**MODELLING OF METABOLIC CONTROL BY SHORT CHAIN FATTY ACIDS AT THE LEVEL OF THE FUNCTIONAL PROTEOME**

**By**

**León Eduardo Sánchez-Mejorada Zúñiga**

**MPhil October 2014**

**University of Sheffield **

**Modelling of metabolic control by Short Chain Fatty Acids at the level of the functional proteome**

**León Eduardo Sánchez-Mejorada Zúñiga**

**Submitted to the University of Sheffield in conformity with the requirements for the degree of Master of Philosophy (MPhil)**

**Molecular Gastroenterology Research Group**

Academic Unit of Surgical Oncology

Department of Oncology  
The Medical School

Beech Hill Road, Sheffield, S10 2RX

**OCTOBER 2014**

**ACKNOWLEDGEMENTS**

I would like to acknowledge and extend my gratitude to:

My family in Mexico. Thank you very much for your unconditional support and help every time I needed it. I know that is very difficult to be far away from home and not be in touch as often as we should or we would like to. Even you are miles away I fell your warmth and love every second.

Thank you very much indeed. I love you Mooomi, Cuchis and Yoyis.

My family in Sheffield. Thank you guys, you made my life here much easier. I would not have been able to do this without you all.

I know some of you will be there for me no matter what, some of you have left; ones for good, others forever, and some might come back, some others might not come back never ever again to my life.

The only thing I want to say is that I will remember and love every single one of you, my Sheffy people.

**CONTENTS**

[LIST OF FIGURES 5](#_Toc413753736)

[LIST OF TABLES 6](#_Toc413753737)

[ABSTRACT 7](#_Toc413753738)

[INTRODUCTION 8](#_Toc413753739)

[I.Short chain fatty acids 9](#_Toc413753740)

[I.1 Butyrate 10](#_Toc413753741)

[I.2 Propionate 12](#_Toc413753742)

[I.3 Acetate 13](#_Toc413753743)

[I.4 Other SCFAs 13](#_Toc413753744)

[II. Metabolic outcomes of the alteration in the acetylation patterns of metabolic enzymes 15](#_Toc413753745)

[II.1 Implications of lysine residue acetylation in metabolic enzyme function 15](#_Toc413753746)

[II.1.i Glycolysis 15](#_Toc413753747)

[III.1.ii The TCA cycle 18](#_Toc413753748)

[II.1.iii β-oxidation 23](#_Toc413753749)

[II.1.iv Pyruvate metabolism 24](#_Toc413753750)

[III.Proteomic analysis 26](#_Toc413753751)

[IV.Modelling overview 32](#_Toc413753752)

[AIMS 36](#_Toc413753753)

[OBJECTIVES 36](#_Toc413753754)

[METHODS 37](#_Toc413753755)

[Mitochondrially enriched fractions 37](#_Toc413753756)

[Acetylation of metabolic pathways 37](#_Toc413753757)

[Functional annotations 38](#_Toc413753758)

[Pathways coverage 39](#_Toc413753759)

[Matching pathways representation 39](#_Toc413753760)

[Modelling 40](#_Toc413753761)

[Validation of the models 42](#_Toc413753762)

[RESULTS 44](#_Toc413753763)

[Representation and acetylation of metabolic pathways 44](#_Toc413753764)

[Percentage coverage of pathways 45](#_Toc413753765)

[Functional annotations 48](#_Toc413753766)

[Matching pathways 57](#_Toc413753767)

[Action of SCFAs on acetyl-CoA and CoA 60](#_Toc413753768)

[Validation of the models 69](#_Toc413753769)

[DISCUSSION 72](#_Toc413753770)

[CONCLUSIONS 79](#_Toc413753771)

[REFERENCES 81](#_Toc413753772)

[APPENDIX I. Abbreviations 85](#_Toc413753773)

[APPENDIX II. Implications of acetylated lysine residues in metabolic enzymes/Variations in response to butyrate in metabolic enzymes 88](#_Toc413753774)

[Appendix III. *p*-values in every functional annotation of each GO term chosen in the global and acetylated proteome datasets. 91](#_Toc413753775)

# LIST OF FIGURES

[Figure 1. Structural differences in butyrate, isobutyrate, valerate and isovalerate. 14](#_Toc413419261)

[Figure 2. Enzymes and reactions in glycolysis. 16](#_Toc413419262)

[Figure 3. Phosphoglycerate mutase regulation by glucose levels. 18](#_Toc413419263)

[Figure 4. Enzymes and reactions in the TCA cycle. 19](#_Toc413419264)

[Figure 5. Isocitrate dehydrogenase regulation by glucose levels. 21](#_Toc413419265)

[Figure 6. Regulation of malate dehydrogenase by glucose levels. 22](#_Toc413419266)

[Figure 7. Enzymes and reactions in β-oxidation. 23](#_Toc413419267)

[Figure 8. Enzymes and reactions in pyruvate metabolism. 24](#_Toc413419268)

[Figure 9. Butyrate\_v3-1 model. 33](#_Toc413419269)

[Figure 10. SCFA Uptake and Oxidation v4 model. 34](#_Toc413419270)

[Figure 11. Representation and acetylation of major metabolic pathways involving acetyl-CoA as a component. 45](#_Toc413419271)

[Figure 12. Percentage Coverage of pathways. 47](#_Toc413419272)

[Figure 13. BP enrichment scores. 50](#_Toc413419273)

[Figure 14. CC enrichment scores. 51](#_Toc413419274)

[Figure 15. MF enrichment scores.. 52](#_Toc413419275)

[Figure 16. CC metabolism-related enrichment scores. 54](#_Toc413419276)

[Figure 17. BP metabolism-related enrichment scores. 55](#_Toc413419277)

[Figure 18. MF metabolism-related enrichment scores. 56](#_Toc413419278)

[Figure 19. Changes in the concentration of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA at different time points using the butyrate\_v3-1 model. 61](#_Toc413419279)

[Figure 20. Changes in the concentration of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA at different time points using the SCFA Uptake and Oxidation v4 model. 63](#_Toc413419280)

[Figure 21. Effect of acetate on acetyl-CoA levels. 65](#_Toc413419281)

[Figure 22. Effect of butyrate on acetyl-CoA levels. 65](#_Toc413419282)

[Figure 23. Effect of propionate on acetyl-CoA levels. 66](#_Toc413419283)

[Figure 24. Effect of glucose on acetyl-CoA levels. 66](#_Toc413419284)

[Figure 25. Effect of acetate on CoA levels. 67](#_Toc413419285)

[Figure 26. Effect of butyrate on CoA levels. 68](#_Toc413419286)

[Figure 27. Effect of propionate on CoA levels. 68](#_Toc413419287)

[Figure 28. Effect of glucose on CoA levels. 69](#_Toc413419288)

[Figure 29. Effect of increasing concentrations of SCFAs on CoA levels. 70](#_Toc413419289)

# LIST OF TABLES

[Table 1. Short Chain Fatty Acids.*.* 9](#_Toc413666988)

[Table 2. Effect of acetylation on phosphofructokinase, fructose-bisphosphate aldolase, glyceraldehyde 3-phosphate dehydrogenase and phosphoglycerate mutase. 16](#_Toc413666989)

[Table 3. Effect of acetylation on isocitrate dehydrogenase, succinate dehydrogenase, and malate dehydrogenase. 19](#_Toc413666990)

[Table 4. Effect of acetylation on acyl-CoA synthetase, enoyl-CoA hydratase, and β-hydroxyacyl-CoA dehydrogenase. 23](#_Toc413666991)

[Table 5. Effect of acetylation on pyruvate kinase, phosphoenolpyruvate carboxykinase, lactate dehydrogenase, and pyruvate carboxylase. 25](#_Toc413666992)

[Table 6. Protein Profiling Techniques. 28](#_Toc413666993)

[Table 7. Summary of findings made using protein profiling techniques. 31](#_Toc413666994)

[Table 8. Models developed by BC. 40](#_Toc413666995)

[Table 9. Initial concentrations of the substrate species defined in the models. 42](#_Toc413666996)

[Table 10. Coverage ratio. 48](#_Toc413666997)

[Table 11. Matching pathways. 58](#_Toc413666998)

[Table 12. Enzymes matching Metabolism pathway.*.* 59](#_Toc413666999)

[Table 13. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on glycolytic enzymes. 88](#_Toc413667000)

[Table 14. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on the TCA cycle enzymes. 89](#_Toc413667001)

[Table 15. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on β-oxidation enzymes. 90](#_Toc413667002)

[Table 16. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on Pyruvate metabolism enzymes. 90](#_Toc413667003)

[Table 17. BP GO term *p-*values. 91](#_Toc413667004)

[Table 18. CC GO term *p-*values. 91](#_Toc413667005)

[Table 19. MF GO term *p-*values. 92](#_Toc413667006)

# ABSTRACT

**Background**: Fibre fermentation by bacteria in the luminal colon leads to the production of short chain fatty acids (SCFAs), which have been thought to alter global protein acetylation patterns depending on their concentration levels through their role as lysine deacetylase inhibitors.

