sets simultaneously in an attempt to identify relevant clusters from the metabolism information. As shown in Figure 31A, four clusters could be identified at the shortest distance in the HCL dendrogram. As was done for the density-specific analysis in the previous section, the clusters were labelled lag- and early exponential growth, exponential growth, stationary phase and decline phase, based on the time-points they contain. The last time point of the densest culture has been confirmed as an outlier from the entire data set. One possibility for the isolation of this time point is the onset of other forms of cell death acting upon severe metabolic stress. Unlike results obtained in the previous section, where it was not possible to clearly relate the last time point of the least dense culture (D1) to stationary or decline phases, the analysis of the entire data set clearly shows that, at the last day, the least dense culture is already in decline phase (Figure 31A). The heatmap resulting from the alignment dictated by HCL (Figure 31B) shows that the transition from lag/early experimental growth into the exponential phase is associated with a depletion in the concentrations of most amino acids. Chiefly, the concentrations of aspartate and glutamate witness a very rapid decrease. These amino acids can be replenished by asparagine and arginine, respectively, whose concentrations remain at average values during the exponential growth validating the hypothesis of arginine-supported growth upon depletion of glutamate. Additionally, lactate accumulates during lag/early exponential and exponential phases and starts reducing on the transition to stationary phase. This lactate metabolic shift had not been identified in batch GS-NS0 cultures with serum

(García-Münzer, Kostoglou, *et al.*, 2015; Kiparissides, Pistikopoulos and Mantalaris, 2015; Quiroga-Campano, 2016). It is interesting to observe that, within stationary phase, two groups can be distinguished, and that lactate is a key difference between them tracking the lactate metabolic shift to mid-stationary phase in all cultures. As cells transition to stationary growth, arginine and asparagine concentrations are lower and alanine, accumulated during earlier stages, starts to be consumed. Finally, the decline phase is characterized by the presence of glutamate, aspartate and glutamine in the supernatant. Additionally, glycine concentrations reach a maximum and ATP concentrations decrease. The last time-point of the densest culture is an outlier characterized by extremely low ATP concentrations as well as full consumption of lactate.

Two distinct groups at the longest distance in HCL can be identified, coincident with the depletion of glutamate in culture, which corroborates the crucial role of this amino acid for GS-NS0 cells.

The clusters obtained by HCL have been confirmed by Principal Component Analysis (PCA). This technique has allowed to identify a space defined by 3 principal components, PC1, PC2 and PC3 describing 58, 25 and 5.4% of the data variance, respectively (Figure 31C). As PC3 only describes 5.4% of the variance, a 2D representation of the PCA space is provided in Figure 31D based on PC1 and PC2 only. With the information obtained from HCL, different areas within this space can be identified, providing an important tool for data visualization. Different colours in the PCA plot show different growth stages identified by HCL and one outlier, as discussed above. The hyperplane shown in Figure 31E and the black bisector line in the 2D PCA plot (Figure 31D) separate the two clusters identified at the furthest distance in the dendrogram of Figure 31A, allowing to differentiate between time-points taken before and after glutamate depletion. Similarly, the hyperplane shown in Figure 31F and the purple bisector line in the 2D PCA plot (Figure 31D) separate productive from non-productive stages of cultures, although this separation could not be identified by HCL.

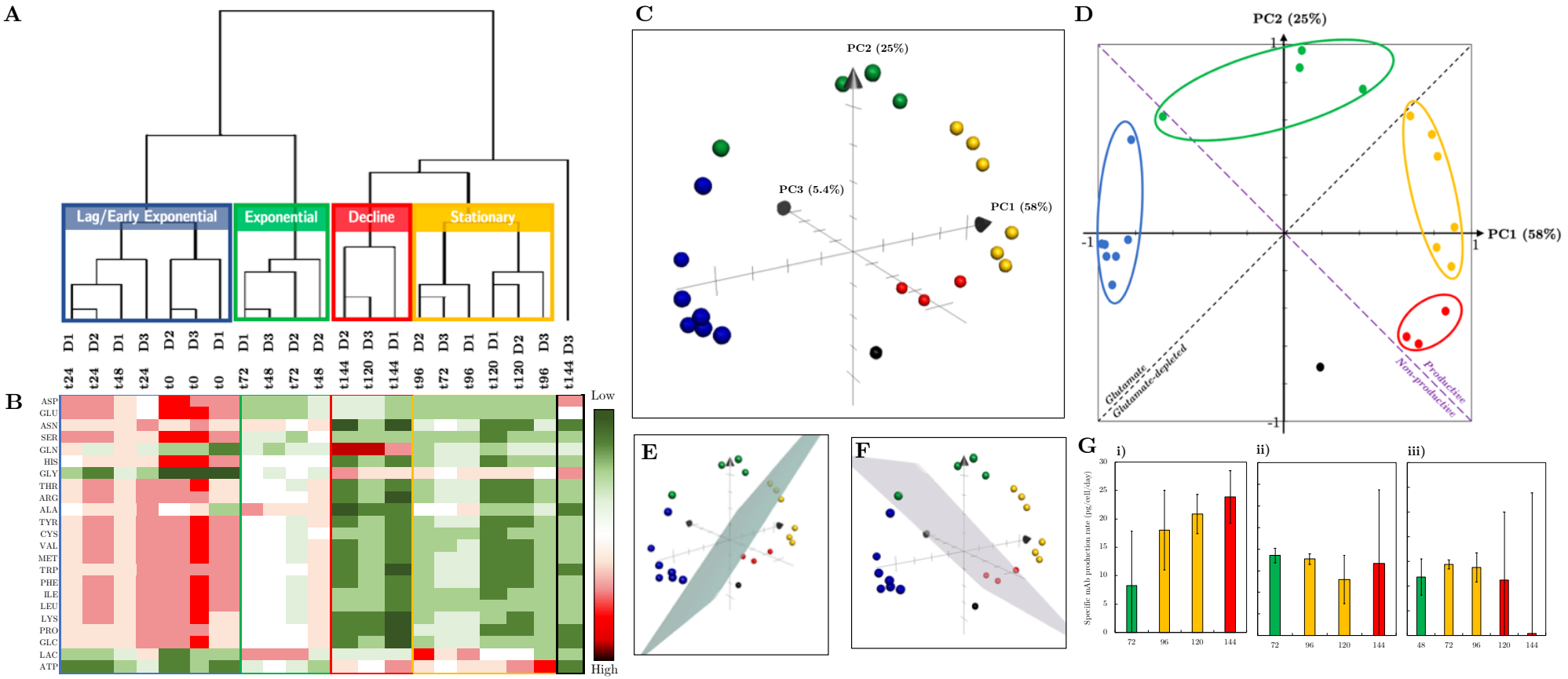


Figure 31: Hierarchical clustering (A), heat map (B) and PCA (C) of metabolic profiles obtained from all GS-NS0 cultures. A 2D PCA plot based on PC1 and PC2 only is also shown (D). Samples in hierarchical clustering are identified by culture time (t=0-144h) and density level (D1, D2 and D3 respectively for inoculation densities of 1,2 and 4x10^5 cells/mL). The last time point of the densest culture (144h, D3) is an outlier confirmed by all techniques. The two clusters identified at the maximum distance in the dendrogram (A) correspond to groups above and below the black dashed bisector line in the 2D PCA plot (D) or the hyperplane in the 3D PCA plot (E). The purple dashed line in the 2D PCA plot (D) and the hyperplane in 3D (F) separate productive and non-productive clusters although this separation cannot be identified from hierarchical clustering (A).

The integrated analysis benefits from having information from different cultures at different conditions contributing for a better understanding of the culture behaviour and the information that can be drawn from the analysis of the metabolic profiles. Integrated HCL has confirmed PCA results and the analysis of the principal component space has allowed to identify growth, metabolism and productivity features in GS-NS0 cell cultures based solely on metabolic profiles. Specifically, exponential, stationary and decline phases have been identified as productive, which has been corroborated by specific productivity calculations. The transition from the non-productive lag/early-exponential phase into the productive stages, beginning at the exponential phase, is dictated by the lactate and glutamate metabolic shifts. However, because both occur simultaneously, it is not straightforward to identify whether both are responsible for this transition or if only one them (and which one) drives the transition into the exponential phase alone.

4.7 CONCLUSIONS

In this chapter, an integrated, kinetic model of cell cycle, metabolism, cell death (apoptosis) and mAb production for GS-NS0 cells is presented. To the best of the author's knowledge, this is the first model able to describe these biological features in an integrated manner making use of biologically significant relationships between growth, metabolism, death and protein production.

The GSA of the model has allowed to identify the significant model parameters, to design meaningful experiments and to obtain relevant information about the system under study and to conduct model-aided experimental development. Specifically, the cycling times were identified as significant parameters for several model outputs, being predominantly important at the early stages of the culture. The cycling times of GS-NS0 cells in chemically defined serum-free conditions were experimentally determined for the first time in this work, obtaining a total cycling time of 22h, 4h longer than experimental estimations

in cultures with serum, corroborating that the presence of serum in culture increases the growth rate. Moreover, in terms of mAb production across the cell cycle, GSA has hinted the importance of later phases of the cycle to mAb production, thereby predicting an experimentally observed feature of this cell line. GSA has shown to clearly identify significant relationships between model parameters and model outputs, which reflect the biology of the system.

Finally, the identification of the important model parameters has significantly reduced the dimension of the parameter estimation problem (from 63 to 30 parameters) allowing for a successful solution using the MLE algorithm available in gProms. Experimental validation conducted with two independent data sets demonstrated the model's ability of predicting batch culture data within a 1-4x10^5cell/mL range of inoculation densities. The model is computationally tractable and has a short computation time making it a suitable tool for model-based optimization – including understanding the impact of arrest techniques, developing fed-batch strategies, optimizing media compositions, delaying the onset of apoptosis – and control applications.

HCL and PCA have been successfully used as unsupervised tools for multivariate metabolism data analysis. Using 23 metabolic profiles simultaneously, and without any further information on the culture conditions or the sampling times, these tools were able to group (cluster) the time points according to the corresponding growth stage of the culture and to do so in chronological manner. In addition, the simultaneous analysis of all data sets by HCL and PCA has allowed to identify regions occurring before and after the glutamate and lactate metabolic shifts occur, corroborating the importance of these two metabolic changes in GS-NS0 cultures. Moreover, two clearly distinguishable clusters related to mAb productivity could be identified by hierarchical clustering of inoculation density-specific profiles but combining data for all cultures, such clustering was not observed. However, PCA of the entire data set has allowed to distinguish productive and non-productive clusters.

The promising results obtained with GS-NS0 metabolic profiles, motivate the application of these techniques to other cell lines as a means to increase process understanding. Moreover, they open the possibility of developing data-driven models expanding the PCA results presented to development of Projection on Latent Structures (PLS) models which can relate metabolic profiles to cell viability, mAb productivity etc. with potentially important predictive capabilities.

The work presented in this chapter has shown that it is possible to use MVA tools to support the development of integrated kinetic models linking several cellular functions in a computationally efficient manner while capturing population heterogeneity. The latter is essential in linking biological kinetics to bioreactor macro environment towards the improvement of culture system design and scale-up.

Chapter 5

Mapping the growth and metabolism of GS-CHO cells

In this chapter, the study of growth, death and metabolism of a GS-CHO cell line is described. The work presented in chapter 4 motivates the development of integrated models able to describe different biological phenomena simultaneously and has identified the application of unsupervised multivariate data analysis techniques as tools to understand culture growth and productivity. These methodologies will be applied to data obtained in batch cultures of GS-CHO cells to investigate whether relevant information pertaining to growth and productivity can be drawn in this case as well. The work reported in this chapter will support the development of a mathematical model of cell-cycle, metabolism and apoptosis of these cells presented in chapter 6.

5.1 INTRODUCTION

In spite of being the workhorse of the biopharmaceutical industry, there are some aspects of CHO cells metabolism and growth which have not been covered in previous studies and are essential to improve process understanding and to support integrated mathematical models aimed at accelerating process development. Among such issues, are detailed studies of CHO cell cycle. The importance of detailed

studies of cell cycle for the biopharmaceutical community derives from reports of different mAb production rates across the different phases of the cell cycle in CHO cells (Dutton, Scharer and Moo-Young, 2006; Park *et al.*, 2015) and has motivated the development of technology for on-line measurement of cell cycle (Fuge *et al.*, 2017). In addition, the need for detailed studies, particularly those aimed at identifying the co-existence of a quiescent cell population (in phase G_0) with proliferative cells in bioreactors, has also been specifically identified (Sitton and Srienc, 2008). The work of Garcia-Münzer *et al.* filled an important gap in cell cycle literature of industrially relevant cell lines, by studying the cyclin blueprint of GS-NS0 cells (García-Münzer *et al.*, 2013). They concluded that the expression of cyclin E1 was a good biomarker for the G_1-S transition, while cyclin B1 was a suitable protein to track the S-G_2/M transition. In addition, they reported the cycling times of GS-NS0 cells in cultures with serum (García-Münzer *et al.*, 2013; García-Münzer, Kostoglou, *et al.*, 2015) and the corresponding values for serum-free cultures are reported in chapter 4 of the present work. However, similar works for other industrially relevant cell lines, namely CHO or Sp2/0, remain unavailable.

Literature on CHO cell metabolism is extensive, covering works with metabolic networks essentially covering central metabolism (Carinhas *et al.*, 2013; Pan *et al.*, 2017; Templeton *et al.*, 2017) and expanding to genome scale models (Hefzi *et al.*, 2016) in batch (Carinhas *et al.*, 2013), fed-batch (Pan *et al.*, 2017; Templeton *et al.*, 2017) or perfusion (Templeton *et al.*, 2017) cultures. While some of these studies link metabolism to cell growth – for example by studying metabolism at different culture phases (Carinhas *et al.*, 2013) – the identification of such culture phases is often based on the observation of the growth kinetics alone (Carinhas *et al.,* 2013; Dean and Reddy, 2013; Robitaille, Chen and Jolicoeur, 2015) or on the lactate metabolic shift (Kyriakopoulos, Polizzi and Kontoravdi, 2013; Templeton *et al.*, 2014). In addition, most available studies are usually based on steady-state techniques, such as Metabolic Flux Analysis (MFA). Therefore, there

is a need for unsupervised methodologies to link metabolism and growth and for tools to describe cell metabolism in a dynamic manner.

5.2 OBJECTIVES

The aim of the present chapter is to study cell growth and death, cell cycle and metabolism of GS-CHO cells and to establish relationships between these biological processes. To this end, the following objectives are set:

- Use the multivariate data analysis (MVA) framework presented in chapter 4 to establish whether metabolic profiles obtained from GS-CHO cultures can provide information related to growth, death and productivity of this cell line;
- Study the metabolism of GS-CHO cells to support the development of a dynamic mathematical model;
- Design and execute experiments that allow for the identification of biomarkers involved in cell cycle regulation and kinetics of GS-CHO cells;
- Understand whether metabolic limitations are important for cell cycle regulation in GS-CHO cells;

5.3 METHODOLOGY

GS-CHO cells were cultured in batch mode and the concentrations of glucose, lactate and ammonia in supernatants collected every 24h were measured. Extracellular concentrations of 20 amino acids and monoclonal antibody were measured by HPLC, using appropriate methods, and ATP concentrations were also measured, using a luminescence assay. Data were analysed using multivariate methods, namely HCL and PCA. The interpretation of metabolic profiles was used to understand the transitions between phases and the metabolism in each stage of the culture.

The proliferation of GS-CHO cells was analysed using two assays, namely the expression of protein Ki-67 and the EdU assay. Ki-67 is a protein expressed only in proliferative cells (Scholzen and Gerdes, 2000) and it was measured by flow cytometry in order to identify the presence of a G_0 population. The EdU method was used to determine the cycling times, in the same manner as described in chapter 4 for GS-NS0 cells.

In order to establish the dynamic cyclin expression profiles of GS-CHO cells, and to identify suitable biomarkers to track cell cycle transitions, three cyclins were studied, namely cyclins E1, D1 and B1, in the course of batch cultures of GS-CHO cell cultures, using flow cytometry. Cyclins E1 and D1 are both considered G_1-cyclins but a constant expression of cyclin D1 throughout the cell cycle has been reported in normally proliferating cells, making it unsuitable as a phase G_1 biomarker. Specifically, this has been identified in the industrially relevant GS-NS0 cell line, supporting the need to study another G_1-cyclin (García-Münzer *et al.*, 2013).

The flow cytometry data analysis procedure included gating viable cells from debris, doublets and non-viable cells. Next, the DNA was deconvoluted using the software tool. Fluorescence-minus-one (FMO) controls were used for cyclins and Ki-67 expression measurements, as was done for EdU analyses. A PI single-stain control was used as FMO to gate the cyclins- or Ki-67-negative regions before analysis of samples stained with DNA+cyclin or DNA+Ki67, respectively. For samples analysed in the course of the EdU proliferation assay, DNA+CyclinE or DNA+CyclinB controls were used to gate the EdU negative regions while the cyclin-negative regions were gated using a DNA+EdU control. Following the methodology suggested by Garcia-Münzer *et al.* for NS0 cells, the expression of cyclins E1, D1 and B1 was assumed to be minimal at phase G_2/M, phase S and phase G_1, respectively (García-Münzer *et al.*, 2013).

5.4 RESULTS AND DISCUSSION

5.4.1 GROWTH AND MAB PRODUCTION PROFILES

Batch cultures of GS-CHO cells were conducted for 10 days (240h) at three inoculation densities, specifically 1, 2 and 4x10^5 cells/mL. Figure 32 shows the corresponding profiles of viable cell density, viability and mAb titres obtained for these cultures.

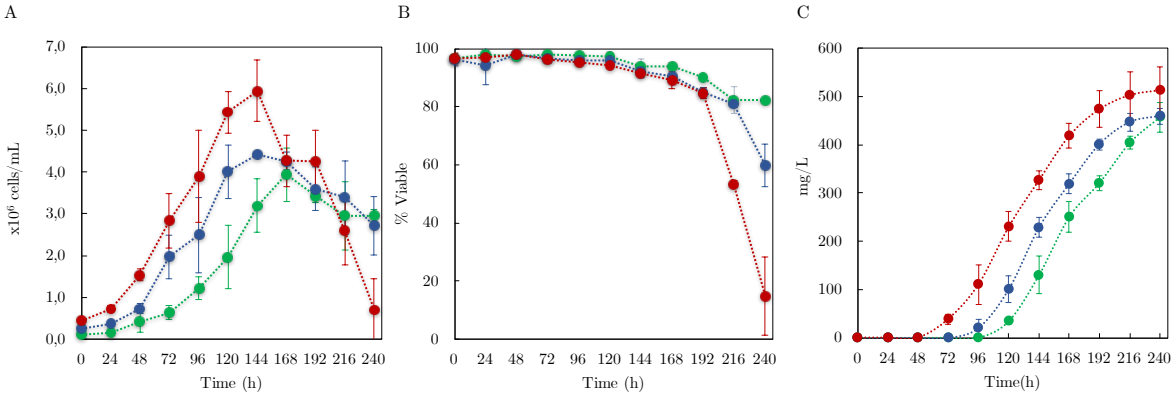


Figure 32: Viable cell density (A), viability (B) and mAb titre (C) measured in the course of 240h of GS-CHO cells batch cultures inoculated at 1 (green), 2 (blue) and 4x10^5 cells/mL (red). All cultures were conducted in triplicates (N=3 biological replicates, error bars show one standard deviation).

Data in Figure 32 show that the densest culture reaches the maximum peak viable cell density while the other two cultures reach a lower peak viable cell density (Figure 32A). The viabilities of all cultures are very comparable (Figure 32B) during the first 192h when they start to decrease sharply. The evaluation of the growth kinetics (Figure 32A) suggests that the lag phase is shorter for the denser culture and longest for the least dense culture. As a consequence, the densest culture starts to accumulate mAb earlier, starting from 72h, whereas a comparable mAb concentration is only seen for the least dense culture at 120h (Figure 32C). The titres reported in this work are comparable to those reported for other GS-CHO clones (Carinhas *et al.*, 2013; Kyriakopoulos, Polizzi and Kontoravdi, 2013).

5.4.2 MULTIVARIATE DATA ANALYSIS OF METABOLIC PROFILES

In this section, the metabolic profiles of glucose, lactate, ammonia, 20 amino acids and ATP obtained for all cultures are analysed by hierarchical clustering and PCA. These results are shown in Figure 33.

The first important result is that, at the shortest distance in the dendrogram obtained from HCL (Figure 33A), the same clusters obtained in chapter 4 for GS-NS0 cells can be identified for GS-CHO cells. Specifically, it is possible to identify one cluster with time-points collected during lag and early exponential growth, another comprising time-points collected during exponential growth, a third one with samples collected during the stationary phase and, finally, a cluster with time-points referring to the decline phase (Figure 33A). From the heat-map of Figure 33B, the lactate metabolic shift appears to occur on the transition between early exponential and exponential phases, while the aspartate metabolic shift occurs concurrently with the exponential/stationary transition.

The clustering results are corroborated by PCA (Figure 33C and D). Three principal components (PC1, PC2 and PC3) could be obtained describing respectively 58.6, 23.5 and 6.4% of the variance in the data and the relative position of the clusters is identical to that obtained for GS-NS0 cells (chapter 4). Clusters furthest apart in the dendrogram of Figure 33A can be separated by a hyperplane in the 3D PCA plot of Figure 33C, as shown in Figure 33E. This hyperplane separates lag and exponential growth from stationary and decline phases. As seen for GS-NS0 cells, here too there is no direct evidence of a separation between productive and non-productive clusters in the hierarchical clustering analysis (Figure 33A) but it is possible to distinguish clusters based on productivity by drawing a hyperplane in the 3D PCA plot as shown in Figure 33F or the bisector line in the 2D PCA plot (Figure 33D). The productivity levels for each cluster and each inoculation density are shown in Figure 33G and confirm these results (no mAb production was observed during lag/early exponential phase).

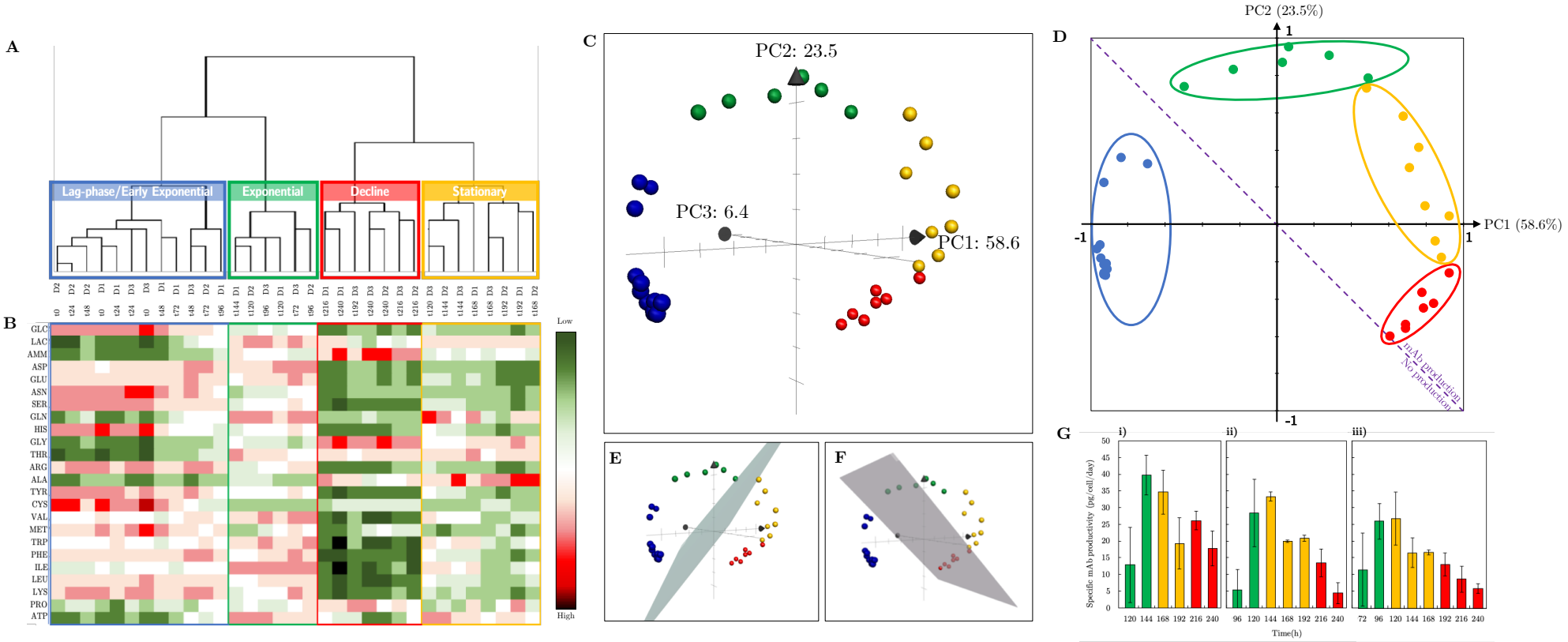


Figure 33: Multivariate metabolic profile analysis. Glucose, lactate, ammonia, 20 amino acids and ATP profiles obtained in the course of 10-day batch cultures of GS-CHO cells were analysed by hierarchical clustering (A), the corresponding heat map (B) and Principal Component Analysis (C). A 2D representation of PCA using only PC1 and PC2 is shown in (D). Samples are identified by time (0-240h) and density level (D1, D2 and D3 for 1, 2 and 4x10^{5}cells/mL). Clusters separated at the longest distance in the dendrogram (A) can be distinguished by a hyperplane in the PCA space (E). Productive and non-productive clusters can be distinguished by drawing a hyperplane in the 3D PCA space (F) or the bisector line in 2D PCA plot (D) and confirmed by specific mAb productivities (G) obtained for inoculation levels (i) D1, (ii) D2 and (iii) D3. Colours in panels C-G refer to clusters identified in panel A.

A detailed analysis of each of the culture phases identified from MVA shows that the growth rates and the nutrient consumption rates observed for all cultures in each phase are comparable (data shown in Appendix II). However, the duration of the different phases depends on the seeding density. Comparable growth and nutrient consumption rates, but different starting densities, explain why the densest culture reaches the maximum peak viable cell density. The different duration of the growth stages explains different final mAb titres identified in section 5.4.1. The densest culture spends 192h in productive stages (exponential, stationary and decline), while the least dense culture has a productive time of only 144h.

Productivity values obtained in this work are different from those reported for cultures of low- and high-producing GS-CHO clones by Kyriakopoulos, Polizzi and Kontoravdi. The authors have used the lactate metabolic shift as criteria to report their data and identified the lactate producing stage as a mAb producing stage, with specific mAb production rates of 12 and 27 pg/cell/day for the low- and high-producing clones, respectively. During the lactate consumption stage, these rates reduced to 4 and 8 pg/cell/day, respectively (Kyriakopoulos, Polizzi and Kontoravdi, 2013). In this work, during lactate-producing stages (lag and early exponential phase and exponential phase) the specific mAb production rates were 0 and 17 pg/cell/day respectively, for the nominal inoculation density, while during stationary and decline phases (lactate consumption stages) the specific mAb production rates increased to 24.2 and 8.97 pg/cell/day (a detailed comparison between growth and metabolic rates obtained in this work and those reported in other studies is shown in Appendix II). This suggests that the relationship between the lactate metabolic shift and the mAb productivity could be clone-specific. The analysis of 243 bioreactor runs at different scales conducted at Genentech reported by Le *et al.* supports the results obtained in this work. The authors concluded that lactate was a good process indicator since the lactate metabolic shift had occurred in nearly all of their top titre runs, while those runs delivering lower titres did not switch to lactate consumption (Le *et al.*, 2012).

The application of unsupervised techniques to analyse metabolic data and identify the time horizons of each growth stage, has led to some important differences in the values of nutrient metabolism rates reported for each stage. For example, in this work asparagine and serine were identified as the most consumed amino acids during the earlier stages of culture, in agreement with previous studies (Carinhas *et al.,* 2013; Kyriakopoulos, Polizzi and Kontoravdi, 2013). However, while those studies identify aspartate (Carinhas *et al.*, 2013) or glutamate (Kyriakopoulos, Polizzi and Kontoravdi, 2013) as the third most consumed amino acid during exponential growth, in this work production of both aspartate and glutamate occurred during the lag/early-exponential phase of culture. The depletion of both aspartate and glutamate has been identified by MVA as being linked to the end of the exponential growth and the beginning of the stationary phase (Figure 33B). Additionally, isoleucine production is also identified in this work during the lag/early exponential phase, while this amino acid is consumed throughout the remaining stages of culture. These are, to the best of the author's knowledge, novel, and possibly clone specific trends not identified before.

Another significant difference is the alanine production/consumption rates. Carinhas *et al.* calculated alanine production rates of 154 and 65 fmol/cell/day, respectively, for exponential and stationary growth of their high producer clone, and 103 and -11 fmol/cell/day for the low producer (Appendix II). These numbers suggest that the cultures of the low-producer experience an alanine metabolic shift on the transition from exponential to stationary phase while, for the high producer, no metabolic shift is observed. Carinhas *et al.* used the observation of growth kinetics to discriminate between exponential and stationary phases and conducted MFA for the different growth stages identified. Their MFA results show alanine production throughout cultures. This is contradicted by their own dynamic profiles showing that the alanine metabolic shift occurs in both cultures (Carinhas *et al.*, 2013). Conversely, in this work, alanine consumption has been identified as a feature of the transition from stationary to decline phase, while this amino acid is

consumed during all previous culture stages. Therefore, it can be inferred that the MVA framework derived in this work provides a stronger support to establish the time horizons of each growth stage in culture, thus being useful to support MFA studies, which require steady state assumptions to hold for the time horizon of the analysis. Following the MVA of metabolic data conducted in this work, four different regimes were identified and hence four MFAs would be necessary to understand the metabolism during each phase, as was done in a prior study (Templeton *et al.*, 2013).

In the majority of prior studies, high-throughput methodologies, such as gas chromatography-mass spectrometry (GC-MS), where several tens of metabolites can be detected, are used alone or in combination (GC-MS and HPLC e.g.) to analyse a broader spectrum of metabolites (Chrysanthopoulos, Goudar and Klapa, 2010; Dietmair *et al.,* 2012; Vernardis, Goudar and Klapa, 2013). In this work, data were acquired using a bioprofile, HPLC analysis with on-line derivatisation and a luminescence assay, all of which have the potential of being performed at-line or even on-line. Results showed that MVA of metabolic data is able to provide information on growth, to distinguish productive and non-productive time-points, to identify culture stages in a chronological manner and to identify essential metabolic changes pertaining to the transitions between different growth stages. In addition, it is possible to distinguish between clusters occurring before and after the depletion of the main nitrogen source which is also acknowledged in literature to be essential for TCA cycle, and hence for energy generation. The metabolic changes responsible for the transition between the different growth stages are discussed in the next section.

5.4.3 METABOLIC SHIFTS: UNDERSTANDING METABOLISM

Once different culture regimes can be identified based on the analysis of the metabolic profiles, it is important to understand what metabolic changes are related to the transition between different culture stages. This information is important for

the optimization of cell culture through the development of feeding strategies, medium design or a combination of both objectives. This analysis will be conducted using data from the culture inoculated at $2x10^5$cells/mL only, as this is the nominal density suggested by the supplier.

5.4.3.1 Lag- and Early-exponential growth

The first culture regime covers time points taken between 0 and 72h of culture. In this period, glucose is consumed, and lactate is produced (Figure 34A). According to previous studies with GS-CHO cells, glucose is mostly degraded through glycolysis, producing pyruvate. Some of the pyruvate produced may be used in the TCA cycle and some will be used to produce lactate. In spite of accumulating a toxic by-product, lactate production regenerates the co-factor NAD^+ required for glycolysis. As discussed above, and in line with other studies (Carinhas *et al.*, 2013; Kyriakopoulos, Polizzi and Kontoravdi, 2013), asparagine is the fastest consumed amino acid (-459 fmol/cell/day) in the first 72h of culture. This amino acid can be converted to aspartate, producing ammonia (Carinhas *et al.*, 2013; Kyriakopoulos and Kontoravdi, 2014), and then to oxaloacetate, a TCA-cycle intermediate, playing an important role in the energy metabolism of GS-CHO cells (Carinhas *et al.*, 2013; Duarte *et al.*, 2014) and other CHO cell lines (Dean and Reddy, 2013). The ammonia-producing conversion of asparagine to aspartate refers to an asparaginase (E.C.3.5.1.1) activity which, although the enzyme is not present in CHO cells, has been included in CHO metabolic networks of different scales (Carinhas *et al.*, 2013; Hefzi *et al.*, 2016). As mentioned by Kyriakopoulos *et al.*, taking this activity into consideration is necessary to describe the metabolism of GS-CHO cells, specifically the ammonia production observed in culture and it may be due to an asparaginase functionality of asparagine synthetase (E.C.6.3.5.4) (Kyriakopoulos and Kontoravdi, 2014). The conversion of aspartate to oxaloacetate is catalysed by aspartate transaminase (E.C.2.6.1.1), which transfers the amino group in aspartate to α-ketoglutarate, producing glutamate. Compartmentalized MFA of CHO-K1 cells suggests that this reaction occurs mostly

in the cytoplasm producing oxaloacetate which can then be converted to acetyl-CoA for energy production (Nicolae *et al.*, 2014). This hypothesis is also supported by compartment-specific enzyme activity studies showing that the aspartate transaminase activity is observed exclusively in the cytoplasm (not in the mitochondrion) in the earlier stages of the culture (Wahrheit, Niklas and Heinzle, 2014). As a result of these reactions, aspartate is produced during this period and production of glutamate, glutamine and isoleucine is also observed (Figure 34B). Glutamate can produced as a result of the conversion of aspartate to oxaloacetate (transaminase reaction) while isoleucine (Figure 34B) can be produced from TCA cycle intermediates (including succinyl-CoA), a reaction considered in a previously published GS-CHO metabolic network (Carinhas *et al.*, 2013).

In order to regenerate the α-ketoglutarate required for the aspartate transamination reaction, cells can either produce alanine from pyruvate or convert glutamate directly to α-ketoglutarate producing ammonia. The increasing concentrations of originally absent alanine, reaching 0.256mM after 72h of culture (Figure 34C), in combination with compartment-specific enzyme activity measurements showing that alanine transaminase is active throughout most of the culture (Wahrheit, Niklas and Heinzle, 2014), suggest that the transaminase reaction with glutamate is occurring. From Figure 34C, it can also be seen that ammonia accumulates in culture. As explained above, this toxic by-product can be produced from asparagine or glutamate degradation and can be consumed for the production of glutamine.

Finally, serine is consumed, and glycine and threonine both accumulate (Figure 34D). This has been suggested to be caused by serine conversion to glycine and from glycine to threonine (Hagrot *et al.*, 2017). However, the methionine cycle, which is linked to histone methylation, and hence gene expression and to the folate cycle, could also be accounted for in explaining these profiles since the conversion of serine to glycine is also part of the folate cycle (Mentch and Locasale, 2016). The suggested metabolic network for this initial phase of culture is shown in Figure 34E.

This metabolic network could be validated with intracellular metabolite measurements.

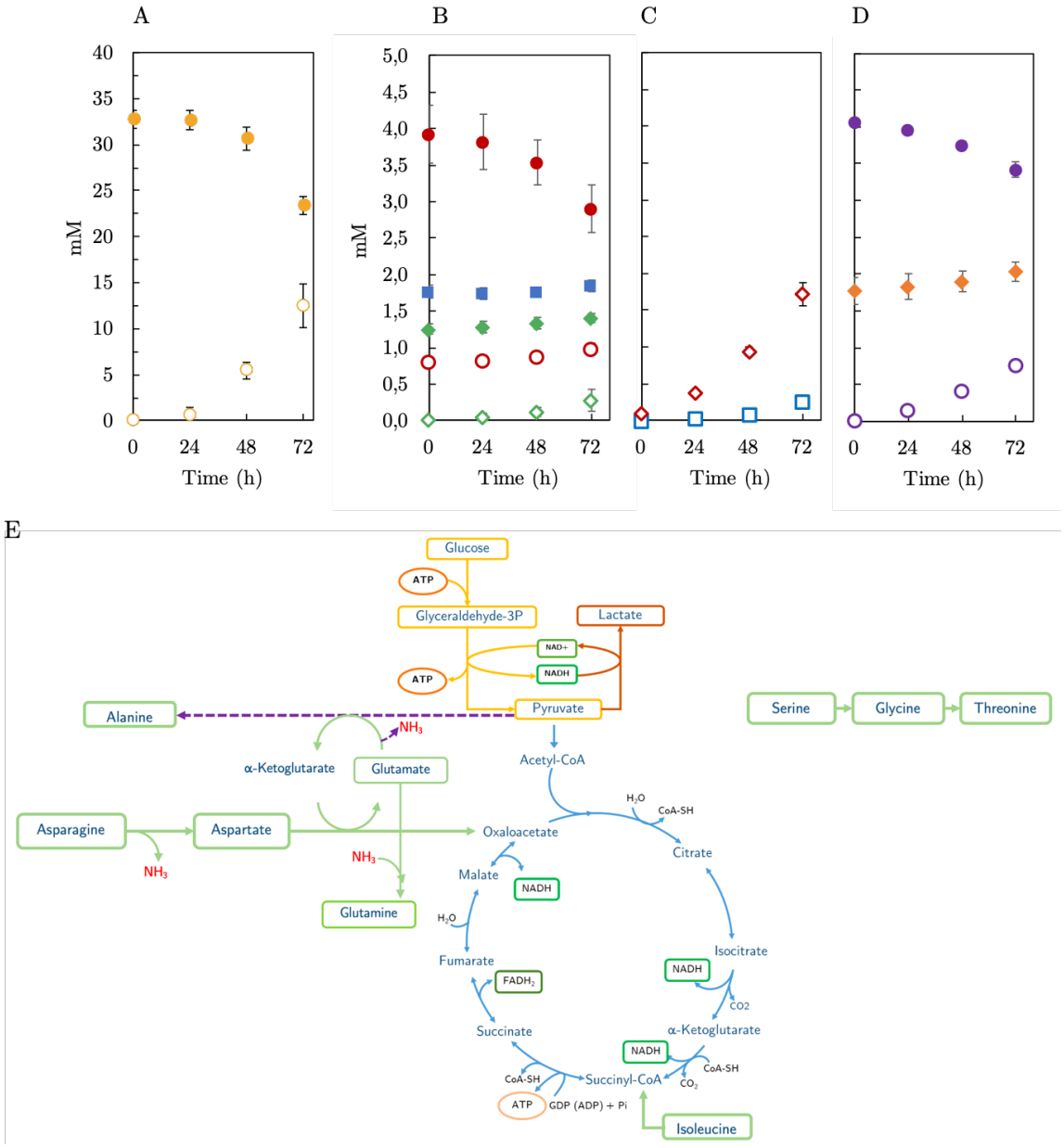


Figure 34: Dynamic profiles of (A) glucose (closed circles) and lactate (open circles); (B) asparagine (closed circles), aspartate (open circles), glutamate (closed diamonds), glutamine (open diamonds) and isoleucine (close squares); (C) alanine (open squares) and ammonia (open diamonds); (D) serine (closed circles), glycine (open circles) and threonine (closed diamonds). Y-axis for plots B-D is the same. (E) Suggested metabolic network for the first 72h of culture. Alternative routes for glutamate to α-ketoglutarate conversion are shown in purple dashed lines (inoculation at 2x10^5 cells/mL; N=3 biological replicates; error bars show 1 standard deviation).

5.4.3.2 Exponential growth

The following culture stage is exponential growth, which begins between 72 and 96h of culture, and finishes between 120h and 144h of culture. During this phase, glucose continues to be consumed at the same rate, but lactate reaches a plateau (Figure 35A) suggesting a lower consumption rate or equilibrium between consumption and production. This feature of the transition to late exponential growth can also be identified from the heat map in Figure 33B. It is also during this phase that aspartate, glutamate and isoleucine (originally produced) start to be consumed (Figure 35B), possibly due to asparagine no longer providing enough carbon to regenerate TCA cycle intermediates. Glutamine concentration reaches a plateau during this stage (Figure 35B). Ammonia continues to accumulate in culture (Figure 35C), while threonine production stops (Figure 35D) and histidine and arginine start to be consumed (Figure 35E). These changes are summarized in Figure 35F.

As one of the important metabolic changes of the transition to exponential growth, it is important to analyse the lactate metabolic shift. In fed-batch cultures of GS-CHO cells, the lactate metabolic shift has been shown to occur independently of the concentrations of asparagine, when the lactate concentration reached approximately 13mM (Kyriakopoulos and Kontoravdi, 2014), consistent with observations in batch cultures (Duarte *et al.*, 2014). The value observed in this work is slightly higher, comparable to another study (Carinhas *et al.*, 2013). Lactate concentrations below 20mM are tolerable for CHO cell cultures (Hartley *et al.*, 2018). Different researchers have attempted to establish mechanisms for the lactate metabolic shift, linking it to glycolysis or to glutaminolysis (Zagari *et al.*, 2013; Hartley *et al.*, 2018), although the exact explanation remains unclear (Hartley *et al.*, 2018). Kyriakopoulos and Kontoravdi report that the lactate metabolic shift always occurs at the same lactate concentration (Kyriakopoulos and Kontoravdi, 2014) although, as recently suggested, the actual drive could be a prior change in

the NAD^+/NADH ratio (Hartley *et al.*, 2018), which has been observed by other authors (Templeton *et al.*, 2013).

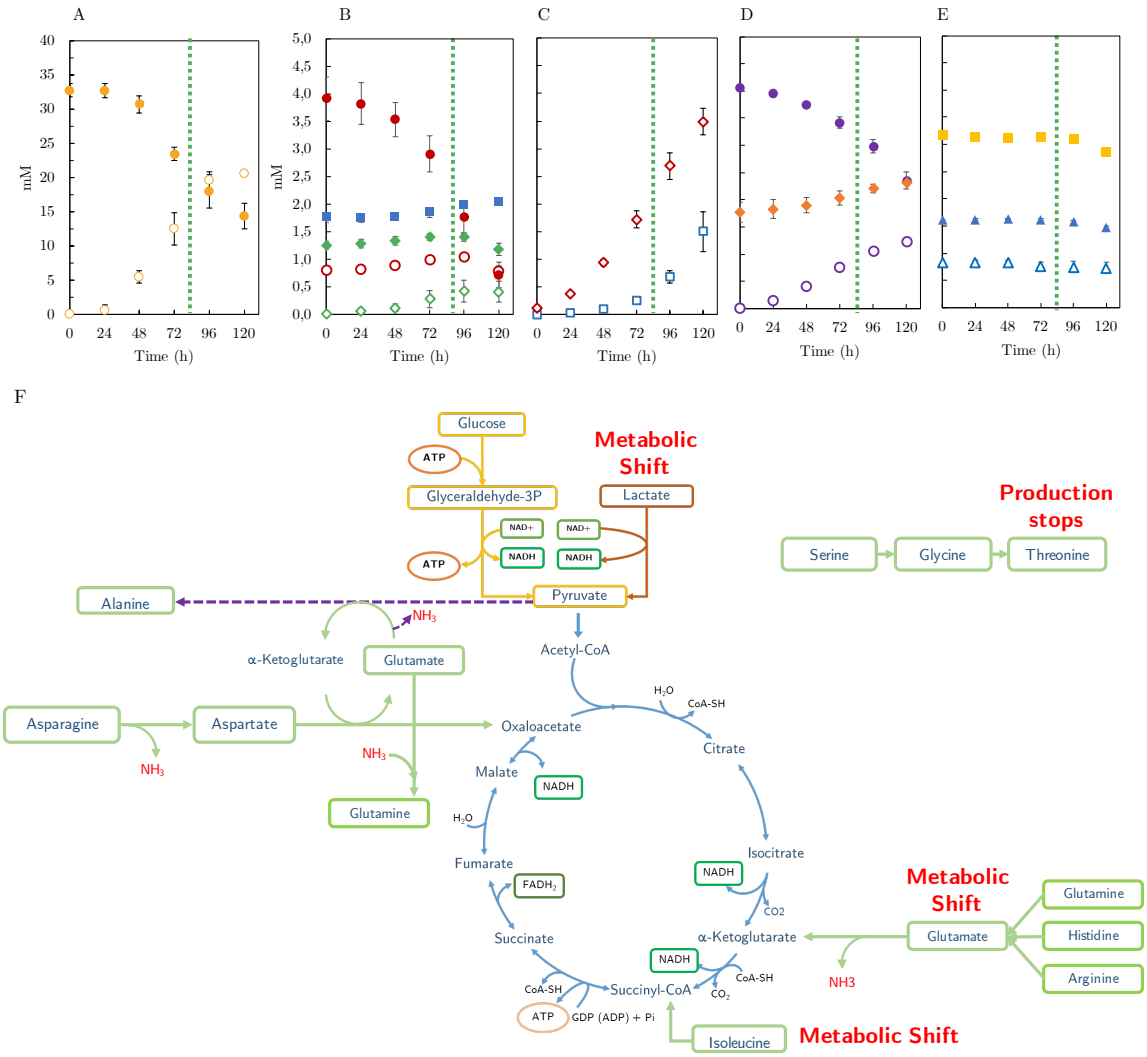


Figure 35: Dynamic profiles in panels (A)-(D) show the same nutrients as in Figure 34. Panel (E) shows arginine (closed triangles), histidine (open triangles) and leucine (closed squares). The transition into exponential growth is shown by a dashed green line in all dynamic profiles. A possible metabolic network and changes related to the transition to exponential growth is suggested (F).

The hypothesis of a glutamine depletion-driven metabolic shift does not appear to explain the results obtained in this work and prior studies for GS-CHO cells. Results show that maximum glutamine concentrations coincide with the occurrence of the lactate metabolic shift and in previously reported fed-batch cultures, the lactate shift occurred despite glutamine concentrations being

consistently over 1.5mM (Kyriakopoulos and Kontoravdi, 2014). Furthermore, in one of the fed-batch cultures reported, all lactate was consumed in spite of glucose concentrations higher than 20mM in culture, also eliminating the possibility of a glucose-controlled metabolic shift.

5.4.3.3 Stationary phase

According to results shown in Figure 33, after 144h cells are already in stationary growth, suggesting that the transition occurred between 120 and 144h of culture. This phase lasts at least 2 days until 192h of culture.

As observed for exponential growth, glucose continues to be consumed at the same rate. During stationary phase, the lactate metabolic shift becomes more evident with decreasing concentrations of this by-product observed throughout this period (Figure 36A). At the beginning of stationary phase, asparagine and aspartate concentrations are low (≈0.2mM) and both nutrients, as well as glutamate, are depleted in the course of this phase, while isoleucine consumption continues (Figure 36B). Glutamine is steadily consumed, a feature also identified by MFA for other GS-CHO clones (Carinhas *et al.*, 2013).

Another relevant feature of the stationary phase is the first ammonia metabolic shift (Figure 36C). As previously suggested, this is probably linked to alanine production occurring until mid-stationary phase (168h) when the alanine metabolic shift occurs (Li *et al.*, 2012). Alanine conversion to pyruvate requires ammonia and, as lactate is consumed producing pyruvate, more alanine is produced. Since aspartate becomes exhausted and glutamate is likely directed to fuelling α-ketoglutarate production to satisfy the energetic needs (Carinhas *et al.*, 2013), it is possible that the α-ketoglutarate/glutamate loop no longer provides enough ammonia for this reaction. Therefore, alanine production requires that ammonia be incorporated, causing the first metabolic shift. At the end of the stationary phase, asparagine, aspartate and glutamate are all exhausted in culture (Figure 36B). Cells change their metabolism again, consuming alanine to produce pyruvate thereby

continuing to fuel the TCA cycle. Alanine conversion to pyruvate produces ammonia, leading to the second ammonia metabolic shift observed at the end of the stationary phase (Figure 36C).

Serine conversion to glycine continues, and the concentration of threonine remains stable (Figure 36D). Leucine consumption continues at the same rate observed during exponential growth and valine consumption begins at this stage. On the other hand, arginine and histidine consumption stop during this phase of the culture (Figure 36E).

5.4.3.4 Decline phase

The last two days of culture are classified as decline phase. The viable cell density decreases and so does viability (Figure 32A and B). During this period, glucose and several amino acids are depleted in culture, while ammonia continues to accumulate. As a consequence of nutrient depletion, the energetic requirements cannot be satisfied as observed by decreased ATP concentration (Figure 33C) and cells die.

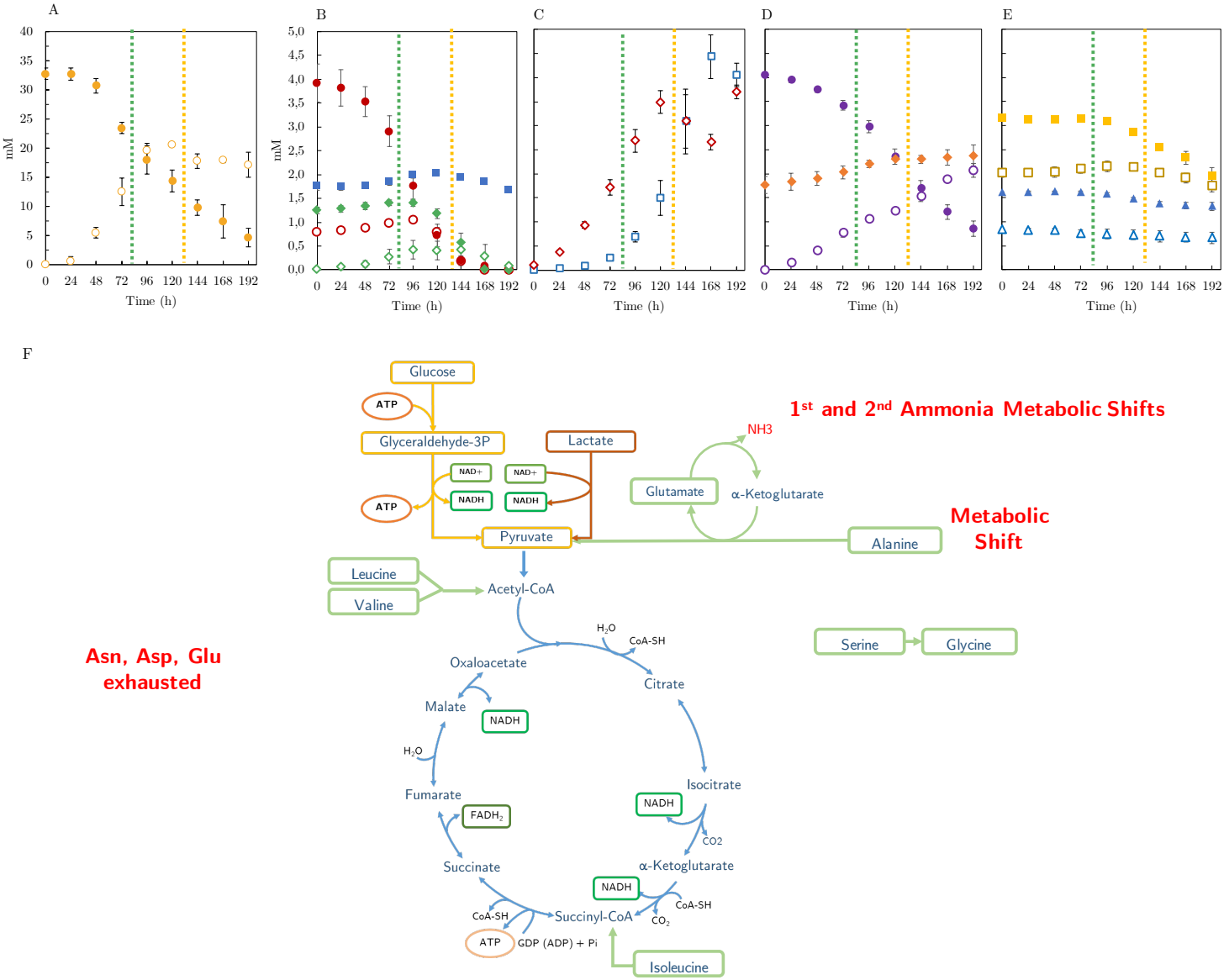


Figure 36: Dynamic profiles of (A) glucose, lactate, (B) asparagine, aspartate, glutamate, glutamine, isoleucine, (C) alanine, ammonia, (D) serine, glycine, threonine, (E) arginine, histidine, leucine (key as in Figure 34 and Figure 35) and valine (open squares) whose consumption starts during stationary phase. The transition into exponential growth and stationary phase are shown by green and yellow dashed line, respectively, in all dynamic profiles. A map of metabolism changes related to the transition to the stationary phase is suggested (F).

5.4.4 ENERGY PRODUCTION

The dynamic energy production profile obtained for all GS-CHO cultures is shown in Figure 37A. Although this profile follows the viable cell densities observed for each inoculation density (Figure 32A), the normalized ATP content per viable cell shows significant differences throughout the cultures (Figure 37B) resembling the dynamic profiles reported for hybridoma cultures (Sonderhoff, Kilburn and Piret, 1992). These differences were confirmed by two-way ANOVA, where inoculation density and time were both identified as significant sources of variance ($p\text{-}value<0.001$) as well as their interaction ($p\text{-}value<0.01$). Estimating the energy requirements of a cell line is important because not satisfying such requirements leads to cell death in culture (Eguchi, Shimizu and Tsujimoto, 1997), but also because it has been shown that this is a useful criterion for medium design (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). In order to estimate the critical specific ATP content for this cell line, hierarchical clustering was used to analyse the data allowing to identify three clusters and two outliers (Figure 37C). The clusters are labelled as "Late-stage", "Early-stage" and "Death" based on the time points that they include (for the purpose of estimating the minimum specific ATP content, the "Death" cluster and the two outliers were not considered).

The "Early-stage and "Late-stage" clusters have the same number of samples and are approximately equally representative of each inoculation density. The box plots obtained for the viability of samples in each cluster (Figure 37D) show that the metabolism of samples in both clusters is able to sustain viability (>80%) and confirms that *t216 D1* is an outlier in the "Early-stage" cluster. One-way ANOVA was used to compare the values within each cluster and no significant differences could be identified ($p\text{-}value>0.05$). Therefore, all samples in each cluster were averaged to calculate the specific ATP contents of 6.78±1.7 and 4.14±1.2 fmol/cell for the "Early-stage" and the "Late-stage" clusters, respectively. These values are in excellent agreement with those reported by Matuszczyk *et al.* for cytosolic ATP contents in CHO cells during different phases of culture. At the beginning of their

cultures, when glutamine and glucose are both available, the authors measured a specific ATP content of 6.28±0.38 fmol/cell, while at the later stages, after glutamine depletion, ATP levels where 4.43±0.41 fmol/cell (Matuszczyk *et al.*, 2015). The specific ATP contents obtained in this work are also in the 4-8 fmol/cell range reported earlier for CHO cells (Sonderhoff, Kilburn and Piret, 1992).

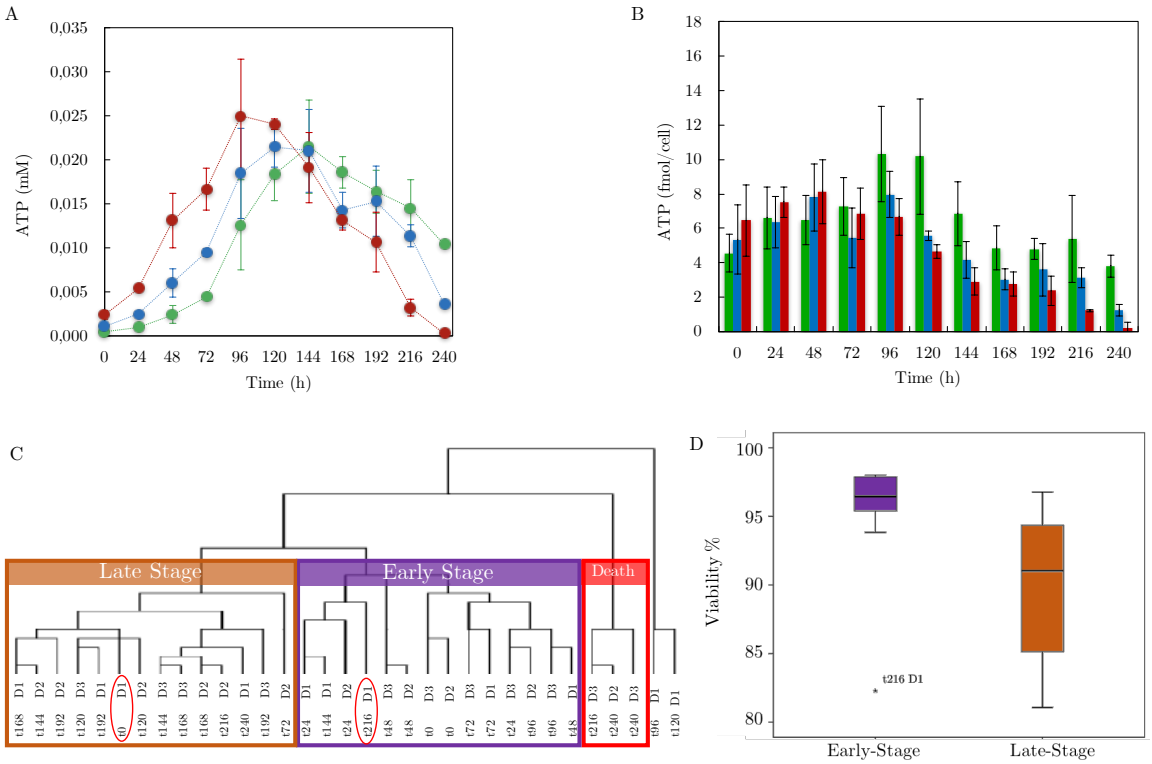


Figure 37: ATP concentrations over 240h of GS-CHO cultures inoculated at 1, 2 and 4x10^5 cells/mL respectively in green, blue and red (A) and the corresponding normalised values (B) which were analysed by hierarchical clustering (C). Time points are labelled according to time (t0-t240h) and inoculation density level (D1-D3). Time points where the clustering does not correspond to the culture stage where they were collected are marked with red ovals. There are also two outliers on the right-most side of panel C. The distributions of viabilities of samples in the "Late-stage" (brown) and "Early-stage" (purple) clusters are shown in a box plot (D) where the in the "Early-stage" cluster is confirmed. (In panels A and B, N=3 biological replicates, error bars show one standard deviation).

Interestingly, the "Early-stage" and "Late-stage" clusters obtained from the specific ATP levels alone correlate with the occurrence of the lactate metabolic shift in culture. Energy generation is another aspect related to the lactate metabolic shift that has not been fully elucidated yet. In the specific case of GS-CHO cells, Carinhas *et al.* have identified that the transition from exponential to stationary

growth phases is associated with a more efficient central metabolism. The authors identify glucose carbon being deviated to glycerol production and to lactate as metabolic inefficiencies during exponential growth. The transition to the stationary phase is accompanied by a more efficient metabolism where carbon deviated from glycolysis to produce glycerol-3-phosphate increases but it is offset by lactate consumed to produce pyruvate (Carinhas *et al.,* 2013) suggesting an important role of the lactate metabolic shift for glycerol production. Chen et al. have also reported that lactate can be consumed to produce lipids, although they suggest that lactate produced is consumed in the mitochondrion to produce citrate, which is mostly deviated to lipid synthesis, although some also remains in the TCA cycle producing α-carbons, such as α-ketoglutarate, which the authors have experimentally confirmed (Chen *et al.,* 2016).

Another hypothesis has been raised by Buchsteiner *et al.* and motivated the stimulation of the PDC (pyruvate dehydrogenase complex) in CHO-K1 cells. Based on the observation that reduced lactate production led to lower ATP levels, the authors suggested that, in the presence of lactate, the maintenance energy requirements to sustain ion gradients are higher, explaining the higher levels of ATP observed before the lactate shift (Buchsteiner *et al.*, 2018). Conflictingly, two research groups observed increased energy generation with decreasing lactate levels (Zhou *et al.*, 2011; Zagari *et al.*, 2013), which has been attributed to lactate inhibition of OxPhos, while also showing that this behaviour is clone- and medium-dependent (Zagari *et al.*, 2013). More evidence that lactate-mediated inhibition of OxPhos cannot be reconciled across all CHO cell lines has been provided by Li *et al.* who studied the impact of feeding lactate to CHO cultures after the shift has occurred. Their results show that, although peak viable cell density occurs simultaneously with the lactate metabolic shift, daily lactate feeding (keeping its concentration at 20mM after the shift) sustained culture viability above 80% for at least 3 days longer than control cultures, while maintaining low ammonia concentrations and comparable viable cell densities (Li *et al.*, 2012).

Consistent with prior reports for GS-CHO cells (Carinhas *et al.*, 2013; Kyriakopoulos and Kontoravdi, 2014), viability of cultures reported in this work was not compromised after the lactate metabolic shift occurred, remaining above 80% for 4 days irrespective of the inoculation density (Figure 32B) and in spite of the reduced ATP content (Figure 37D). Furthermore, for all inoculation densities, ATP production profiles follow those of lactate accumulation (Figure 38) suggesting that, from the previously mentioned hypotheses, those of lactate being used to produce lipids or that of more ATP being required as lactate is produced appear more likely to support the observations herein reported. Conversely, the hypothesis of lactate inhibiting OxPhos – which would lead to increased energy after the metabolic shift – does not appear plausible.

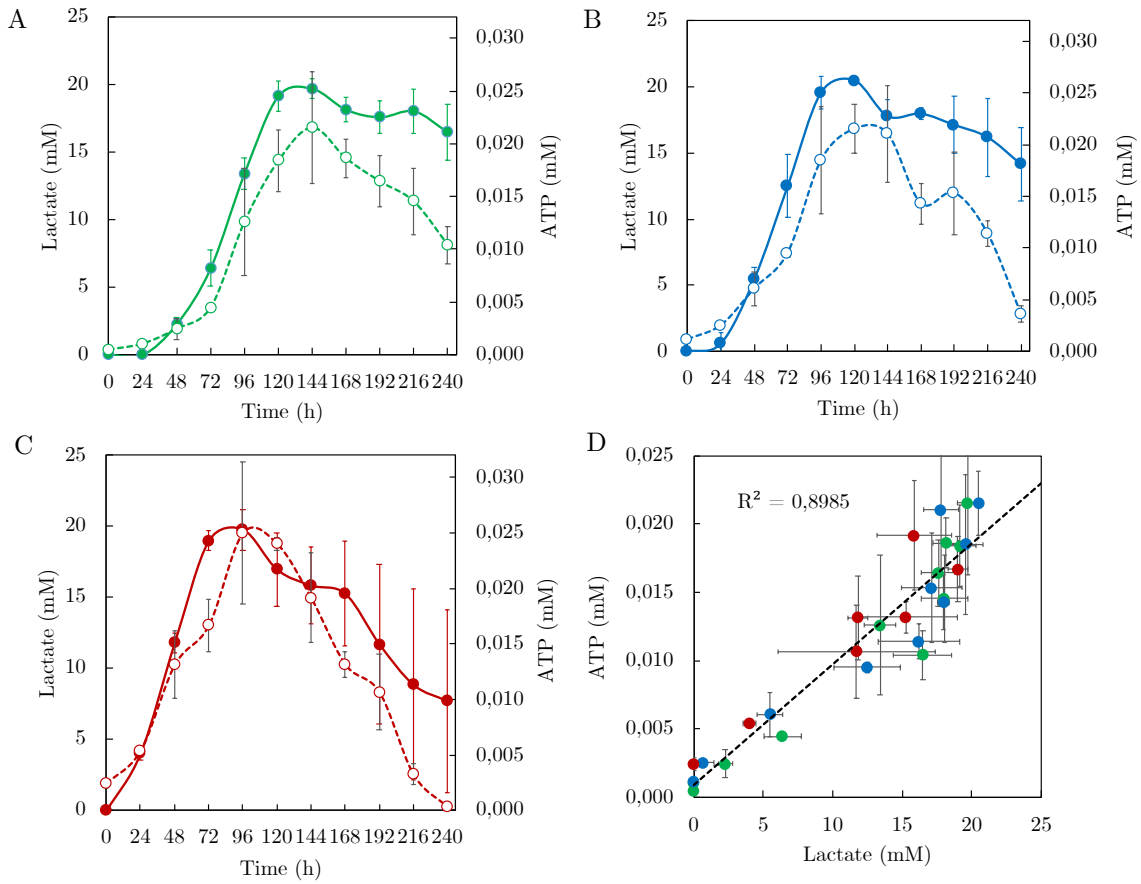


Figure 38: Concentrations of ATP (filled circles, full lines) and lactate (empty circles, dashed lines) obtained in the course of 240h GS-CHO cultures inoculated at (A) 1, (B) 2 and (C) 4x10^5 cells/mL and correlation of ATP and lactate contents obtained for all cultures (D) where points relating to plots A, B and C are shown in green, blue and red, respectively. (N=3 biological replicates, error bars show one standard deviation).

As less energy is produced at later stages of cultures conducted in this work, the specific ATP content obtained for the "Late-stage" cluster is considered to be a better estimation of the minimum energetic requirements of this clone. As shown in Figure 39, samples with specific ATP contents below the minimum observed for this cluster have been classified as decline phase, according to the metabolic profile analysis shown in Figure 33A.

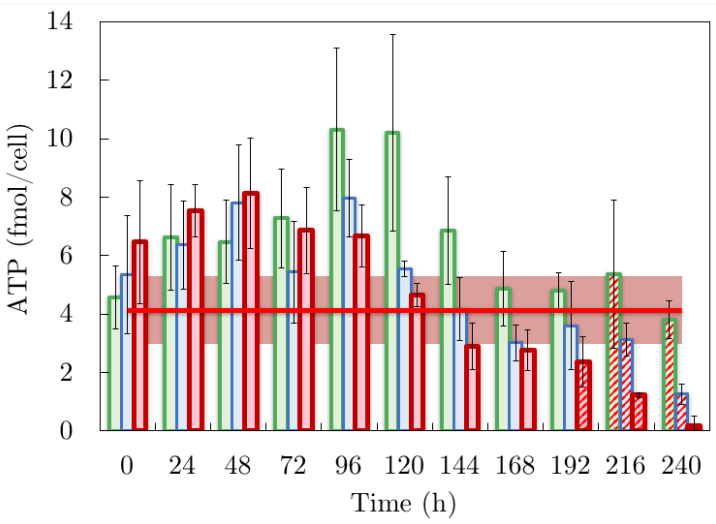


Figure 39: Specific ATP contents obtained over 240h of GS-CHO cultures inoculated at 1, 2 and 4x10^5 cells/mL respectively in green, blue and pink columns. The red line and the shaded region show the critical ATP content of 4.14±1.2 fmol ATP/cell estimated in this work. Columns with diagonal red stripes correspond to samples classified as "Decline phase" in the metabolic profile analysis of Figure 33A. (N=3 biological replicates, error bars show one standard deviation).

5.4.5 MAPPING THE CELL CYCLE

5.4.5.1 Determination of cycling times

From batch cultures of GS-CHO cells, the apparent growth rate during exponential growth was estimated to be 0.0221h^{-1} corresponding to an apparent doubling time of 31h. For this reason, samples were taken during 30h post-EdU exposure to ensure sampling over an entire cell cycle. The cell cycle distributions obtained over 30h of culture for the EdU-positive and EdU-negative populations are shown in Figure 40.

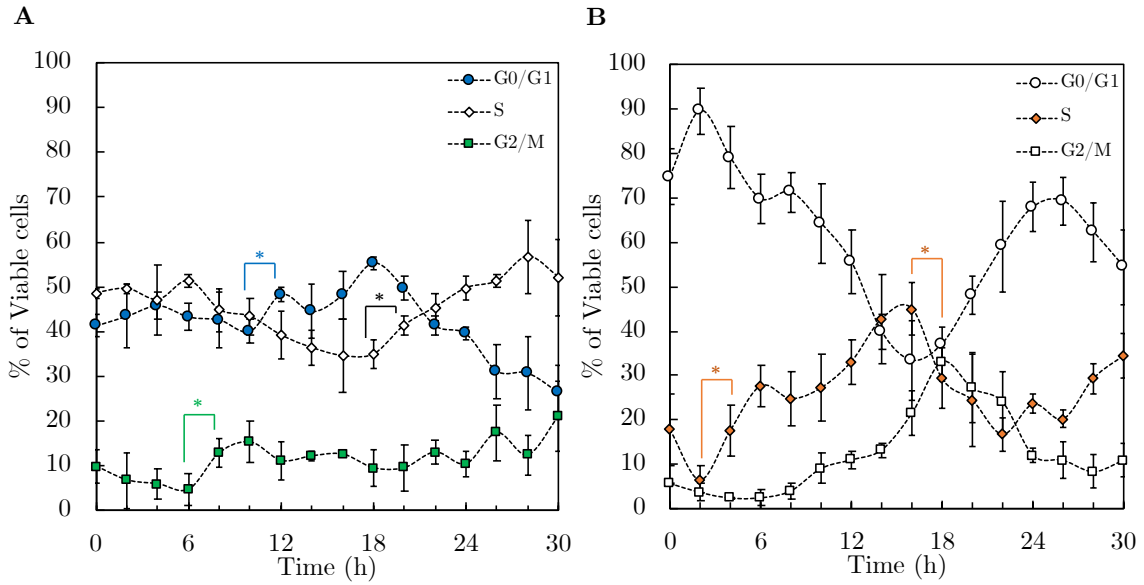


Figure 40: Cell cycle distributions of EdU-positive (A) and EdU-negative (B) populations over 30 hours of GS-CHO cell cultures inoculated at 2x10^5 cells/mL (N=3 biological replicates, error bars show one standard deviation, $^*p\text{-}value<0.05$ one-way ANOVA).

The cell cycle distributions of the EdU-positive population (Figure 40A) show a significant increase ($p\text{-}value<0.05$) in the G_2/M fraction between 6-8h, suggesting cells entered this phase. At 10-12h, the G_0/G_1 population has a significant increase ($p\text{-}value<0.05$), suggesting that cells have transited through phase G_2/M in approximately 4h. The population that has entered phase G_0/G_1 at 12h transits through this phase in 8h, which is supported by a significant decrease ($p\text{-}value<0.05$) at 18-20h. Finally, the S-phase time is estimated from the EdU-negative population (Figure 40B). A significant increase ($p\text{-}value<0.05$) in S-phase is observed between 2 and 4h and cells appear to remain in this phase until 16-18h of culture, where a significant degrease ($p\text{-}value<0.05$) is observed suggesting a total transit time of 14h. In total, the doubling time of GS-CHO cells is estimated to be approximately 26h, corresponding to a maximum growth rate of 0.0267 h^{-1}. This value is in good agreement with the range reported for nine other GS-CHO clones from the same supplier, where specific growth rates between 0.0242 and 0.0338h^{-1} were reported (Kyriakopoulos, Polizzi and Kontoravdi, 2013; Zou and Al-Rubeai, 2016)

5.4.5.2 Dynamic Cyclin Expression

Identification of suitable cyclin biomarkers

The study of the dynamic cyclin expression profiles for GS-CHO cells is important for the identification of suitable cell-cycle biomarkers, which can support the development of mathematical models of the cell cycle for this cell line. To this end, the dynamic expressions of two G_1-cyclins (E1 and D1) and one G_2-cyclin (B1) were measured by flow cytometry in the course of batch cultures.

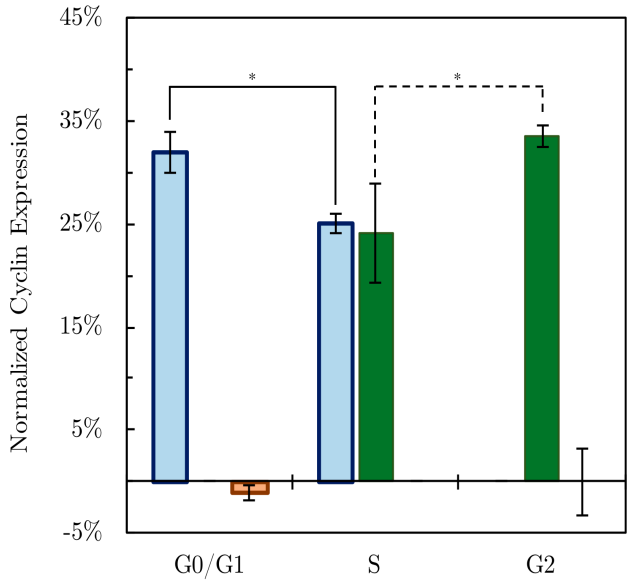


Figure 41: Expression of cyclins E1 (blue, normalized to phase G_2/M), B1 (green, normalized to phase G_1) and D1 (orange, normalized to phase S) in GS-CHO cell cultures inoculated at $2x10^5$ cells/mL in different phases of the cell cycle after 48h of batch culture (N=3 biological replicates, error bars show one standard deviation, $^*p\text{-}value<0.05$ one-way ANOVA).

Analysis of cells collected after 48h of culture (early exponential growth) revealed that the expression of cyclin E1 is significantly higher in phase G_1 when compared to phase S ($p\text{-}value<0.05$) and that cyclin B1 is significantly more expressed in phase G_2 than in phase S ($p\text{-}value<0.05$). A consistent expression of cyclin D1 was found throughout cell cycle with no significant differences encountered between phases ($p\text{-}value>0.05$). These results are shown in Figure 41 and show that the results obtained earlier for GS-NS0 cells (García-Münzer *et al.*,

2013) are also true for GS-CHO cells and allow to identify cyclins E1 and B1 as biomarkers of phase G_1 and phase G_2 of the cell cycle, respectively.

The expression of biomarker cyclins was also analysed in the course of the EdU proliferation assay. Cyclin E1 expression in the EdU-positive population (Figure 42A), initially in phase S, shows decreasing levels between 0 and 6h, consistent with the transition to phase G_2 and an increased expression on the transition to phase G_1 until the transition to S-phase (30%). This threshold expression is confirmed for the EdU-negative population (Figure 42B). Initially in phase G_1, the expression of cyclin E1 peaks at the transition to phase S at 2h and decreases throughout this phase increasing again after cells transit through phase G_2. The expression of cyclin B1 in the EdU-positive population starts during phase S (Figure 42C) reaching a maximum at 6h (15-20%) before the transition to phase G_2, where it is maximal. This threshold is consistent with EdU-positive population data (Figure 42D) where expression is seen from 2h until the transition to phase G_2 at 16h and a peak during phase G2 (30%) before transition to phase G_1. After the transition to phase S, cyclin B1 expression is observed again.

These cyclin profiles are consistent with the observations of Garcia *et al.* for GS-NS0 cells (García-Münzer *et al.*, 2013), showing that cyclin E1 expression is maximal during phase G_1 decreasing gradually throughout phase S while cyclin B1 is expressed during phase S, reaching a maximum during phase G_2. In addition to their work, this report provides more evidence that cyclins E1 and B1 are suitable biomarkers for cell cycle progression in industrially relevant cell lines where cyclin D1 appears to be consistently expressed throughout the cell cycle. Therefore, monitoring these cyclins can support the development of mathematical models of the cell cycle in this cell line.

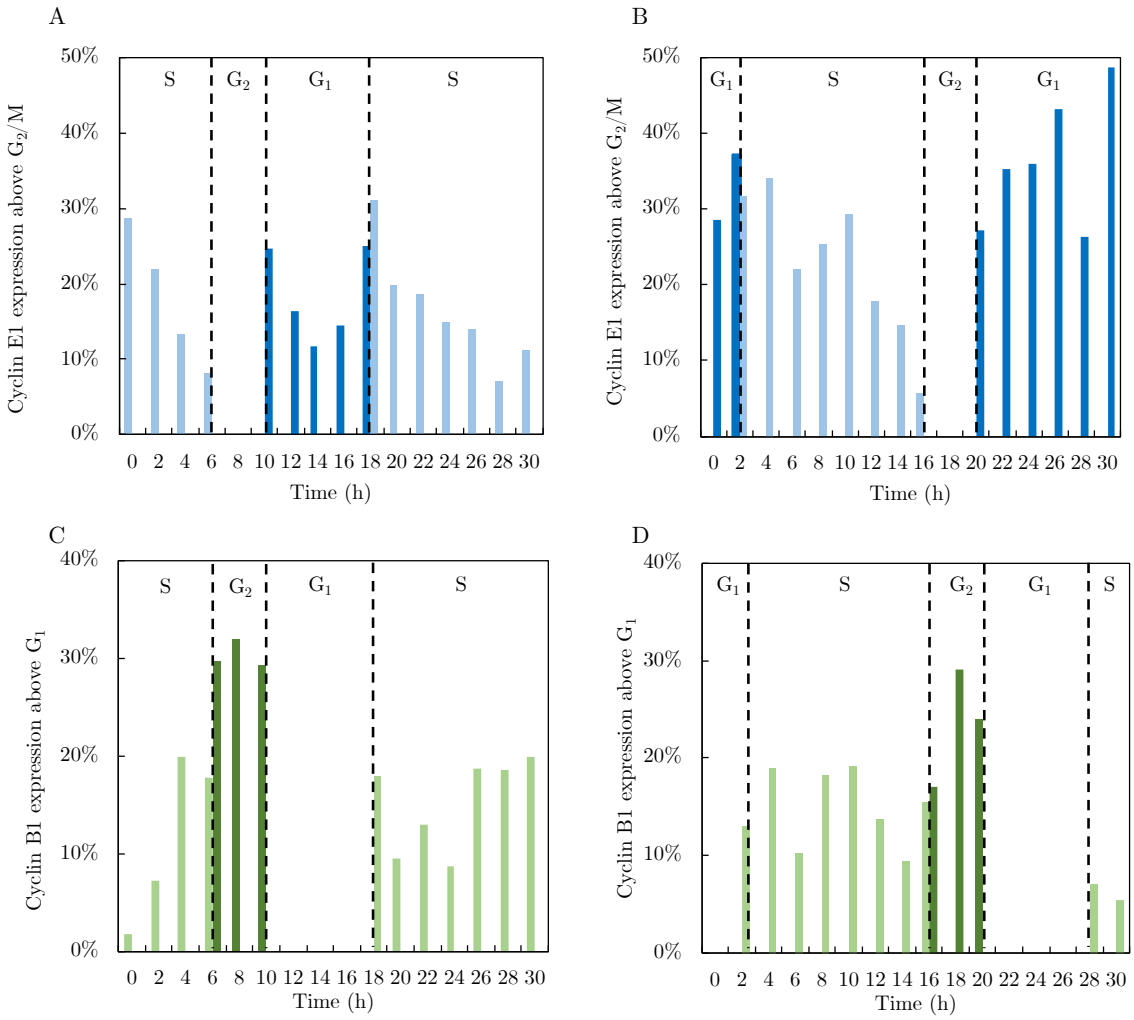


Figure 42: Expression of cyclins E1 (A-B) and B1 (C-D) for EdU positive (A, C) and EdU negative (B, D) populations. Lighter colours show expression in phase S.