**Aim**: Perform a representation analysis on a proteomic dataset. Model the action of increasing concentrations of acetate, butyrate, propionate and glucose on acetyl-CoA and CoA and to simulate the effect of a time course on the levels of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA.

**Methods**: An existing mitochondrial proteomic dataset was examined to detect which metabolic enzymes in glycolysis, the TCA cycle, β-oxidation and pyruvate metabolism are acetylated. Free access software (DAVID, Uniprot identifier, KEGG, and Reactome) was used to generate a representation and acetylation map of these metabolic pathways. Enrichment scores of functional annotations in gene ontology terms, pathway percentage coverage and matching representation of these pathways were also elucidated by using this software. Two models developed by Bernard Corfe, butyrate\_v3-1 and SCFA uptake and oxidation v4, were used to simulate the effect of a time course and to model the action of increasing concentrations of SCFAs and glucose in COPASI.

**Results**: Changes in the acetylation patterns of metabolic enzymes were reflected in terms of enrichment scores, percentage coverage, and matching representation. The best model to simulate the effect of a time course was SCFA Uptake and Oxidation v4 model, whereas butyrate\_v3-1 model was the best to simulate the action of increasing concentrations of SCFAs and glucose.

**Conclusions**: SCFAs produce changes in the acetylation patterns of metabolic enzymes. There is an effect of a time course on the levels of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA. Increasing concentrations of SCFAs and glucose also have an effect on CoA and acetyl-CoA levels. Since there are discrepancies between the simulations and *in vitro* data, further studies are needed to elucidate the role of a time course and increasing concentrations of SCFAs and to validate the models used.

Keywords: metabolic control, butyrate, short chain fatty acids, acetylation.

# INTRODUCTION

The importance of short chain fatty acids (SCFAs) is related to their roles in cancer and in regulating metabolism. For example, it has been shown that the short chain fatty acid (SCFA) butyrate stimulates cell cycle arrest and the terminal differentiaion of colorectal cancer cells ([1](#_ENREF_1)). This SCFA has also been linked to a shift in the expression of tumour growth proteins in human colon tumour tissues ([2](#_ENREF_2)) and in inducing apoptosis in human colon cancer cells ([3](#_ENREF_3)). The role of butyrate in controlling cell metabolism state ([4](#_ENREF_4)) could be associated to the Warburg effect, which is the change from oxidative metabolism to aerobic glycolysis in cancer cells ([5](#_ENREF_5)). Another SCFA, propionate, is also able to induce apoptosis in cancer cells in a synergic fashion with butyrate and pyruvate ([6](#_ENREF_6), [7](#_ENREF_7)).

SCFAs could also potentially produce a change in metabolic pathways by altering their expression, function and/or activity ([8](#_ENREF_8)) as well as their transcription levels ([9](#_ENREF_9)) by modifying global protein acetylation patterns ([10](#_ENREF_10)), altering cellular metabolism, and hence controlling cell fate ([11](#_ENREF_11)), since acetylation might work as a cellular resource to deactivate metabolic enzymes to regulate dynamic cellular processes ([12](#_ENREF_12)). Furthermore, the concentration of metabolites needed to start metabolic pathways, for example SCFAs, can lead to a change in the acetylation patterns of metabolic enzymes as well, pointing to an essential role of acetylation in regulating metabolism in response to the extracellular environment ([11](#_ENREF_11)).

The action of SCFAs can be elucidated using different approaches, like bioinformatics analysis and modelling. This report describes two different investigations: bioinformatics analysis of proteomic data and metabolic modelling. The bioinformatics analysis will be done on a mitochondrial proteomic data set which will be mined in terms of which proteins are detected to be acetylated by qualitatively examining the data. On the other hand, the metabolic modelling will be mined in terms of the action of SCFAs and glucose increasing concentrations on acetyl-CoA and CoA levels and the effect of physiological ranges of SCFAs on a time course.

# Short chain fatty acids

Anaerobic bacteria in the colon regulate the fermentation of dietary components ([13](#_ENREF_13), [14](#_ENREF_14)). SCFAs act as the main anions in colonic contents, and are one of the main products of fibre fermentation ([15](#_ENREF_15)). Another function of SCFAs is the control of gene expression and cell proliferation by lowering pH levels ([16-18](#_ENREF_16)). The main role of SCFAs is related to histone acetylation ([19](#_ENREF_19)) as lysine deacetylase inhibitors (KDACi). In terms of lysine deacetylases (KDACs) inhibition, butyrate is the most effective short-chain fatty acid (SCFA) ([19-21](#_ENREF_19)), followed by propionate, and acetate ([19](#_ENREF_19)). In the colon, acetate is the most abundant SCFA, followed by butyrate and propionate. Table 1 summarises physiological ranges, and the function of the three most abundant SCFAs.

|  |  |  |  |
| --- | --- | --- | --- |
| SCFA | Physiological ranges | Structure | Physiological functions |
| Butyrate | 1-20mM |  | Stimulates cell cycle arrest and terminal differentiation of colorectal cancer cells ([1](#_ENREF_1)) by inhibiting KDACs ([4](#_ENREF_4), [22](#_ENREF_22))andDNA\* synthesis ([23](#_ENREF_23)).  Down-regulates Sp1 binding and up-regulates Sp3 binding ([24](#_ENREF_24)) leading to cell cycle arrest ([22](#_ENREF_22)), and apoptosis ([25](#_ENREF_25)).  The “Butyrate paradox” is related to contrasting effects on the proliferation of normal and cancerous colonocytes. *In vitro* butyrate restrains cell proliferation enhancing differentiation and apoptosis in colon cancer cells ([26-30](#_ENREF_26)). *In vivo* butyrate is related to tumour development ([31](#_ENREF_31), [32](#_ENREF_32)) and incidence ([33](#_ENREF_33), [34](#_ENREF_34)). |
| Propionate | 1-20mM |  | Cytotoxic in the nucleus at a higher than physiological concentrations (68.3mM) ([20](#_ENREF_20)).  Its KDACs inhibitory effect is not as potent as butyrate ([20](#_ENREF_20), [21](#_ENREF_21)).  Inhibits HDAC8\*, HDAC6\*, and FB188 HDAH\* in preparations derived from rat, human and bacterial sources ([35](#_ENREF_35)). |
| Acetate | 1-52mM |  | Not very potent HDACi ([19-21](#_ENREF_19)).  Despite this, it has been shown that it is able to inhibit HDAC8, HDAC6 and FB188 HDAH in preparations derived from rat, human and bacterial sources ([35](#_ENREF_35)). This effect is synergistic with propionate ([35](#_ENREF_35)).  Acetylates rat brain histones, specifically H4\* and H3\* ([36](#_ENREF_36)). |

Table 1. Short Chain Fatty Acids. *This table summarises the known action of each SCFA at physiological ranges. \*See Appendix 1 for abbreviations.*

# I.1 Butyrate

In the distal colon, butyrate, a non-toxic KDACi ([20](#_ENREF_20), [25](#_ENREF_25)), is the main source of energy for colonocytes ([37](#_ENREF_37)), and inhibits deoxyribonucleic acid (DNA) synthesis ([23](#_ENREF_23)) and KDACs ([4](#_ENREF_4), [22](#_ENREF_22)) at physiological ranges, 1-20mM. At these ranges, butyrate also stimulates cell cycle arrest and terminal differentiation of cancer cells ([1](#_ENREF_1)). By inhibiting KDACs, butyrate is also related to down-regulation of specific protein-1 (Sp1) and up-regulation of specific protein-3 (Sp3) binding ([24](#_ENREF_24)). A change in the binding patterns of Sp1 and Sp3 induces p21 upregulation ([25](#_ENREF_25)), whilst enhances Bcl antagonist killer (Bak) expression. Cell cycle arrest is induced by the enhancement of p21 expression ([22](#_ENREF_22)), whereas apoptosis is activated by an enhancement of Bak expression ([25](#_ENREF_25)). Butyrate is not only related to cell cycle regulation and activation of apoptosis, since its concentration levels, as well as its KDACi role could both alter global protein acetylation patterns ([10](#_ENREF_10)). This SCFA is also associated with a shift in the expression of tumour growth proteins in human colon tumour tissues ([2](#_ENREF_2)) and apoptosis in human colon cancer cells ([3](#_ENREF_3)).