Identification of metabolic limitations for cyclin expression

The expression of cyclins E1 and B1 was analysed at 120, 144 and 168h of culture and compared to those observed at 48h. It was observed that expression of both cyclins is significantly lower at 120-168h ($p\text{-}value<0.05$) compared to that observed at 48h but not significantly different among time points beyond 120h (Figure 43).

In order to establish the importance of metabolic stresses on cell cycle, Garcia-Münzer *et al.* compared cyclin expressions before and after glutamate was depleted and observed significant changes (García-Münzer, Kostoglou, *et al.*, 2015). However, in their cultures, glutamate depletion coincides with the lactate metabolic

shift and hence the two effects cannot be decoupled. Nevertheless, the authors do acknowledge the importance of lactate for cyclin expression in the formulation of their mathematical model of GS-NS0 cell cycle (García-Münzer, Kostoglou, *et al.*, 2015).

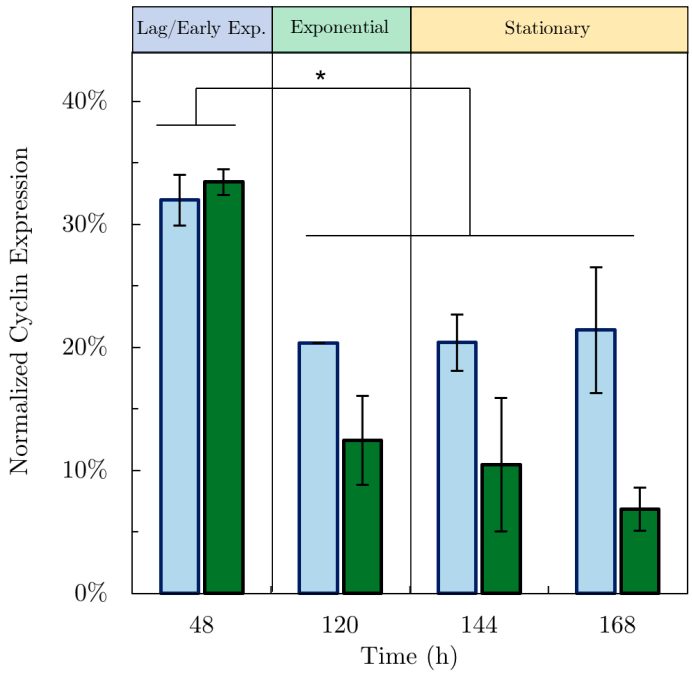


Figure 43: Expressions of cyclins E1 (blue) and B1 (green) at 48 (early exponential growth), 120 (exponential growth), 144 and 168h (stationary phase) of culture. Growth stages refer to in Figure 33A (N=3 biological replicates, error bars show one standard deviation, $^{*}p\text{-}value<0.05$ one-way ANOVA).

The heat map obtained from the simultaneous analysis of all metabolic profiles obtained from GS-CHO cell cultures shows differences in the lactate profiles between early exponential and exponential phases, but no differences between exponential and stationary phases (Figure 33B). On the contrary, the concentrations of asparagine, aspartate and glutamate all show differences between exponential and stationary growth but not between early exponential and exponential phases (Figure 33B). As the lactate and the aspartate metabolic shifts occur almost simultaneously, decoupling both effects is also challenging in GS-CHO cultures. It can be noted that the significant differences observed in cyclin expression between 48 and 120h of culture coincide with the lactate metabolic shift, whereas the concentrations of aspartate at 48h (0.88mM) and at 120h (0.79mM)

are comparable. In addition, data from GS-CHO fed-batch cultures conducted by Kyriakopoulos *et al.* suggest that growth is more closely linked to the lactate metabolic shift than to asparagine concentrations. Indeed, peak viable cell densities observed for different feeding strategies tested by the authors occur shortly after the maximum lactate concentration is reached, irrespective of asparagine concentrations in culture, which were kept between 2-10mM (Kyriakopoulos and Kontoravdi, 2014). Taken together, the analysis of metabolic profiles obtained in this work and previously published fed-batch data corroborate the important role of the lactate metabolic shift for cell growth. Additionally, if, as discussed above (section 5.4.3.2), a NAD^+/NADH imbalance, such as that observed by Templeton *et al.* (Templeton *et al.*, 2013), is confirmed to drive the lactate metabolic shift in GS-CHO cells, such imbalance may also be behind a reduced cyclin expression observed after the lactate metabolic shift. It has been reported that a NAD^+/NADH imbalance induces the expression of p27, which, in turn, inhibits expressions of cyclins A, B, D and E (Kaplon, Dam and Peeper, 2015).

5.4.6 IDENTIFICATION OF NON-PROLIFERATIVE CELLS

The distinction between proliferative and non-proliferative cells was made based on the expression of the Ki-67 protein, only present in proliferating cells (not in phase G_0). The dynamic profile of Ki-67 expression is shown in Figure 44. The data show that, throughout cultures, there is a 5-10 percentage-point difference between viable and Ki-67-positive cells. This suggests that, apart from apoptotic/dead cells, there are further 5-10% of the cells in culture that are not proliferating (Ki-67-negative).

The addition of G_0/G_1 cell cycle arresting agents to CHO cell cultures is often used as a means to increase productivity. Sodium butyrate is a very common choice (Carinhas *et al.,* 2013), but others have also been tested (Park *et al.*, 2015). It is therefore likely that the non-proliferating population identified in this work

contributes significantly to the overall productivity of the culture, although this could not be assessed.

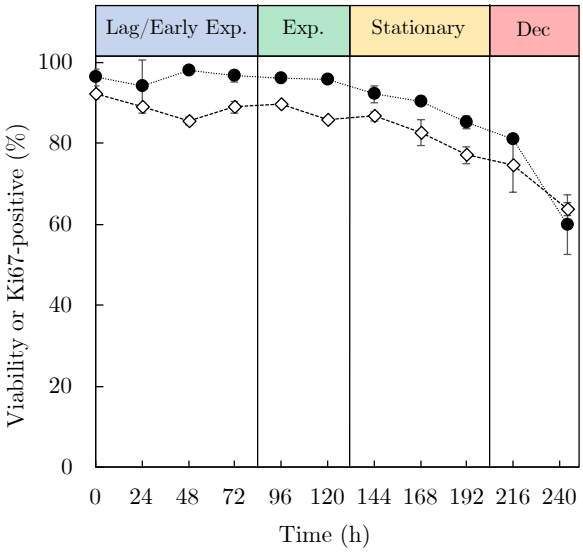


Figure 44: Expression of Ki-67 (open diamonds) and culture viability (closed circles) over 240h of batch cultures of GS-CHO cells inoculated at 2x10^5 cells/mL. Growth stages refer to in Figure 33A (N=3 biological replicates, error bars show one standard deviation).

5.4.7 EXPRESSION OF CASPASE-3, CASPASE-8, BCL-2 AND BAX GENES

In order to characterize cell death in batch cultures of GS-CHO cells, the expression of caspase-3, caspase-8, bcl-2 and bax was analysed at the gene level by measuring the transcripts. Caspase-8 is an initiator caspase which will be active at the early stages of apoptosis (McIlwain, Berger and Mak, 2015) and its inhibition has been shown to improve culture performance in batch and fed-batch cultures of CHO-K1 cells (Yun *et al.*, 2007a). In the later stages of apoptosis, caspase-3, an executioner caspase (McIlwain, Berger and Mak, 2015) will be activated by upstream caspases including caspase-8, leading to cell death (Wong *et al.*, 2006). Bcl-2 encodes for a protein with anti-apoptotic activity and its inhibition has been attempted in various mammalian cell lines as a means to improve culture performance (Lee and Lee, 2003; Templeton *et al.*, 2014). Its overexpression has been shown to direct pyruvate to mitochondrial oxidation, reducing lactate accumulation in culture and leading to higher cell densities (Templeton *et al.*, 2014). Bax is a bcl-2 family member with pro-apoptotic activity whose inhibition

has been shown to contribute to improved performance of CHO cell culture (Cost *et al.*, 2010).

The apoptotic and dead cell densities obtained in the course of 240h of GS-CHO batch cultures, inoculated at $2x10^5$cells/mL, and the corresponding expression of apoptotic-related genes are shown in Figure 45. The average expression of caspase-3 and caspase-8 show an increasing trend throughout the culture, decreasing at the end of the stationary phase (192h) until the end of the culture. The expression of caspase-3 is significantly increased at 120h and 168h compared to 24h ($p>0.05$, ANOVA). The expression of bcl-2 and bax is significantly higher as cells enter the decline phase (216h) compared to lag/early exponential time-points. More experimental data would be necessary in an attempt to reduce the variance of the data so that more meaningful conclusions could be drawn. Towards the later stages of the culture, as previously reported (Weinlich *et al.*, 2016), the expression of bax and bcl-2 increase, as a consequence of their activation by caspase-8. This is accompanied by an increase in both apoptotic and dead cell densities. It is possible that the activation of bax, leading to cytochrome-c release drives apoptosis in the latest stages of culture. A significant increase in apoptotic and dead cell populations is observed between lag/early exponential and exponential phases (0-96h) and late stationary/decline phases (168-216h) and also between 120/144h and 192/216h. The apoptotic cell population is also significantly different at 168h vs. 216h .

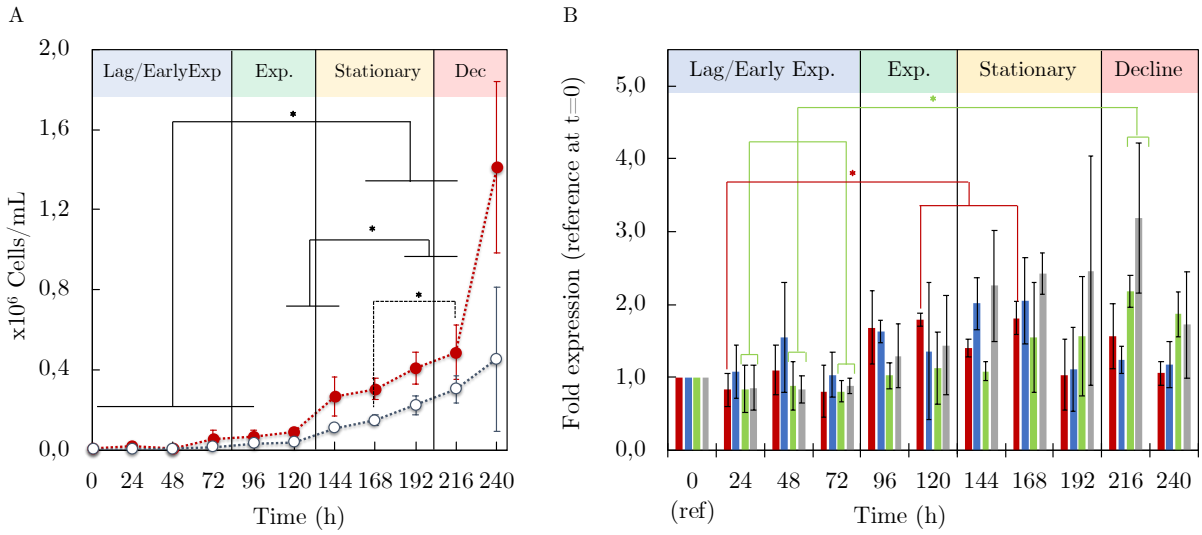


Figure 45: (A) Apoptotic (empty circles) and dead (closed circles) cell densities and (B) expression of caspase-3 (red), caspase-8 (blue), bcl-2 (green) and bax (grey) in the course of 240h of GS-CHO batch cultures inoculated at $2x10^5$ cells/mL. Growth stages refer to Figure 33A. (N=3 biological replicates, error bars show one standard deviation; *p-value<0.05, one-way ANOVA. Dashed line in A depicts a difference in apoptotic population only).

5.5 CONCLUSIONS

In this chapter, growth, metabolism and death of GS-CHO cells were studied. Firstly, hierarchical clustering and PCA were used as unsupervised tools to map growth and productivity of GS-CHO cells, based solely on metabolic profiles obtained in the course of batch cultures. It was possible to identify four different growth phases, based on hierarchical clustering, and to map the obtained clusters using PCA. This technique has allowed to obtain a chronological order of the clusters and to distinguish between productive/non-productive phases of the culture. In addition, the MVA methodology herein presented was shown to provide scientific support for the identification of the time horizons of the different growth stages in culture. This is important for the implementation of MFA studies as well as for the application of kinetic models of cell culture.

Some metabolic changes pertaining to growth regimes transitions appear to be conserved for GS-NS0 and GS-CHO cells. These include the alanine and lactate metabolic shifts and the depletion of amino acids essential for energy metabolism.

After successful application to these two cell lines, the MVA methodology suggested is deemed promising for the analysis of culture data and could potentially be used to improve process understanding.

The metabolic changes identified by MVA were explored in detail for each of the growth stages of the culture. Several important metabolic shifts were identified – lactate, aspartate, glutamate, alanine and ammonia – and possible maps of the active metabolic pathways at each stage of the culture were suggested by combining observations from this work with previous reports in literature. It has been found that the transition from early to late exponential growth appears to be linked to the lactate metabolic shift, whereas the depletion of amino acids essential for energy metabolism (asparagine, aspartate and glutamate) mediates the transition to the stationary phase. The MVA analysis conducted was based on relatively simple analytical techniques which can be employed for bioreactor online or at-line monitoring. Hence, MVA offers the possibility to monitor culture kinetics and productivity instead of simply providing data on individual metabolites.

In this work, the minimum ATP content of a GS-CHO viable cell has been determined to be 4.12 ± 1.2 fmol/cell and it has been confirmed that, when cells can no longer satisfy their energetic requirements, transition into the decline phase occurs. ATP production in GS-CHO cell cultures decreases after the lactate metabolic shift. Therefore, hypotheses of lactate-induced mitochondrial activity impairment do not explain the metabolism of GS-CHO cells. Instead, the hypothesis of TCA cycle carbon being replenished by lactate but being deviated to lipids production would justify a reduction in ATP production. Alternatively, a variation on the NAD^+/NADH ratio could be behind a lactate metabolic shift. Decreasing lactate concentration leads to lower ATP requirements to maintain gradients, and hence to reduced ATP production.

In order to understand the cell cycle of GS-CHO cells, and how it is affected by metabolism, dynamic profiles of cyclins E1, D1 and B1 were obtained. Cyclins E1 and B1 were identified as suitable biomarkers for G_1-S and G_2-G_1 transitions,

respectively. Moreover, combining cyclin expression data with metabolic profile analysis has suggested that the lactate and the aspartate metabolic shifts may be behind a decreased cyclin expression beyond exponential growth. Combining data obtained in this work and previously published studies, the lactate metabolic shift may have a dominant role in the reduced cyclin expression observed beyond the exponential phase. Perturbations on the NAD^+/NADH ratio are known to induce cell cycle arrest by activating p27 and reducing cyclin expression, suggesting a possible common mechanism could be behind the lactate metabolic shift and the cyclin expression reduction observed almost simultaneously.

Using the EdU method, the cell cycle times for this cell line were estimated for the first time, to the author's best knowledge, to be 8h for phase G_0/G_1, 14h for phase S and 4h for phase G_2/M, to a total doubling time of 26h, in agreement with values reported for other GS-CHO clones. Furthermore, dynamic profiles of the expression of Ki-67 over 10 days of cultures revealed the presence of a quiescent population of approximately 5-10% of viable cells.

The findings described in this chapter have allowed to characterize batch cultures of GS-CHO cells and can underpin the development of a predictive mathematical model to support process development. The identification of suitable cell cycle biomarkers (cyclins E1 and B1), and evidence of metabolic changes suppressing their expression on the transition between exponential and stationary phases of culture, are essential to develop a mathematical model of the cell cycle of GS-CHO cells. Exo-metabolome profiling has been successfully used as a tool to predict growth kinetics and proliferative stages, and to understand the metabolic changes associated with transitions between growth phases. Therefore, the use of a dynamic exo-metabolome-based model also emerges as a suitable tool to track metabolism and to potentially optimise cell culture.

Chapter 6

An integrated mathematical model of GS-CHO cell culture

In this chapter the structure of the integrated mathematical model of GS-NS0 cells developed in Chapter 4 is combined with the detailed analysis of GS-CHO cell cultures presented in Chapter 5 to develop an integrated mathematical model of cell-cycle, metabolism and apoptosis of GS-CHO cell culture.

6.1 INTRODUCTION

Chinese Hamster Ovary (CHO) cells are the most widely used host organism for the production of biologics, in particular mAbs (Lalonde and Durocher, 2017). This industry is experiencing an unprecedented growth driven by faster approvals and an increasing number of biosimilars competing with branded mAbs (Grilo and Mantalaris, 2019b).

The wide use of this cell line has motivated a significant number of scientific publications using it as model organism. In 2017 alone, over 250 studies on cell line development and engineering were published (Hong, Lakshmanan, *et al.*, 2018) and, throughout the previous decades, several other works were aimed at understanding different aspects related to the biology and the process performance of this cell line including energy production (Chapman, Webb and Borsa, 1971; Matuszczyk *et al.*,

2015), the cell cycle and its relationship to productivity (Renner *et al.*, 1995; Dutton, Scharer and Moo-Young, 2006; Sitton and Srienc, 2008; Makimoto, Takahashi and Takasugi, 2013; Park *et al.*, 2015; Fuge *et al.*, 2017), the effects of toxic by-products (Lao and Toth, 1997), the effects of different carbon and nitrogen sources (Altamirano *et al.*, 2000; Zhang *et al.*, 2006), metabolic flux analysis (MFA) to understand metabolism differences throughout culture (Ahn and Antoniewicz, 2011; Sengupta, Rose and Morgan, 2011; Selvarasu *et al.*, 2012; Carinhas *et al.*, 2013), minimizing lactate production (Zhou *et al.*, 2011) and understanding the lactate metabolic shift and its relationship to productivity (Le *et al.*, 2012; Li *et al.*, 2012; Martínez *et al.*, 2013; Hong, Nargund, *et al.*, 2018), understanding amino acid transport and metabolism (Kontoravdi, Wong, *et al.*, 2007; Kyriakopoulos, Polizzi and Kontoravdi, 2013; Duarte *et al.*, 2014; Fomina-Yadlin *et al.*, 2014; Hagrot *et al.*, 2017), studying enzymes involved in central carbon metabolism (Zou and Al-Rubeai, 2016), impact of process conditions on the proteome throughout culture (Park *et al.*, 2017; Albrecht *et al.*, 2018).

In spite of its wide use and extensive literature available, there is no integrated mathematical model of CHO cells describing growth, cell cycle, metabolism, mAb production and apoptosis in an integrated manner. Such model would provide an important tool for cell culture optimization. Galbraith *et al.* have described the ideal scenario of cell culture optimization as that where stoichiometric models would be used to understand cell physiology and kinetic models would be used for optimization (Galbraith *et al.*, 2018). The authors also describe this ideal scenario as *"impossible in the near-term"*.

In this work (Chapter 5) a multivariate data analysis methodology was used to analyse samples from GS-CHO batch cell cultures. This analysis has allowed to identify important metabolic shifts related to different growth and productivity stages of culture. Among the advantages of the framework proposed is the use of simple techniques for data acquisition and, as described before using Alanine as an example, its ability to provide dynamic results. However, it cannot be used to

perform dynamic optimization of cell culture. Instead, the knowledge acquired in Chapter 5 can be used to build a kinetic mathematical model of cell culture which could be used for optimization.

6.2 OBJECTIVES

The aim of the present chapter is to provide an integrated and experimentally validated mathematical model of cell-cycle, metabolism and apoptosis in GS-CHO cells and, to achieve it, the following objectives were set:

- Formulate an integrated dynamic mathematical model of GS-CHO cell culture based on the structure derived in Chapter 4 and the information obtained in Chapter 5;
- Use GSA to identify the significant model parameters;
- Use GS-CHO cell culture data to re-estimate the significant model parameters;
- Use independent experimental data-sets to validate the model.

6.3 MODEL DEVELOPMENT

As explained above, the model framework developed in Chapter 4 is to be combined with experimental data and conclusions derived in Chapter 5 for GS-CHO batch culture in order to develop an integrated mathematical model for this cell line.

6.3.1 A CYCLIN AND DNA-BASED CELL CYCLE MODEL

In Chapter 5, cyclins E1 and B1 have been shown to be good biomarkers for tracking, respectively, phase G_1 and phase G_2/M of the cell cycle in GS-CHO cells. Therefore, the same equations used in chapter 4 for GS-NS0 cells can be used to model the cell cycle in GS-CHO cells. These equations are repeated below.

$$\frac{\partial N_{G1}(cycE,t)}{\partial t}+\frac{\partial}{\partial cycE,}[r_{G1}(cycE,S)N_{G1}(cycE,t)]=-\Gamma_{G1}(cycE,S)N_{G1}(cycE,t)$$

$$-\ [\mu_A(g)+\mu_D(E)]N_{G1}(cycE,t)\ +2\delta(cycE)\int_{cycB_{,min}}^{cycB_{,max}}\Gamma_{G2}(cycB,S)N_{G2}(cycB,t)\,dcycB$$

$$\frac{\partial N_S(DNA,t)}{\partial t}+\frac{\partial}{\partial DNA,}[r_S(DNA,S)N_S(DNA,t)]=-\ [\mu_A(g)+\mu_D(E)]N_S(DNA,t)\ +$$

$$+\ \delta(DNA-1)\int_{cycE_{,min}}^{cycE_{,max}}\Gamma_{G1}(cycE,S)N_{G1}(cycE,t)\,dcycE$$

$$\frac{\partial(cycB,t)}{\partial t}+\frac{\partial}{\partial cycB,}[r_{G2}(cycB,S)N_{G2}(cycB,t)]=-\ \Gamma_{G2}(cycB,S)N_{G2}(cycB,t)$$

$$-\ [\mu_A(g)+\mu_D(E)]N_{G2}(cycB,t)\ +\ \delta(cycB-cycB_S)r_S(DNA=2,S)N_S(DNA=2,t)$$

Additionally, the times spent by this cell line in each phase of the cycle in exponential growth have also been estimated, providing important information for model development. Finally, as was also discussed in Chapter 5, cyclin expression is reduced after the lactate metabolic shift which coincides with the depletion of the main amino acids fuelling the TCA cycle in this cell line, namely, aspartate (initially produced from asparagine) and glutamate. Therefore, both effects are considered as affecting cell cycle rates and, ultimately, the growth rate. This has been accounted for by replacing equations used in Chapter 4 and using equations (46) to (50).

$$r_s=\frac{1}{t_{S,min}}\cdot f_{limAspGlu}\cdot f_{limLac}\cdot f_{limAmm} \tag{46}$$

$$r_{i=G1,G2}=\frac{1}{\tau_{k=t_{G1,min},t_{G2,min}}}\cdot f_{limAspGlu}\cdot f_{limLac}\cdot f_{limAmm} \tag{47}$$

$$f_{lim}AspGlu = \left(\frac{[Asp]}{[Asp]+K_{Asp}} + \frac{[Glu]}{[Glu]+K_{Glu}}\right) \tag{48}$$

$$f_{limLac}\left(\frac{K_{i,Lac}}{[Lac]+K_{i,Lac}}\right) \tag{49}$$

$$f_{limAmm}\left(\frac{K_{Amm}}{[Amm]+K_{Amm}}\right) \tag{50}$$

The rates of transition between phases (Γ_i) have been calculated using the same equations used in Chapter 4, which are repeated below.

$$\Gamma_{i=G_1,G_2} = \begin{cases} 0 & below\ cyclin\ threshold \\ p_i.\, f_{lim}Glc & above\ \ cyclin\ threshold \end{cases}$$

$$f_{lim}Glc = \left(\frac{[Glc]}{[Glc]+K_{Glc}}\right)$$

The times spent in each phase are calculated from equation (17) and used in equation (18) to compute the growth rate, as in Chapter 4.

$$t_{i=G1,G2} = \frac{cyclin\ threshold}{r_{i=G1,G2}} \quad ; \quad t_S = \frac{1}{r_S}$$

$$\mu = \frac{ln(2)}{t_{G_0/G_1} + t_S + t_{G_2/M}}$$

The equations used to calculate the total number of viable cells, the viable cell density and the apoptotic and dead cell densities mass balances have also been kept and are shown below.

$$N_{i=G1,S,G2} = \int_{x_{k,min}}^{x_{k,max}} N_i(x_k)dx_k$$

$$X_V = \frac{N_{G_1} + N_S + N_{G_2}}{V}$$

$$\frac{d(X_A V)}{dt} = \mu_A X_V V - \mu_{AD} X_A V$$

$$\frac{d(X_D V)}{dt} = \mu_{AD} X_A V + \mu_D X_V V - k_{lys} X_D V$$

As was done in Chapter 4, the rate of early apoptosis was linked to the expression of caspase-8 but, in the case of GS-CHO cells, the expression of the *bax* gene, which increases towards the end of the culture, was also included (equation (51)). The rates of late apoptosis and death through other pathways were calculated using the same equations used in Chapter 4.

$$\mu_A = \mu_{A,max}\left(\frac{mRNA_{c8}}{mRNA_{c8} + K_{c8}} + \frac{mRNA_{bax}}{mRNA_{bax} + K_{bax}}\right)$$

$$\mu_{AD} = \mu_{AD,max}\frac{mRNA_{c3}}{mRNA_{c3} + K_{c3}} \tag{51}$$

$$\mu_D = \mu_{D,max}\frac{[ATP]}{[ATP] + ATP_{crit}}$$

6.3.2 A KINETIC METABOLIC MODEL OF GS-CHO CELLS

As reported in several works, and discussed in Chapter 5, the metabolism of GS-CHO cells is significantly different from of GS-NS0 cells. Therefore, a kinetic model able to describe the metabolism of GS-CHO cells in culture had to be developed *de novo.*

As before, glucose and amino acids are assumed to be used for growth, energy generation and mAb production. The amino acid (AA) composition of CHO cells experimentally obtained by Servarasu *et al.*, was used to calculate the growth-related requirements (Selvarasu *et al.*, 2012) for any nutrient *"j"* as shown in equation (31) (Chapter 4), repeated below.

$$Q_{x,j} = \frac{\mu}{Y_{x/j}}$$

For the purpose of calculating the biomass requirements, $Y_{x/j}$ (cell/mmol) is the reciprocal of the compositions obtained by Servarasu *et al.* The composition of the antibody produced by this particular GS-CHO cell line provided (by the manufacturer) in mmol of nutrient *j* per milligram of mAb ($M_{mAb,j}$) was used to estimate the AA requirements to produce mAb as a function of the mAb production rate ($M_{mAb,x}$), as expressed in equation (32), Chapter 4, repeated below..

$$\mathrm{Q_{mAb,j}} = \mathrm{M_{mAb,x}} \cdot \mathrm{M_{mAb,j}}$$

The energy yield of each amino acid was, as suggested before, estimated based on the NADH and $FADH_2$ co-factors produced (details are provided in Appendix III). The rates at which each amino acid was consumed ($Q_{tca,AAj}$, equation (33), repeated below) were estimated as a fraction (γ_j) of an overall value ($Q_{tca,AA}$, equation (34), repeated below) in order to reduce the number of parameters in the model, as done for GS-NS0 cells (Chapter 4).

$$\mathrm{Q_{tca,AAj}} = \gamma_j \cdot \mathrm{Q_{tca,AA}}$$

$$\mathrm{Q_{tca,AA}} = \frac{\mu}{\mathrm{Y_{tca/AA}}} + \mathrm{M_{tca/AA}}$$

Two significant differences between the metabolism of GS-CHO and GS-NS0 cells are the production of lactate and ammonia in appreciable quantities, and the fact that both by-products show metabolic shifts in culture, as described in Chapter 5.

6.3.2.1 Glucose, Lactate and Alanine

The glucose mass balance is shown in equation (52). The $Q_{x,glc}$, $Q_{tca,glc}$ and $Q_{mAb,glc}$ terms describe the use of glucose for biomass production, energy generation

in the TCA cycle and mAb production, respectively, and are calculated as described (Chapter 4) in equations (25)to (27).

$$\frac{d(V.[Glc])}{dt} = -\ X_V.V.(Q_{x,Glc} + \ Q_{mAb,Glc} + \ Q_{tca,Glc} + Q_{lac,Glc} + Q_{ala,Glc} \ + \ Q_{ile,Glc}) \quad (52)$$

As mentioned above, and discussed in Chapter 5, lactate production in GS-CHO is significant. As lactate is a by-product of glucose metabolism, its production rate is accounted for in the glucose mass balance. As was also discussed in Chapter 5, a lactate metabolic shift occurs in culture and, as observed in previous studies with GS-CHO cells, this shift occurs always when a given lactate concentration is reached in culture (Carinhas *et al.*, 2013; Kyriakopoulos and Kontoravdi, 2014). After this point, lactate is consumed at a constant rate. Therefore, the lactate mass balance is given by

$$\frac{d(V.[Lac])}{dt} = \ X_V.V.Q_{Lac} \quad (53)$$

where the lactate metabolism rate (Q_{Lac}) is calculated as shown in equation (54). Although the actual drive of the lactate metabolic shift remains unclear, a lactate threshold-driven shift is assumed as this is a feature of all cultures conducted in this work and identified in literature for GS-CHO cell lines (Carinhas *et al.*, 2013; Kyriakopoulos and Kontoravdi, 2014). The threshold has been determined experimentally to be approximately 20mM (Chapter 5).

$$Q_{Lac} = \begin{cases} 2Q_{Lac,Glc} = 2.\left(\frac{\mu}{Y_{Lac/Glc}} + M_{Lac/Glc}\right) & [Lac] \ < \ threshold \\ -\ Q_{cons,Lac} & [Lac] \ \geq \ threshold \end{cases} \quad (54)$$

As shown in Chapter 5, alanine is produced throughout GS-CHO culture, until the decline, phase and irrespective of growth and therefore, in the kinetic model of GS-CHO cells in culture, alanine production was considered to be independent of growth. However, alanine production occurs until aspartate is exhausted in culture. From this point onwards, it is consumed to produce pyruvate as shown in the work of Li *et al.* who showed that feeding pyruvate to CHO cultures delays the lactate switch and prevents the alanine metabolic shift hence leading to lower ammonia levels in culture, while not affecting viability nor titre (Li *et al.*, 2012). Additionally, starting CHO-K1 cultures without alanine reduces the glycine production suggesting that serine carbon, used to produce this amino acid under control conditions, could be directed to pyruvate production instead (Hagrot *et al.*, 2017) showing the importance of alanine in replenishing pyruvate levels. In addition, the same study shows that starting cultures with twice as much alanine as that present in control cultures has shown no reduced ammonia production as it led to unchanged alanine production rate. The authors have also shown that in the absence of asparagine, CHO-K1 cells produce more alanine, more glutamate and less aspartate. Taken together, these results suggest that aspartate and glutamate feed the TCA cycle and alanine production has an important role in the metabolism of these amino acids, possibly for transaminase reactions as shown in Figure 46.

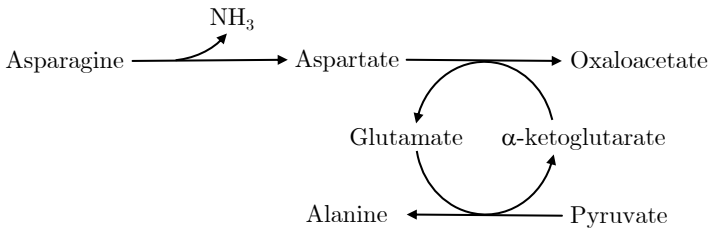


Figure 46: Asparagine, aspartate, glutamate, and alanine metabolic interplay.

Therefore, alanine production is assumed to occur after aspartate is exhausted at a rate $Q_{Ala,cons}$. The alanine mass balance is given by equation (55) where the alanine usage for biomass and mAb production are calculated as explained above and the

alanine production rate is given by equation (56) and the alanine consumption rate is calculated from equation (57).

$$\frac{d([Ala].V)}{dt} = X_V.V.(Q_{Ala,Glc} - Q_{x,Ala} - Q_{mAb,Ala} - Q_{tca,Ala} - Q_{Ala,cons}) \quad (55)$$

$$Q_{Ala,Glc} = \begin{cases} 2.M_{Ala/Glc} & [Asp] > 0 \\ 0 & [Asp] = 0 \end{cases} \quad (56)$$

$$Q_{Ala,cons} = \begin{cases} 0 & [Asp] > 0 \\ Q_{Ala,cons} & [Asp] = 0 \end{cases} \quad (57)$$

In GS-CHO cultures, production of isoleucine has been observed until the end of exponential growth (Chapter 5). During exponential growth, glutamate and aspartate are needed for energy production in the TCA cycle, and isoleucine production stops. It has been assumed that carbon deviated from pyruvate (Nelson and Cox, 2013) or from the TCA cycle (Xie and Wang, 1996; Nelson and Cox, 2013) is used to produce this amino acid and, hence, it is attributed to a glucose consumption at a rate $Q_{ile,glc}$ calculated from equation (58).

$$Q_{Ile,Glc} = \begin{cases} \frac{\mu}{Y_{Ile/Glc}} & [Glu] > 1mM \\ 0 & [Glu] \leq 1mM \end{cases} \quad (58)$$

6.3.2.2 Aspartate and Asparagine

From the beginning of the culture, asparagine is the fastest consumed amino acid. As identified by MFA of GS-CHO cultures (Carinhas *et al.*, 2013), asparagine is degraded to aspartate producing ammonia (Kyriakopoulos and Kontoravdi, 2014) and aspartate yields its amino group to α-ketoglutarate producing oxaloacetate (for the TCA cycle) and glutamate. The α-ketoglutarate/glutamate loop can be regenerated by pyruvate to alanine conversion (Figure 46). These results are also confirmed by the work of Hagrot *et al.* showing that cultures with twice the starting asparagine concentration produce significantly more (approximately twice as much)

aspartate and consume significantly more asparagine compared to the control cultures (Hagrot *et al.*, 2017).

The mass balances to asparagine and aspartate are calculated from equations (59) and (60).

$$\frac{d(V.[Asn])}{dt} = - X_V.V.(Q_{x,Asn} + Q_{mAb,Asn} + Q_{Asp,Asn}) \tag{59}$$

$$\frac{d(V.[Asp])}{dt} = X_V.V.(Q_{Asp,Asn} - Q_{x,Asp} - Q_{mAb,Asp} - Q_{tca,Asp}) \tag{60}$$

Aspartate is produced from asparagine at a rate $Q_{Asp,Asn}$ calculated as shown in equation (61).

$$Q_{Asp,Asn} = \frac{\mu}{Y_{Asp/Asn}} + M_{Asp/Asn} \tag{61}$$

Both amino acids are required for biomass – $Q_{x,Asn}$ and $Q_{x,Asp}$ (equation (31)) – mAb – $Q_{mAb,Asn}$ and $Q_{mAb,Asp}$ (equation(32)) – and energy production $Q_{tca,Asn}$ and $Q_{tca,Asp}$ (equation (33)).

6.3.2.3 Glutamate and Glutamine

The metabolism of glutamate and glutamine is particularly important for organisms engineered with the GS system. Extracellular glutamine concentration is minimal, if existent, in GS-NS0 cell cultures, where glutamate has been considered to be the main energy yielding amino-acid (García-Münzer, Kostoglou, *et al.*, 2015) and the first amino acid to be consumed (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). In GS-CHO cells, asparagine has been observed to be the first amino acid to be consumed (Carinhas *et al.*, 2013). As shown in Chapter 5, during exponential phase, after asparagine concentrations have significantly dropped, aspartate and glutamate are consumed, fuelling energy production.

As discussed in Chapter 5, glutamate accumulates in culture during the lag/early-exponential phase and so does glutamine. As cells enter exponential growth, glutamate is immediately consumed in the TCA cycle for energy production, but glutamine is not re-uptaken until cells enter stationary growth i.e. after glutamate and aspartate are exhausted. The mass balances to glutamate and glutamine are shown in equations (62) and (63).

$$\frac{d(V.[Glu])}{dt} = X_V.V.(Q_{prod,Glu} - Q_{x,Glu} - Q_{mAb,Glu} - Q_{tca,Glu} - Q_{Gln,Glu}) \tag{62}$$

$$\frac{d(V.[Gln])}{dt} = X_V.V.(Q_{Gln,Glu} - Q_{x,Gln} - Q_{mAb,Gln} - Q_{tca,Gln}) \tag{63}$$

As described for other nutrients, biomass requirements on glutamate and glutamine are accounted for by terms $Q_{x,Glu}$ and $Q_{x,Gln}$ (equation (31)); mAb requirements are accounted for by $Q_{mAb,Glu}$ and $Q_{mAb,Gln}$ (equation (32)) and energy production requirements are accounted for by $Q_{tca,Glu}$ and $Q_{tca,Gln}$ (equation (33)).

As described above, and discussed in Chapter 6, glutamate accumulates in culture as a result of the transamination of aspartate to oxaloacetate (Figure 46) during lag/early-exponential growth while, during exponential growth, glutamate no longer accumulates. Therefore, the glutamate production rate is calculated as shown in equation (64) and glutamine is produced throughout the culture growth stages at a rate from equation (65).

$$Q_{prod,Glu} = \begin{cases} Q_{tca,Asp} & [Asn] > 1mM \\ 0 & [Asn] \leq 1mM \end{cases} \tag{64}$$

$$Q_{Gln,Glu} = \frac{\mu}{Y_{Gln/Glu}} \tag{65}$$

The use of glutamine in the TCA cycle occurs only during the stationary phase, being therefore calculated as shown in equation (66).

$$Q_{tca,Gln} = \begin{cases} 0 & [Asp] > 0 \\ Q_{tca,Gln} & [Asp] = 0 \end{cases} \quad (66)$$

6.3.2.4 Ammonia

Ammonia is a by-product of amino acid metabolism, particularly of glutamine and asparagine degradation. As mentioned above, ammonia is produced from the beginning of the culture as a by-product of asparagine degradation to aspartate (Figure 46) and can be incorporated to produce glutamine from glutamate. Glutamine usage in the TCA cycle produces two molecules of ammonia per molecule of glutamine consumed, as this amino acid is first converted to glutamate and then to α-ketoglutarate. Until the end of exponential growth, these are assumed to be the main ammonia producing reactions.

During stationary growth, the first ammonia metabolic shift occurs. Glutamate concentrations reduce significantly but alanine production from pyruvate continues, requiring cells to uptake ammonia from the medium to produce alanine. Alanine production only stops at the end of the stationary phase. The consequence of these reactions is that an overall consumption of ammonia is observed in culture during stationary growth.

The beginning of the decline phase is characterized by the alanine metabolic shift. Alanine is consumed producing pyruvate, which can supply several cellular activities and this reaction produces ammonia. This leads to the second ammonia metabolic shift.

The ammonia mass balance is given by equation (67) where Q_{amm} is the ammonia metabolism rate calculated as shown in equation (68).

$$\frac{d(V.[Amm])}{dt} = X_V.V.Q_{amm} \quad (67)$$

$$Q_{amm} = \begin{cases} 2Q_{tca,Gln} + Q_{Asp,Asn} - Q_{Glu,Gln} & [Asn] \geq 1mM \\ 2Q_{tca,Gln} + Q_{tca,Glu} - 2Q_{Ala,Glc} & 0 < [Asn] < 1mM \\ 2Q_{tca,Gln} + Q_{Tca,Ala} + Q_{ala} & [Asn] = 0mM \end{cases} \quad (68)$$

No thermal degradation of glutamine to ammonia was considered in this model as GS-CHO cells are cultured in glutamine-free medium for 10 days and glutamine half-lives in several serum-free media have been reported to be between 6 and 9 days for the fastest cases reaching over 17 days for the longest half-lives. Glutamine decay rates were found to be dependent on media composition and pH being slower at lower pH (Ozturk and Palsson, 1990; Murphy and Young, 2013). In this work, glutamine is produced during the first 3 days until its concentration reaches a plateau and it is nearly all consumed within 2 days. Therefore, although some thermal decomposition may occur, it is neglected in face of metabolic rates.

6.3.2.5 Serine, Glycine and Threonine

As discussed in Chapter 5, serine is consumed from the beginning of the culture to produce glycine which, during the growth stages of culture (lag/early exponential and exponential phases identified in Chapter 5), can be used to produce threonine. The work of Hagrot *et al.* shows that in the absence of serine, cells will consume glycine (instead of producing it, as under control conditions) and serine will be produced. If the initial threonine concentration is twice that used in control cultures, glycine production is significantly lower. These results suggest that under normal metabolism, serine is used to produce glycine which is, in turn, used to produce threonine. The mass balances to these amino acids are given by equations (69) to (71).

$$\frac{d(V.[Ser])}{dt} = -X_V.V.(Q_{x,Ser} + Q_{mAb,Ser} + Q_{Gly,Ser}) \tag{69}$$

$$\frac{d(V.[Gly])}{dt} = X_V.V.(Q_{Gly,Ser} - Q_{x,Gly} - Q_{mAb,Gly} - Q_{Thr,Gly}) \tag{70}$$

$$\frac{d(V.[Thr])}{dt} = X_V.V.(Q_{Thr,Gly} - Q_{x,Thr} - Q_{mAb,Thr} - Q_{tca,Thr}) \tag{71}$$

Glycine production occurs throughout the culture at a rate $Q_{Gly,Ser}$ calculated from equation (72) and threonine production occurs during the lag/early-exponential phase of culture at a rate $Q_{Thr,Gly}$ (equation (73)).

$$Q_{Gly,Ser} = \frac{\mu}{Y_{Gly/Ser}} \tag{72}$$

$$Q_{Thr,Gly} = \begin{cases} \frac{\mu}{Y_{Thr/Gly}} & [Glu] \geq 1mM \\ 0 & [Glu] < 1mM \end{cases} \tag{73}$$

6.3.2.6 Other amino acids

All essential and non-essential amino acids are tracked by the present metabolic model using the same approach presented for GS-NS0 cells. The generic mass balance to amino acids is given by equation (74) where the rates terms are calculated using equations (31)to (33).

$$\frac{d(V.[AA_j])}{dt} = - X_V.V.(Q_{x,AAj} + Q_{mAb,AAj} + Q_{tca,AAj}) \tag{74}$$

In the case of isoleucine, as discussed above, there is an additional production term ($Q_{Ile,Glc}$, equation (58)).

In the case of cysteine and tryptophan, the rate of usage for energy production, $Q_{tca,AA,}$ is zero throughout culture as these amino acids are not used to produce TCA cycle intermediates (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

6.3.2.7 ATP

Energy generated from all nutrients is calculated based on the cofactors that their metabolism generates through the TCA cycle (Appendix III), the most important energy production pathway for GS-CHO cells (Zou and Al-Rubeai, 2016). As was done for GS-NS0 cells, the ATP mass balance is calculated as expressed in equation (38), repeated below.

$$\frac{d(V.[ATP])}{dt} = C_{glc}ATP_{glc} + \sum_j C_j ATP_j - \mathrm{X_V}.V.\left(\frac{\mu}{Y_{x/ATP}} + M_{x/ATP} + M_{mAb,ATP}.MmAbx\right)$$

In the ATP mass balance the "C" terms account for the total usage of a given nutrient in energy production (equation (35), repeated below) and the terms "ATP_j" represent the number of molecules of ATP produced per molecule of nutrient "j" consumed in the TCA cycle (Appendix IV).

$$C_j = \mathrm{X_V}.V.Q_{tca,AAj}$$

6.3.2.8 Monoclonal Antibody

The monoclonal antibody mass balance equation is exactly the same used in Chapter 4 (equation (39), repeated below) and all terms therein are also exactly the same described in Chapter 4 (equations (40)and (41)).

$$\frac{d(V.[mAb])}{dt} = \mathrm{X_V}.V.mAb_{release} + (\mu_{ad} + \mu_d).xD.V(mAb_{acc} - mAb_{deg})$$

6.3.3 APOPTOSIS

The same four genes used to track apoptosis in GS-NS0 cells are considered in this model of GS-CHO cells. Caspases 3 and 8 were identified as important apoptosis mediators in CHO cells and that caspase-3 positive cells are also caspase-8- and caspase-9- positive (Yun *et al.*, 2007b; Meshram *et al.*, 2012). Therefore, caspase-8 was used to calculate the rate of early apoptosis and caspase-3 was used to calculate the rate of late apoptosis. For GS-CHO cells, ammonia has been added as a caspase activator, based on prior studies showing that 8mM ammonia reduced growth by 50% and increases caspase expression (Kurano *et al.*, 1990; Lao and Toth, 1997; Ha and Lee, 2015). The mass balances to caspase-8 and caspase-3 are shown in equations (75) and (76), respectively, where the β terms are the basal gene expression rates, which can be increased by nutritional stresses, and

the "α" terms are the transcript degradation rates. The anti-apoptotic gene *bcl-2* has been considered an inhibitor of both caspase-8 and caspase-3, considering first order kinetics with rates $k_{bcl2,c8}$ and $k_{bcl2,c3}$, respectively.

$$\frac{dmRNA_{c8}}{dt} = \beta_{c8} \sum_{w=Asp,Glu,Amm} \frac{km_{w,c8}}{[w] + km_{w,c8}} - k_{bcl2,c8}.mRNA_{bcl2} - \alpha_{c8} mRNA_{c8} \tag{75}$$

$$\frac{dmRNA_{c3}}{dt} = \beta_{c3} \left(\frac{km_{Asp,c3}}{[Asp] + km_{Asp,c3}} + \frac{[Amm]}{[Amm] + km_{AMM,c3}} \right) + \\ + \sum_{g=c3,c8} k_{c3,g} mRNA_g - k_{bcl2,c3} mRNA_{bcl2} - \alpha_{c3}\mathrm{mRNA}_{c3} \tag{76}$$

The expression of anti-apoptosis gene *bcl-2* in CHO cells has shown to prolong culture longevity (Grilo and Mantalaris, 2019a) and the pro-apoptotic gene *bax* has also been considered important for the intrinsic pathway of apoptosis in CHO cells (Wong *et al.*, 2006). The mass balances to these two genes are shown in equations (44) (repeated below from Chapter 4) and (77), respectively.

$$\frac{dmRNA_{bcl2}}{dt} = \beta_{bcl2} \left[\sum_{w=Glc,Arg} \frac{km_{w,bcl2}}{[w] + km_{w,bcl2}} + \left(\frac{mAb}{mAb + km_{mAB,bcl2}} \right)^2 \right] - \alpha_{bcl2}.mRNA_{bcl2}$$

$$\frac{d\ mRNA_{bax}}{dt} = \beta_{bax} + k_{bax,c8} mRNA_{c8} - k_{bcl2} mRNA_{bcl-2} - \alpha_{bax}.mRNA_{bax} \tag{77}$$

6.4 MODEL ANALYSIS

The flow of information in the present model is the same illustrated in Chapter 4 (Figure 14). With the formulation described above, the model has 78 parameters and, just as identified in Chapter 4, a conservation study is necessary to ensure no loss of entities (cells) in the cell cycle model which remains probabilistic. The set-up for the conservation study was the same used in chapter 5 and the criterion (99% conservation at peak density) was also the same. Figure

47 shows the results of this study where it can be seen that the number of bins in the cyclin domains was varied from 50 to 300 and 100 bins in each cyclin domain is enough to satisfy the conservation criterion. With this discretization strategy, the model has 845 equations comprising 558 algebraic equations and 257 differential equations and the computational time to solve it is 2 seconds.

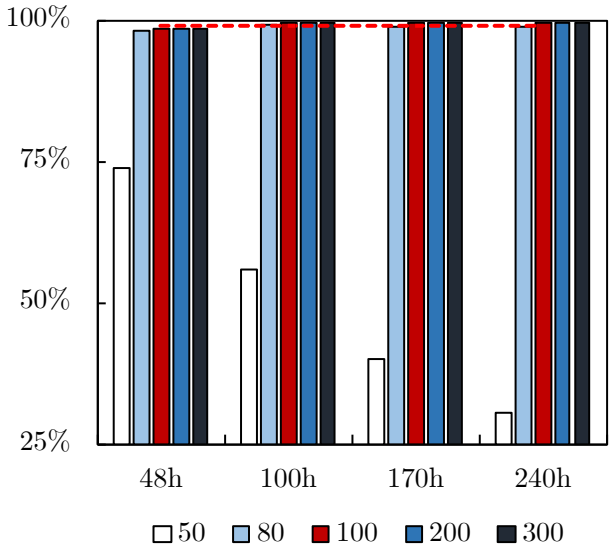


Figure 47: Conservation study results. In both cyclin domains the number of bins was varied from 50 to 300 while keeping 20 bins in the DNA domain. The red dashed line shows 99% conservation.

6.5 GLOBAL SENSITIVITY ANALYSIS

Global sensitivity analysis has been used to identify the significant model parameters. As was done in chapter 4, the Sobol method was used to calculate the sensitivity indices (SIs) for all model parameters setting a discrimination criterion for the important ones at 10%. All model parameters were varied 50% from their nominal values and 5000 model simulations were used to compute the SIs. The SIs of the 78 model parameters were calculated at 60h (exponential growth), 160h (stationary growth) and 200h (decline phase) for 35 model outputs – viable, apoptotic and dead cells, glucose, lactate, ammonia, ATP, mAb, three cell cycle fractions, 20 amino acids and 4 genes. The full GSA results for this model are presented in Appendix III. Figure 48 shows the GSA results for viable cells, fraction of cells in phase G_1, apoptotic cells, ATP, mAb, caspase-8 transcripts and aspartate.

As explained before, global sensitivity analysis is an important tool to assess model structure based on the biological significance of the information obtained. From Figure 48, it can be understood that the duration of phase S is the bottleneck of GS-CHO cell growth. This has been experimentally observed, as phase S is where these cells spend most of their replicating time. GSA results show that viable cell density (xV) predictions are most significantly affected by the duration of phase S ($timeS_{min}$) and the critical cyclin B threshold for transition to phase G_2/M (critB). Finally, for the later stages of the culture, it is the availability of aspartate that most greatly affects viable cell density, as shown by the importance of parameter "CoeffAsp", which describes aspartate consumption rate in the TCA cycle. This result is supported by observations in Chapter 5, and in previous literature reports, regarding the importance of asparagine/aspartate metabolism for this cell line (Carinhas et al., 2013; Kyriakopoulos, Polizzi and Kontoravdi, 2013).

With regards to energy production, it is noteworthy that GSA predicts glucose to be relevant for ATP generation throughout culture while amino acids are only predicted to be relevant during the growth stages of culture. Production of mAb is predicted to be very much related to viable cells throughout culture as some parameters affecting this output also govern cell growth. Additionally, during growth stages the mAb release rate is most important while, during stationary and decline phases, death rates gain an increasing significance. This is likely because the model formulation accounts for the mAb released to the medium as a consequence of loss of membrane integrity during cell death. Finally, aspartate production is dominated by its production from asparagine and its use in the TCA cycle, which is consistent with experimental findings in this work and previous literature, as discussed in Chapter 5.

In total, only 37 out of the 78 model parameters were found significant (Appendix III).

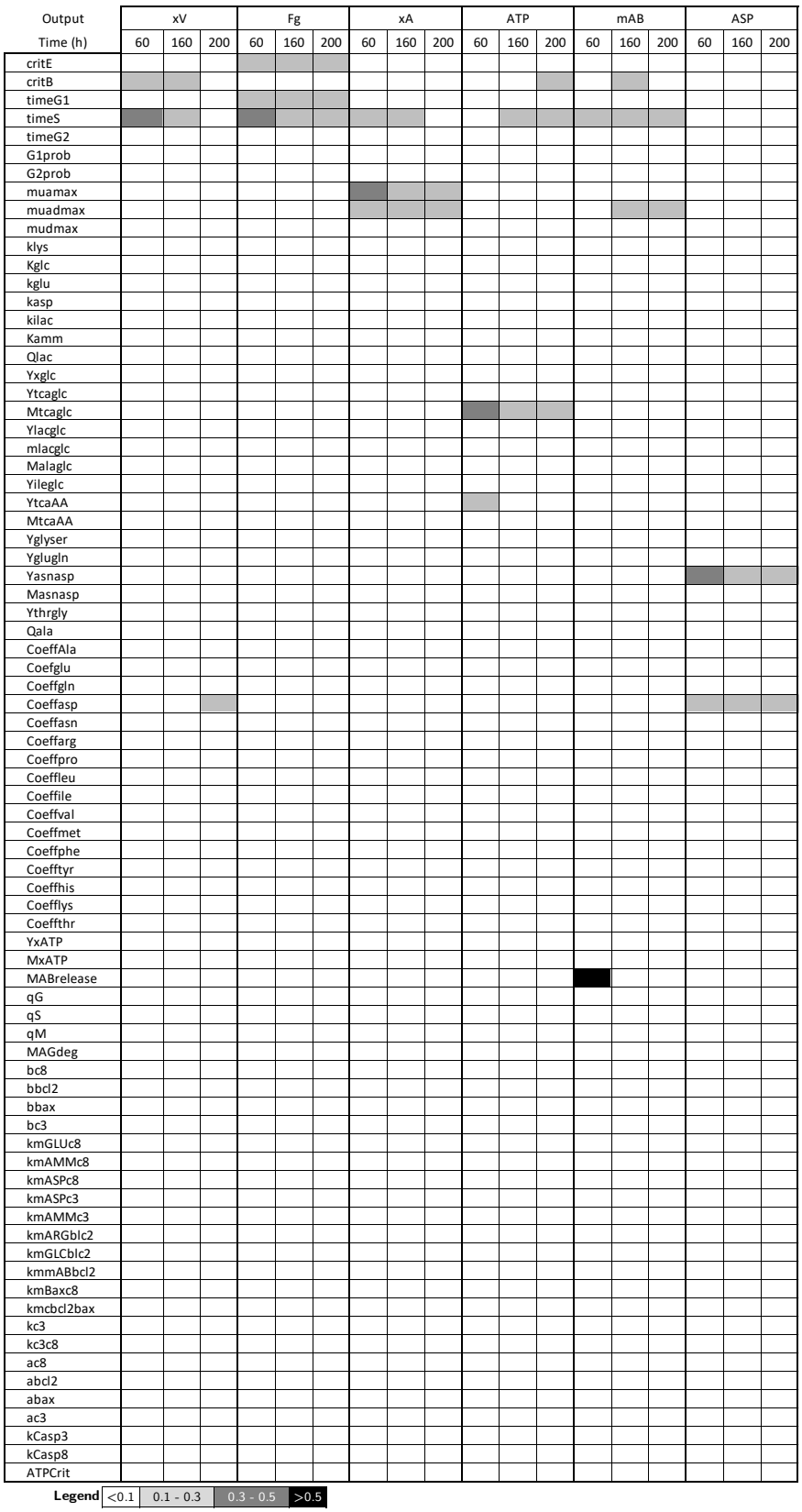


Figure 48: Global Sensitivity Analysis results. The significance of each parameter was assessed using the Sobol method at 60, 160 and 200h on viable cells (xV), fraction of cells in phase G1 (Fg), apoptotic cells (xA), energy production (ATP), monoclonal antibody (mAb), caspase-8 transcripts (casp-8) and aspartate concentration (Asp). A minimum threshold of 10% was set to discriminate important parameters. Significant parameters are shaded according to their SIs in light grey (0.1 – 0.3), dark grey (0.3 – 0.5) or black (>0.5).

6.6 PARAMETER ESTIMATION

After identification of the significant model parameters, those were re-estimated using experimental data from batch cultures of GS-CHO cells conducted at the density recommended by the manufacturer ($2x10^8$cells/L). The results are presented in Table 5. In order to allow for the use of deterministic PE algorithms, branched kinetics were implemented using on/off switches based on the *tanh* function as described in Appendix IV.

The MLE algorithm used solved the parameter estimation problem and estimated most parameters within a reasonable confidence interval. The times spent in each cell cycle phase, 7.98±0.37, 11.28±0.43 and 3.72±0.46 hours were estimated in good agreement with the experimental values, respectively 8, 14 and 4 hours for phases G_1, S and G_2/M (Chapter 5)

When attempting to estimate the coefficients related to the use of arginine, proline, methionine and valine in the TCA cycle, the lowest bound allowed in the algorithm (zero) was hit and, hence, a confidence interval could not be obtained. These results suggest that these amino acids are not used in the TCA cycle. This is in agreement with a previously published FBA analysis of GS-CHO cells predicting that none of these amino acids are used in the TCA cycle (Kyriakopoulos and Kontoravdi, 2014). Additionally, these parameter estimation results suggest that some isoleucine and some leucine are used in the TCA cycle. While the FBA analysis mentioned above suggests that some isoleucine is used in the TCA cycle, it does not suggest any usage of leucine for energy production. The very low coefficients herein obtained are also in agreement with these FBA predictions. Aspartate, glutamate and glutamine are predicted to be used in energy production which is also the prediction of the parameter estimation results herein obtained.

Table 5: Parameter estimation results. Significant model parameters were re-estimated with experimental data obtained in this work and the MLE algorithm implemented in gProms.

Parameter (units)	**Biological Significance**	**Final Value**	**95% CI**
critE (--)	Cyclin E threshold for G_1 to S transition	12.9	0.3
critB (--)	Cyclin B threshold for S to G_2/M transition	36.6	0.97
$timeG1_{min}$ (h)	Time spent in phase G_0/G_1	7.98	0.37
$timeS_{min}$ (h)	Time spent in phase S	11.28	0.43
$timeG2_{min}$ (h)	Time spent in phase G_2/M	3.72	0.46
$\mu_{a,max}$ (h^{-1})	Maximum rate of early apoptosis	$1.82x10^{-3}$	$2.2x10^{-4}$
$\mu_{ad,max}$ (h^{-1})	Maximum rate of late apoptosis	$1.06\ x10^{-1}$	$1.9x10^{-2}$
coeffAla (--)	Relation between total rate of AA consumption for energy production and the rate of consumption of each AA in the TCA cycle	$9.01\ x10^{-2}$	$4.7x10^{-2}$
coeffArg (--)		0	--
coeffAsp (--)		$5.12x10^{-1}$	$7.3x10^{-2}$
coeffGln (--)		$9.00x10^{-2}$	$2.1x10^{-2}$
coeffGlu (--)		$1.05x10^{-1}$	$9.7x10^{-3}$
coeffIle (--)		$5.23x10^{-2}$	$5.1x10^{-3}$
coeffLeu (--)		$4.56x10^{-2}$	$1.8x10^{-3}$
coeffPro (--)		0	--
coeffMet (--)		0	--
coeffVal (--)		0	--
$m_{Ala,Glc}$(mmol/cell)	Non-growth related glucose used to alanine	$6.04x10^{-12}$	$1.9x10^{-13}$
$m_{tca,Glc}$(mmol/cell)	TCA usage of glucose for maintenance	$2.15x10^{-11}$	$9.4x10^{-13}$
$Q_{cons,Ala}$ (mmol/cell)	Alanine consumption rate	$4.41x10^{-12}$	$1.3x10^{-12}$
$Q_{cons,Lac}$ (mmol/cell)	Lactate production rate	$1.01x10^{-11}$	$4.7x10^{-12}$
$Y_{Asp,Asn}$ (mol/mol)	Yield of aspartate on asparagine	$1.40x10^{9}$	$4.6x10^{7}$
$Y_{Gln,Glu}$ (mol/mol)	Yield of glutamine on glutamate	$4.95x10^{9}$	$2.9x10^{8}$
$Y_{Gly,Ser}$ (mol/mol)	Yield of glycine on serine	$1.86x10^{9}$	$2.1x10^{8}$
$Y_{Ile,Glc}$ (mol/mol)	Yield of isoleucine on glucose	$1.06x10^{10}$	$6.8x10^{9}$
$Y_{Lac,Glc}$ (mol/mol)	Yield of lactate on glucose	$3.25x10^{8}$	$2.6x10^{7}$
$Y_{tca,AA}$ (mol/mol)	Growth-related AA use in TCA cycle	$1.36x10^{9}$	$4.0x10^{8}$
$Y_{Thr,Gly}$ (mol/mol)	Yield of threonine on Glycine	$9.01x10^{9}$	$9.4x10^{9}$
$mAb_{release}$	Rate of mAb release from viable cells	$1.12x10^{-9}$	$1.1x10^{-10}$
α_{bax} (h^{-1})	Bax mRNA degradation rate	$1.00x10^{-1}$	$7.1x10^{-2}$
$\alpha_{bcl\text{-}2}$ (h^{-1})	Bcl-2 mRNA degradation rate	$9.61x10^{-2}$	$8.4x10^{-2}$
α_{casp3} (h^{-1})	Casp-3 mRNA degradation rate	$2.06x10^{-1}$	$1.8x10^{-2}$
α_{casp8} (h^{-1})	Casp-8 mRNA degradation rate	$3.00x10^{-3}$	$1.0x10^{-2}$
β_{bax} (h^{-1})	Bax mRNA base production rate	$1.00\ x10^{-3}$	$8.4x10^{-2}$
$\beta_{bcl\text{-}2}$ (h^{-1})	Bcl-2 mRNA base production rate	$8.00x10^{-2}$	$6.7x10^{-2}$
β_{casp3} (h^{-1})	Casp-3 mRNA base production rate	$1.00x10^{-1}$	$2.0x10^{-2}$
β_{casp8} (h^{-1})	Casp-8 mRNA base production rate	$4.00x10^{-2}$	$8.2x10^{-3}$

In spite of solving the estimation problem, there were three parameters for which the algorithm found 95% confidence intervals larger than the parameter value. This happened for the rate of degradation of caspase 8 transcripts (α_{casp8}), the basal expression rate of bax (β_{bax}) and the yield of threonine on glycine

($Y_{Thr/Gly}$). These results mean that these parameters cannot be accurately estimated from the experimental data. This problem has also been identified in Chapter 5 for gene expression parameters related to the bax gene. It is deemed that more experimental data could improve the parameter estimation problem. Apart from these three parameters, in all remaining cases parameter values could be estimated with confidence intervals in several cases one order of magnitude lower than the value of the parameter.

6.7 RESULTS AND DISCUSSION

In this section the model predictions of three sets of experimental data are shown. One data set, related to cultures conducted at the nominal density recommended by the manufacturer was used to estimate the relevant model parameters. The other two data sets were predicted without changing any parameter values and are therefore used for model validation. Throughout this section, the data set used for parameter estimation is shown in blue and those used for model predictions are shown in green (lower density) and pink (higher density).

6.7.1 VIABLE, APOPTOTIC AND DEAD CELL DENSITIES

In Figure 49 model predictions of viable, apoptotic and dead cell densities are shown together with the corresponding model predictions. Viable cell densities (Figure 49A) are closely predicted during the growth stages but they are overestimated during the stationary and decline phases for all cell densities. In particular, there appears to be an overprediction of the growth rate between 120 and 168h leading to the overestimation of the viable cell density, especially for the nominal and the lower inoculation densities. The trends observed in apoptotic cell densities (Figure 49B) are well captured by the model but the magnitudes are generally overpredicted, for all conditions. The dead cell densities (Figure 49C) are closely predicted during the initial stages of culture but overestimated during stationary and decline phases of culture, for all densities.

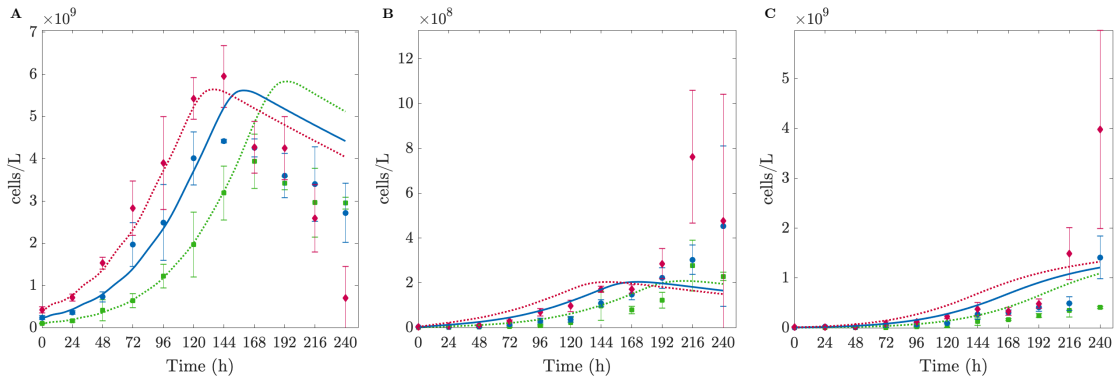


Figure 49. Model predictions (lines) and experimental data (dots, N=3 biological replicates, error bars show one standard deviation) of viable (A), apoptotic (B) and dead (C) cell densities obtained in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

Figure 50 shows model predictions and experimental data of cell cycle for all conditions tested. Phase G_2/M is closely predicted for the nominal density (Figure 50A) and relatively well predicted for the two validation data sets (Figure 50B and C). The predictions for phases G_1 and S are less closely predicted suggesting that the transition between phase G_1 and phase S is not well predicted by the model. This is likely a parametric problem, but it was not possible to further improve the parameter estimation results. Another possibility to improve these predictions is to conduct a detailed analysis of distribution of the cells in the cyclin and DNA domains in each phase of the cell cycle (García-Münzer *et al.*, 2013). Another strategy could be to increase the resolution of the measurements in order to improve the parameter estimation solution.

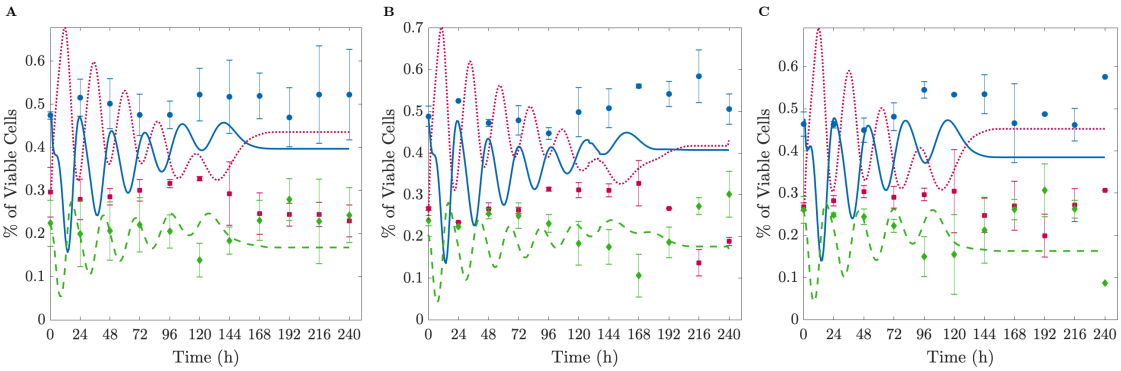


Figure 50. Cell cycle predictions (lines) and experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for phases G_1 (blue circles, solid lines), S (red squares and dotted lines) and G_2/M (green diamonds and dashed line) obtained in GS-CHO batch cultures inoculated at 2 (A), 1 (A) and 4x10^8 cells/L (C).

6.7.2 MONOCLONAL ANTIBODY

Monoclonal antibody production in GS-CHO cell cultures is well predicted by the model for all cell densities. For the nominal and the highest inoculation densities, mAb titre is slightly overpredicted throughout while, for the lowest inoculation density, the titre is very closely predicted throughout the culture duration.

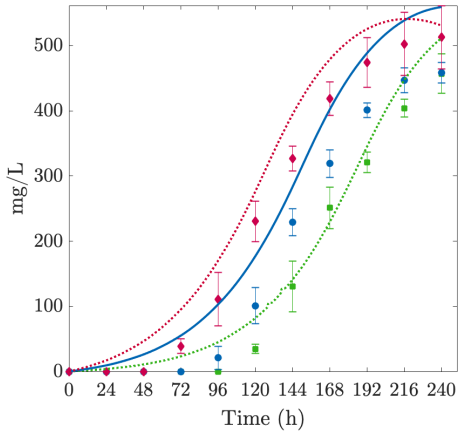


Figure 51. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for mAb production in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

6.7.3 GLUCOSE, LACTATE AND ALANINE METABOLISM

As mentioned above, glucose is metabolised to satisfy different requirements and some carbon is diverted to lactate production. Furthermore, some pyruvate produced in culture is not directed to the TCA cycle but to alanine production. Figure 52A shows the overall glucose consumption profiles obtained for all conditions. For cultures inoculated at 2 and 4x10^{8}cells/mL, the glucose consumption profile is closely predicted throughout culture although it does not account for the dramatic reduction in glucose consumption during the decline phase. For the lowest inoculation density, the glucose consumption rate is underestimated leading to an overprediction in the extracellular concentration of glucose.

In Figure 52B, the lactate profiles are shown. Model predictions follow experimental data closely for all inoculation densities. The time and magnitude of the metabolic shift is very well predicted for all inoculation densities.

Finally, in Figure 52C, the alanine profiles are shown to be well predicted by the model for all inoculation densities including alanine production from glucose (before the metabolic shift) and its usage to produce pyruvate during the decline phase.

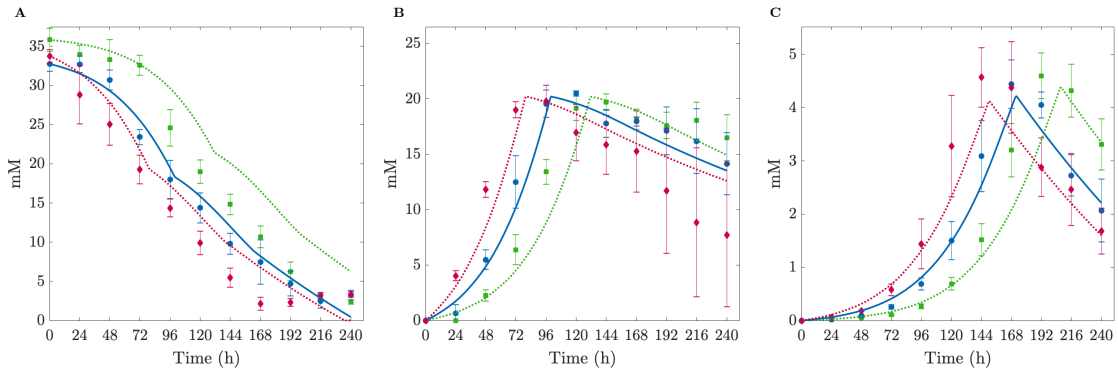


Figure 52. Model predictions (lines) and experimental data (dots, N=3 biological replicates, error bars show 1 standard deviation) of glucose (A), lactate (B) and alanine (C) profiles obtained in the course of GS-CHO batch cultures inoculated at 1 (green squares, dotted lines), 2 (blue circles and solid lines) and 4x108cells/mL (red diamonds and dashed lines).

6.7.4 AMMONIA

As discussed in Chapter 6, ammonia metabolism presents two shifts. These have been accounted for the model structure presented above.

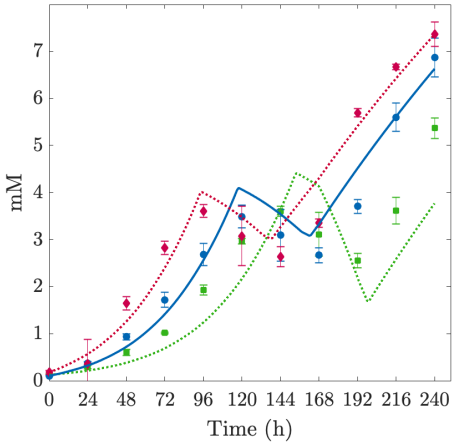


Figure 53. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for ammonia production in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

Figure 53 shows the model predictions and experimentally measured ammonia concentrations in batch cultures. The trends and magnitudes are well predicted before the first ammonia metabolic shift. Time points in between the two

metabolic shifts are well predicted for the nominal and the densest inoculation densities. For the lowest inoculation density, the model overestimates ammonia concentrations before the first metabolic shift and underestimates the time points after the second metabolic shift. For the other conditions, the model performs well after the second metabolic shift.

6.7.5 NON-ESSENTIAL AMINO ACIDS

The undisturbed growth of GS-CHO cells in culture has been shown to be correlated with the metabolism of asparagine (Carinhas *et al.*, 2013; Kyriakopoulos, Polizzi and Kontoravdi, 2013; Duarte *et al.*, 2014). While Duarte *et al.* have suggested that CHO cells could be auxotrophic for asparagine, genome-scale studies conducted with CHO cell lines suggest that this is, likely, a feature unique to the cell line used by Duarte *et al.* (Hefzi *et al.*, 2016). The MFA work of Carinhas *et al.* shows that asparagine is consumed to produce aspartate which is consistent with the FBA work of Kyriakopoulos and Kontoravdi where nearly all asparagine consumed is predicted to be directed to aspartate (Kyriakopoulos and Kontoravdi, 2014). As shown above, this has been accounted for in the development of the kinetic model of GS-CHO cell culture.

Additionally, aspartate is considered to be converted to oxaloacetate thereby entering the TCA cycle, as shown in the MFA of Carinhas *et al.* In the FBA work of Kyriakopoulos and Kontoravdi, aspartate is considered to be transaminated to oxaloacetate at the cost of producing glutamate which is predicted to be the most significant contributor to the TCA cycle. In Chapter 5, it has been shown that both aspartate and glutamate accumulate in culture during the lag/early-exponential phase being consumed only during exponential growth.

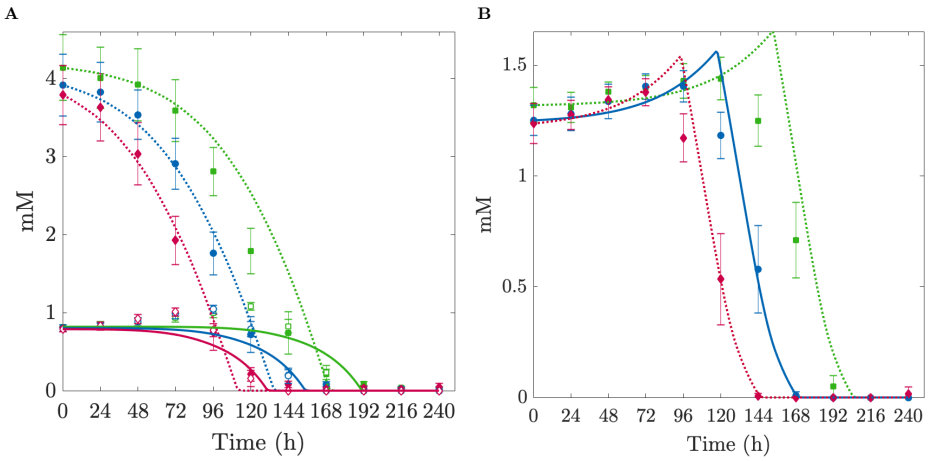


Figure 54. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) asparagine (solid symbols, dotted lines) and aspartate (empty symbols and solid lines) and (B) glutamate metabolism in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

Figure 54 shows the model predictions and corresponding experimental data for asparagine, aspartate, glutamine and glutamate in culture. Results in Figure 54A show that the model can closely predict asparagine concentration in culture while aspartate accumulation is not captured by the model although concentrations are well predicted. In Figure 54B, it is shown that the glutamate profiles are well predicted for all densities throughout culture.

Glutamate can also be used to produce glutamine, which accumulates in cultures, being only consumed during stationary growth (Figure 55A). The standard deviations in glutamine are attributed to the oscillating behaviour of glutamine uptake in GS-CHO cell lines (Kyriakopoulos, Polizzi and Kontoravdi, 2013). As mentioned above, parameter estimation results and a previously published FBA (Kyriakopoulos and Kontoravdi, 2014) suggest that arginine (Figure 55A), phenylalanine and proline (Figure 55B), tyrosine and cysteine (Figure 55C) are not required to fuel the TCA cycle.

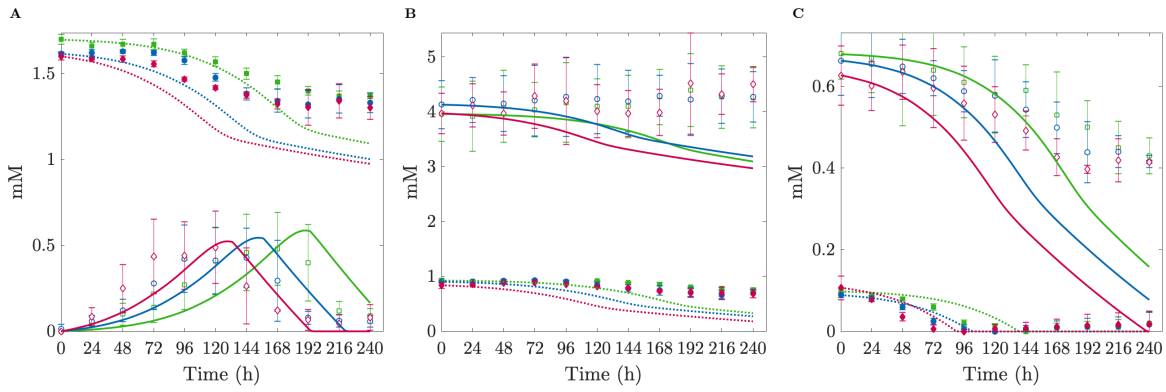


Figure 55. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) arginine (solid symbols, dotted lines) and glutamine (empty symbols and solid lines), (B) proline (solid symbols, dotted lines) and phenylalanine (empty symbols and solid lines) and (C) tyrosine (solid symbols, dotted lines) and cysteine (empty symbols and solid lines) metabolism in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

The profiles of glutamine, Figure 55A, are captured although there is a plateau in glutamine concentrations during the exponential phase that is not predicted by the model.

All other amino acid profiles shown in Figure 55, except cysteine show a significant deviation from the early stages of the culture. As none of these amino acids is used in energy production, their usage includes only terms accounting for biomass and mAb production. As the composition of the exact monoclonal antibody produced by this cell line was provided by the manufacturer, deviations between model and experimental data can be due to differences in the CHO cell composition experimentally obtained by Selvarasu *et al.* (and used in this model) and the composition of the particular cell line used in this work. A different cell line composition affects the model predictions for all amino acids and adds to the importance of having a clone-specific model. If a better estimate of the amino acid requirements for this cell line becomes available, parameters related to amino acid metabolism should also be re-estimated.

Serine, glycine and threonine (essential amino acid) profiles are shown in Figure 56. Although the trends are captured during the growth stages, as was discussed above for other amino acids, all profiles are underpredicted which is,

likely, to the differences in the CHO composition used and the real composition of this particular clone.