As reviewed by Burgess, 2012 ([5](#_ENREF_5)) and Gonçalves and Martel, 2013 ([38](#_ENREF_38)) there are different effects of butyrate in the proliferation of normal and cancerous colonocytes. The contradiction in the effects of butyrate are also related to its different effects *in vivo* and *in vitro* as reviewed by Csordas, 1996 ([30](#_ENREF_30)) and Lupton, 2004 ([39](#_ENREF_39)). The “butyrate paradox” explains the contradictory effects of butyrate, since *in vitro* butyrate increases differentiation and apoptosis in colon cancer cells ([26-29](#_ENREF_26)), whereas increases colon neoplasia ([31](#_ENREF_31)), or reduces the number of colon tumours in rats ([32](#_ENREF_32)) *in vivo,* depending on its concentration.

It is interesting that the findings described above point butyrate to have a “chemopreventive role” in colon cancer cells *in vitro*, whilst *in vivo* it has the opposite effect. Discrepancies in the effect of butyrate are not only related to *in vitro* and/or *in vivo* conditions but also to its concentration, which varies depending on the type of fibre consumed ([32](#_ENREF_32)). For example, when the source of fibre is either oat bran or wheat bran, butyrate production is higher in rats consuming oat bran ([32](#_ENREF_32)). This high production of butyrate (18mM) enhances colon neoplasia as well as the incidence of colon tumours when azoxymethane (AOM), an intestinal carcinogenesis inducer, is used ([32](#_ENREF_32)). Alternatively, when the fibre source is wheat bran, there is a low production of butyrate, resembling the physiological ranges (1-20mM), which decreases the incidence of colon tumours ([32](#_ENREF_32)). The effect of butyrate when AOM is used is also controversial, as it has been suggested that there is no effect on inhibiting intestinal carcinogenesis induced by AOM ([34](#_ENREF_34)). This observation has also been made based on findings that no changes in p21 expression or the level of apoptosis were observed ([34](#_ENREF_34)). These findings contrast to what happens *in vitro*, as previously shown by Chirakkal *et al.,* 2006 ([25](#_ENREF_25)). Another intestinal neoplasia inducer, 1, 2-dimethylhydrazine was used by Freeman H, 1986 ([31](#_ENREF_31)) to show the induction of colonic neoplastic development by butyrate with no incidence in small bowel tumours in rats. A lack of effect in terms of tumour incidence and number was also shown by Deschner *et al.,* 1990 ([33](#_ENREF_33)) by using tributyrin.

The Warburg effect, the change from oxidative metabolism to aerobic glycolysis in cancer cells, might be the reason of the “butyrate paradox” ([5](#_ENREF_5)). Burgess D, 2012 ([5](#_ENREF_5)) propose this based on Donohoe *et al*., 2012 study ([4](#_ENREF_4)). This study was focused in the role of butyrate on histone acetylation. According to their findings, this role is more complex than what was established before. For example, whilst butyrate is being metabolised to acetyl-CoA in the mitochondria, it stimulates lysine acetyltransferases (KATs) activity, enhancing the acetylation of histones. This role of butyrate stimulating KATs is combined with its traditional role as KDACi. The new role of increasing KATs activity is coupled to another remarkable finding, that butyrate can control cell metabolism state. They suggested that based on the metabolic state of the cell, presence or absence of the Warburg effect, the control of cellular metabolism can be achieved by butyrate and acetyl-CoA, depending on their concentrations in the nucleus. The concentrations of butyrate in the nucleus determine whether butyrate inhibits KDACs or enhances KATs, controlling the expression of different target genes. Not only the concentration, or if butyrate is used *in vivo* or *in vitro*,or the mitochondrial metabolism are the only features related to the “butyrate paradox”. An alternative feature that might help to elucidate this paradox is the effect of butyrate on cancer development which depends on the time of administration, since there are different outcomes in the development of the polyps depending on the time point they are treated. For example, the effect of butyrate on cancer development is an alternative feature that might help to elucidate the “butyrate paradox”. This effect depends on the time of administration, since there are different outcomes in the development of the polyps depending on the time point they are treated. As proposed by Lupton J, 2004 ([39](#_ENREF_39)) at early stages of polyp formation a protective effect of butyrate is possible, whereas there is no protective effect in polyps turning into carcinoma.

Corfe B, 2012 ([40](#_ENREF_40)) proposed that the inhibitory mechanism of butyrate is related to the inhibition of deacetylation rather than to the inhibition of deacetylase enzymes, based on Candido *et al.,* 1978 ([41](#_ENREF_41)) original study. The way in which the inhibition of deacetylation takes place is due to a negative feedback, or a product inhibition event. The combination of a fatty acid and its ligated acyl-CoA form are substrates for protein acylation and are the products of deacetylation. The abundance of these acyl-CoA forms will consequently change the acylation and deacetylation equilibrium to the acylated state.

# I.2 Propionate

Propionate is a SCFA that usually has a low cytotoxic effect, but when used at a higher than physiological concentrations (68.3mM) has been shown to have a cytotoxic effect, damaging DNA ([20](#_ENREF_20)). The effect of propionate in terms of KDACs inhibition *in vitro* is not as potent as butyrate ([20](#_ENREF_20), [21](#_ENREF_21)). Because of this low effect in terms if KDACs inhibition propionate is often used at high concentrations, producing its cytotoxic effects in the nucleus ([20](#_ENREF_20)). Even its KDACs inhibitory effect is not as potent as butyrate, it is possible for propionate to reach levels high enough as to inhibit KDACs in the colon ([20](#_ENREF_20)). Moreover, if acyl moieties and lysine residues are modified to alter the interactions among proteins, the inhibitory role of propionate is achieved ([35](#_ENREF_35)). In order to identify the active sites of histone deacetylases (HDACs) and HDACs-like enzymes, Riester *et al.,* 2004 ([35](#_ENREF_35)) did these modifications in four different preparations derived from rat, human and bacterial sources. Their findings showed that HDACs, specifically HDAC8, HDAC6, and FB188 HDAH (KDAC6-like enzyme from *Bordetella/Alcaligenes* strain FB188) were inhibited by sodium propionate and sodium acetate at higher than physiological concentration (50mM). Propionate also induces apoptosis in a synergic fashion with butyrate and pyruvate in the Michigan Cancer Foundation-7 (MCF-7) breast cancer cell line ([6](#_ENREF_6)). The synergistic effects of propionate and butyrate are also able to induce apoptosis in Caco-2 cells, producing G2-M arrest ([42](#_ENREF_42)) by shifting D-glucose metabolism and redox state.

# I.3 Acetate

Even acetate has been broadly described as a not very potent KDACi ([19-21](#_ENREF_19)), when peptidic substrates, lysine residues and acyl moieties are modified, an inhibitory effect of acetate on HDACs is achieved in four different preparations derived from rat, human and bacterial sources ([35](#_ENREF_35)). Riester *et al.,* 2004 ([35](#_ENREF_35)) showed that there is a synergistic effect of acetate and propionate at 50mM in terms of HDAC8, HDAC6, and FB188 HDAH inhibition.

The inhibitory effect of acetate on HDACs has also been shown to be *in vivo*, as Soliman and Rosenberger (2011) ([36](#_ENREF_36)) showed an increase in the acetylation state of rat brain histones at lysine residue 8 and 16 of histone-4 (H4), and at lysine residue 9 of histone-3 (H3). The effect of acetate is exclusively related to HDACs inhibition, since it does not have any effect on KATs, enhancing histone acetylation by a reduction in HDACs activity and expression due to acetyl-CoA production derived from this SCFA.

# I.4 Other SCFAs

Valerate, isobutyrate and isovalerate are other SCFAs to consider. The position of the last carbon is how isobutyrate and isovalerate differ structurally from butyrate and valerate. This carbon is branched and attached to either the second carbon in isobutyrate, or to the third carbon in isovalerate, and not in a straight line as in butyrate and valerate (Figure 1).