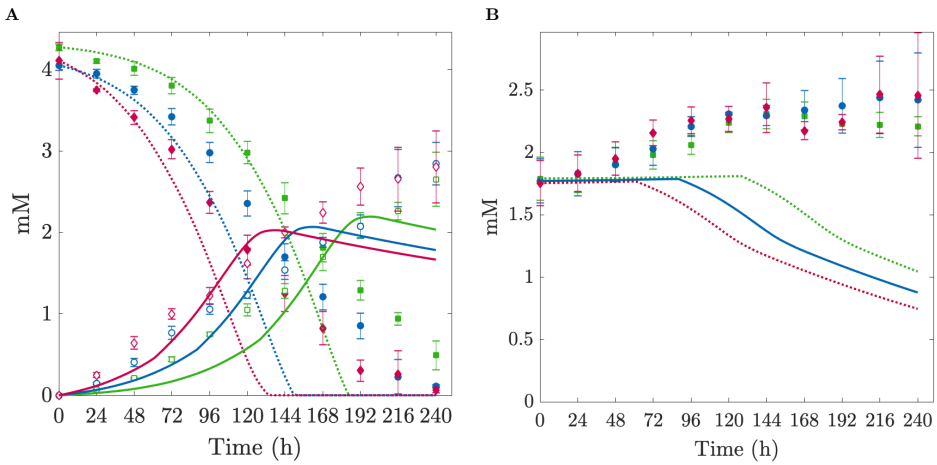


Figure 56. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) serine (solid symbols, dotted lines) and glycine (empty symbols and solid lines), (B) threonine metabolism in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

6.7.6 ESSENTIAL AMINO ACIDS

The depletion of non-essential amino acids is generally related to a loss in viability. In GS-CHO cell cultures, none of the non-essential amino acids was depleted in culture (Figure 57) at the end of the culture, although viability did decrease significantly. This is a significant difference compared to GS-NS0 cells (Chapter 5).

The metabolism of leucine and histidine (Figure 57A) is well captured by the model although the issue of cell composition appears to affect model predictions of all essential amino acids. The metabolism of isoleucine (Figure 57B), including the metabolic shift, is well predicted by the model for all conditions. Model predictions of the profiles of tryptophan (B), lysine and valine (Figure 57C) and methionine (Figure 57D) appear to be affected by the cell composition used in this model, as discussed above. The differences observed could also be, at least partly, caused by the growth rate over prediction observed between 120 and 168h of culture.

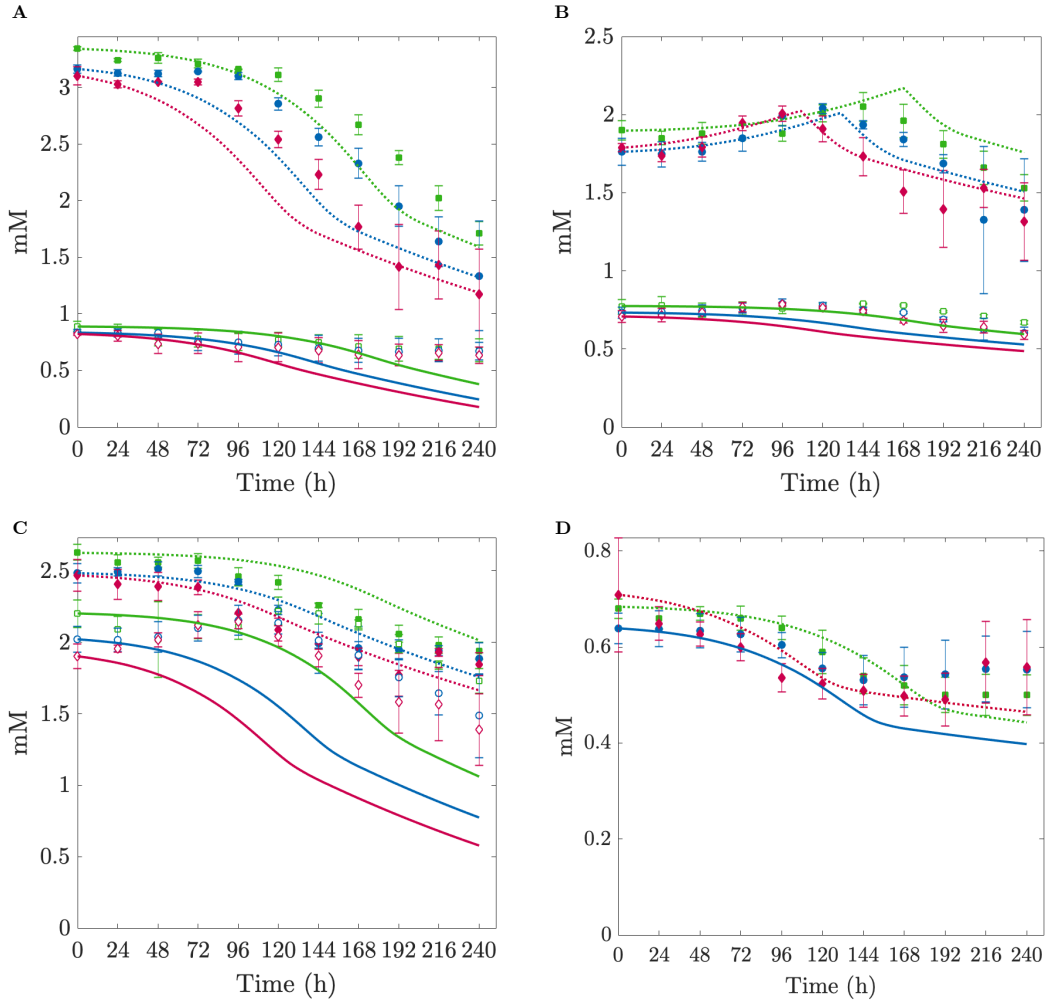


Figure 57: Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for (A) leucine (solid symbols, dotted lines) and histidine (empty symbols and solid lines), (B) isoleucine (solid symbols, dotted lines) and tryptophan (empty symbols and solid lines), (C) lysine (solid symbols, dotted lines) and valine (empty symbols and solid lines) and (D) methionine (solid symbols, dotted lines) in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

6.7.7 ATP

Energy production in GS-CHO cell culture, in the form of ATP, can be generated from glucose and amino acids, as described while introducing the model structure. The ATP profiles for GS-CHO cells are shown in Figure 58. While the ATP profiles are followed closely during growth and decline phases, the plateau seen during the stationary phase is not well predicted by the model, which is particularly evident for the culture inoculated at the lowest density.

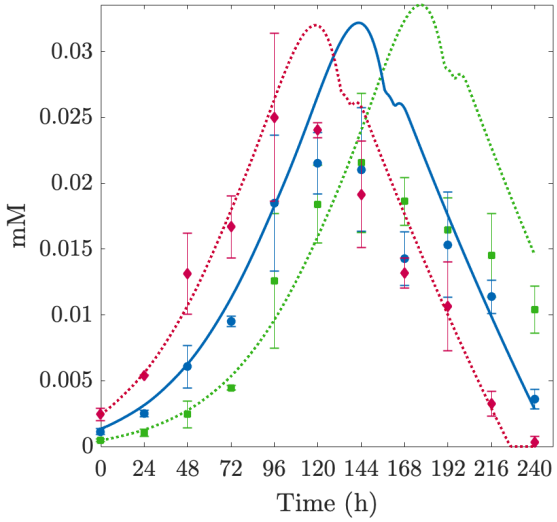


Figure 58. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for ATP production in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

6.7.8 APOPTOSIS-RELATED GENE EXPRESSION

While the rate of death through processes other than apoptosis is calculated as a function of the minimum ATP requirements of a cell (estimated in Chapter 5), the rate of early apoptosis (viable to apoptotic) and late apoptosis (apoptotic to dead) are linked to the experimentally measured expression of genes related to these processes. Specifically, the rate of early-stage apoptosis is linked to the expression of caspase-8 and bax (a downstream target of caspase-8 continuously expressed throughout culture) and the rate of late apoptosis is calculated as a function of the expression of caspase-3. Finally, the anti-apoptotic effect of bcl-2 is also accounted for as explained in the model structure. The expression of these genes and the corresponding model predictions are shown in Figure 59.

The initial activation followed by deactivation of caspase-8 (Figure 59A) is well predicted by the model presenting an improvement compared to the model obtained for GS-NS0 cells. However, in spite of predicting the trends, the magnitudes of the experimental data are not very well captured. The expression of the bax (Figure 59B) gene shows a significant uncertainty in the data points inherent to the technique used and the model predictions are also not able to track

the oscillatory behaviour of the expression of this gene. The trends observed in the expression of caspase-3 (Figure 59C) are captured by the model until the decline phase where the decrease in transcripts is not captured by the model. Finally, the expression of bcl-2 (Figure 59D) is well predicted by the model for all inoculation densities.

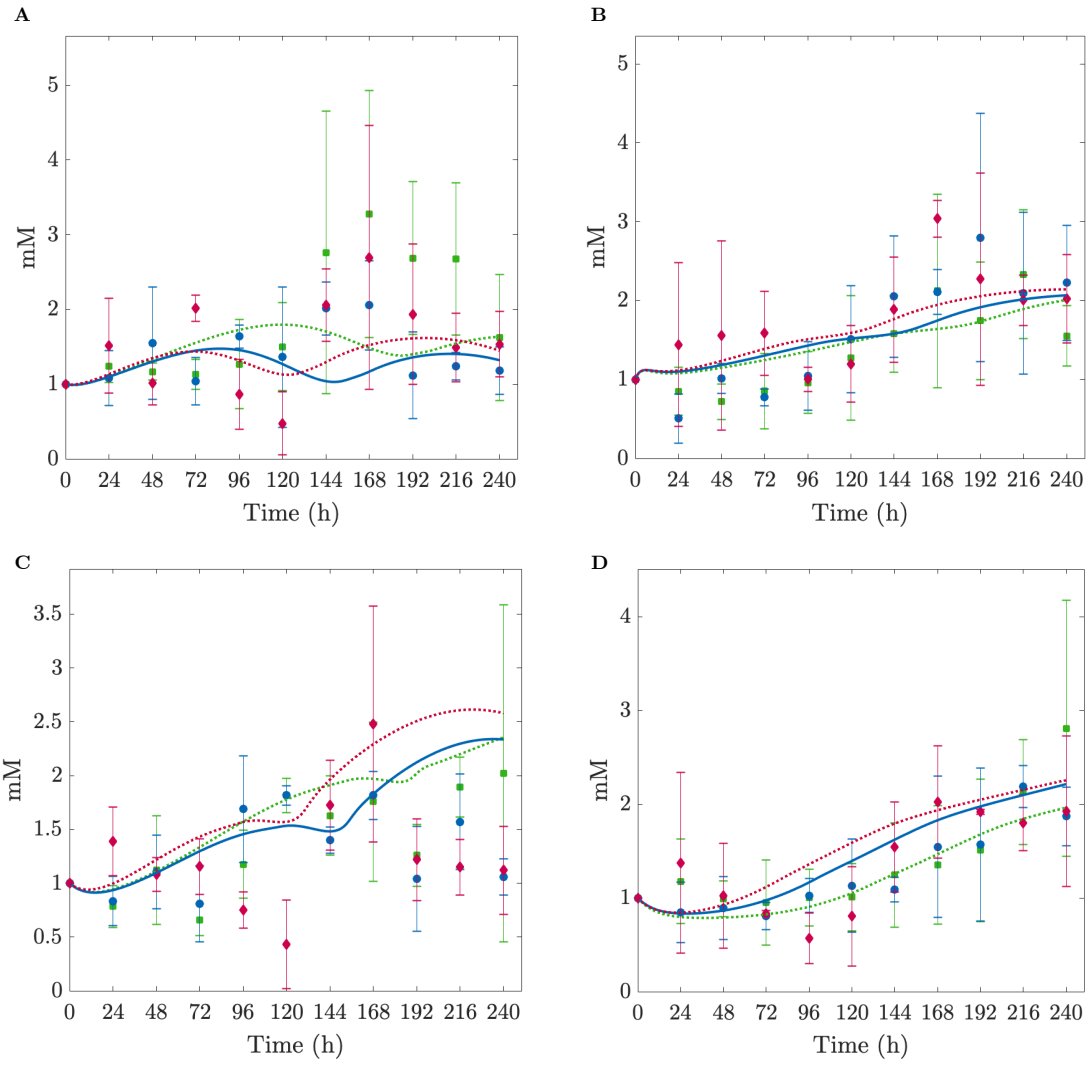


Figure 59. Experimental data (dots, N=3 biological replicates, error bars show one standard deviation) for the expression of (A) caspase-8, (B) bax, (C) caspase-3 and (D) bcl-2 in GS-CHO batch cultures inoculated at 1 (green squares and dotted lines), 2 (blue circles and solid lines) and 4x10^8 cells/L (red diamonds and dashed lines).

6.8 CONCLUSIONS

Following the integrated model structure presented in Chapter 4 and the detailed mapping of metabolism and cell cycle of GS-CHO cells obtained in Chapter 5, an integrated mathematical model of cell cycle, apoptosis and metabolism is suggested in this chapter.

Following the same model development framework used in Chapter 4, GSA was used to identify the model parameters most significantly affecting important model outputs. From the total 78 parameters of the model, only 37 model parameters were found significant, i.e. there is at least one time point for at least one output where such parameter has a sensitivity index higher than 0.1. GSA has also allowed to derive biologically significant information showing that the model structure has the potential to describe the biology of the system.

Parameter estimation was conducted in order to estimate the important model parameters identified by GSA. Experimental data obtained in this work from batch cultures conducted following the manufacturer's instructions was used for this purpose. Results show that the model can describe metabolism and growth of this cell line, predict energy generation during most of the culture duration and capture the main trends of cell cycle and gene expression. Finally, model predictions of two independent data sets were also used to assess model predictive capabilities. For most outputs, model performance at the nominal inoculation density (used for parameter estimation), is as good as for the validation data sets.

For most amino acids, the consumption rates are overestimated by the model leading to lower concentration predictions. Amino acid usage is considered for energy (TCA cycle), mAb production and biomass production. For amino acids not used in the TCA cycle, usage for mAb production is given by the mAb composition obtained by the cell line manufacturer and the usage for biomass production is calculated based on the experimental composition of CHO cell lines obtained by Selvarasu *et al.* It is likely that a more accurate determination of the amino acid

composition will lead to a significant improvement in the performance of the present metabolic model.

The work presented in this chapter presents an integrated kinetic model of GS-CHO cell cultures which can be used for optimization of medium compositions, feeding strategies (requires validation in fed-batch mode) or bioreactor geometry (requires validation and integration with CFD models). In conjunction with the work presented in chapter 6 and previous literature, this work addresses the hopes of other authors that the use of steady-state techniques, such as MFA, for metabolism understanding be used as sources of information to build kinetic models of cell culture, which could be used for optimization (Carinhas *et al.*, 2013; Galbraith *et al.*, 2018).

Chapter 7

Future work

This work has delivered kinetic models of GS-CHO and GS-NS0 cell cultures aimed at improving process development of biologics manufactured using these two cell lines. The metabolic sub-models presented herein can be used alone, combined with a simpler growth and death model, for medium design. The metabolism analysis framework presented in Chapter 4 and Chapter 5 allows to identify different growth regimes and the main metabolic changes occurring at the transitions between them. In Chapter 6, a kinetic model of GS-CHO cultures has successfully been developed based on the information of the MVA presented in Chapter 5. Additionally, MFA may be employed in order to further elucidate metabolism during each culture phase identified by MVA.

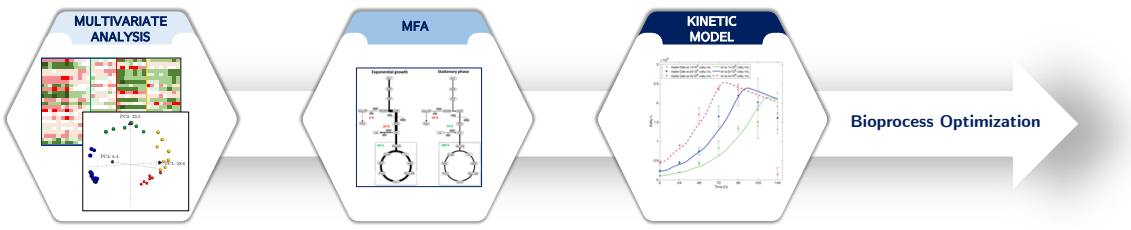


Figure 60. Bioprocess optimization route combining multivariate analysis, steady state techniques and kinetic models.

This analysis would greatly benefit from including both intra- and extra-cellular metabolites in order to validate metabolic networks describing cell metabolism in culture. In conjunction with the MVA framework developed in this work, MFA studies can support the development of kinetic models, which can be used for medium design and feeding strategies optimization as illustrated in Figure 60. It is envisaged that the combination of MVA - to identify the main culture stages and gain understanding of main metabolic shifts and productive culture stages – with MFA for metabolic network understanding and finally kinetic dynamic models, which can be used for optimization, will allow for the implementation of the long-term vision, expressed by other authors, that the integration of steady-state and kinetic models of metabolism should be attempted for cell culture optimization.

Another important extension to the present work would be the incorporation of product quality predictions in the models developed. In particular, it is interesting to understand how apoptosis affects product quality and how the optimization of feeding regimes, media composition or bioreactor design may contribute to improved product quality by minimizing apoptotic cell death. An on-going collaboration with Bayer, under the scope of the ModLife ITN, is looking into the impact of apoptotic stress on product quality of mAbs in GS-NS0 cultures. This research path has the potential to deliver integrated control strategies for cell culture improvements on titre and product quality (Figure 61).

Furthermore, it would be important to identify if the mAb produced during different culture stages has the same quality or whether metabolic limitations to growth affect product quality. This could be an important criteria for the development of feeding strategies for fed-batch cultures.

The mathematical models developed in this work cover the two most widely used cell lines used industrially for mAbs manufacturing. Their ability to capture population heterogeneity provides a unique opportunity to perform cell culture optimization taking into consideration both cell biology and bioreactor

(environment) effects. In order to accomplish this, it is necessary to couple the mathematical models developed in this work with computational fluid dynamics (CFD) of the bioreactors. This has been done during this project, in collaboration with Mr. Ke Meng (MSc. candidate), by combining simpler biological models with CFD descriptions of a bench-scale bioreactor in order to show the potential of this approach. Therefore, the author believes that this is a path worth exploring. Such integrated optimization tool (Figure 62) has the potential of optimizing design and operation of bioreactors in an integrated fashion.

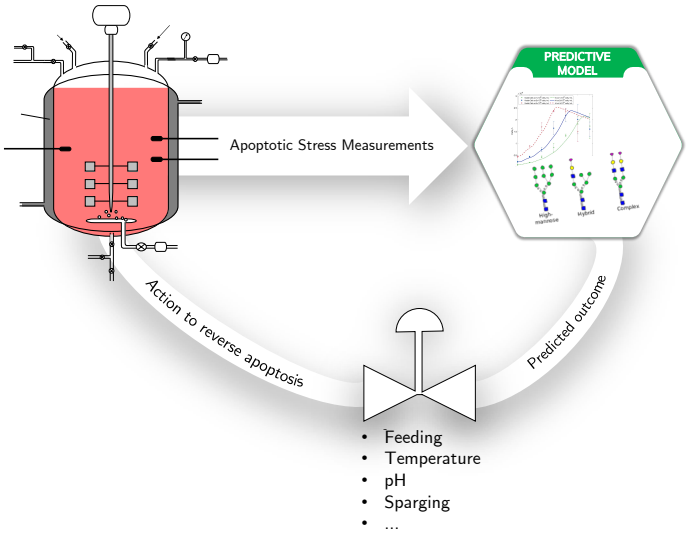


Figure 61: Model based control of product quality based on apoptosis reversal.

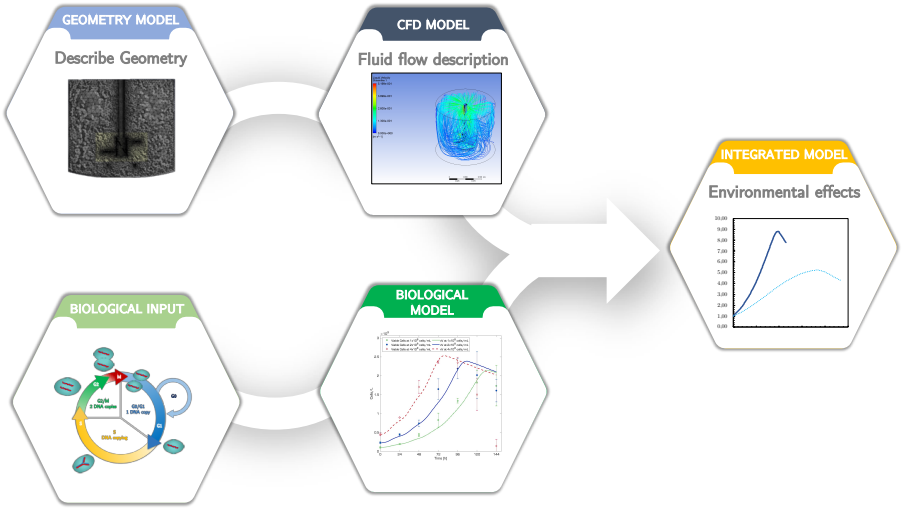


Figure 62: Integration of biological models and CFD model for integrated bioreactor optimization

References

A. Hjortso, M. and Nielsen, J. (1995) 'Population balance models of autonomous microbial oscillations', *Journal of Biotechnology*, 42(3), pp. 255–269. doi: 10.1016/0168-1656(95)00086-6.

Abt, V. *et al.* (2018) 'Model-based tools for optimal experiments in bioprocess engineering', *Current Opinion in Chemical Engineering*, pp. 244–252. doi: 10.1016/j.coche.2018.11.007.

Abu-Absi, S. *et al.* (2013) 'Cell Culture Process Operations for Recombinant Protein Production', *Adv. Biochem. Eng. Biotech.* Springer, 139, pp. 35–68. doi: 10.1007/10_2013_252.

Ahn, W. S. and Antoniewicz, M. R. (2011) 'Metabolic flux analysis of CHO cells at growth and non-growth phases using isotopic tracers and mass spectrometry', *Metabolic Engineering*. doi: 10.1016/j.ymben.2011.07.002.

Al-Rubeai, M. *et al.* (1991) 'Cell cycle, cell size and mitochondrial activity of hybridoma cells during batch cultivation', *Cytotechnology*, 7(3), pp. 179–186. doi: 10.1007/BF00365929.

Al-Rubeai, M. *et al.* (1992) 'Specific monoclonal antibody productivity and the cell cycle-comparisons of batch, continuous and perfusion cultures', *Cytotechnology*, 9(1–3), pp. 85–97. doi: 10.1007/BF02521735.

Al-Rubeai, M. and Emery, A. N. (1993) 'Flow cytometry in animal culture', *Nature biotechnology*, 11(5), pp. 572–4, 577–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7763604.

Albrecht, S. *et al.* (2018) 'Proteomics in biomanufacturing control: Protein dynamics of CHO-K1 cells and conditioned media during apoptosis and necrosis', *Biotechnology and Bioengineering*, 115(6), pp. 1509–1520. doi: 10.1002/bit.26563.

Almquist, J. *et al.* (2014) 'Kinetic models in industrial biotechnology - Improving cell factory performance', *Metabolic Engineering*. Elsevier, 24, pp. 38–60. doi: 10.1016/j.ymben.2014.03.007.

Altamirano, C. *et al.* (2000) 'Improvement of CHO cell culture medium formulation: Simultaneous substitution of glucose and glutamine', *Biotechnology Progress*. doi: 10.1021/bp990124j.

Altman, B. J. and Rathmell, J. C. (2012) 'Metabolic stress in autophagy and cell death pathways', *Cold Spring Harbor Perspectives in Biology*, 4(9), pp. 8763–8763. doi: 10.1101/cshperspect.a008763.

Antonsson, B. *et al.* (1997) 'Inhibition of Bax channel-forming activity by Bcl-2.', *Science (New York, N.Y.)*, 277(5324), pp. 370–372. doi: 10.1126/science.277.5324.370.

Ashkenazi, A. *et al.* (2017) 'From basic apoptosis discoveries to advanced selective BCL - 2 family inhibitors', *Nature Publishing Group*. Nature Publishing Group, 16(4), pp. 273–284. doi: 10.1038/nrd.2016.253.

Barnes, L. M., Bentley, C. M. and Dickson, A. J. (2000) 'Advances in animal cell recombinant protein production: GS-NS0 expression system', *Cytotechnology*, 32(2), pp. 109–123. doi: 10.1023/A:1008170710003.

Batchelor, E. *et al.* (2008) 'Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage.', *Molecular cell*, 30(3), pp. 277–89. doi: 10.1016/j.molcel.2008.03.016.

Bebbington, C. R. *et al.* (1992) 'High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker.', *Bio/technology (Nature Publishing Company)*, 10(2), pp. 169–75. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1369477.

Bentele, M. *et al.* (2004) 'Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis', *Journal of Cell Biology*, 166(6), pp. 839–851. doi: 10.1083/jcb.200404158.

van den Berg, R. A. *et al.* (2006) 'Centering, scaling, and transformations: improving the biological information content of metabolomics data.', *BMC genomics*, 7(19), p. 142. doi: 10.1186/1471-2164-7-142.

Bibila, T. *et al.* (1994) 'Investigation of NS0 cell metabolic behavior in monoclonal antibody producing clones.', *Annals of the New York Academy of Sciences*, 745, pp. 277–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7832518.

Bibila, T. A. and Flickinger, M. C. (1991) 'A model of interorganelle monoclonal antibody transport and secretion in mouse hybridoma cells', *Biotechnology and Bioengineering*, 38(7), pp. 767–780. doi: 10.1002/bit.260380711.

Bibila, T. A. and Flickinger, M. C. (1992a) 'Use of a structured kinetic model of antibody synthesis and secretion for optimization of antibody production systems: I. Steady-state analysis.', *Biotechnology and bioengineering*, 39(3), pp. 251–61. doi: 10.1002/bit.260390302.

Bibila, T. A. and Flickinger, M. C. (1992b) 'Use of a structured kinetic model of

antibody synthesis and secretion for optimization of antibody production systems: I. Steady-state analysis.', *Biotechnology and bioengineering*, 39(3), pp. 251–61. doi: 10.1002/bit.260390302.

Blackstone, C. and Chang, C.-R. (2011) 'Mitochondria unite to survive.', *Nature Cell Biology*. Nature Publishing Group, 13(5), pp. 521–522. doi: 10.1038/ncb0511-521.

Bloemkolk, J.-W. *et al.* (1992) 'Effect of temperature on hybridoma cell cycle and MAb production', *Biotechnology and Bioengineering*, 40(3), pp. 427–431. doi: 10.1002/bit.260400312.

Bobak, Y. *et al.* (2016) 'Arginine deprivation induces endoplasmic reticulum stress in human solid cancer cells', *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd, 70, pp. 29–38. doi: 10.1016/j.biocel.2015.10.027.

Browne, S. M. and Al-Rubeai, M. (2011) 'Analysis of an artificially selected GS-NS0 variant with increased resistance to apoptosis', *Biotechnology and Bioengineering*, 108(4), pp. 880–892. doi: 10.1002/bit.22994.

Buchsteiner, M. *et al.* (2018) 'Improving culture performance and antibody production in CHO cell culture processes by reducing the Warburg effect', *Biotechnology and Bioengineering*. doi: 10.1002/bit.26724.

Butler, M. (2005) 'Animal cell cultures : recent achievements and perspectives in the production of biopharmaceuticals', *Appl. Microbiol. Biot.*, 68, pp. 283–291. doi: 10.1007/s00253-005-1980-8.

Butler, M. and Jenkins, H. (1989) 'Nutritional aspects of the growth of animal cells in culture', *Journal of Biotechnology*, 12(2), pp. 97–110. doi: 10.1016/0168-1656(89)90009-6.

Cacciuttolo, M. (2007) 'Fed-batch or perfusion? A matter of perspective', in Buttler, M. (ed.) *Cell culture and upstream processing*. New York: Taylor and Francis, pp. 173–184.

Campano, A. Q. *et al.* (2016) 'A Predictive Model for Energy Metabolism and ATP Balance in Mammalian Cells: Towards the Energy-Based Optimization of mAb Production', in *Computer Aided Chemical Engineering*, pp. 1581–1586. doi: 10.1016/B978-0-444-63428-3.50268-X.

Canales, M. *et al.* (2010) 'Bioprocess design and economics of recombinant BM86/BM95 antigen production for anti-tick vaccines', *Biochemical Engineering Journal*. Elsevier B.V., 52(1), pp. 79–90. doi: 10.1016/j.bej.2010.07.008.

Carinhas, N. *et al.* (2013) 'Metabolic signatures of GS-CHO cell clones associated with butyrate treatment and culture phase transition', *Biotechnology and Bioengineering*, 110(12), pp. 3244–3257. doi: 10.1002/bit.24983.

Chapman, J. D., Webb, R. G. and Borsa, J. (1971) 'ATP pool levels in synchronously growing Chinese hamster cells.', *The Journal of cell biology*, 49(1), pp. 229–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/5555578.

Chen, C. *et al.* (2007) 'Modeling of the role of a Bax-activation switch in the mitochondrial apoptosis decision.', *Biophysical journal*, 92(12), pp. 4304–15. doi: 10.1529/biophysj.106.099606.

Chen, Y. J. *et al.* (2016) 'Lactate metabolism is associated with mammalian mitochondria', *Nature Chemical Biology*, 12(11), pp. 937–943. doi: 10.1038/nchembio.2172.

Chong, S. L. *et al.* (2008) 'Cell growth, cell-cycle progress, and antibody production in hybridoma cells cultivated under mild hypothermic conditions.', *Hybridoma*, 27(2), pp. 107–111. doi: 10.1089/hyb.2007.0548.

Chrysanthopoulos, P. K., Goudar, C. T. and Klapa, M. I. (2010) 'Metabolomics for high-resolution monitoring of the cellular physiological state in cell culture engineering', *Metabolic Engineering*. Elsevier, 12(3), pp. 212–222. doi: 10.1016/j.ymben.2009.11.001.

Cohen, G. M. (1997) 'Caspases: the executioners of apoptosis', *Biochemical Journal*, 326(1), pp. 1–16. doi: 10.1042/bj3260001.

Cooper, G. M. and Hausman, R. E. (2007) *The Cell: a molecular approach*. 4th edn. Washington: ASM Press.

Cost, G. J. *et al.* (2010) 'BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells', *Biotechnology and Bioengineering*, 105(2), pp. 330–340. doi: 10.1002/bit.22541.

Cote, R. J. (2013) 'Microbial Media Composition', in Flickinger, M. C. (ed.) *Upstream Industrial Biotechnology: Expression Systems and process development*. Hoboken, New Jersy: J. Wiley and Sons., pp. 413–435.

Darzynkiewicz, Z. *et al.* (1996) 'Cytometry of cyclin proteins.', *Cytometry*, 25(1), pp. 1–13. doi: 10.1002/(SICI)1097-0320(19960901)25:1<1::AID-CYTO1>3.0.CO;2-N.

Dean, J. and Reddy, P. (2013) 'Metabolic analysis of antibody producing CHO cells in fed-batch production', *Biotechnology and Bioengineering*, 110(6), pp. 1735–1747. doi: 10.1002/bit.24826.

deZengotita, V. M. *et al.* (2000) 'Phosphate feeding improves high-cell-concentration NS0 myeloma culture performance for monoclonal antibody production', *Biotechnology and Bioengineering*, 69(5), pp. 566–576. doi: 10.1002/1097-0290(20000905)69:5<566::AID-BIT11>3.0.CO;2-4.

Dietmair, S. *et al.* (2012) 'Metabolite profiling of CHO cells with different growth characteristics', *Biotechnology and Bioengineering*, 109(6), pp. 1404–1414. doi: 10.1002/bit.24496.

Du, Z. *et al.* (2015) 'Use of a small molecule cell cycle inhibitor to control cell growth and improve specific productivity and product quality of recombinant proteins in CHO cell cultures', *Biotechnology and Bioengineering*, 112(1), pp. 141–155. doi: 10.1002/bit.25332.

Duarte, T. M. *et al.* (2014) 'Metabolic responses of CHO cells to limitation of key amino acids', *Biotechnology and Bioengineering*, 111(10). doi: 10.1002/bit.25266.

Dumont, J. *et al.* (2016) 'Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives', *Critical Reviews in Biotechnology*, 36(6), pp. 1110–1122. doi: 10.3109/07388551.2015.1084266.

Dutton, R. L., Scharer, J. and Moo-Young, M. (2006) 'Cell cycle phase dependent productivity of a recombinant Chinese hamster ovary cell line', *Cytotechnology*. doi: 10.1007/s10616-006-9041-4.

Eckhart, L. *et al.* (2008) 'Identification of novel mammalian caspases reveals an important role of gene loss in shaping the human caspase repertoire', *Molecular Biology and Evolution*, 25(5), pp. 831–841. doi: 10.1093/molbev/msn012.

Eguchi, Y., Shimizu, S. and Tsujimoto, Y. (1997) 'Intracellular ATP levels determine cell death fate by apoptosis or necrosis.', *Cancer research*, 57(10), pp. 1835–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9157970.

Eissing, T. *et al.* (2004) 'Bistability analyses of a caspase activation model for receptor-induced apoptosis.', *The Journal of biological chemistry*, 279(35), pp. 36892–7. doi: 10.1074/jbc.M404893200.

Elmore, S. (2007) 'Apoptosis: a review of programmed cell death.', *Toxicologic pathology*, 35(4), pp. 495–516. doi: 10.1080/01926230701320337.

Everhart, L. P. (1972) 'Effects of deprivation of two essential amino acids on DNA synthesis in chinese hamster cells', *Experimental Cell Research*, 74(2), pp. 311–318. doi: 10.1016/0014-4827(72)90382-5.

Fava, L. L. and Bock, F. J. (2012) 'Caspase-2 at a glance', *Journal of Cell Science*. doi: 10.1242/jcs.115105.

Ferrer-Miralles, N. *et al.* (2009) 'Microbial factories for recombinant pharmaceuticals.', *Microb. Cell Fact.*, 8, p. 17. doi: 10.1186/1475-2859-8-17.

Figueroa, B. *et al.* (2007) 'Enhanced cell culture performance using inducible anti-apoptotic genes E1B-19K and Aven in the production of a monoclonal antibody with Chinese hamster ovary cells', *Biotechnology and Bioengineering*, 97(4), pp. 877–892. doi: 10.1002/bit.21222.

Fischer, S., Handrick, R. and Otte, K. (2015) 'The art of CHO cell engineering: A comprehensive retrospect and future perspectives', *Biotechnology Advances*. Elsevier Inc., 33(8), pp. 1878–1896. doi: 10.1016/j.biotechadv.2015.10.015.

Fomina-Yadlin, D. *et al.* (2014) 'Cellular responses to individual amino-acid depletion in antibody-expressing and parental CHO cell lines', *Biotechnology and Bioengineering*, 111(5), pp. 965–979. doi: 10.1002/bit.25155.

Fredrickson, A. G. and Tsuchiya, H. M. (1963) 'Continuous propagation of

microorganisms', *AIChE Journal*, 9(4), pp. 459–468. doi: 10.1002/aic.690090410.

Freydell, E. J. *et al.* (2011) 'Techno-economic evaluation of an inclusion body solubilization and recombinant protein refolding process', *Biotechnology progress*, 27(5), pp. 1315–28. doi: 10.1002/btpr.652.

Fuentes-Garí, M. *et al.* (2015) 'A mathematical model of subpopulation kinetics for the deconvolution of leukaemia heterogeneity.', *Journal of the Royal Society Interface*, 12(108), pp. 20150276-. doi: 10.1098/rsif.2015.0276.

Fuge, G. *et al.* (2017) 'CHO cells engineered for fluorescence read out of cell cycle and growth rate in real time', *Biotechnology Progress*, 33(5), pp. 1408–1417. doi: 10.1002/btpr.2491.

Fulcher, D. and Basten, A. (1997) 'B cell life span: A review', *Immunology and Cell Biology*, 75(5), pp. 446–455. doi: 10.1038/icb.1997.69.

Fussenegger, M., Bailey, J. E. and Varner, J. (2000) 'A mathematical model of caspase function in apoptosis.', *Nature biotechnology*, 18(7), pp. 768–74. doi: 10.1038/77589.

Galbraith, S. C. *et al.* (2018) 'Media formulation optimization: current and future opportunities', *Current Opinion in Chemical Engineering*, 22, pp. 42–47. doi: 10.1016/j.coche.2018.08.004.

García-Münzer, D. G. *et al.* (2013) 'Developing a cyclin blueprint as a tool for mapping the cell cycle in GS-NS0', *Biochemical Engineering Journal*. Elsevier B.V., 81, pp. 97–107. doi: 10.1016/j.bej.2013.10.008.

García-Münzer, D. G. (2014) *Cell cycle heterogeneity study by an integrated modelling and experimental approach*. Imperial College London.

García-Münzer, D. G., Ivarsson, M., *et al.* (2015) 'An unstructured model of metabolic and temperature dependent cell cycle arrest in hybridoma batch and fed-batch cultures', *Biochemical Engineering Journal*. Elsevier B.V., 93, pp. 260–273. doi: 10.1016/j.bej.2014.10.013.

García-Münzer, D. G., Kostoglou, M., *et al.* (2015) 'Cyclin and DNA Distributed Cell Cycle Model for GS-NS0 Cells', *PLOS Computational Biology*. Edited by D. A. Beard, 11(2), p. e1004062. doi: 10.1371/journal.pcbi.1004062.

García Münzer, D. G. *et al.* (2014) *A Cyclin Distributed Cell Cycle Model in GS-NS0*, *Computer Aided Chemical Engineering*. Elsevier. doi: 10.1016/B978-0-444-63456-6.50004-1.

Gaughan, C. L. (2016) 'The present state of the art in expression, production and characterization of monoclonal antibodies', *Molecular Diversity*. Springer International Publishing, 20(1), pp. 255–270. doi: 10.1007/s11030-015-9625-z.

González Hernández, Y. and Fischer, R. W. (2007) 'Serum-free culturing of mammalian cells-adaptation to and cryopreservation in fully defined media.', *ALTEX : Alternativen*

zu Tierexperimenten, 24(2), pp. 110–116.

Goswami, J. *et al.* (1999) 'Apoptosis in Batch Cultures of Chinese Hamster Ovary Cells', *Biotech Bioeng*, 62(6). doi: 10.1002/(SICI)1097-0290(19990320)62.

Grilo, A. L. *et al.* (2017) 'Monoclonal antibodies production platforms: an opportunity study of a non protein A chromatographic platform based on process economics', *Biotechnology Journal*, 12(12). doi: 10.1002/biot.201700260.

Grilo, A. L. and Mantalaris, A. (2019a) 'Apoptosis: A mammalian cell bioprocessing perspective.', *Biotechnology advances*, 37(3), pp. 459–475. doi: 10.1016/j.biotechadv.2019.02.012.

Grilo, A. L. and Mantalaris, A. (2019b) 'The Increasingly Human and Profitable Monoclonal Antibody Market', *Trends in Biotechnology*. Elsevier Ltd, 37(1), pp. 9–16. doi: 10.1016/j.tibtech.2018.05.014.

Ha, T. K. and Lee, G. M. (2015) 'Glutamine substitution: the role it can play to enhance therapeutic protein production', *Pharmaceutical Bioprocessing*, 3(3), pp. 249–261. doi: 10.4155/pbp.15.6.

Häcker, G. (2000) 'The morphology of apoptosis', *Cell and Tissue Research*, 301(1), pp. 5–17. doi: 10.1007/s004410000193.

Hagrot, E. *et al.* (2017) 'Poly-pathway model, a novel approach to simulate multiple metabolic states by reaction network-based model – Application to amino acid depletion in CHO cell culture', *Journal of Biotechnology*, 259, pp. 235–247. doi: 10.1016/j.jbiotec.2017.05.026.

Hand, D. J. (2014) 'Wonderful Examples, but Let's not Close Our Eyes', *Statistical Science*, 29(1), pp. 98–100. doi: 10.1214/13-STS446.