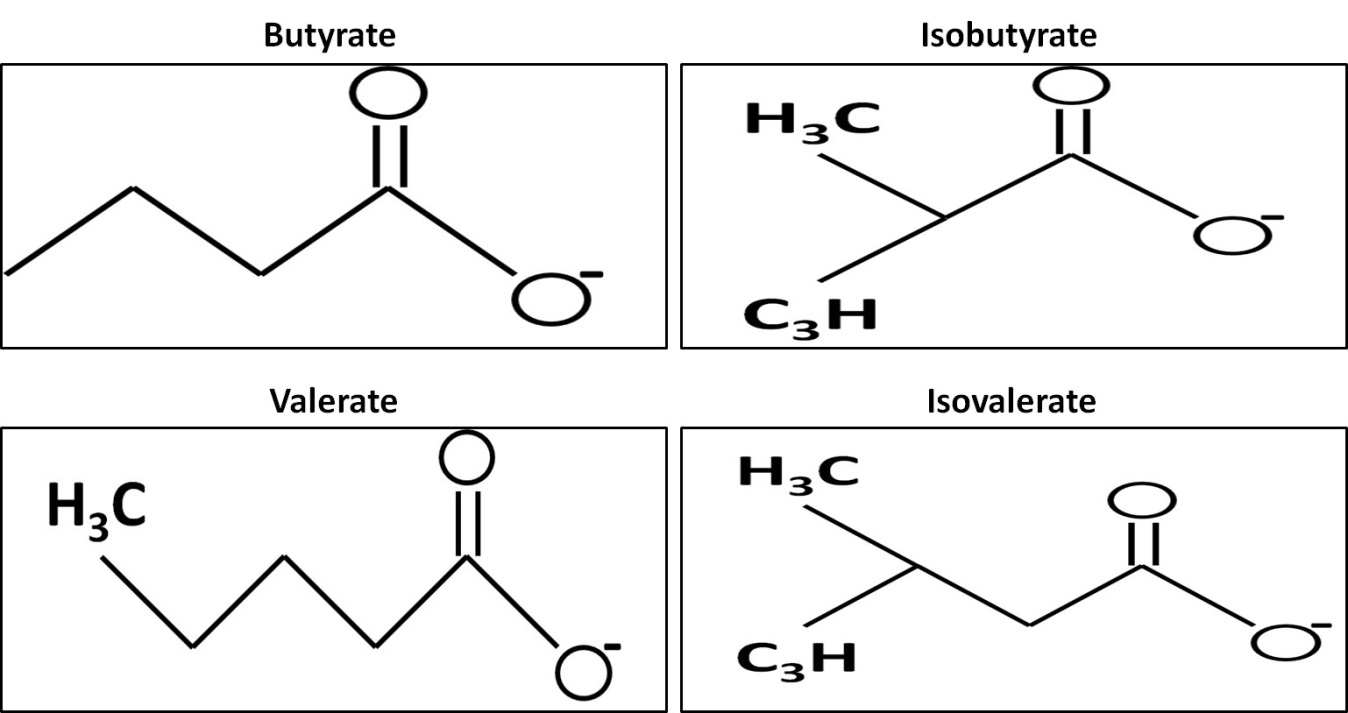


Figure 1. Structural differences in butyrate, isobutyrate, valerate and isovalerate. *Butyrate (top left), isobutyrate (top right), valerate (bottom left) and isobutyrate (bottom right). Note that butyrate and isobutyrate have the same formula C4H7O2, as wells as valerate and isovalerate C5H9O2.*

The effects of valerate on HCT116 human colorectal carcinoma cells are different from the effects of butyrate. Kilner *et al.,* 2012 ([43](#_ENREF_43)) showed that valerate is related to the control of protein expression and cell cycle arrest (cell cycle control), enhancing a shift in the β-tubulin isotypes expression, which modifies the microtubules ([43](#_ENREF_43)), because the correct assembly of these microtubules depends on the composition and expression of their α and β-tubulin subunits and is essential for their function ([44](#_ENREF_44)). This is important since microtubules are a component of the cytoskeleton related to different cellular processes.

Unlike those SCFAs previously mentioned, isobutyrate seems to have no effect, or not to play an important role as a KDACi. Nevertheless, at 8mM concentrations it is possible for isobutyrate to inhibit the induction of egg white messenger ribonucleic acid (mRNA) ([45](#_ENREF_45)). This concentration is considerably higher, more than ten times, than that needed to cause a 50% inhibition by butyrate or propionate ([45](#_ENREF_45)). On the other hand, the availability of volatile fatty acids increases the concentrations of isovalerate and acetate, suggesting a dose-dependent manner to control and up-regulate bacteria and enzymes in the steer gut ([46](#_ENREF_46)).

# II. Metabolic outcomes of the alteration in the acetylation patterns of metabolic enzymes

There is a delicate balance amongst metabolic pathways that could be altered by a shift in their acetylation patterns ([9](#_ENREF_9), [11](#_ENREF_11)), since the enzymes in these pathways are substrates of acetylation ([11](#_ENREF_11)). A decrease in histone acetylation, the inhibition of the enzymatic activity and metabolic outcomes, are the effects of an alteration in the acetylation patterns of metabolic enzymes ([11](#_ENREF_11)). SCFAs could potentially produce a change in metabolic pathways by altering their expression and/or function ([8](#_ENREF_8)) as well as their transcription levels ([9](#_ENREF_9)), altering cellular metabolism, and hence controlling cell fate ([11](#_ENREF_11)). Not just the aforementioned features can shift the acetylation patterns of metabolic enzymes, since the concentration of metabolites needed to start metabolic pathways, like SCFAs, can also lead to a change in the acetylation patterns of metabolic enzymes, pointing to an essential role of acetylation in regulating metabolism in response to the extracellular environment ([11](#_ENREF_11)). For example, it has been shown that acetate and butyrate are the most important substrates for de novo lipid synthesis in colonic epithelial cells in rats ([47](#_ENREF_47)). SCFAs are also oxidized to CO2 and ketone bodies through β-oxidation ([48](#_ENREF_48)) and affect glucose metabolism (gluconeogenesis) in rat hepatocytes ([49](#_ENREF_49)). Moreover, the acetylation of substrate binding sites could be a cellular mechanism to control the activity of metabolic enzymes, since acetylation might work as a cellular resource to deactivate metabolic enzymes to regulate dynamic cellular processes ([12](#_ENREF_12)), similar to the regulation by phosphorylation.

# II.1 Implications of lysine residue acetylation in metabolic enzyme function

# II.1.i Glycolysis

The metabolic pathway that converts glucose to pyruvate is called glycolysis. In healthy cells, its main function is the production of energy in the form of adenosine triphosphate (ATP) and the reduced form of nicotinamide adenine dinucleotide (NADH). The ten steps in glycolysis involve ten major enzymes, with many isoforms of each enzyme having being identified. The most common isoforms, their substrates and products are summarised in Figure 2.

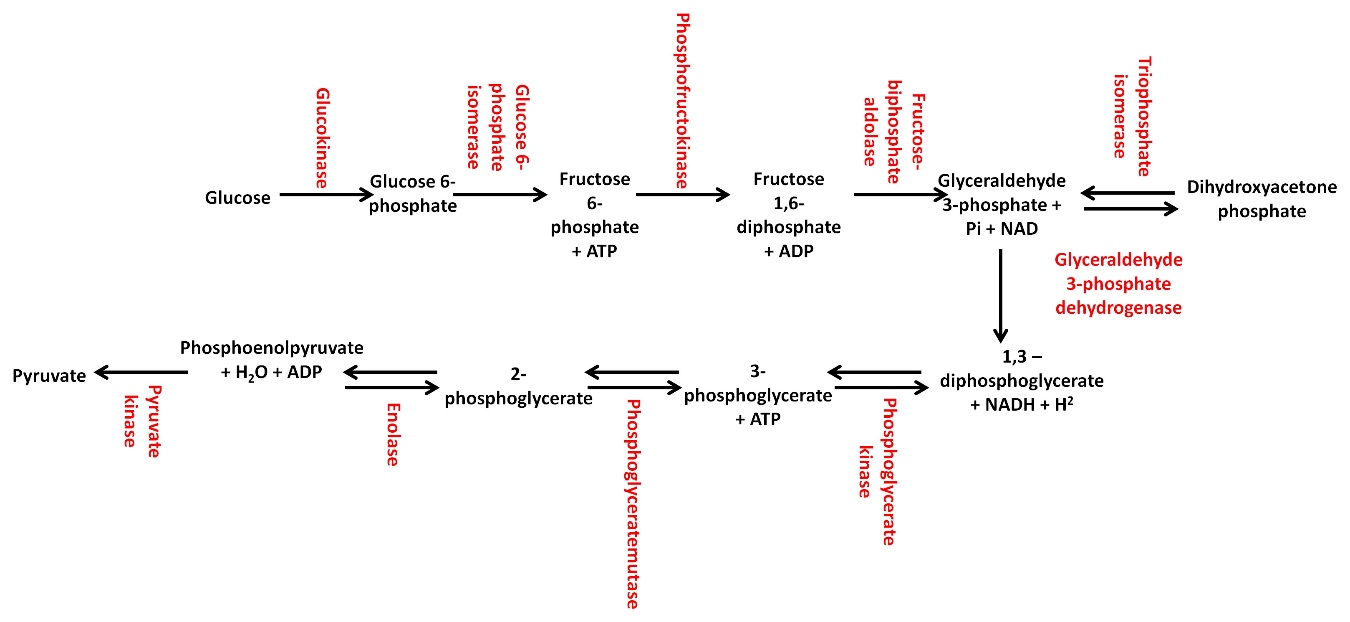


Figure 2. Enzymes and reactions in glycolysis. *Metabolic enzymes in this pathway are shown in red bold font; metabolic products in this pathway are shown in black bold font.*

The function of the glycolytic enzymes can be regulated by post-translational modifications (PTMs), which include acetylation, phosphorylation, glycosylation and ubiquitination, among many others. A change in the acetylation patterns of 4 glycolytic enzymes, phosphofructokinase, fructose-bisphosphate aldolase, glyceraldehyde 3-phosphate dehydrogenase and phosphoglycerate mutase have been shown to have an effect on their activity and/or expression (see Table 2 for a summary of the effects of acetylation on these enzymes). All these enzymes are annotated by the PhosphoSitePlus database, [www.phosphosite.org](http://www.phosphosite.org), accessed on 22-05-2014 to be acetylated. See Appendix 2 for all the acetylated lysine residues.

|  |  |  |
| --- | --- | --- |
| Enzyme | Acetylated Lys residues | Effect |
| Phosphofructokinase | 3 ([50](#_ENREF_50)) | Enhances activity ([50](#_ENREF_50)) |
| Fructose-bisphosphate aldolase | 147 ([12](#_ENREF_12)) | Abolishes activity ([12](#_ENREF_12)) |
| Glyceraldhyde 3-phosphate dehydrogenase | 120 ([12](#_ENREF_12)), 117 ([51](#_ENREF_51)), 160 ([52](#_ENREF_52)), 227 ([51](#_ENREF_51)), 251 ([51](#_ENREF_51)), 254 ([53](#_ENREF_53)) | Stimulates catalytic activity ([52](#_ENREF_52))  Regulates activity and function ([53](#_ENREF_53)) |
| Phosphoglycerate mutase | 251 ([54](#_ENREF_54)), 253 ([54](#_ENREF_54)), 254 ([54](#_ENREF_54)) | Regulatory sites ([54](#_ENREF_54)) |
| Phosphoglycerate kinase | 149 ([55](#_ENREF_55)), 169 ([55](#_ENREF_55)), 291 ([55](#_ENREF_55)) | Affects its activity ([55](#_ENREF_55)) |

Table 2. Effect of acetylation on phosphofructokinase, fructose-bisphosphate aldolase, glyceraldehyde 3-phosphate dehydrogenase and phosphoglycerate mutase.

Some glycolytic enzymes are suitable of being regulated by glucose levels and deacetylation, for example phosphoglycerate mutase. This enzyme is acetylated at lysine residues 251, 253 and 254, but any of these seem to be the active site of this enzyme, since when the aforementioned residues were mutated to arginine (amino acid with a positive charge, mimicking a non-acetylated lysine residue) or glutamine (amino acid with a neutral charge, mimicking an acetylated lysine residue) there were no changes observed in the acetylation profile ([54](#_ENREF_54)). In this enzyme, the acetylations at lysine 251, 253, and 254 residues in human embryonic kidney (HEK) 293 cells are enhanced by the NAD dependant deacetylase sirtuin-1 (Sirt1), pointing these acetylated residues as the regulatory acetylation sites ([54](#_ENREF_54)). The role of Sirt1 in deacetylating phosphpoglycerate mutase was elucidated by knocking-down this and other NAD dependant deacetylase sirtuins, Sirt2 and Sirt3; the activity of this enzyme was 45% higher when Sirt1 was blocked, but not when Sirt2 and Sirt3 were knocked-down. The deacetylation via sirt1 of phosphoglycerate mutase in the aforementioned lysine residues, also down-regulates its catalytic activity. Hallows *et al.,* 2012 ([54](#_ENREF_54)) propose that the mechanism of this regulation is due to an enhancement of the flux through glycolysis via the over-expression of an active variant of phospholgycerate mutase. An accessible 3-phosphoglycerate pool is preserved by attenuating phosphoglycerate mutase activity and restricting flux through glycolysis, enhancing Sirt1 levels (11 fold), leading to a Sirt1-mediated deacetylation and reduction of the activity of this enzyme under glucose restriction. Glucose levels actually control the acetylation of phosphoglycerate mutase since at a high concentration of glucose (5 g/litre for 6 hours) the activity of this enzyme was nearly two fold higher than with none glucose (0 g/litre for 6 hours). This mechanism provides more evidence of the role of acetylation in the activation of metabolic enzymes, offers biochemical evidence of the modulation of glycolysis by the opposed acetylation/deacetylation processes, and reveals Sirt1 deacetylation to be directly controlling phospholgycerate mutase activity. Furthermore, it also contributes to elucidate dynamics of acetylation in terms of regulating acetylation of glycolytic enzymes, energy metabolism, metabolic syndromes and cancer (Warburg effect) via deacetylation by Sirt1. Figure 3 summarises the regulation of phosphoglycerate mutase by glucose concentrations.

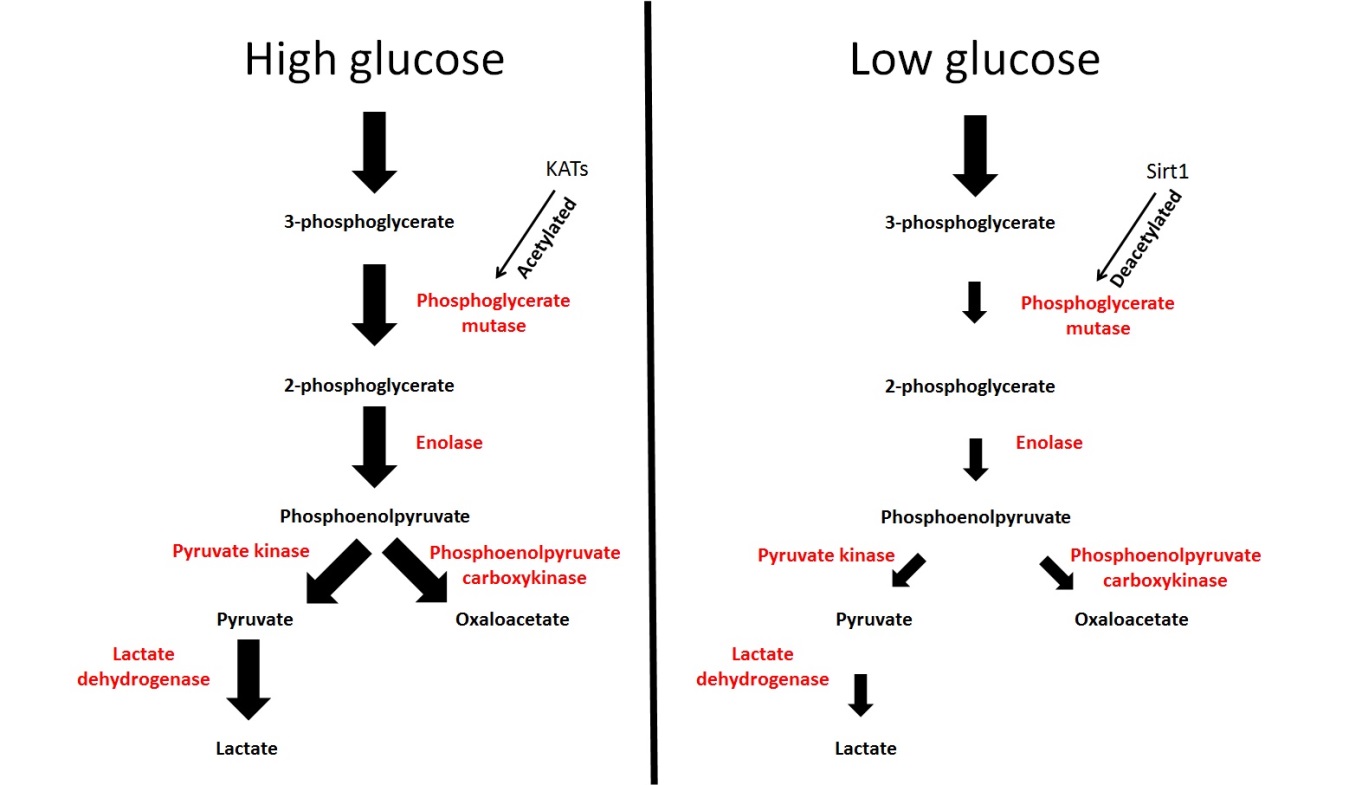


Figure 3. Phosphoglycerate mutase regulation by glucose levels. *Under high glucose conditions, phosphoglycerate mutase is acetylated via KATs enhancing its activity and the production of lactate and the tricarboxylic acid (TCA) cycle substrate oxaloacetate. Under low glucose conditions phosphoglycerate mutase is deacetylated via Sirt1 decreasing its activity and the production of lactate and the TCA cycle substrate oxaloacetate. Metabolic enzymes are shown in red bold font; metabolic products are shown in black bold font. Corregir esta figura, poner acetylated en KATs y deacetylated en Sirt1.*