Haringa, C. *et al.* (2018) 'Computational fluid dynamics simulation of an industrial P. chrysogenum fermentation with a coupled 9-pool metabolic model: Towards rational scale-down and design optimization', *Chemical Engineering Science*, 175, pp. 12–24. doi: 10.1016/j.ces.2017.09.020.

Hartley, F. *et al.* (2018) 'Mechanisms driving the lactate switch in Chinese hamster ovary cells', *Biotechnology and Bioengineering*, 115(8), pp. 1890–1903. doi: 10.1002/bit.26603.

Hefzi, H. *et al.* (2016) 'A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism', *Cell Systems*, 3(5), pp. 434-443.e8. doi: 10.1016/j.cels.2016.10.020.

Heiden, M. G. Vander, Cantley, L. C. and Thompson, C. B. (2009) 'Understanding the warburg effect: The metabolic requirements of cell proliferation', *Science*. doi: 10.1126/science.1160809.

Heinzle, E., Biwer, A. P. and Cooney, C. L. (2006) *Development of Sustainable*

Bioprocesses Modeling and Assessment. John Wiley & Sons, Inc.

Hendrick, V. *et al.* (2001) 'Increased productivity of recombinant tissular plasminogen activator (t-PA) by butyrate and shift of temperature: A cell cycle phases analysis', in *Cytotechnology*. doi: 10.1023/A:1014088919546.

Higel, F. *et al.* (2016) 'N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins', *European Journal of Pharmaceutics and Biopharmaceutics*. Elsevier B.V., 100, pp. 94–100. doi: 10.1016/j.ejpb.2016.01.005.

Ho, Y., Varley, J. and Mantalaris, A. (2006) 'Development and Analysis of a Mathematical Model for Antibody-Producing GS-NS0 Cells Under Normal and Hyperosmotic Culture Conditions', *Biotechnology Progress*, 22(6), pp. 1560–1569. doi: 10.1021/bp060032b.

Hong, J. (1989) 'Yield coefficients for cell mass and product formation', *Biotechnology and Bioengineering*, 33(4), pp. 506–507. doi: 10.1002/bit.260330416.

Hong, J. K., Nargund, S., *et al.* (2018) 'Comparative phenotypic analysis of CHO clones and culture media for lactate shift', *Journal of Biotechnology*, 283, pp. 97–104. doi: 10.1016/j.jbiotec.2018.07.042.

Hong, J. K., Lakshmanan, M., *et al.* (2018) 'Towards next generation CHO cell line development and engineering by systems approaches', *Current Opinion in Chemical Engineering*. Elsevier Ltd, 22, pp. 1–10. doi: 10.1016/j.coche.2018.08.002.

Hua, F. *et al.* (2005) 'Effects of Bcl-2 Levels on Fas Signaling-Induced Caspase-3 Activation: Molecular Genetic Tests of Computational Model Predictions', *The Journal of Immunology*, 175(2), pp. 985–995. doi: 10.4049/jimmunol.175.2.985.

Huber, H. J. *et al.* (2011) 'Mathematical modelling of the mitochondrial apoptosis pathway.', *Biochimica et biophysica acta*. Elsevier B.V., 1813(4), pp. 608–15. doi: 10.1016/j.bbamcr.2010.10.004.

Ichim, G. and Tait, S. W. G. (2016) 'A fate worse than death: apoptosis as an oncogenic process', *Nature Reviews Cancer*, 16(8), pp. 539–548. doi: 10.1038/nrc.2016.58.

Jacobi, A. *et al.* (2014) 'Process Development and Manufacturing of Therapeutic Antibodies', in Dubel, S. and Reichert, J. (eds) *Handbook of Therapeutic Antibodies: Second Edition*. 2nd edn. doi: 10.1002/9783527682423.ch22.

Jandt, U. *et al.* (2015) 'Synchronized mammalian cell culture: Part II-population ensemble modeling and analysis for development of reproducible processes', *Biotechnology Progress*, 31(1), pp. 175–185. doi: 10.1002/btpr.2006.

Jang, J. D. and Barford, J. P. (2000) 'An unstructured kinetic model of macromolecular metabolism in batch and fed-batch cultures of hybridoma cells producing monoclonal antibody', *Biochemical Engineering Journal*, 4(2), pp. 153–168. doi: 10.1016/S1369-703X(99)00041-8.

Jayme, D. (2007) 'Development and optomization of serum-free and protein-free media', in Davis, J. and Stacey, G. (eds) *Medicines from Animal Cell Culture*. Wiley, Chichester, pp. 29–44.

Jiang, Z. *et al.* (2012) 'Fed-batch cell culture process optimization: a rationally integrated approach', *BioProcess International*, 10(3), pp. 40–45. Available at: http://www.bioprocessintl.com/journal/2012/March/Fed-Batch-Cell-Culture-Process-Optimization-328055.

Kang, S. *et al.* (2014) 'Proteomics analysis of altered cellular metabolism induced by insufficient copper level', *Journal of Biotechnology*. doi: 10.1016/j.jbiotec.2014.08.001.

Kaplon, J., Dam, L. Van and Peeper, D. (2015) 'Two-way communication between the metabolic and cell cycle machineries : the molecular basis', *Cell Cycle*, 14(13), pp. 2022–2032. doi: 10.1080/15384101.2015.1044172.

Kaufmann, H. *et al.* (2001) 'Comparative analysis of two controlled proliferation strategies regarding product quality, influence on tetracycline-regulated gene expression, and productivity', *Biotechnology and Bioengineering*, 72(6), pp. 592–602. doi: 10.1002/1097-0290(20010320)72:6<592::AID-BIT1024>3.0.CO;2-J.

Ke, N., Godzik, A. and Reed, J. C. (2001) 'Bcl-B, a Novel Bcl-2 Family Member That Differentially Binds and Regulates Bax and Bak', *Journal of Biological Chemistry*, 276(16), pp. 12481–12484. doi: 10.1074/jbc.C000871200.

Khoo, S. H. G. and Al-Rubeai, M. (2007) 'Metabolomics as a complementary tool in cell culture', *Biotechnology and Applied Biochemistry*, 47(2), p. 71. doi: 10.1042/ba20060221.

Khoo, S. H. G. and Al-Rubeai, M. (2009) 'Metabolic characterization of a hyper-productive state in an antibody producing NS0 myeloma cell line', *Metabolic Engineering*, 11(3), pp. 199–211. doi: 10.1016/j.ymben.2009.02.001.

Kiparissides, A. *et al.* (2009) 'Global Sensitivity Analysis Challenges in Biological Systems Modeling', *Ind. Eng. Chemical Research*, 28(15), pp. 7168–7180. doi: 10.1021/ie900139x.

Kiparissides, A. *et al.* (2011) '"Closing the loop" in biological systems modeling-from the in silico to the in vitro', *Automatica*. Elsevier Ltd, 47(6), pp. 1147–1155. doi: 10.1016/j.automatica.2011.01.013.

Kiparissides, A., Pistikopoulos, E. N. and Mantalaris, A. (2010) *Towards energy-based dynamic optimization of monoclonal antibody producing GS-NS0 Cultures*, *Computer Aided Chemical Engineering*. Elsevier B.V. doi: 10.1016/S1570-7946(10)28099-9.

Kiparissides, A., Pistikopoulos, E. N. and Mantalaris, A. (2015) 'On the model-based optimization of secreting mammalian cell (GS-NS0) cultures', *Biotechnology and Bioengineering*, 112(3), pp. 536–548. doi: 10.1002/bit.25457.

Klontzas, M. E. *et al.* (2017) 'Metabolomics Analysis of the Osteogenic Differentiation of Umbilical Cord Blood Mesenchymal Stem Cells Reveals Differential Sensitivity to

Osteogenic Agents', *Stem Cells and Development*, 26(10), pp. 723–733. doi: 10.1089/scd.2016.0315.

Kober, L., Zehe, C. and Bode, J. (2012) 'Development of a novel ER stress based selection system for the isolation of highly productive clones', *Biotech. Bioeng.*, 109(10), pp. 2599–2611. doi: 10.1002/bit.24527.

Kontoravdi, C., Asprey, S. P., *et al.* (2007) 'Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring', *Computers & Chemical Engineering*, 31(5–6), pp. 392–400. doi: 10.1016/j.compchemeng.2006.04.009.

Kontoravdi, C., Wong, D., *et al.* (2007) 'Modeling Amino Acid Metabolism in Mammalian Cells-Toward the Development of a Model Library', *Biotechnology Progress*, 23(6), pp. 1261–1269. doi: 10.1021/bp070106z.

Kontoravdi, C. *et al.* (2008) 'Application of Global Sensitivity Analysis to Determine Goals for Design of Experiments: An Example Study on Antibody-Producing Cell Cultures', *Biotechnology Progress*, 21(4), pp. 1128–1135. doi: 10.1021/bp050028k.

Kontoravdi, C. and Jimenez del Val, I. (2018) 'Computational tools for predicting and controlling the glycosylation of biopharmaceuticals', *Current Opinion in Chemical Engineering*. doi: 10.1016/j.coche.2018.08.007.

Kontoravdi, C., Pistikopoulos, E. N. and Mantalaris, A. (2010) 'Systematic development of predictive mathematical models for animal cell cultures', *Computers & Chemical Engineering*. Elsevier Ltd, 34(8), pp. 1192–1198. doi: 10.1016/j.compchemeng.2010.03.012.

Kunert, R. and Reinhart, D. (2016) 'Advances in recombinant antibody manufacturing', *Applied Microbiology and Biotechnology*, 100(8), pp. 3451–3461. doi: 10.1007/s00253-016-7388-9.

Kung, P. *et al.* (1979) 'Monoclonal Antibodies Defining Distinctive Human T Cell Surface Antigens', *Science*, 206(4416), pp. 347–349. Available at: http://ukpmc.ac.uk/abstract/MED/314668 (Accessed: 10 February 2013).

Kurano, N. *et al.* (1990) 'Growth behavior of Chinese hamster ovary cells in a compact loop bioreactor. 2. Effects of medium components and waste products', *Journal of Biotechnology*. doi: 10.1016/0168-1656(90)90055-G.

Kyparissidis, A.-D. (2012) *Development of a combined mathematical and experimental framework for the control and optimisation of mammalian cell culture systems*. Imperial College London.

Kyriakopoulos, S., Ang, K. S., Lakshmanan, M., Huang, Z., Yoon, S., Gunawan, R. and Lee., D.-Y. (2018) 'Kinetic Modeling of Mammalian Cell Culture Bioprocessing: the quest to advance biomanufacturing', *Biotechnology Journal*, 13(3).

Kyriakopoulos, S., Ang, K. S., Lakshmanan, M., Huang, Z., Yoon, S., Gunawan, R. and

Lee, D.-Y. (2018) 'Kinetic Modeling of Mammalian Cell Culture Bioprocessing: The Quest to Advance Biomanufacturing', *Biotechnology Journal*, 13(3), p. 1700229. doi: 10.1002/biot.201700229.

Kyriakopoulos, S. and Kontoravdi, C. (2014) 'A framework for the systematic design of fed-batch strategies in mammalian cell culture', *Biotechnology and Bioengineering*, 111(12), pp. 2466–2476. doi: 10.1002/bit.25319.

Kyriakopoulos, S., Polizzi, K. M. and Kontoravdi, C. (2013) 'Comparative analysis of amino acid metabolism and transport in CHO variants with different levels of productivity', *Journal of Biotechnology*, 168(4), pp. 543–551. doi: 10.1016/j.jbiotec.2013.09.007.

Lalonde, M. E. and Durocher, Y. (2017) 'Therapeutic glycoprotein production in mammalian cells', *Journal of Biotechnology*. doi: 10.1016/j.jbiotec.2017.04.028.

Lamb, J. and Wheatley, D. N. (2000) 'Single amino acid (arginine) deprivation induces G1 arrest associated with inhibition of cdk4 expression in cultured human diploid fibroblasts.', *Experimental cell research*, 255(2), pp. 238–49. doi: 10.1006/excr.1999.4779.

Lao, M. S. and Toth, D. (1997) 'Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture', *Biotechnology Progress*. doi: 10.1021/bp9602360.

Le, H. *et al.* (2012) 'Multivariate analysis of cell culture bioprocess data—Lactate consumption as process indicator', *Journal of Biotechnology*, 162, pp. 210–223. doi: 10.1016/j.jbiotec.2012.08.021.

Lee, I. H. and Finkel, T. (2013) 'Metabolic regulation of the cell cycle', *Current Opinion in Cell Biology*, 25(6), pp. 724–729. doi: 10.1016/j.ceb.2013.07.002.

Lee, S. K. and Lee, G. M. (2003) 'Development of apoptosis-resistant dihydrofolate reductase-deficient Chinese hamster ovary cell line', *Biotechnology and Bioengineering*, 82(7), pp. 872–876. doi: 10.1002/bit.10633.

Lencastre Fernandes, R. *et al.* (2011) 'Experimental methods and modeling techniques for description of cell population heterogeneity', *Biotechnology Advances*. Elsevier Inc., 29(6), pp. 575–599. doi: 10.1016/j.biotechadv.2011.03.007.

Ley, K. D. (1973) 'Synchronization of mitochondrial DNA synthesis in chinese hamster cells (line CHO) deprived of isoleucine', *The Journal of Cell Biology*, 58(2), pp. 340–345. doi: 10.1083/jcb.58.2.340.

Li, F. *et al.* (2010) 'Cell culture processes for monoclonal antibody production', *mAbs*, 2(5), pp. 466–479. doi: 10.4161/mabs.2.5.12720.

Li, J. *et al.* (2012) 'Feeding lactate for CHO cell culture processes: Impact on culture metabolism and performance', *Biotechnology and Bioengineering*, 109(5), pp. 1173–1186. doi: 10.1002/bit.24389.

Li, J. *et al.* (2013) 'Deficiency of b-arrestin1 ameliorates collagen-induced arthritis with impaired TH17 cell differentiation', *Proceedings of the National Academy of Sciences*, 110(18), pp. 7395–7400. doi: 10.1073/pnas.1221608110.

Li, Y. *et al.* (2016) 'MiR-182-5p protects inner ear hair cells from cisplatin-induced apoptosis by inhibiting FOXO3a', *Cell Death & Disease*, 7(9), pp. e2362–e2362. doi: 10.1038/cddis.2016.246.

Lim, S. F. *et al.* (2006) 'RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells', *Metabolic Engineering*, 8(6), pp. 509–522. doi: 10.1016/j.ymben.2006.05.005.

Ling, W. L. W. *et al.* (2003) 'Improvement of Monoclonal Antibody Production in Hybridoma Cells by Dimethyl Sulfoxide', *Biotechnology Progress*, 19(1), pp. 158–162. doi: 10.1021/bp020068d.

Liu, Y.-H. *et al.* (2007) 'A population balance model describing the cell cycle dynamics of myeloma cell cultivation.', *Biotechnology progress*, 23(5), pp. 1198–1209. doi: 10.1021/bp070152z.

Lu, F. *et al.* (2013) 'Automated dynamic fed-batch process and media optimization for high productivity cell culture process development', *Biotechnology and Bioengineering*, 110(1), pp. 191–205. doi: 10.1002/bit.24602.

Lynch, D. H. and Yang, X.-D. (2002) 'Therapeutic potential of ABX-EGF: A fully human anti-epidermal growth factor receptor monoclonal antibody for cancer treatment', *Semin. Oncol.*, 29(1), pp. 47–50. doi: 10.1053/sonc.2002.31522.

Makimoto, Y., Takahashi, E. and Takasugi, H. (2013) 'Relationship Between Cell Cycle Phases and Monoclonal Antibody Production in Microcarrier Perfusion Culture of Recombinant CHO Cells', in *Animal Cell Technology: Basic & Applied Aspects*. doi: 10.1007/978-94-017-0728-2_19.

Mantzaris, N. V., Daoutidis, P. and Srienc, F. (2001) 'Numerical solution of multi-variable cell population balance models: I. Finite difference methods', *Computers & Chemical Engineering*, 25(11–12), pp. 1411–1440. doi: 10.1016/S0098-1354(01)00709-8.

Marquardt, W. (2005) 'Model-based experimental analysis of kinetic phenomena in multi-phase reactive systems', *Chemical Engineering Research and Design*. The Institution of Chemical Engineers, 83(6 A), pp. 561–573. doi: 10.1205/cherd.05086.

Martens, D. E. *et al.* (1995) 'A combined cell-cycle and metabolic model for the growth of hybridoma cells in steady-state continuous culture', *Biotechnology and Bioengineering*, 48(1), pp. 49–65. doi: 10.1002/bit.260480109.

Martínez, V. S. *et al.* (2013) 'Flux balance analysis of CHO cells before and after a metabolic switch from lactate production to consumption', *Biotechnology and Bioengineering*. doi: 10.1002/bit.24728.

Matuszczyk, J.-C. *et al.* (2015) 'Compartment-specific metabolomics for CHO reveals

that ATP pools in mitochondria are much lower than in cytosol', *Biotechnology Journal*, 10(10), pp. 1639–1650. doi: 10.1002/biot.201500060.

McAtee Pereira, A. G. *et al.* (2018) '13C Flux Analysis Reveals that Rebalancing Medium Amino Acid Composition can Reduce Ammonia Production while Preserving Central Carbon Metabolism of CHO Cell Cultures', *Biotechnology Journal*. doi: 10.1002/biot.201700518.

McIlwain, D. R., Berger, T. and Mak, T. W. (2015) 'Caspase Functions in Cell Death and Disease', *Cold Spring Harbor Perspectives in Biology*, 7(4), p. a026716. doi: 10.1101/cshperspect.a026716.

Mentch, S. J. and Locasale, J. W. (2016) 'One-carbon metabolism and epigenetics: understanding the specificity', *Annals of the New York Academy of Sciences*, 1363(1), pp. 91–98. doi: 10.1111/nyas.12956.

Meshram, M. *et al.* (2012) 'Population-based modeling of the progression of apoptosis in mammalian cell culture', *Biotechnology and Bioengineering*, 109(5), pp. 1193–1204. doi: 10.1002/bit.24392.

Mulukutla, B. C., Gramer, M. and Hu, W. S. (2012) 'On metabolic shift to lactate consumption in fed-batch culture of mammalian cells', *Metabolic Engineering*. Elsevier, 14(2), pp. 138–149. doi: 10.1016/j.ymben.2011.12.006.

Murphy, T. A. and Young, J. D. (2013) 'ETA: Robust software for determination of cell specific rates from extracellular time courses', *Biotechnology and Bioengineering*, 110(6), pp. 1748–1758. doi: 10.1002/bit.24836.

Murtagh, F. and Contreras, P. (2012) 'Algorithms for hierarchical clustering: An overview', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1), pp. 86–97. doi: 10.1002/widm.53.

Naderi, S. *et al.* (2011) 'Development of a mathematical model for evaluating the dynamics of normal and apoptotic Chinese hamster ovary cells', *Biotechnology Progress*, 27(5), pp. 1197–1205. doi: 10.1002/btpr.647.

Nagata, S. *et al.* (2003) 'Degradation of chromosomal DNA during apoptosis.', *Cell death and differentiation*, 10(1), pp. 108–16. doi: 10.1038/sj.cdd.4401161.

Nelson, D. L. and Cox, M. M. (2013) *Lehninger Principles of Biochemistry*. 6th edn, *Book*. 6th edn.

Nicolae, A. *et al.* (2014) 'Non-stationary 13C metabolic flux analysis of Chinese hamster ovary cells in batch culture using extracellular labeling highlights metabolic reversibility and compartmentation', *BMC Systems Biology*, 8(1), p. 50. doi: 10.1186/1752-0509-8-50.

Noh, S. M., Shin, S. and Lee, G. M. (2018) 'Comprehensive characterization of glutamine synthetase-mediated selection for the establishment of recombinant CHO cells producing monoclonal antibodies', *Scientific Reports*, 8(1), p. 5361. doi: 10.1038/s41598-018-23720-9.

Okazaki, N. *et al.* (2008) 'Simple computational models of type I/type II cells in Fas signaling-induced apoptosis', *Journal of Theoretical Biology*, 250(4), pp. 621–633. doi: 10.1016/j.jtbi.2007.10.030.

Orrenius, S., Zhivotovsky, B. and Nicotera, P. (2003) 'Calcium: Regulation of cell death: the calcium–apoptosis link', *Nature Reviews Molecular Cell Biology*, 4(7), pp. 552–565. doi: 10.1038/nrm1150.

Ozturk, S. S. and Palsson, B. O. (1990) 'Chemical Decomposition of Glutamine in Cell Culture Media: Effect of Media Type, pH, and Serum Concentration', *Biotechnology Progress*, 6(2), pp. 121–128. doi: 10.1021/bp00002a005.

Pace, V. *et al.* (2010) 'Experimental testing of a mathematical model relevant to the extrinsic pathway of apoptosis.', *Cell stress & chaperones*, 15(1), pp. 13–23. doi: 10.1007/s12192-009-0118-9.

Pan, X. *et al.* (2017) 'Metabolic characterization of a CHO cell size increase phase in fed-batch cultures', *Applied Microbiology and Biotechnology*, 101(22), pp. 8101–8113. doi: 10.1007/s00253-017-8531-y.

Park, J. H. *et al.* (2015) 'Valeric acid induces cell cycle arrest at G1 phase in CHO cell cultures and improves recombinant antibody productivity', *Biotechnology Journal*, pp. 1–10. doi: 10.1002/biot.201500327.

Park, J. H. *et al.* (2017) 'Proteomic Analysis of Host Cell Protein Dynamics in the Culture Supernatants of Antibody-Producing CHO Cells', *Scientific Reports*, 7(1), p. 44246. doi: 10.1038/srep44246.

Pereira, S., Kildegaard, H. F. and Andersen, M. R. (2018) 'Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients', *Biotechnology Journal*. doi: 10.1002/biot.201700499.

Pioch, D. *et al.* (2007) 'At-line Monitoring of Bioprocess-Relevant Marker Genes', *Engineering in Life Sciences*, 7(4), pp. 373–379. doi: 10.1002/elsc.200720197.

Pollock, J., Ho, S. V. and Farid, S. S. (2013) 'Fed-batch and perfusion culture processes: Economic, environmental, and operational feasibility under uncertainty', *Biotech. Bioeng.*, 110(1), pp. 206–219. doi: 10.1002/bit.24608.

Popp, O. *et al.* (2016) 'A hybrid approach identifies metabolic signatures of high-producers for chinese hamster ovary clone selection and process optimization', *Biotechnology and Bioengineering*, 113(9), pp. 2005–2019. doi: 10.1002/bit.25958.

Potgieter, T. I. *et al.* (2009) 'Production of monoclonal antibodies by glycoengineered Pichia pastoris', *Journal of Biotechnology*, 139(4), pp. 318–325. doi: 10.1016/j.jbiotec.2008.12.015.

Quiroga-Campano, A. *et al.* (2016) 'A predictive model for energy metabolism and ATP balance in mammalian cells: towards the energy-based optimization of mAB production', in Kravanja, Z. and Bogataj, M. (eds) *Proceedings of the 26th European Symposium on*

Computer Aided Process Engineering. Elsevier B.V., pp. 1581–1586. doi: http://dx.doi.org/10.1016/B978-0-444-63428-3.50268-X.

Quiroga-Campano, A. L. (2016) *'Mathematical modelling and experimental validation for the optimization and control of mammalian cell culture systems'*. Imperial College London.

Quiroga-Campano, A. L., Panoskaltsis, N. and Mantalaris, A. (2018) 'Energy-based culture medium design for biomanufacturing optimization: A case study in monoclonal antibody production by GS-NS0 cells', *Metabolic Engineering*, 47, pp. 21–30. doi: 10.1016/j.ymben.2018.02.013.

Reinhart, D. *et al.* (2015) 'Benchmarking of commercially available CHO cell culture media for antibody production', *Applied Microbiology and Biotechnology*, 99(11), pp. 4645–4657. doi: 10.1007/s00253-015-6514-4.

Reinhart, D. *et al.* (2018) 'Differential gene expression of a feed-spiked super-producing CHO cell line', *Journal of Biotechnology*, 285, pp. 23–37. doi: 10.1016/j.jbiotec.2018.08.013.

Renner, W. A. *et al.* (1995) 'Recombinant cyclin E expression activates proliferation and obviates surface attachment of chinese hamster ovary (CHO) cells in protein-free medium', *Biotechnology and Bioengineering*, 47(4), pp. 476–482. doi: 10.1002/bit.260470409.

Reynolds, J. E. *et al.* (1996) 'BCL-2 and MCL-1 expression in Chinese hamster ovary cells inhibits intracellular acidification and apoptosis induced by staurosporine.', *Experimental cell research*, 225(2), pp. 430–6. doi: 10.1006/excr.1996.0194.

Rigual-González, Y. *et al.* (2016) 'Application of a new model based on oxygen balance to determine the oxygen uptake rate in mammalian cell chemostat cultures', *Chemical Engineering Science*. Elsevier, 152, pp. 586–590. doi: 10.1016/j.ces.2016.06.051.

Robitaille, J., Chen, J. and Jolicoeur, M. (2015) 'A single dynamic metabolic model can describe mAb producing CHO cell batch and fed-batch cultures on different culture media', *PLoS ONE*, 10(9). doi: 10.1371/journal.pone.0136815.

Rogers, A. and Ierapetritou, M. (2015) 'Challenges and opportunities in modeling pharmaceutical manufacturing processes', *Computers and Chemical Engineering*. Elsevier Ltd, 81, pp. 32–39. doi: 10.1016/j.compchemeng.2015.03.018.

Rogers, R. S. *et al.* (2015) 'Development of a quantitative mass spectrometry multi-attribute method for characterization, quality control testing and disposition of biologics', *mAbs*, 7(5), pp. 881–890. doi: 10.1080/19420862.2015.1069454.

Salazar, A., Keusgen, M. and Von Hagen, J. (2016) 'Amino acids in the cultivation of mammalian cells', *Amino Acids*. Springer Vienna, 48(5), pp. 1161–1171. doi: 10.1007/s00726-016-2181-8.

Sanderson, C. S. (1997) *The Development and Application of a Structured Model for*

Animal Cell Metabolism. University of Sidney, Australia.

Schneider, M., Marison, I. W. and Von Stockar, U. (1996) 'The importance of ammonia in mammalian cell culture', *Journal of Biotechnology*. doi: 10.1016/0168-1656(95)00196-4.

Scholzen, T. and Gerdes, J. (2000) 'The Ki-67 protein: From the known and the unknown', *Journal of Cellular Physiology*, 182(3), pp. 311–322. doi: 10.1002/(SICI)1097-4652(200003)182:3<311::AID-JCP1>3.0.CO;2-9.

Schroeder, H. W. and Cavacini, L. (2010) 'Structure and function of immunoglobulins.', *J. Allergy Clin. Immun.* Elsevier Ltd, 125(2 Suppl 2), pp. S41-52. doi: 10.1016/j.jaci.2009.09.046.

Selvarasu, S. *et al.* (2012) 'Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture', *Biotechnology and Bioengineering*. doi: 10.1002/bit.24445.

Sengupta, N., Rose, S. T. and Morgan, J. A. (2011) 'Metabolic flux analysis of CHO cell metabolism in the late non-growth phase', *Biotechnology and Bioengineering*. doi: 10.1002/bit.22890.

Seo, J. S. *et al.* (2013) 'Effect of culture pH on recombinant antibody production by a new human cell line, F2N78, grown in suspension at 33.0 C and 37.0 C', *Applied Microbiology and Biotechnology*, 97(12), pp. 5283–5291. doi: 10.1007/s00253-013-4849-2.

Shi, M. *et al.* (2005) 'Controlled Growth of Chinese Hamster Ovary Cells and High Expression of Antibody-IL-2 Fusion Proteins by Temperature Manipulation', *Biotechnology Letters*, 27(23–24), pp. 1879–1884. doi: 10.1007/s10529-005-3897-9.

Sidoli, F. R., Asprey, S. P. and Mantalaris, A. (2006) 'A coupled single cell-population-balance model for mammalian cell cultures', *Industrial and Engineering Chemistry Research*, 45(16), pp. 5801–5811. doi: 10.1021/ie0511581.

Sidoli, F. R., Mantalaris, A. and Asprey, S. P. (2005) 'Toward global parametric estimability of a large-scale kinetic single-cell model for mammalian cell cultures', *Industrial and Engineering Chemistry Research*, 44(4), pp. 868–878. doi: 10.1021/ie0401556.

Sieck, J. B. *et al.* (2013) 'Development of a Scale-Down Model of hydrodynamic stress to study the performance of an industrial CHO cell line under simulated production scale bioreactor conditions', *Journal of Biotechnology*. Elsevier B.V., 164(1), pp. 41–49. doi: 10.1016/j.jbiotec.2012.11.012.

Singh, R. P. *et al.* (1994) 'Cell death in bioreactors: A role for apoptosis', *Biotechnology and Bioengineering*, 44(6), pp. 720–726. doi: 10.1002/bit.260440608.

Sitton, G. and Srienc, F. (2008) 'Growth dynamics of mammalian cells monitored with automated cell cycle staining and flow cytometry', *Cytometry Part A*, 73A(6), pp. 538–545. doi: 10.1002/cyto.a.20567.

Siu, K. T., Rosner, M. R. and Minella, A. C. (2012) 'An integrated view of cyclin E function and regulation', *Cell Cycle*, 11(1), pp. 57–64. doi: 10.4161/cc.11.1.18775.

Sobol', I. M. *et al.* (2011) 'Construction and Comparison of High-Dimensional Sobol' Generators', *Wilmott*, 2011(56), pp. 64–79. doi: 10.1002/wilm.10056.

Sobol, I. M. (2001) 'Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates', *Mathematics and Computers in Simulation*. doi: 10.1016/S0378-4754(00)00270-6.

Sokolov, M. *et al.* (2017) 'Enhanced process understanding and multivariate prediction of the relationship between cell culture process and monoclonal antibody quality', *Biotechnology Progress*, 33(5), pp. 1368–1380. doi: 10.1002/btpr.2502.

Sonderhoff, S. A., Kilburn, D. G. and Piret, J. M. (1992) 'Analysis of mammalian viable cell biomass based on cellular ATP', *Biotechnology and Bioengineering*, 39(8), pp. 859–864. doi: 10.1002/bit.260390807.

Spencer, S. L. and Sorger, P. K. (2011) 'Measuring and modeling apoptosis in single cells', *Cell*. Elsevier Inc., 144(6), pp. 926–939. doi: 10.1016/j.cell.2011.03.002.

Steinebach, F. *et al.* (2016) 'Model based adaptive control of a continuous capture process for monoclonal antibodies production', *Journal of Chromatography A*. Elsevier B.V., 1444, pp. 50–56. doi: 10.1016/j.chroma.2016.03.014.

Stucki, J. W. and Simon, H.-U. (2005) 'Mathematical modeling of the regulation of caspase-3 activation and degradation.', *Journal of theoretical biology*, 234(1), pp. 123–31. doi: 10.1016/j.jtbi.2004.11.011.

Sureshkumar, G. K. and Mutharasan, R. (1991) 'The Influence of Temperature on a Mouse-Mouse Hybridoma Growth and Monoclonal Antibody Production', *Biotechnology and Bioengineering*, 37, pp. 292–295.

Suzuki, E. . and Ollis, D. F. (1989) 'Cell cycle model for antibody production kinetics', *Biotechnology and Bioengineering*, 34(11), pp. 1398–1402. doi: 10.1002/bit.260341109.

Swiderek, H. and Al-Rubeai, M. (2007) 'Functional genome-wide analysis of antibody producing NS0 cell line cultivated at different temperatures.', *Biotechnology and bioengineering*, 98(3), pp. 616–30. doi: 10.1002/bit.21445.

Talemi, S. R. *et al.* (2015) 'Development of a robust DNA damage model including persistent telomere-associated damage with application to secondary cancer risk assessment.', *Scientific reports*. Nature Publishing Group, 5(July 2015), p. 13540. doi: 10.1038/srep13540.

Tang, H. L. *et al.* (2009) 'Reversibility of apoptosis in cancer cells.', *British journal of cancer*, 100(1), pp. 118–122. doi: 10.1038/sj.bjc.6604802.

Tang, H. L. *et al.* (2012) 'Cell survival , DNA damage , and oncogenic transformation after a transient and reversible apoptotic response', *Molecular Biology of the Cell*, 23.

doi: 10.1091/mbc.E11-11-0926.

Taylor, R. C., Cullen, S. P. and Martin, S. J. (2008) 'Apoptosis: controlled demolition at the cellular level.', *Nature reviews. Molecular cell biology*, 9(3), pp. 231–241. doi: 10.1038/nrm2312.

Templeton, N. *et al.* (2013) 'Peak antibody production is associated with increased oxidative metabolism in an industrially relevant fed-batch CHO cell culture', *Biotechnology and Bioengineering*, 110(7), pp. 2013–2024. doi: 10.1002/bit.24858.

Templeton, N. *et al.* (2014) 'The impact of anti-apoptotic gene Bcl-2?? expression on CHO central metabolism', *Metabolic Engineering*. Elsevier, 25, pp. 92–102. doi: 10.1016/j.ymben.2014.06.010.

Templeton, N. *et al.* (2017) '13C metabolic flux analysis identifies limitations to increasing specific productivity in fed-batch and perfusion', *Metabolic Engineering*. doi: 10.1016/j.ymben.2017.09.010.

Tey, B. T. and Al-Rubeai, M. (2004) 'Suppression of apoptosis in perfusion culture of Myeloma NS0 cells enhances cell growth but reduces antibody productivity', *Apoptosis*, 9(6), pp. 843–852. doi: 10.1023/B:APPT.0000045792.63249.5a.

Tey, B. T. and Al-Rubeai, M. (2005) 'Effect of Bcl-2 overexpression on cell cycle and antibody productivity in chemostat cultures of myeloma NS0 cells.', *Journal of bioscience and bioengineering*, 100(3), pp. 303–10. doi: 10.1263/jbb.100.303.

Tyson, J. J. *et al.* (2011) 'Dynamic modelling of oestrogen signalling and cell fate in breast cancer cells.', *Nature reviews. Cancer*. Nature Publishing Group, 11(7), pp. 523–532. doi: 10.1038/nrc3081.

Usaku, C. (2015) *An integrated experimental and modelling approach for the study of apoptosis in GS-NS0 cell cultures*. Imperial College London.

Vakifahmetoglu-Norberg, H. and Zhivotovsky, B. (2010) 'The unpredictable caspase-2 : what can it do ?', *Trends in Cell Biology*. Elsevier Ltd, 20(3), pp. 150–159. doi: 10.1016/j.tcb.2009.12.006.

Del Val, I. J., Polizzi, K. M. and Kontoravdi, C. (2016) 'A theoretical estimate for nucleotide sugar demand towards Chinese Hamster Ovary cellular glycosylation', *Scientific Reports*. Nature Publishing Group, 6(January), pp. 1–15. doi: 10.1038/srep28547.

Vernardis, S. I. *et al.* (2017) 'Human embryonic and induced pluripotent stem cells maintain phenotype but alter their metabolism after exposure to ROCK inhibitor', *Scientific Reports*, 7(1), p. 42138. doi: 10.1038/srep42138.

Vernardis, S. I., Goudar, C. T. and Klapa, M. I. (2013) 'Metabolic profiling reveals that time related physiological changes in mammalian cell perfusion cultures are bioreactor scale independent', *Metabolic Engineering*. Elsevier, 19, pp. 1–9. doi: 10.1016/j.ymben.2013.04.005.

Wahrheit, J., Niklas, J. and Heinzle, E. (2014) 'Metabolic control at the cytosol–mitochondria interface in different growth phases of CHO cells', *Metabolic Engineering*, 23, pp. 9–21. doi: 10.1016/j.ymben.2014.02.001.

Waldmann, T. a (2003) 'Immunotherapy: past, present and future', *Nat. Med*, 9(3), pp. 269–77. doi: 10.1038/nm0303-269.

Wamelink, M. M. C., Struys, E. A. and Jakobs, C. (2008) 'The biochemistry, metabolism and inherited defects of the pentose phosphate pathway: A review', *Journal of Inherited Metabolic Disease*, 31(6), pp. 703–717. doi: 10.1007/s10545-008-1015-6.

Wang, W. *et al.* (2017) 'Suppressing P16 Ink4a and P14 ARF pathways overcomes apoptosis in individualized human embryonic stem cells', *The FASEB Journal*, 31(3), pp. 1130–1140. doi: 10.1096/fj.201600782R.

Wei, Y.-Y. C. *et al.* (2011) 'Proteomics analysis of chinese hamster ovary cells undergoing apoptosis during prolonged cultivation', *Cytotechnology*, 63(6), pp. 663–677. doi: 10.1007/s10616-011-9385-2.

Weinlich, R. *et al.* (2016) 'Necroptosis in development, inflammation and disease', *Nature Publishing Group*. Nature Publishing Group, 18(2), pp. 127–136. doi: 10.1038/nrm.2016.149.

Wong, D. C. F. *et al.* (2006) 'Targeting early apoptotic genes in batch and fed-batch CHO cell cultures', *Biotechnology and Bioengineering*, 95(3), pp. 350–361. doi: 10.1002/bit.20871.

Wu, M. *et al.* (2004) 'The effect of hyperosmotic pressure on antibody production and gene expression in the GS-NS0 cell line', 46, pp. 41–46.

Wurm, F. M. (2004) 'Production of recombinant protein therapeutics in cultivated mammalian cells', *Nat Biotechnol*, 22(11), pp. 1393–1398. doi: 10.1038/nbt1026.

Xiao, F. *et al.* (2016) 'Leucine deprivation inhibits proliferation and induces apoptosis of human breast cancer cells via fatty acid synthase', *Oncotarget*, 7(39). doi: 10.18632/oncotarget.11626.

Xie, L. and Wang, D. I. (1996) 'Material balance studies on animal cell metabolism using a stoichiometrically based reaction network.', *Biotechnology and bioengineering*, 52(5), pp. 579–90. doi: 10.1002/(SICI)1097-0290(19961205)52:5<579::AID-BIT5>3.0.CO;2-G.

Xie, L. and Wang, D. I. C. (1994) 'Applications of improved stoichiometric model in medium design and fed-batch cultivation of animal cells in bioreactor', *Cytotechnology*, 15(1–3), pp. 17–29. doi: 10.1007/BF00762376.

Yang, O., Qadan, M. and Ierapetritou, M. (2019) 'Economic Analysis of Batch and Continuous Biopharmaceutical Antibody Production: a Review', *Journal of Pharmaceutical Innovation*. doi: 10.1007/s12247-018-09370-4.

Yuan, H.-X., Xiong, Y. and Guan, K.-L. (2013) 'Nutrient Sensing, Metabolism, and Cell

Growth Control', *Molecular Cell*, 49(3), pp. 379–387. doi: 10.1016/j.molcel.2013.01.019.

Yun, C. Y. *et al.* (2007a) 'Specific inhibition of caspase-8 and -9 in CHO cells enhances cell viability in batch and fed-batch cultures', *Metabolic Engineering*, 9(5–6), pp. 406–418. doi: 10.1016/j.ymben.2007.06.001.

Yun, C. Y. *et al.* (2007b) 'Specific inhibition of caspase-8 and -9 in CHO cells enhances cell viability in batch and fed-batch cultures', *Metabolic Engineering*. doi: 10.1016/j.ymben.2007.06.001.

Zagari, F. *et al.* (2013) 'Lactate metabolism shift in CHO cell culture: The role of mitochondrial oxidative activity', *New Biotechnology*, 30(2), pp. 238–245. doi: 10.1016/j.nbt.2012.05.021.

Zha, H. *et al.* (1996) 'Proapoptotic protein Bax heterodimerizes with Bcl-2 and homodimerizes with Bax via a novel domain (BH3) distinct from BH1 and BH2', *Journal of Biological Chemistry*, 271(13), pp. 7440–7444. doi: 10.1074/jbc.271.13.7440.

Zhang, F. *et al.* (2006) 'Metabolic characteristics of recombinant Chinese hamster ovary cells expressing glutamine synthetase in presence and absence of glutamine', *Cytotechnology*, 51(1), pp. 21–28. doi: 10.1007/s10616-006-9010-y.

Zhang, J. (2010) 'Mammalian cell culture', in Baltz, R. H., Davies, J. E., and Demian, A. L. (eds) *Manual of Industrial Biotechnology*. Washington, D.C.: ASM Press, pp. 157–178.

Zhang, T., Brazhnik, P. and Tyson, J. J. (2009) 'Computational analysis of dynamical responses to the intrinsic pathway of programmed cell death', *Biophysical Journal*. Biophysical Society, 97(2), pp. 415–434. doi: 10.1016/j.bpj.2009.04.053.

Zhou, M. *et al.* (2011) 'Decreasing lactate level and increasing antibody production in Chinese Hamster Ovary cells (CHO) by reducing the expression of lactate dehydrogenase and pyruvate dehydrogenase kinases', *Journal of Biotechnology*, 153(1–2), pp. 27–34. doi: 10.1016/j.jbiotec.2011.03.003.

Ziehn, T. (2008) *GUI-HDMR User Documentation*.

Ziehn, T. and Tomlin, A. S. (2009) 'GUI–HDMR – A software tool for global sensitivity analysis of complex models', *Environmental Modelling & Software*, 24(7), pp. 775–785. doi: 10.1016/j.envsoft.2008.12.002.

Zou, W. and Al-Rubeai, M. (2016) 'Understanding central carbon metabolism of rapidly proliferating mammalian cells based on analysis of key enzymatic activities in GS-CHO cell lines', *Biotechnology and Applied Biochemistry*, 63(5), pp. 642–651. doi: 10.1002/bab.1409.

Appendices

Appendix I

Global Sensitivity Analysis of the mathematical model developed in Chapter 5.

The GSA results of the model presented in Chapter 5 have been divided into three figures. Figure I presents the results related to viable, apoptotic and dead cells, cell cycle distribution, ATP and mAb; Figure II shows the results related to metabolism of amino acids and glucose and the results related to gene expression are in Figure III. As mentioned in Chapter 5, GSA was conducting using the Sobol method and sensitivity indices were calculated at three biologically relevant time points – 48h (end of lag-phase), 72h (mid-exponential phase) and 120h (decline phase) – and a threshold of 0.1 has been set to discriminate between significant and non-significant parameters.

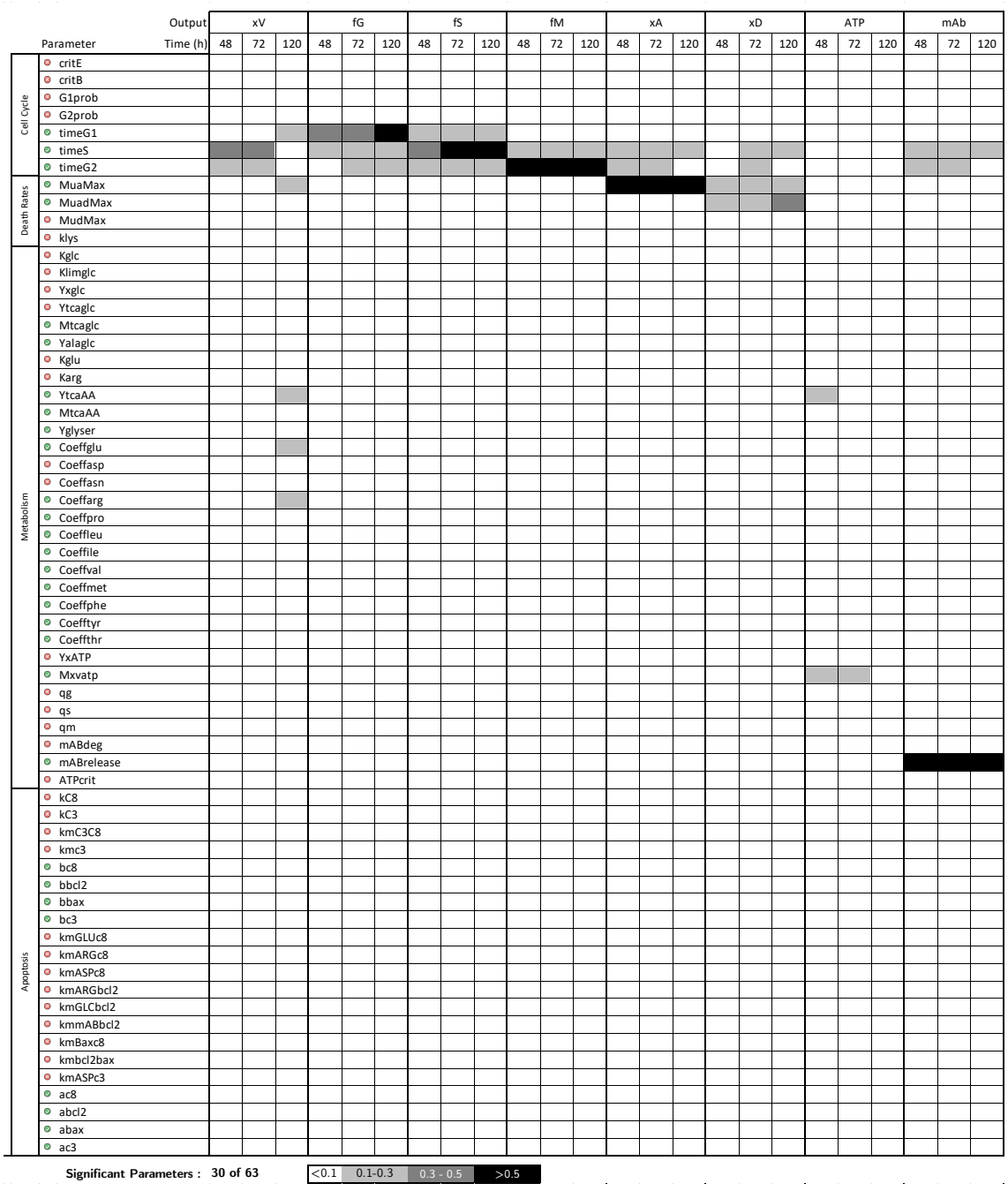


Figure I. GSA results for viable cells (xV), fraction of cells in phase G1 (fG), phase S (fS) and phase G2/M (fM), apoptotic cells (xA), dead cells (xD), ATP and mAb.

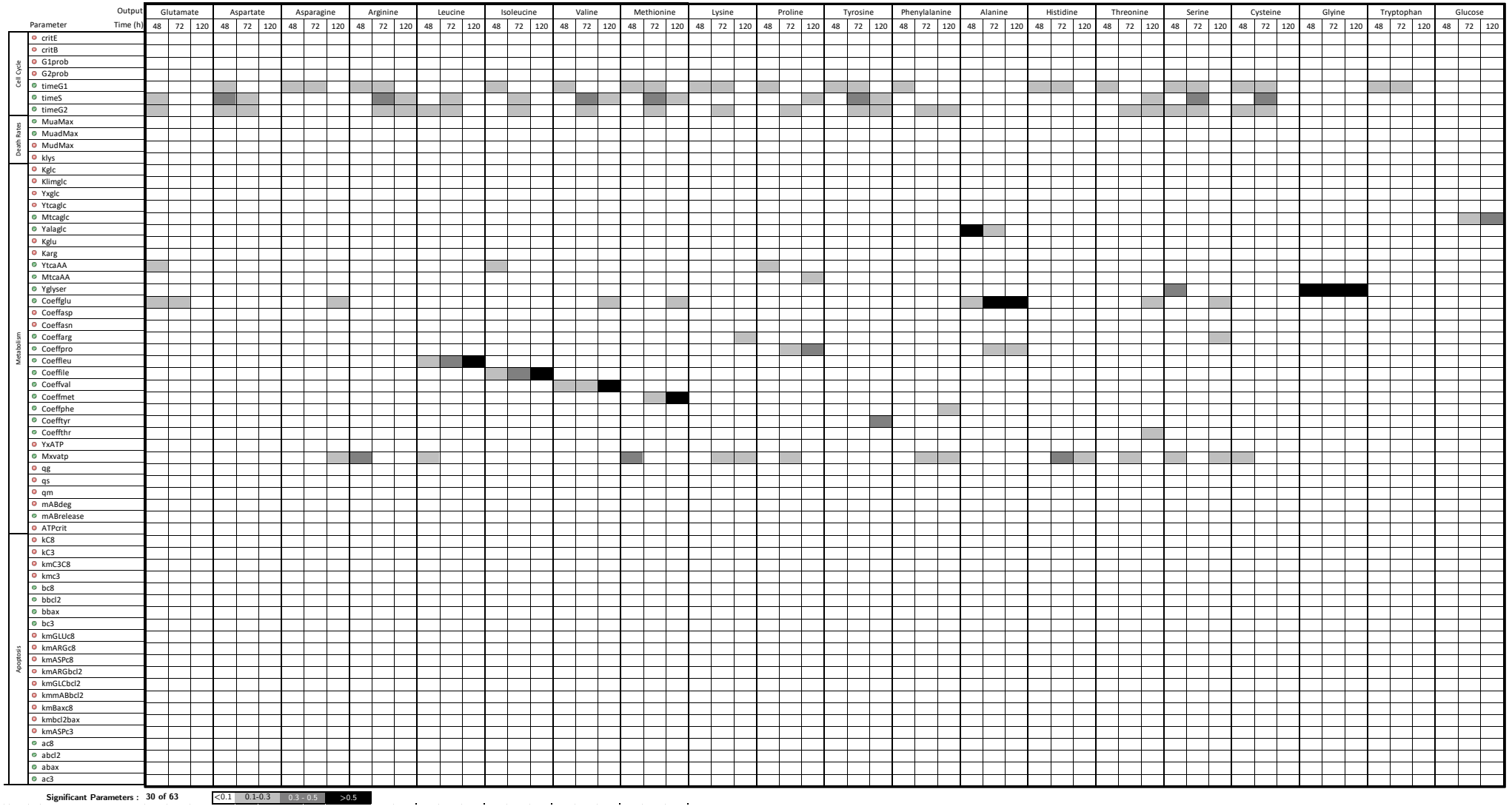


Figure II. GSA results for the metabolism of amino acids and glucose

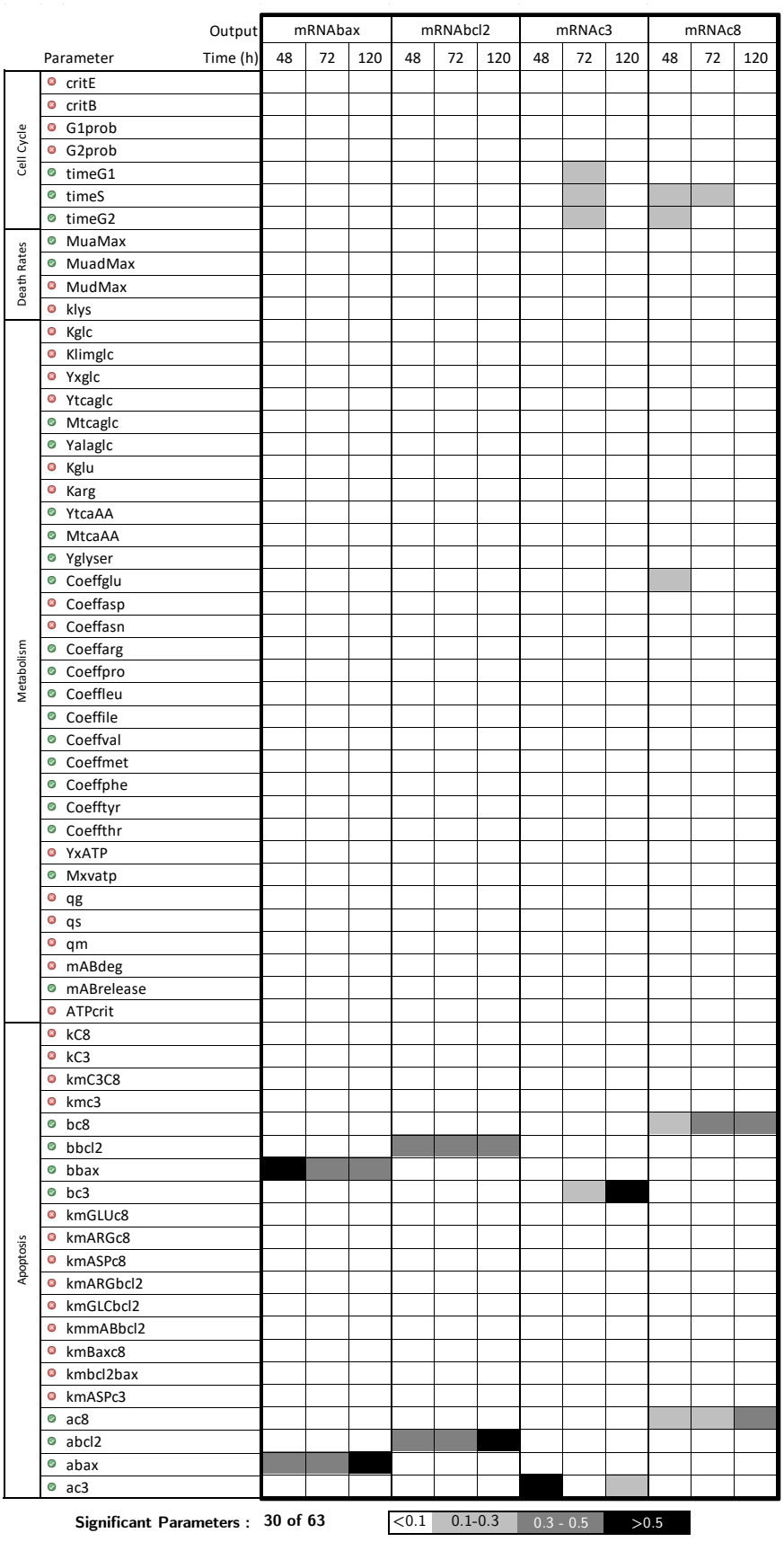


Figure III: GSA results for gene expression of bax (mRNAbax), bcl2 (mRNAbcl2), caspase-3 (mRNAc3) and caspase-8 (mRNAc8).

Appendix II

Growth, metabolism rates and productivity of GS-CHO cells

In this appendix, the growth and metabolism rates and the mAb productivity of GS-CHO cells obtained in this work is showed and compared to other studies for the same cell line (Table I).

Table 1. Specific growth rate (h^{-1}), productivity (pg/cell/day) and metabolism rates (fmol/cell/day) observed for GS-CHO cells in this study and other works with the same cell line. Negative rates indicate consumption.

	This work (nominal density)[1]				**Kyriakopoulos, Polizzi and Kontoravdi, 2013[2]**				**This work (high density)[3]**				**Carinhas et al., 2013[4]**			
					Low Producer		High Producer						High Producer		Low Producer	
	Lag/Early Exponential	Exponential	Stationary	Decline	Lactate Producing	Lactate Consuming	Lactate Producing	Lactate Consuming	Lag/Early Exponential	Exponential	Stationary	Decline	Exponential	Stationary	Exponential	Stationary
μ (day-1)	0.70	0.36	-0.02	-0.15	0.89	-0.1	0.81	-0.05	0.64	0.46	0.11	-0.43	n.a.	n.a.	n.a.	n.a.
qmAB (pg/cell/day)	0.00	16.96	24.20	8.97	12	4	27	8	0.00	18.77	22.73	9.69	20	12.5	7.5	5
Aspartate	94	-25	-45	0	-109	-7	-110	-7	77	-13	-69	0	-164	-60	-129	-65
Glutamate	89	-34	-65	0	-180	-8	-163	-9	65	-23	-93	0	-81	-39	-86	-14
Asparagine	-459	-440	-45	0	-729	-46	-412	-75	-428	-454	-63	-2	-486	-50	-386	-56
Serine	-339	-204	-92	-111	-294	-59	-313	-56	-495	-195	-106	-52	-280	-63	-233	-70
Glutamine	132	33	-18	-3	51	-4	32	-3	147	46	-13	-8	18	-8	2	3
Histidine	-29	-5	-4	1	-35	-3	-38	-2	-49	-5	-3	-1	-42	-3	-24	-6
Glycine	429	98	52	117	96	73	114	53	407	119	72	16	45	52	48	74
Threonine	146	56	0	10	-72	-6	-83	-7	123	64	-6	58	-100	-34	-36	-19
Arginine	9	-26	-8	2	-52	10	-112	1	-16	-21	-9	-2	-72	-31	-75	-9
Alanine	119	227	122	-317	150	-25	77	15	112	221	246	-223	154	65	103	-11
Tyrosine	-22	-9	-8	-4	-14	-11	-38	-4	-7	-16	-7	4	-34	-11	-29	-9
Cysteine	-33	-6	1	1	n.a.	n.a.	n.a.	n.a.	-46	-8	1	1	n.a.	n.a.	n.a.	n.a.
Valine	29	14	-22	-41	-59	-30	-143	-17	74	30	-27	-19	-98	-27	-80	-28
Methionine	-6	-14	-2	2	-19	-2.1	-34	-0.1	-66	-17	-2	8	-31	-5	-32	-5
Tryptophan	14	5	-6	-14	19	-3	-11	0	14	11	-7	-12	-21	-11	-21	-6
Phenylalanine	15	-14	-7	-3	-10	-4	-49	-1	27	-3	-8	-4	-42	-11	-32	-8
Isoleucine	20	40	-23	-44	-22	-21	-81	-12	-24	46	-30	0	-77	-30	-80	-27
Leucine	-38	-48	-58	-94	-57	-51	-173	-30	-63	-36	-64	-45	-131	-54	-127	-48
Lysine	20	-49	-18	-9	n.a.	n.a.	n.a.	n.a.	-72	-30	-20	6	-82	-32	-100	-18
Proline	65	7	14	3	-6	-5	-112	4	82	58	-15	26	-76	-27	-89	-19
Glucose	-3546	-1727	-594	-142	n.a.	n.a.	n.a.	n.a.	-5917	-2117	-833	154	-1763	-693	-2014	-535
Lactate	5711	1810	-379	-496	2044	-456	2592	-374	7315	1881	-349	-823	882	-292	874	-378
Ammonia	858	351	148	525	n.a.	n.a.	n.a.	n.a.	382	357	34	453	287	109	150	121

[1] *Inoculum at 2x10^5 cells/mL; final mAb titre 459±19 mg/L.*

[2] *Inoculum at 2x10^5 cells/mL; final mAb titre for low producer was 230±10 mg/L and for high producer was 550±50mg/L.*

[3] *Inoculum at 4x10^5 cells/mL; final titre 513±49 mg/L*

[4] *Inoculum at 4x10^5 cells/mL; final mAb titre for high producer was 450 mg/L and for low producer was 200 mg/L.*

Appendix III

Global Sensitivity Analysis of the mathematical model developed in Chapter 7.

The GSA results of the model presented in Chapter 7 have been divided into three figures. Figure IV presents the results related to viable, apoptotic and dead cells, cell cycle distribution, ATP and mAb; Figure V shows the results related to metabolism of amino acids, glucose, lactate and ammonia and the results related to gene expression are in Figure VI. As mentioned in Chapter 7, GSA was conducted using the Sobol method and sensitivity indices were calculated at 60h (exponential growth), 160h (stationary growth) and 200h (decline phase) setting a threshold of 0.1 to discriminate between significant and non-significant parameters.

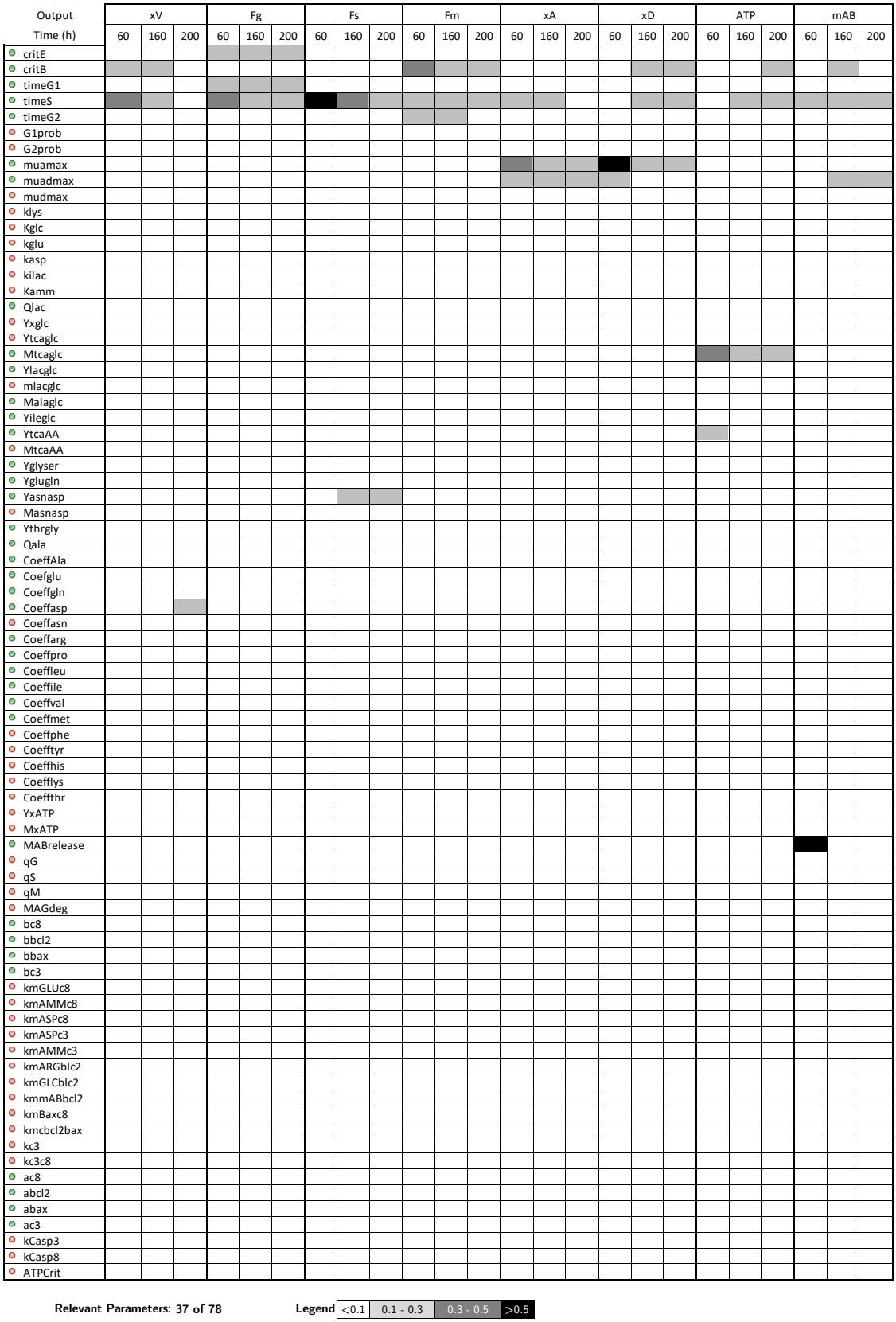


Figure IV. GSA results for viable cells (xV), fraction of cells in phase G1 (fG), phase S (fS) and phase G2/M (fM), apoptotic cells (xA), dead cells (xD), ATP and mAb.

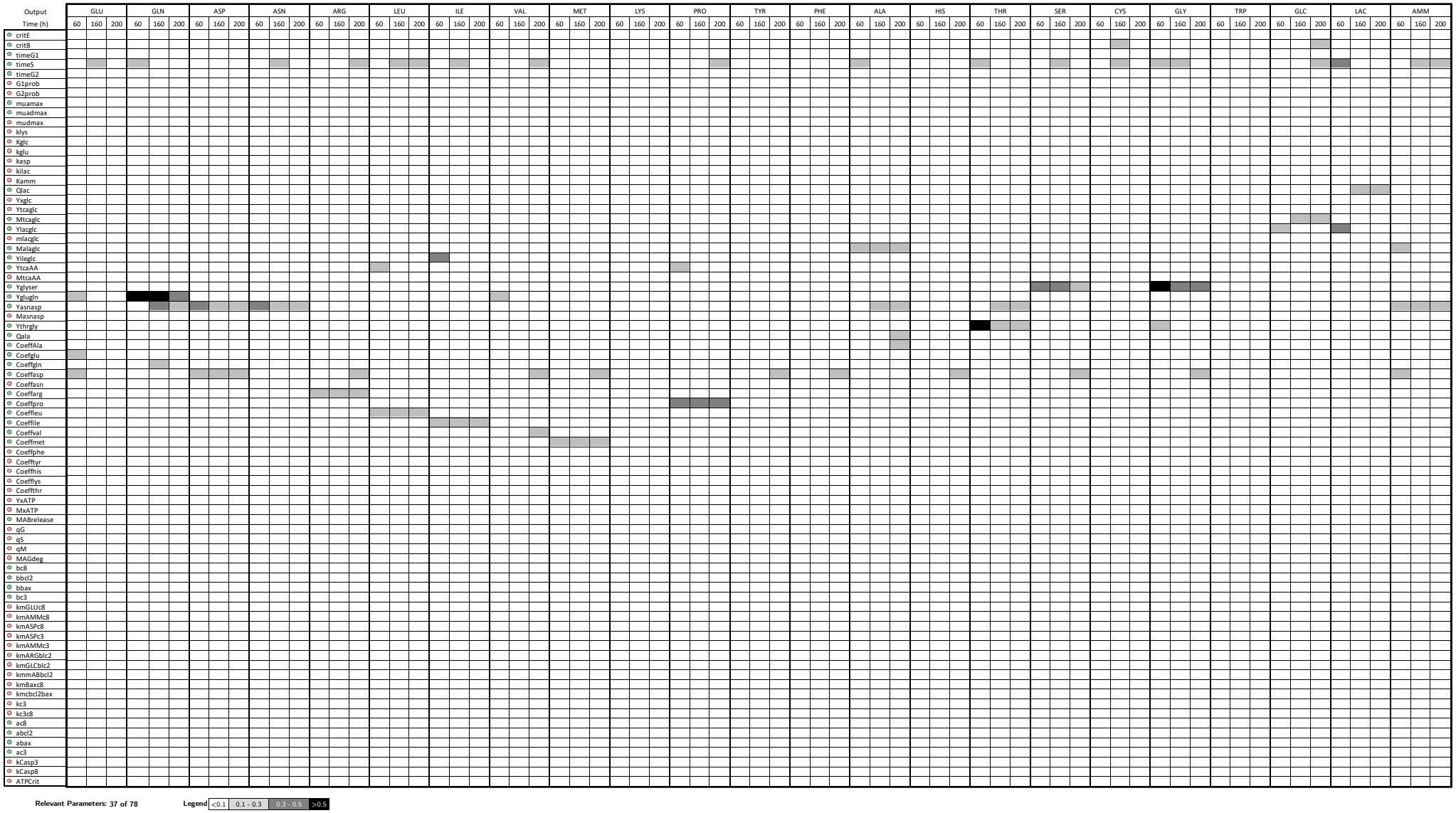


Figure V: GSA results for the metabolism of amino acids, glucose, lactate and ammonia.

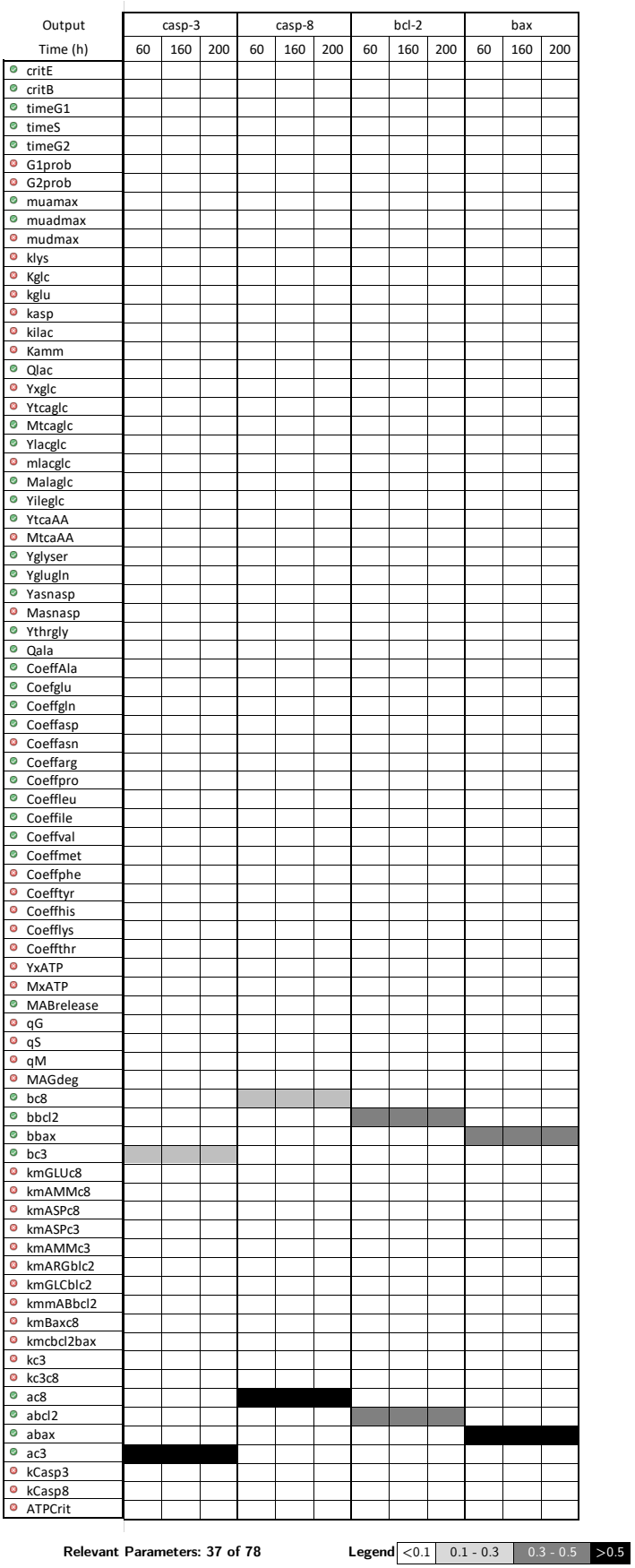


Figure VI. GSA results for gene expression of bax (mRNAbax), bcl2 (mRNAbcl2), caspase-3 (mRNAc3) and caspase-8 (mRNAc8).

Appendix IV

ATP-yielding coefficients used in the CHO metabolic model

In this appendix, the detailed calculation of the ATP yield from each nutrient (glucose or amino acid) following the methodology suggested by Quiroga-Campano, Panoskaltsis and Mantalaris (2018) is shown in detail. For the reactions where the Kegg database was used as a reference, a detailed explanation of the pathways and the derivation of the stoichiometric reaction are provided.

The calculations assume acetyl-CoA to be the limiting reagent for TCA cycle regeneration (Quiroga-Campano, Panoskaltsis and Mantalaris, 2018). ATP production in the TCA cycle is accounted for using the yields of the co-factors NADH and FADH2 assumed to be 2.5 and 1.5 mole ATP/mole, respectively, for mitochondrial co-factors and 1.5 mole ATP/mole for both co-factors, if produced in the cytoplasm. The stoichiometric reactions considered are shown in Table II and the ATP coefficient calculation details are shown in Table III.

Table II. Stoichiometric reactions considered for each nutrient (adapted from Quiroga-Campano, Panoskaltsis and Mantalaris, 2018)

	Nutrient	**Stoichiometric Reaction**	**Reference**
1	Glucose (to Lactate)	$Glucose + 2ADP + 2Pi \rightarrow 2Lactate + 2ATP + 2H_2O$	(Xie and Wang, 1996)
2	Glucose (to Alanine)	$Glucose + 2ADP + 2Pi + 2NAD^+ + Glutamate \rightarrow 2Alanine + 2ATP + 2NADH + 2H^+ + 2H_2O + \alpha Ketoglutarate$	
3	Glucose (to Ac.CoA)	$Glucose + 2CoASH + 2ADP + 2Pi + 4NAD^+ \rightarrow 2AcetylCoA + 2ATP + 2NADH + 2H^+ + 2NADH_{(M)} + 2H_2O + 2CO_2$	
4	Alanine	$Alanine + CoASH + NAD^+ + \alpha Ketoglutarate \rightarrow AcetylCoa + CO_2 + NADH_{(M)} + H^+ + Glutamate$	(Nelson and Cox, 2013)
5	Lysine	$Lysine + 3NAD^+ + FAD + 2H_2O \rightarrow 2Acetyl\ CoA + 3NADH + FADH_2 + 2CO_2 + 2NH_{2-}$	Kegg Database[1]
6	Leucine	$Leucine + FAD + ATP + H_2O + CoA + HCO_3^- \rightarrow 3AcetylCoA + FADH_2 + ADP + Orthophosphate + H^+ + NH_{2-}$	
7	Isoleucine	$Isoleucine + 3NAD^+ + 2FAD + 2CoA + ATP + HCO_3^- + GDP + Pi + 2H_2O \rightarrow AcetylCoA + Oxaloacetate + CoASH + CO_2 + 2NADH + NADH_{(M)} + 2FADH_{2\ (M)} + NH_2^- + GTP + ADP + Orthophosphate + ADP + H^+$	
8	Glutamate	$Glutamate + 2H_2O + 4NAD^+ + FAD + GDP + Pi \rightarrow Oxaloacetate + CoA + NADH + 3NADH_M + FADH_{2\ (M)} + 2CO_2 + GTP + NH_4^+$	(Xie and Wang, 1996)
9	Arginine	$Arginine + 3H_2O + 5NAD^+ + FAD + GDP + P_i \rightarrow Oxaloacetate + CoA + 2NADH + 3NADH_M + FADH_{2\ (M)} + 2CO_2 + GTP$	(Nelson and Cox, 2013)
10	Proline	$Proline + 2H_2O + 5NAD + FAD + GDP + P_i \rightarrow Oxaloacetate + CoA + 2NADH + 3NADH_M + FADH_{2\ (M)} + 2CO_2 + GTP$	
11	Histidine	$Histidine + Tetrahydrofolate \rightarrow Glutamate + NH_3 + 5\text{-}Formiminotetrahydrofolate$	Kegg Database[1]
12	Phenylalanine	$Phenylalanine + Tetrahydrobiopterin + O_2 \rightarrow Tyrosine + H_2O + Dihydrobiopterin$	(Nelson and Cox, 2013)
13	Tyrosine	$Tyrosine + 2O_2 + H_2O + NAD^+ \rightarrow 2AcetylCoA + Oxaloacetate + NADH_{(M)} + NH_{2-} + AMP + 2Pi + CO_2$	
14	Aspartate	$Aspartate + \alpha Ketoglutarate \rightarrow Oxaloacetate + Glutamate$	(Xie and Wang, 1996)
15	Asparagine	$Asparagine + H_2O + \alpha Ketoglutarate \rightarrow Oxaloacetate + Glutamate + NH_4^+$	
16	Valine	$Valine + 3NAD^+ + 2FAD + ATP + GDP \rightarrow Oxaloacetate + CoASH + 2NADH + NADH_M + 2FADH_{2\ (M)} + CO_2 + NH_{2-} + AMP + GTP + Pi + 2H^+$	Kegg Database[1]
17	Threonine	$Threonine + 2\,NAD^+ + FAD + CO_2 + GDP + P_i \rightarrow Oxaloacetate + NADH + NADH_M + FADH_{2\ (M)} + NH_{2-} + GTP$	
18	Methionine	$Methionine + 2ATP + NAD + FAD + 2H_2O + GDP \rightarrow Oxaloacetate + PP_i + NADH + NADH_M + FADH_{2\ (M)} + NH_4 + 2ADP + GTP$	(Nelson and Cox, 2013)
19	TCA Cycle	$AcetylCoA + 2H_2O + GDP + P_i + 3NAD^+ + FAD \rightarrow CoASH + GTP + 3NADH_{(M)} + FADH_{2\ (M)} + 2CO_2$	

[1] Kyoto Encyclopaedia of Genes and Genomes database accessed at www.kegg.jp in December 2018. Organism selected to be the Chinese Hamster (*Cricetulus griseus*).

Table III. ATPj coefficient calculation based on the stoichiometric reactions in Table I (adapted from Quiroga-Campano, Panoskaltsis and Mantalaris, 2018).

	ATP (C)	**NADH (C)[1]**	**NADH (M)[1]**	**$FADH_2$ (C)[1]**	**$FADH_2$ (M)[1]**	**GTP**	**ATP from TCA[2]**	**ATP Balance[3]**	**ATPj[4]**
1	2							2	2
2	2	2						5	5
3	2	2	2				20	30	24
4			1				10	10	9.5
5		3		1			20	26	20
6	-1			1			30	30.5	21.5
7	-1	1	1	1	1	1	10	20	17
8			3		1	1		10	10
9			4		1	1		12.5	12.5
10			4		1	1		12.5	12.5
11			3		1	1		10	10
12	-1		1				20	21.5	15.5
13	-1		1				20	21.5	15.5
14[5]							10	10	7
15									0
16	-1	2	1	1	1	1		8.5	8.5
17	-1	1	1		1	1		5.5	5.5
18		1	1		1	1		6.5	6.5

[1] It has been assumed that each mole NADH or $FADH_2$ produced in the cytoplasm yields 1.5 mole of ATP. NADH and $FADH_2$ produced in the mitochondrion are assumed to yield 2.5 and 1.5 moles of ATP, respectively (Quiroga-Campano, Panoskaltsis and Mantalaris 2018).

[2] ATP from the TCA cycle is calculated based on co-factors and GTP produced until Ac.CoA is regenerated. In total, the TCA cycle yields 10 mole ATP per mole of Ac.CoA used.

[3] ATP Balance refers to the sum of all components in previous columns.

[4] Due to Ac.CoA deviations to lipid metabolism, only 30% of the TCA yield is considered (Quiroga-Campano, Panoskaltsis and Mantalaris 2018).

[5] The reaction of aspartate to oxaloacetate occurs in the cytoplasm, mostly (Nicolae *et al.*, 2014) and this carbon can be used in the TCA.