# III.1.ii The TCA cycle

The TCA cycle is a series of chemical reactions to obtain energy from acetyl-CoA by means of the aerobic catabolism of carbohydrates, amino acids, and fatty acids. In this cycle, energy is produced in the mitochondria by oxidisation of acetate into CO2. The sources of acetate in the TCA cycle are carbohydrates, amino acids and fatty acids. There are ten steps in the TCA cycle, each of them catalysed by different enzymes. Although there are ten major enzymes, many isoforms of each enzyme have been identified. The ten main enzymes that take part in the TCA cycle, and the reactions catalyzed by them are summarised in Figure 4.

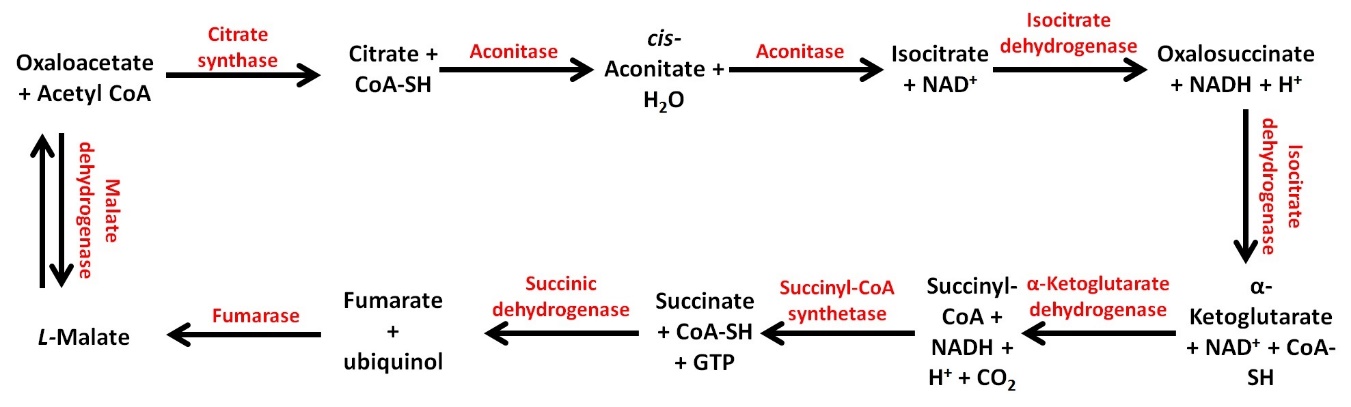


Figure 4. Enzymes and reactions in the TCA cycle. *Metabolic enzymes in this pathway are shown in red bold font; metabolic products in this pathway are shown in black bold font.*

Although the activity of the enzymes in the TCA cycle is controlled by acetylation, just the acetylation patterns of 3 enzymes in this pathway have been shown to have an effect on their activity and/or expression (see Table 3 for a summary of the effects of acetylation on these enzymes). These enzymes are isocitrate dehydrogenase, succinate dehydrogenase, and malate dehydrogenase, annotated by the PhosphoSitePlus database, [www.phosphosite.org](http://www.phosphosite.org), accessed on 22-05-2014 to be acetylated. See Appendix 2 for all the acetylated lysine residues.

|  |  |  |
| --- | --- | --- |
| Enzyme | Acetylated Lys residues | Effect |
| Isocitrate dehydrogenase | 413 ([56](#_ENREF_56)) | Inhibits activity ([56](#_ENREF_56), [57](#_ENREF_57)) |
| Succinate dehydrogenase | 179 ([58](#_ENREF_58)), 182 ([58](#_ENREF_58)), 250 ([58](#_ENREF_58)), 335 ([58](#_ENREF_58)), 480 ([58](#_ENREF_58)), 485([58](#_ENREF_58)), 498 ([58](#_ENREF_58)), 547 ([58](#_ENREF_58)), 550 ([58](#_ENREF_58)), 598 ([58](#_ENREF_58)), 608 ([58](#_ENREF_58)), 624 ([58](#_ENREF_58)), 633 ([58](#_ENREF_58)) | Slows down the TCA cycle ([59](#_ENREF_59))  Regulates activity ([58](#_ENREF_58)) |
| Malate dehydrogenase | 118 ([60](#_ENREF_60), [61](#_ENREF_61)), 121 ([60](#_ENREF_60), [61](#_ENREF_61)), 298 ([60](#_ENREF_60), [61](#_ENREF_61)), 185 ([62](#_ENREF_62)), 301 ([60-62](#_ENREF_60)), 307 ([60-62](#_ENREF_60)), 314 ([60-62](#_ENREF_60)) | Activates the enzyme ([62](#_ENREF_62))  Regulates adipogenic differentiation ([60](#_ENREF_60), [61](#_ENREF_61)) |

Table 3. Effect of acetylation on isocitrate dehydrogenase, succinate dehydrogenase, and malate dehydrogenase.

As with glycolytic enzymes, enzymes in the TCA cycle are suitable of being regulated by calorie restriction (CR) and glucose levels, for example isocitrate dehydrogenase. The acetylation status of isocitrate dehydrogenase is related to its activity ([56](#_ENREF_56), [57](#_ENREF_57)). According to Yu *et al.,* 2012 ([56](#_ENREF_56)) *in vitro* and *in vivo* findings, Sirt3 deacetylation of this enzyme at lysine residue 413, pointed as a regulatory site, stimulates its catalytic activity by 25-fold under calorie restriction (CR) and glucose restriction. On the other hand, its acetylation at the same lysine residue decreases its activity as much as 44-fold under the same dietary restrictions. In HEK293 cells, lowering glucose concentrations (5, 3, 1, and 0 g/litre for 6 hours) produced a decrease in the acetylation levels of lysine residue 413, and an increase in Sirt3 levels, which was inversely related to isocitrate dehydrogenase activity. These findings link glucose levels, nutrient status, and isocitrate dehydrogenase activity directly. Under CR, the acetylation levels of lysine residue 413 is decreased by 3-fold when compared to control diet, matching the 3-fold increase observed in Sirt3 levels under CR. When lysine residues were mutated to glutamine, a 22-fold decrease in isocitrate dehydrogenase activity was observed, whereas when it was mutated to arginine there was a slight change in the activity of this enzyme, suggesting that the positive charge at lysine residue 413 is essential for its catalytic activity. NAD-dependant Sirt3-mediated deacetylation of this enzyme fully restores its catalytic activity, which is important because of the role of Sirt3 as cellular protection from oxidative stress produced by reactive oxygen species (ROS) throughout periods of enhanced oxidative metabolism, which is dependent on isocitrate dehydrogenase activity. When isocitrate dehydrogenase was knocked-down these protective effects of Sirt3 were abolished or considerably reduced, whereas when lysine residue was mutated to arginine, it was able to protect Sirt3 knocked-down mouse embryonic fibroblasts (MEF) from oxidative stress via increased levels of reduced glutathione, pointing isocitrate dehydrogenase as having a key role in preserving the redox equilibrium in mitochondria and a regulatory role of reversible acetylation of lysine residue 413.

Similarly to Yu *et al.,*2012 ([56](#_ENREF_56)), Schlicker *et al.,* 2008 ([57](#_ENREF_57)) also found that Sirt3 deactylation is NAD-dependent and activates isocitrate dehydrogenase, an important point to regulate flux through the TCA cycle. They found out by an enzyme-linked immunosorbent assay (ELISA) system that Sirt3 efficiently deacetylates the substrate of this enzyme when it was isolated from mitochondria. Another mitochondrial member of the sirtuins deacetylases family, Sirt5, did not significantly decrease the acetylation levels of isocitrate dehydrogenase, even when used at high concentrations. Isocitrate dehydrogenase enhanced activity is Sirt3 concentration-dependent, since increasing concentrations of this deacetylase increase the activity of this enzyme. Based on the mitochondrial colocalization of Sirt3 and isocitrate dehydrogenase, a role of this enzyme as a physiological substrate for Sirt3 leading to its subsequent activation by deacetylation is very likely.

The activity of isocitrate dehydrogenase is also controlled by deacetylation, the opposite process of acetylation. In HEK293 and MEF cells deacetylation of isocitrate dehydrogenase, via Sirt3 inhibits its activity under CR and glucose restriction ([56](#_ENREF_56)). Under the aforementioned conditions, the lysine 413 residue is deacetylated leading to the inhibition of the enzyme ([56](#_ENREF_56)). Based on this finding, it is very likely that there is a link between cancer cell metabolism and the Sirt3-induced isocitrate dehydrogenase inhibition, since this is an important mechanism to control metabolic enzymes. Figure 5 summarises the regulation of isocitrate dehydrogenase by glucose concentrations.

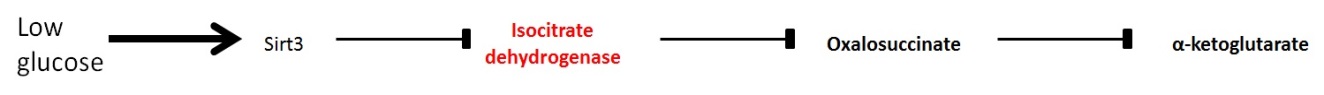


Figure 5. Isocitrate dehydrogenase regulation by glucose levels. *Under low glucose conditions, the activity of Sirt3 is enhanced, increasing the deacetylation of isocitrate dehydrogenase, hence decreasing its enzymatic activity (production of TCA cycle substrates oxalosuccinate and α-ketoglutarate). Metabolic enzyme is shown in red bold font; products are shown in black bold font.*

Another TCA cycle enzyme, malate dehydrogenase, is suitable of being regulated by the energy state of the cell. Kim *et al.,* 2012 and 2013 ([60](#_ENREF_60), [61](#_ENREF_61)) showed that malate dehydrogenase expression is enhanced, two-fold, during adipogenesis and that acetylation at lysine residue 118 is affected by the energy state of the cell and subsequently plays a role in the regulation of adipogenic differentiation, as there is a dramatic increase, up to 6-fold, in the acetylation levels of this enzyme in the last stage of adipogenesis. The overexpression of wild-type malate dehydrogenase in 3T3-L1 preadipocytes induced a significant increase in the cells going through adipogenesis as well as in lipid accumulation, whereas when acetylation at putative lysine acetylation sites 301, 307 and 314 is blocked, there was a significant reduction in adipogenic differentiation, showing once again the importance of the acetylation of this enzyme in adipogenic differentiation. The acetylation of ectopic malate dehydrogenase had an increase of 50% during adipogenesis, being correlated with its increase in enzymatic activity. As well as its acetylation levels, ectopic expression of malate dehydrogenase was also increased in 3T3-L1 preadipocytes, being also related to induced adipogenesis, since an enhanced differentiation was observed in the cells; this enhancement is in terms of the time that cells take to be fully differentiated which downs from 12 to 8 days after differentiation is initiated. Interestingly, they found that glucose increasing concentrations, 5.5 to 50mM, are also related to adipogenic differentiation and malate dehydrogenase acetylation, since there is a correlation amongst adipogenic differentiation, acetylation levels of this enzyme and the higher the extracellular concentrations of glucose are. When lysine acetylation sites 118, 121, and 298 were mutated to arginine or glutamine, a loss in the capacity of the cells to go through adipogenic differentiation was induceoxidatived but had no effect on its enzymatic activity, pointing malate dehydrogenase acetylation to be closely related to adipogenic differentiation. Malate dehydrogenase acetylation also increased its enzymatic activity and subsequently enhancement in the intracellular levels of the reduced form of nicotinamide adenine dinucleotide phosphate (NADPH) and acetyl-CoA in the synthesis of fatty acids by accelerating the shuttle of citrate, suggesting that these processes are also regulated by acetylation during adipogenesis. Altogether, these findings show that adipogenic differentiation may be controlled by the acetylation patterns of malate dehydrogenase and the acetylation of this enzyme is one of the cross-talk mechanisms between the energy levels within the cell, fatty acid biosynthesis, and adipogensis ([61](#_ENREF_61)). Figure 6 summarises the regulation of isocitrate dehydrogenase by glucose concentrations.

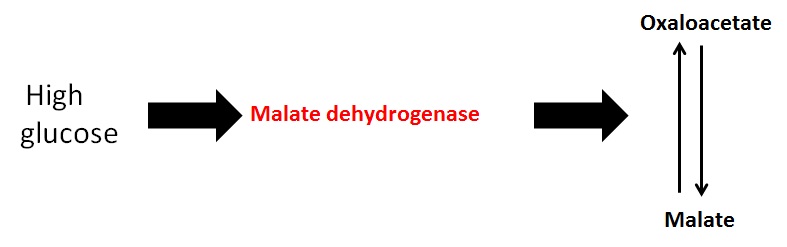


Figure 6. Regulation of malate dehydrogenase by glucose levels. *Under high glucose conditions, the activity of malate dehydrogenase is enhanced, increasing its enzymatic activity (production of malate and oxaloacetate).*

# II.1.iii β-oxidation

β-oxidation is the metabolic pathway by which fatty acids are transformed into acetyl-CoA. Once acetyl-CoA is produced, it enters the TCA cycle. There are five steps in β-oxidation, each of them taken by different enzymes. Although there are five major enzymes, many isoforms of each enzyme have been identified, showing different chain-length specificities. The five main enzymes that take part in β-oxidation and the reactions catalyzed by them are summarised in Figure 7.

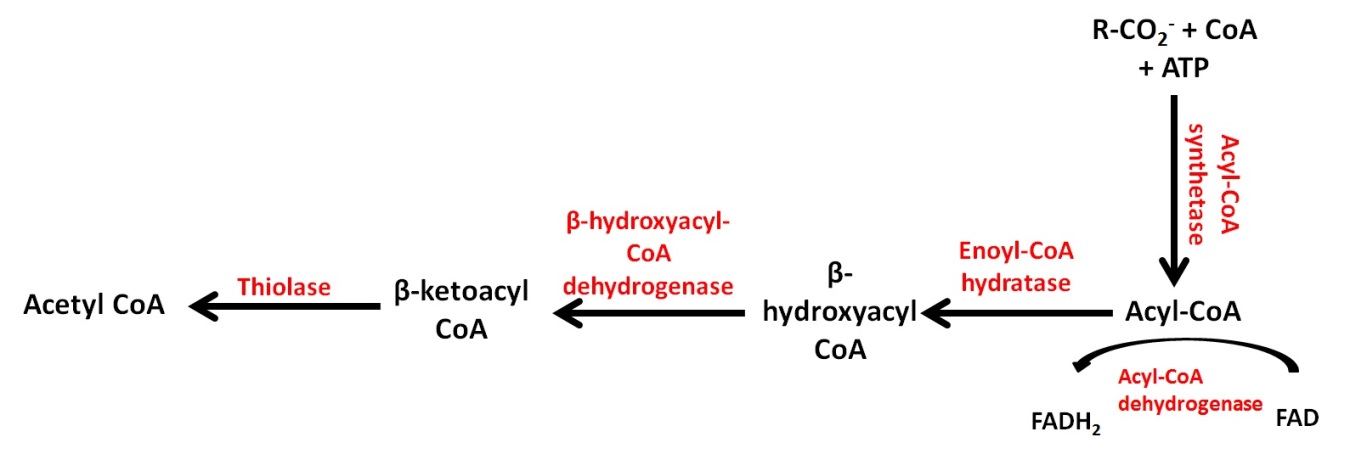


Figure 7. Enzymes and reactions in β-oxidation. *Metabolic enzymes in this pathway are shown in red bold font; metabolic products in this pathway are shown in black bold font.*

Even the activity of the β-oxidation enzymes is controlled by acetylation, just the acetylation patterns of 3 enzymes have been shown to have an effect on their activity and/or expression (see Table 4 for a summary of the effects of acetylation on these enzymes). These enzymes are acyl-CoA synthetase, enoyl-CoA hydratase and β-hydroxyacyl-CoA dehydrogenase and are annotated by the PhosphoSitePlus database, [www.phosphosite.org](http://www.phosphosite.org), accessed on 22-05-2014 to be acetylated. See Appendix 2 for all the acetylated lysine residues.