Reactions from Kegg database

Lysine metabolism (details in Figure VII)

Lysine + 2 2-Oxoglutarate + $3NAD^+$ + FAD + $2H_2O$ → 2AcetylCoA + 2 Glutamate + $2CO_2$ + $3H^+$ + **3NADH** + **$FADH_2$**

Lysine + NADPH + 2-Oxoglutarate + H^+ → Saccharopine + $NADP^+$ + H_2O	[E.C. 1.5.1.8]
Saccharopine + NAD^+ + H_2O → L-2-aminoadipate-6-semialdehyde + Glutamate + H^+ + **NADH**	[E.C. 1.5.1.9]
L-2-aminoadipate-6-semialdehyde + NAD^+ + H_2O → L-2-aminoadipate + H^+ + **NADH**	[E.C. 1.2.1.31]
L-2-aminoadipate + 2-Oxoglutarate → 2-oxoadipate + Glutamate	[E.C. 2.6.1.39]
2-oxoadipate + CoA + NAD^+ → GlutarylCoA + CO_2 + NADH + H^+	[E.C. 1.2.4.2]
GlutarylCoA + FAD → CrotonoylCoA + CO_2 + **$FADH_2$**	[E.C. 1.3.8.6]
CrotonoylCoA + H_2O → (S)-3-hydroxylbutanoylCoA	[E.C. 4.2.1.17]
(S)-3-hydroxylbutanoylCoA + NAD^+ → AcetoacetylCoA + H^+ + **NADH**	[E.C. 1.1.1.35]
AcetoacetylCoA → 2AcetylCoA + CoA	[E.C. 2.3.1.9]

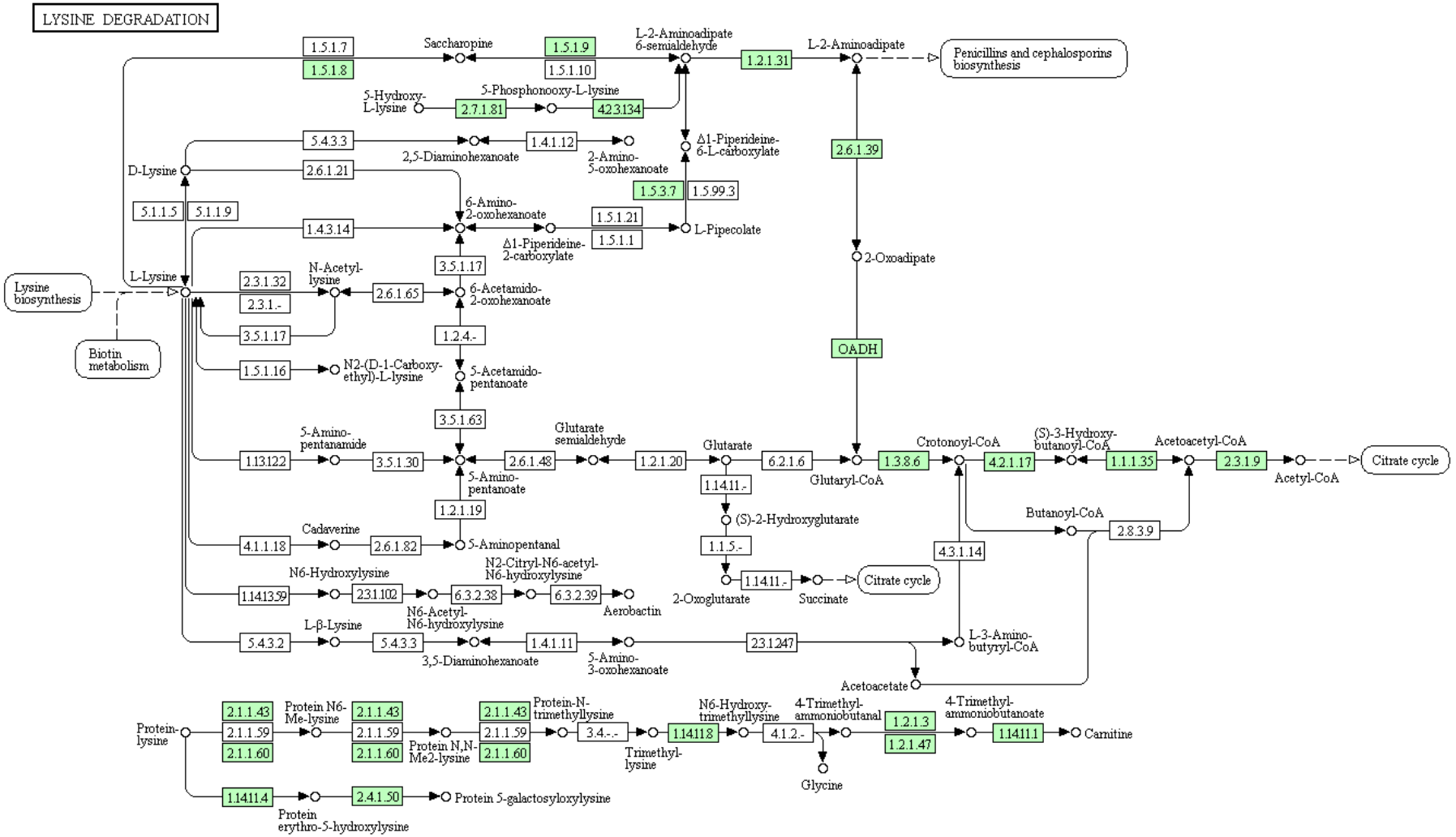


Figure VII. Lysine degradation pathways showing intermediates and enzymes involved. Enzymes shown in green are present in CHO, enzymes in white are absent (www.kegg.jp).

Leucine metabolism (details in Figure VIII)

Leucine + 2-Oxoglutarate + FAD + CoA + ATP + H_2O + HCO_3^- → 3AcetylCoA + Glutamate + CO_2 + $3H^+$ + ADP + Orthophosphate + **$FADH_2$**

Leucine + 2-Oxoglutarate → 4-Methyl-2-oxopentanoate + Glutamate	[E.C. 2.6.1.42]
4-Methyl-2-oxopentanoate + Thiamin diphosphate → 3-Methyl-1-hydroxybutyl-ThPP + CO_2	[E.C. 1.2.4.4]
3-Methyl-1-hydroxybutyl-ThPP + Enzyme N6-(lipoyl)lysine → S-(3-methylbutanoyl)dihydrolipoyllysine + ThPP	[E.C. 1.2.4.4]
S-(3-methylbutanoyl)dihydrolipoyllysine + CoA → 3-Methylbutanoyl-CoA + Enzyme N6-(dihydrolipoyl)lysine	[E.C. 2.3.1.168]
3-Methylbutanoyl-CoA + FAD → 3-Methylcrotonyl-CoA + **$FADH_2$**	[E.C. 1.3.8.7]
3-Methylcrotonyl-CoA + ATP + HCO_3^- → ADP + Orthophosphate + 3-Methylglutaconyl-CoA	[E.C. 6.4.1.4]
3-Methylglutaconyl-CoA + H_2O → (S)-3-Hydroxy-3-methylglutaryl-CoA	[E.C. 4.2.1.18]
(S)-3-Hydroxy-3-methylglutaryl-CoA + CoA → Acetyl-CoA + H_2O + Acetoacetyl-CoA	[E.C. 2.3.3.10]
AcetoacetylCoA → 2AcetylCoA + CoA	[E.C. 2.3.1.9]

Isoleucine metabolism (details in Figure VIII)

Isoleucine + 2-Oxoglutarate + 2FAD + $2H_2O$ + $2CoA^+$ + $3NAD^+$ + **ATP** + HCO_3^- + GDP + Pi → AcetylCoA + Oxaloacetate + Glutamate +CoA-SH + CO_2 + H^+ + ADP + Orthophosphate + **GTP** + **$2FADH_2$** + **NADH(m)** + **2NADH(c)**

Leucine + 2-Oxoglutarate → (S)3-Methyl-2-oxopentanoate + Glutamate	[E.C. 2.6.1.42]
3-Methyl-2-oxopentanoate + Thiamin diphosphate → 2-Methyl-1-hydroxybutyl-ThPP + CO_2	[E.C. 1.2.4.4]
2-Methyl-1-hydroxybutyl-ThPP + Enzyme N6-(lipoyl)lysine → S-(2-methylbutanoyl)dihydrolipoyllysine + ThPP	[E.C. 1.2.4.4]
S-(2-methylbutanoyl)dihydrolipoyllysine + CoA → 2-Methylbutanoyl-CoA + Enzyme N6-(dihydrolipoyl)lysine	[E.C. 2.3.1.168]
2-Methylbutanoyl-CoA + FAD → 2-Methyl-but-2-enoyl-CoA + **$FADH_2$**	[E.C. 1.3.8.7]
2-Methyl-but-2-enoyl-CoA + H_2O → 3-Hydroxy-2-methylbutyril-CoA	[E.C. 4.2.1.17]
3-Hydroxy-2-methylbutanoyl-CoA + NAD^+ → 2-Methylacetoacetyl-CoA + **NADH** + H+	[E.C. 1.1.1.35]
2-Methylacetoacetyl-CoA + CoA → Acetyl-CoA + Propanoyl-CoA	[E.C. 2.3.1.16]
Propanoyl-CoA + **ATP** + HCO_3^- → ADP + Orthophosphate + (S)-Methylmalonyl-CoA	[E.C. 6.4.1.3]
(S)-Methylmalonyl-CoA → SuccinylCoA	[E.C. 5.1.99.1 and 5.4.99.2
SuccinylCoA + GDP + Pi + FAD + H_2O + NAD^+ → Oxaloacetate + CoA-SH + **GTP** + **NADH(m)** + **$FADH_2$(m)**	TCA Cycle

Valine metabolism (details in Figure VIII)

Valine + 2-Oxoglutarate + 2FAD + $2H_2O$ + $2CoA^+$ + $3NAD^+$ + ATP + GDP → Oxaloacetate + Glutamate + CoA-SH + CO_2 + $2H^+$ + AMP + Pi + GTP + **$2FADH_2$** + **NADH(m)** + **2NADH(c)**

Valine + 2-Oxoglutarate → 3-Methyl-2-oxopentanoate + Glutamate	[E.C. 2.6.1.42]
3-Methyl-2-oxopentanoate + Thiamin diphosphate → 2-Methyl-1-hydroxybutyl-ThPP + CO_2	[E.C. 1.2.4.4]
2-Methyl-1-hydroxybutyl-ThPP + Enzyme N6-(lipoyl)lysine → S-(2-methylbutanoyl)dihydrolipoyllysine + ThPP	[E.C. 1.2.4.4]
S-(2-methylbutanoyl)dihydrolipoyllysine + CoA → Isobutyryl-CoA + Enzyme N6-(dihydrolipoyl)lysine	[E.C. 2.3.1.168]
Isobutyryl-CoA + FAD → Metacrylyl + **$FADH_2$**	[E.C. 1.3.8.7]
Metacrylyl-CoA + H_2O → 3-Hydroxy-isobutyryl-CoA	[E.C. 4.2.1.17]
3-Hydroxy-isobutyryl-CoA + NAD^+ → 3-Hydroxy-isobutyrate + CoA	[E.C. 3.1.2.4]
3-Hydroxy-isobutyrate + NAD^+ → (S)-Methylmalonate semialdehyde + **NADH** + H^+	[E.C. 1.1.1.35]
(S)-Methylmalonate semialdehyde + NAD^+ + H_2O → Methylmalonate + **NADH** + H^+	[E.C. 1.2.3.1]
Methylmalonate + CoA + **ATP** → (R)-Methylmalonyl-CoA + AMP + Diphosphate	[E.C. 6.2.1.-]
(R)-Methylmalonyl-CoA → Succinyl-CoA	[E.C. 5.4.99.2]
SuccinylCoA + GDP + Pi + FAD + H_2O + NAD^+ → Oxaloacetate + CoA-SH + **GTP** + **NADH(m)** + **$FADH_2$(m)**	TCA Cycle

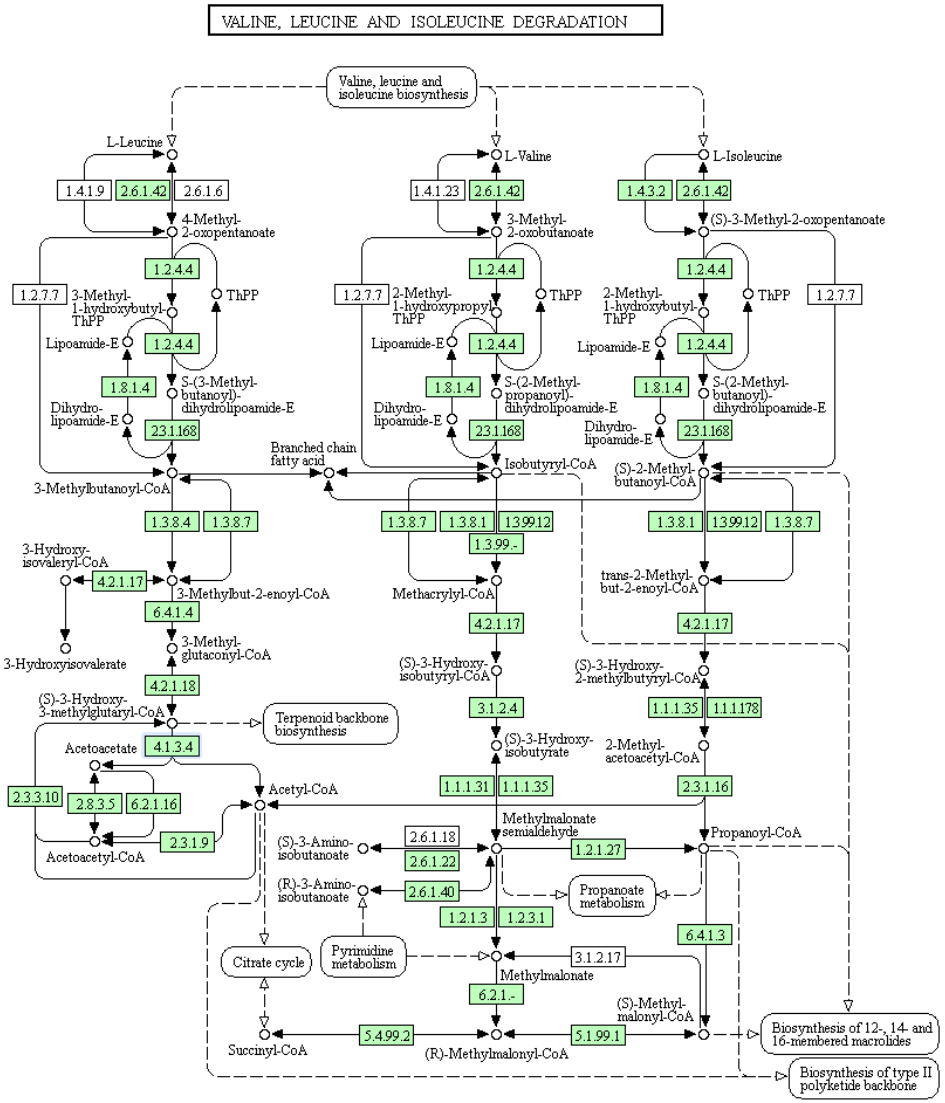


Figure VIII..Leucine, Isoleucine and Valine degradation pathways (www.kegg.jp).

Histidine metabolism (details in Figure IX)

Histidine can be converted to Glutamate, as follows, which is then converted to α-ketoglutarate and enters the TCA cycle.

Histidine + Tetrahydrofolate → Glutamate + NH_3 + 5-Formiminotetrahydrofolate

Histidine → Urocanate + NH_3	[E.C. 4.3.1.3]
Urocanate → 4-Imidazolone-5-propanoate + H_2O	[E.C. 4.2.1.49]
4-Imidazolone-5-propanoate + H_2O → N-Formimino-L-glutamate	[E.C. 3.5.2.7]
N-Formimino-L-glutamate + Tetrahydrofolate → 5-Formiminotetrahydrofolate + Glutamate	[E.C. 2.1.5.2]

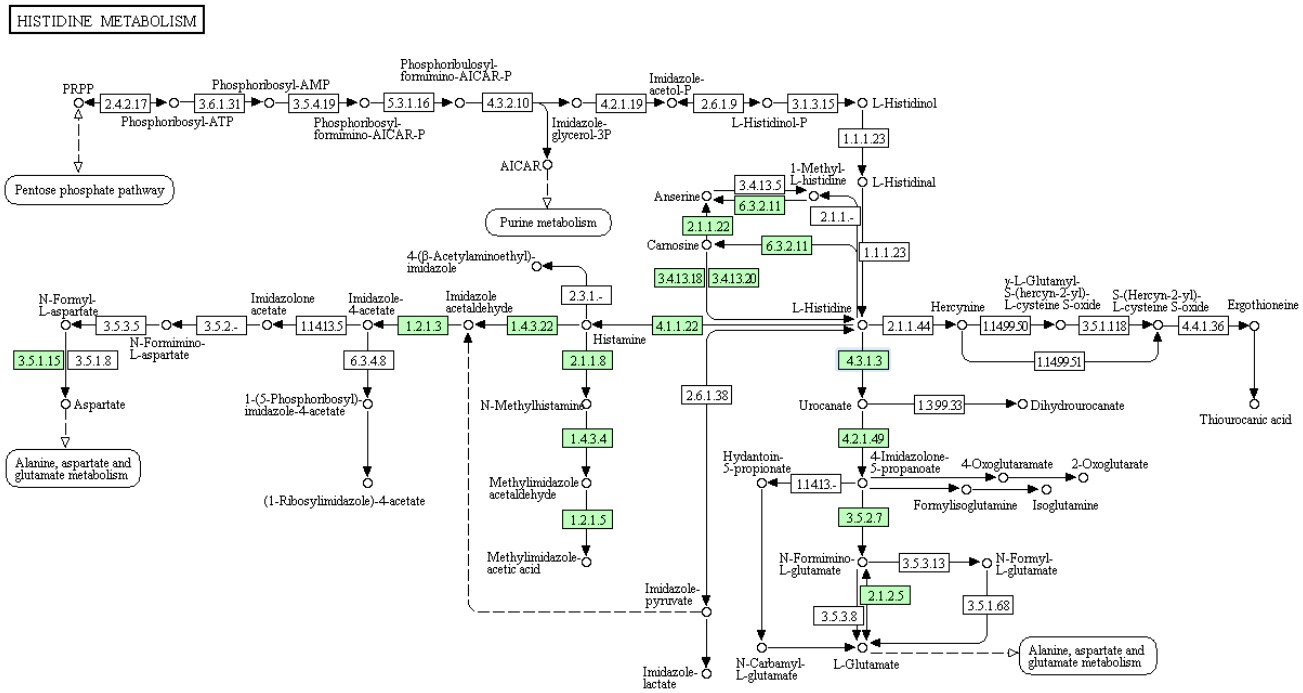


Figure IX: Histidine degradation pathways (www.kegg.jp).

Tyrosine and Phenylalanine metabolism (details in Figure X)

Phenylalanine undergoes conversion to tyrosine via:

Phenylalanine + Tetrahydrobiopterin + O_2 → Tyrosine + H_2O + Dihydrobiopterin [E.C. 1.14.16.1]

Tyrosine metabolism is as follows:

Tyrosine + 2-Oxoglutarate + $2O_2$ + H_2O + ATP + NAD^+ → 2AcetylCoA + Oxaloacetate + Glutamate + AMP + 2Pi + CO_2 + **NADH(m)**

Tyrosine + 2-Oxoglutarate → 3-(4-Hydroxyphenyl)pyruvate + Glutamate	[E.C. 2.6.1.1]
3-(4-Hydroxyphenyl)pyruvate + O_2 → Homogentisate + CO_2	[E.C. 1.13.11.27]
Homogentisate + O_2 → 4-Maleylacetoacetate	[E.C. 1.13.11.5]
4-Maleylacetoacetate → 4-Fumarylacetoacetate	[E.C. 5.2.1.2]
4-Fumarylacetoacetate + H_2O → Acetoacetate + Fumarate	[E.C. 3.7.1.2]
Fumarate + H_2O + NAD^+ → Oxaloacetate + **NADH(m)**	TCA Cycle
Acetoacetate + ATP + CoA → AcetoacetylCoA + AMP + 2Pi	[E.C. 6.2.1.16]
AcetoacetylCoA → 2AcetylCoA + CoA	[E.C. 2.3.1.9]

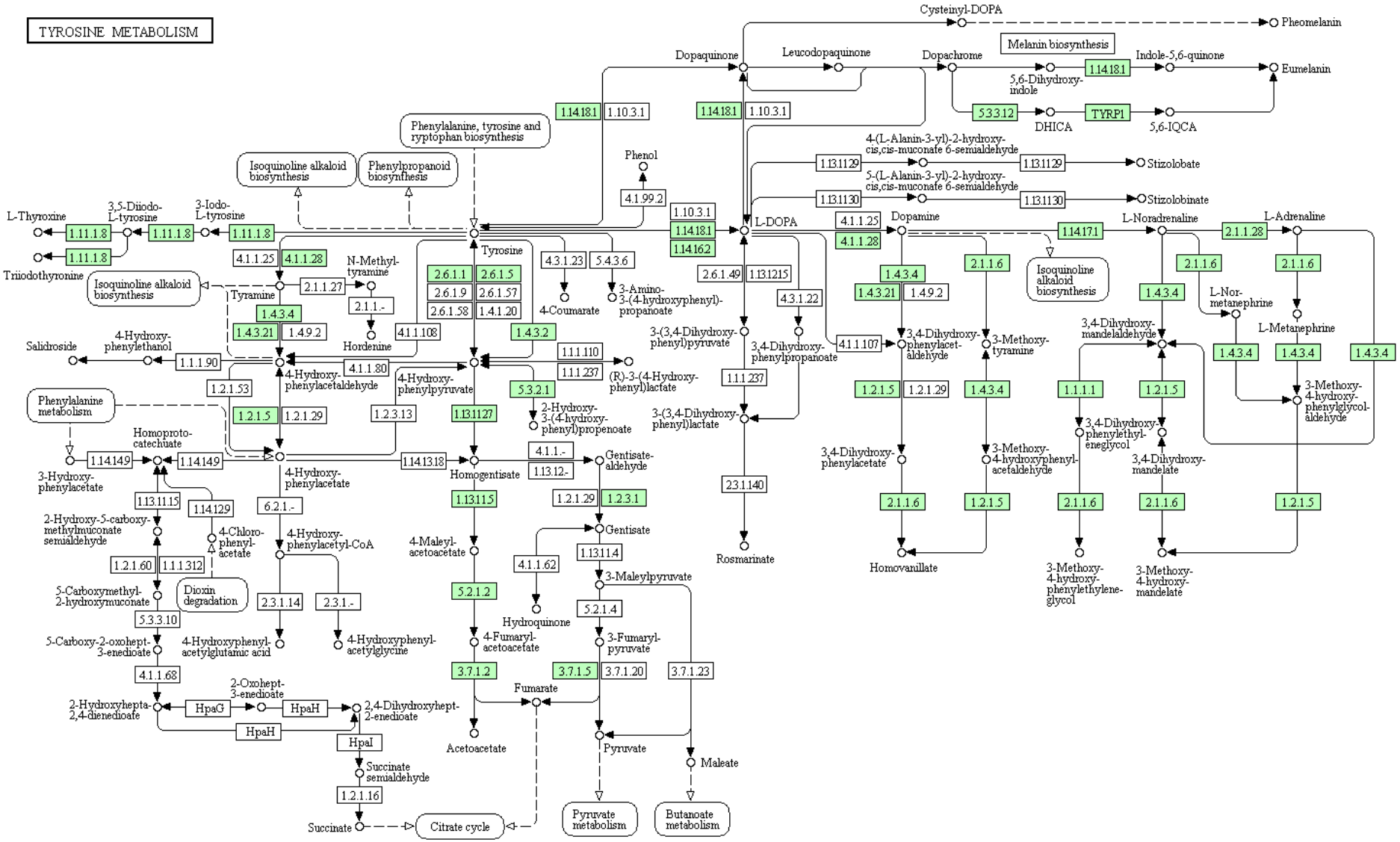


Figure X. Tyrosine degradation pathways (www.kegg.jp).

Appendix IV

The use of hyperbolic tangent functions as on/off switches in mathematical modelling of biological systems

In the mathematical models developed in chapters 4 and 6 of this work, branched kinetics were used to account for the consumption of nutrients to various ends. Likewise, transitions between cell cycle functions are determined by whether the levels of internal variables (cyclin or DNA) are above or below a pre-set (experimentally determined) threshold. Branched functions are required to account for cell cycle transitions, to model metabolic shifts and, in the general case, are important to model any "on/off" switch in a biological system.

The main metabolic shift in GS-NS0 cells is triggered by the depletion of glutamate. It has been hypothesized that arginine is used to satisfy the biomass requirements while arginine, alanine, proline, phenylalanine and tyrosine satisfy the energy requirements. Similarly, the GS-CHO metabolism is characterized by multiple metabolic shifts on lactate, ammonia, alanine, isoleucine, asparagine/aspartate, glutamate/glutamine (equations 52, 54, 55, 56, 62, 64, 66) .

A key challenge on the implementation of mathematical models employing branched kinetics is to find an implementation that does not compromise the continuous nature of the functions used. Any discontinuities increase the computational cost and may compromise the use of parameter estimation algorithms. Therefore, the use of on/off switches is an important aspect of the implementation of mathematical models of biological systems.

In this work, this problem was overcome using *tanh* functions and this appendix describes the implementation of this strategy.

The hyperbolic tangent (*tanh*) function

The *tanh* function is defined as the ratio of the hyperbolic sine to the hyperbolic cosine, similarly to the ordinary trigonometric tangent function. The hyperbolic tangent is hence defined as

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{\frac{e^{x}-e^{-x}}{2}}{\frac{e^{x}+e^{-x}}{2}} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

and it is graphically represented as shown in Figure XI.

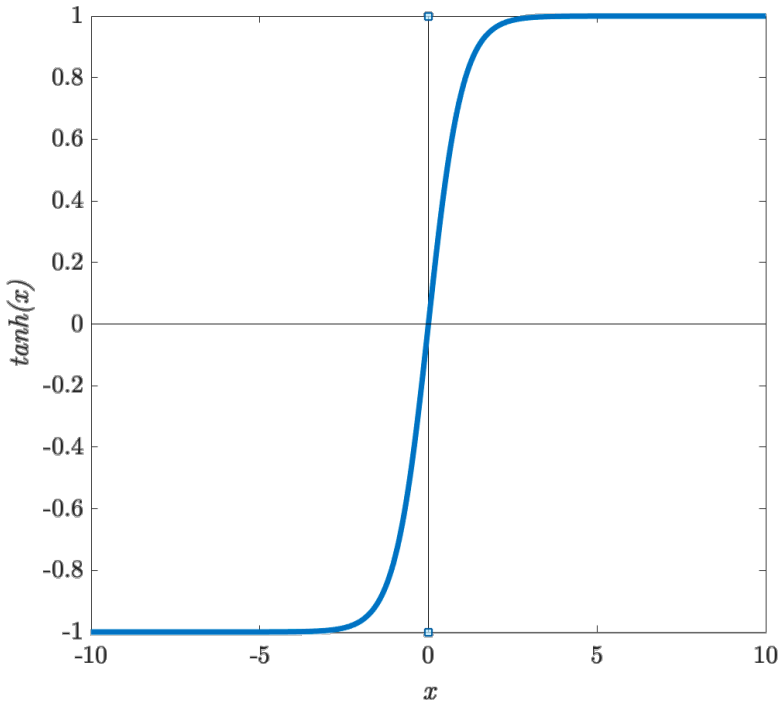


Figure XI: Graphical representation of *tanh*(*x*) vs. x

In its original form, the *tanh* function could be used as a smooth (and therefore slow) -1/1 switch. Therefore, its use as a steep on/off (0/1) switch requires some modification to the original function. First, let us consider the sharpness of the switch, i.e., how steep the step from -1 to 1 is in the original form of the *tanh* function. To this end, and without loss of generality, it helps to take a closer look at the [-4; 4] interval

in the domain. If the argument of the *tanh* function is divided by a scaler in the [0;1[interval, steeper functions are obtained as the denominator increases, i.e.

$$\lim_{p \to 0} \tanh\left(\frac{x}{p}\right) = \frac{x}{|x|}$$

as shown in Figure XII. A suitable denominator is determined based on two key factors: the first is the user definition of how steep the switch should be, and the second is the tolerance of the solver used which dictates its ability to handle very small quantities without numerical issues.

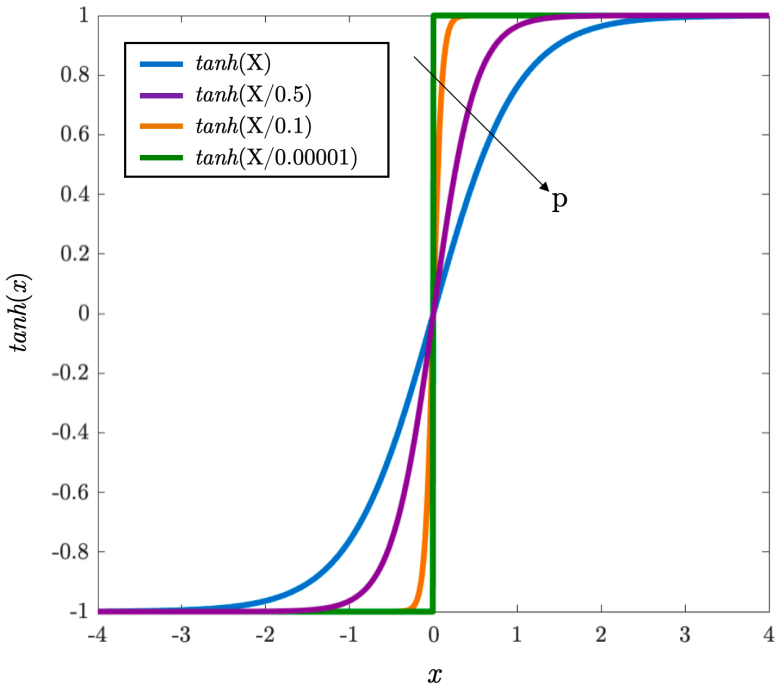


Figure XII: Graphical representation of *tanh(x/p)* vs. x for p=1 (blue), p=0.5 (purple), p=0.1 (orange) and p=0.00001 (green).

Once a suitably steep switch *tanh* function has been decided upon (i.e. the value of p has been chosen), it is necessary to shift the codomain interval from [-1; 1] to [0; 1]. This is trivially solved by using

$$\frac{1}{2}\left[\tanh\left(\frac{x}{p}\right) + 1\right]$$

and, lastly, to control the shift at a particular value of x=a, the function takes the form defined by *f(x)*, if the switch happens as x is increasing or that of *g(x)* if the switch happens as x is decreasing.

$$f(x) = \frac{1}{2}\left[\tanh\left(\frac{x-a}{p}\right)+1\right] \qquad g(x) = \frac{1}{2}\left[\tanh\left(\frac{a-x}{p}\right)+1\right]$$

Examples of the application of *tanh*-switches in biological systems.

Cell cycle transitions

In mathematical models of cell cycle using population balance equations, such as those presented in this work (chapters 4 and 6) and those previously published (Fuentes-Garí *et al.*, 2015; García-Münzer, Kostoglou, *et al.*, 2015) transitions between phases depend on the level of the variable used to track each phase (segregation or distribution variable). In this example, let us consider the transition between phase G_1 to phase S. Phase G_1 is monitored by tracking the accumulation of cyclin E1. Cells begin with zero cyclin E1 and accumulate this protein as they progress through phase G_1. Above a certain cyclin E1 threshold, they may either transition to phase S or remain in phase G_1, transitioning at a later stage. The rate of transition (Γ_{G1}) is therefore given by

$$\Gamma_{G_1} = \begin{cases} 0 & \text{below cyclin E1 threshold} \\ p_i . f_{lim} & \text{above cyclin E1 threshold} \end{cases}$$

where p_i is a probability of transition and f_{lim} is a metabolic limitation function. In order to avoid the use of *if* loops, the branched function can be implemented using a *tanh* switch by coding

$$\Gamma_{G_1} = (p_i . f_{lim}) . \frac{1}{2}\left[\tanh\left(\frac{cycE - cycE_{crit}}{p}\right)+1\right]$$

where $cycE_{crit}$ is the cyclin threshold and which produces the output shown graphically in Figure XIII.

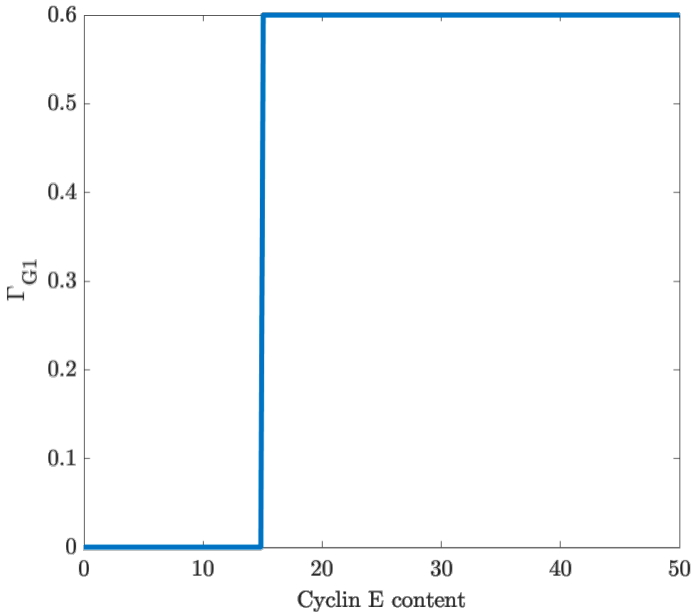


Figure XIII: Representation of Γ_{G1} as a function of the cyclin E1 content as a continuous function using a *tanh*-based switch ($cycE_{crit}$=15).

Another interesting case to consider is that where cells enter a phase of the cell cycle at a precise point in the corresponding domain. Let us consider the transition between phase S and phase G_2/M of the cell cycle where the latter is tracked using cyclin B1. It has been experimentally demonstrated that cyclin B1 accumulation starts during phase S (García-Münzer *et al.*, 2013) and, therefore, cells enter phase G_2/M with a cyclin B1 defined by the accumulation in phase S ($cycB_S$), instead of entering with $cycB$=0. In chapters 4 and 6 of the present work, as well as was done in previous studies, the population balance equation that describes the entrance into phase G_2/M is

$$\frac{\partial(cycB,t)}{\partial t}+\frac{\partial}{\partial cycB,}[r_{G2}(cycB,S)N_{G2}(cycB,t)]=-\Gamma_{G2}(cycB,S)N_{G2}(cycB,t)-$$
$$[\mu_A(g)+\mu_D(E)]N_{G2}(cycB,t)+\delta(cycB\text{-}cycB_S)r_S(DNA=2,S)N_S(DNA=2,t)$$

where the last term describes the inflow of cells from phase S into phase G_2/M. This term should only be active at $cycB=cycB_S$, as is indicated by the Dirac delta function.

This can be implemented with two *tanh* switches combined, yielding the continuous r_s' function defined by

$$\mathrm{r}_S{}' = \mathrm{r}_S \cdot \frac{1}{2}\left[\tanh\left(\frac{\mathrm{cycB} - (\mathrm{cycB}_S + 1 - \Delta cycB)}{\mathrm{p}}\right) - \tanh\left(\frac{\mathrm{cycB} - (\mathrm{cycB}_S + 1 + \Delta cycB)}{\mathrm{p}}\right)\right]$$

where ΔcycB is a small quantity to allow for one single non-zero solution (in $\mathbb{R}^+$) at $cycB = cycB_S + 1$. This function has the graphical representation shown in Figure XIV.

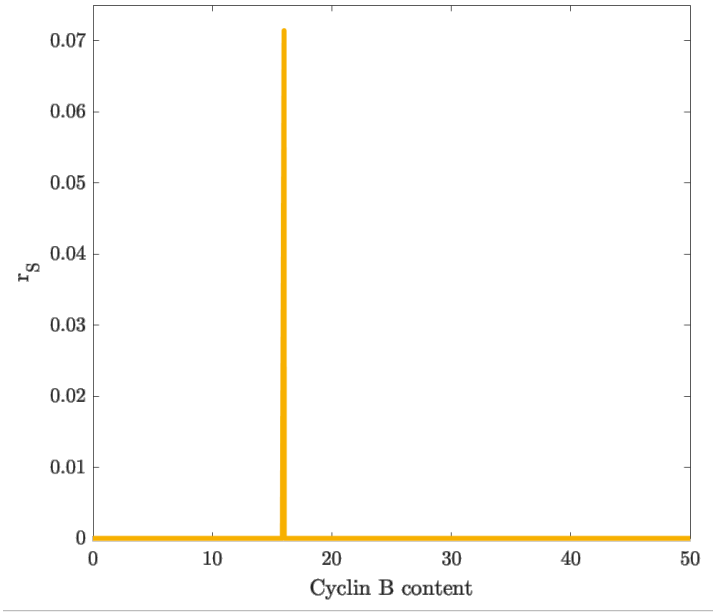


Figure XIV: Representation of r_S as a function of the cyclin B1 content as a continuous function using two *tanh*-based switches ($cycB_S$=15; ΔcycB=0.01).

Metabolic shifts

Several metabolic shifts occur in mammalian cell cultures. In this example, let us consider the GS-CHO cell cultures described in chapter 5 and modelled in chapter 6. In the course of these cultures, two ammonia metabolic shifts occur and therefore the ammonia kinetics has three branches. Let us consider a simplified description of this kinetics, as a function of the concentration of another metabolite (M) as

$$Q_{amm} = \begin{cases} Q_1 & [M] \geq m_1 \\ Q_2 & m_2 \leq [M] < m_1 \\ Q_3 & [M] < m_2 \end{cases}$$

This branched kinetics can be expressed by a continuous function described by

$$Q_{amm} = Q_1 \cdot \frac{1}{2}\left[\tanh\left(\frac{[M]-m_1}{p}\right)+1\right] + Q_2 \cdot \frac{1}{2}\left[\tanh\left(\frac{m_1-[M]}{p}\right)+\tanh\left(\frac{[M]-m_2}{p}\right)\right] + Q_3 \cdot \frac{1}{2}\left[\tanh\left(\frac{m_2-[M]}{p}\right)+1\right]$$

which has the graphical representation shown in Figure XV.

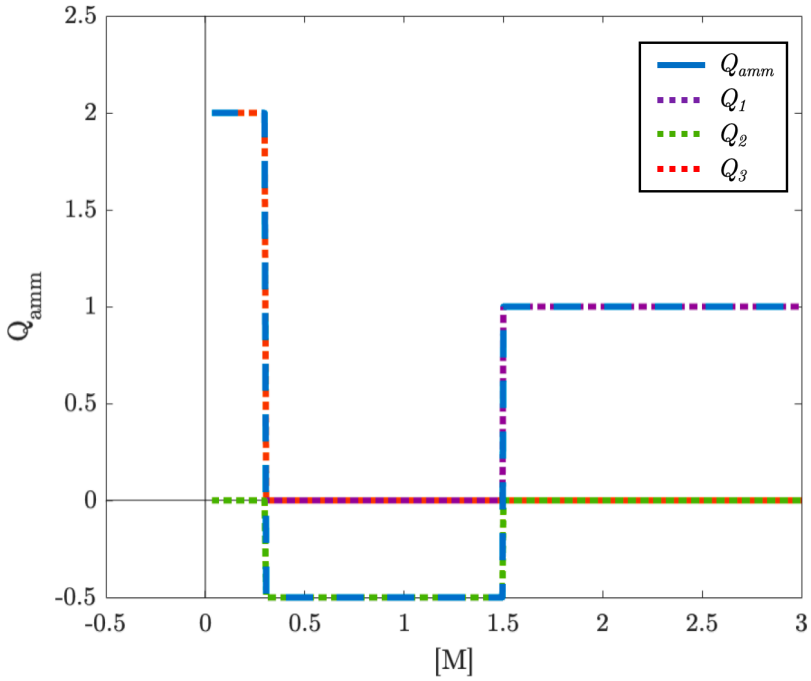


Figure XV: Representation of Q_{amm} as a function of the concentration of metabolite (M) as a continuous function using three *tanh*-based switches (Q_1=1; Q_2=-0.5; Q_3=2; m_1=1.5; m_2=0.3).

In conclusion, the hyperbolic tangent (*tanh*) function provides a suitable straightforward mathematical tool for the implementation of continuous on/off switches in biological systems. It is simple to implement as the *tanh* function is available in most modelling packages (gProms, MatLab e.g.) and highly customizable.