|  |  |  |
| --- | --- | --- |
| Enzyme | Acetylated Lys residues | Effect |
| Acyl-CoA synthetase | 49 ([63](#_ENREF_63)), 52 ([63](#_ENREF_63)), 217 ([63](#_ENREF_63)), 223 ([63](#_ENREF_63)), 241 ([63](#_ENREF_63)), 250 ([63](#_ENREF_63)), 341 ([63](#_ENREF_63)), 387 ([63](#_ENREF_63)), 418 ([64](#_ENREF_64)), 428 ([63](#_ENREF_63)), 504 ([63](#_ENREF_63)), 533 ([63](#_ENREF_63)), 534 ([65](#_ENREF_65)) , 544 ([63](#_ENREF_63)), 562 ([63](#_ENREF_63)), 641([63](#_ENREF_63)), and 649 ([63](#_ENREF_63)) | Inactivates the enzyme ([65](#_ENREF_65))  Enhances β-oxidation pathway ([63](#_ENREF_63)) |
| Enoyl-CoA hydratase | 165 ([11](#_ENREF_11)), 171 ([11](#_ENREF_11)), 346 ([11](#_ENREF_11)), 584 ([11](#_ENREF_11)) | Activates this enzyme ([11](#_ENREF_11)) |
| β-hydroxyacyl-CoA dehydrogenase |  | Activates this enzyme ([11](#_ENREF_11)) |

Table 4. Effect of acetylation on acyl-CoA synthetase, enoyl-CoA hydratase, and β-hydroxyacyl-CoA dehydrogenase.

# II.1.iv Pyruvate metabolism

Pyruvate, along with ATP is the final product of glycolysis. After being produced through glycolysis, pyruvate is aerobically converted by the pyruvate dehydrogenase complex to acetyl-CoA and enters the TCA cycle. Also, when there is not enough oxygen available, and therefore pyruvate cannot be converted into acetyl-CoA, pyruvate is converted to lactate in a reversible reaction by lactate dehydrogenase. The main enzymes involved in pyruvate metabolism are summarised in Figure 8.

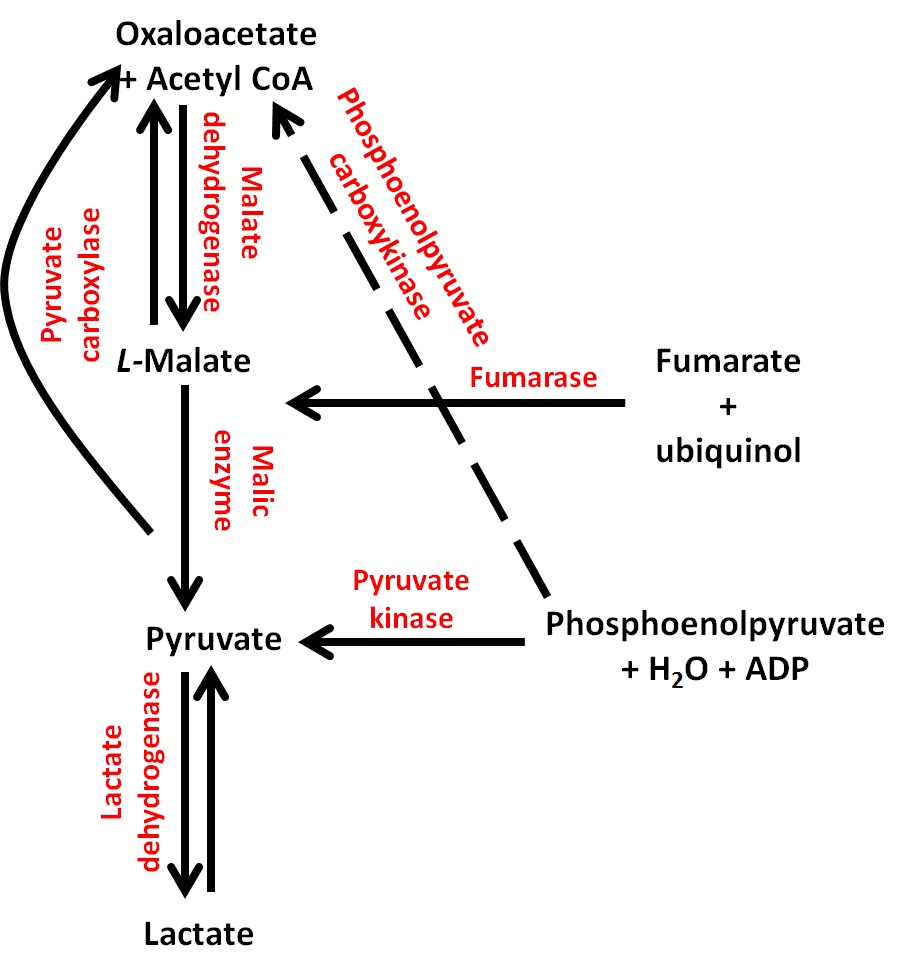


Figure 8. Enzymes and reactions in pyruvate metabolism. *Metabolic enzymes in this pathway are shown in red bold font; metabolic products in this pathway are shown in black bold font.*

Even the activity of the pyruvate metabolism pathway is controlled by acetylation, just the acetylation patterns of 4 enzymes in this pathway have been shown to have an effect on their activity and/or expression (see Table 5 for a summary of the effects of acetylation on these enzymes). These enzymes are pyruvate kinase, lactate dehydrogenase, phosphoenolpyruvate carboxykinase, and pyruvate carboxylase, annotated by the PhosphoSitePlus database, [www.phosphosite.org](http://www.phosphosite.org), accessed on 22-05-2014 to be acetylated. See Appendix 2 for all the acetylated lysine residues.

|  |  |  |
| --- | --- | --- |
| Enzyme | Acetylated Lys residues | Effect |
| Pyruvate kinase | 62 ([66](#_ENREF_66)), 305 ([66](#_ENREF_66)) | Regulates stability and activity ([66](#_ENREF_66)) |
| Phosphoenolpyruvate carboxykinase | 19 ([67](#_ENREF_67)), 70 ([11](#_ENREF_11)), 71 ([11](#_ENREF_11)), 514 ([67](#_ENREF_67)), 594 ([11](#_ENREF_11)) | Regulates stability ([11](#_ENREF_11), [68](#_ENREF_68)) |
| Lactate dehydrogenase | 5 ([69](#_ENREF_69)), 318 ([69](#_ENREF_69)) | Inhibits the enzyme ([69](#_ENREF_69)) |
| Pyruvate carboxylase | 35 ([70](#_ENREF_70)), 39 ([70](#_ENREF_70)), 79 ([70](#_ENREF_70)), 148 ([70](#_ENREF_70)), 152 ([70](#_ENREF_70)), 237 ([70](#_ENREF_70)), 241 ([70](#_ENREF_70)), 316 ([70](#_ENREF_70)), 434 ([70](#_ENREF_70)), 589 ([70](#_ENREF_70)), 717 ([70](#_ENREF_70)), 748 ([70](#_ENREF_70)), 892 ([70](#_ENREF_70)), 969 ([70](#_ENREF_70)), 992 ([70](#_ENREF_70)), and 1909 ([70](#_ENREF_70)) | Regulates activity ([70](#_ENREF_70)) |

Table 5. Effect of acetylation on pyruvate kinase, phosphoenolpyruvate carboxykinase, lactate dehydrogenase, and pyruvate carboxylase.

# Proteomic analysis

The entire set of proteins expressed at a certain time in a tissue, cell or genome is known as the proteome, which is suitable of being studied. Proteomics is the study of this proteome. There are some fields related to proteomics, and functional proteomics is a developing research field in proteomics. The description of cellular mechanisms at the molecular level, as well as the elucidation of the biological role of unidentified proteins are approaches related to the functional proteomics field ([71](#_ENREF_71)). The description of signalling pathways, the complexes in which some proteins are involved, and the activity of proteins are areas in which functional proteomics can be a useful tool in the proteomic field ([72-75](#_ENREF_72)). The fluxes involved in molecular networks, as well as their spatial and temporal properties can also be screened and study in living cells by using functional proteomics ([76](#_ENREF_76)).

The gathering of information via quantitative protein profiling is a useful and important proteomic technique ([77](#_ENREF_77), [78](#_ENREF_78)), particularly when used in its high-throughput mode ([77](#_ENREF_77)). Quantitative protein profiling allows to comprehensibly, accurately and reproducibly identify and quantify proteins contained in any biological sample, relying on quite new technologies ([78](#_ENREF_78)). Reference databases for cells as much as for tissues can be enriched by the information collected and stored using quantitative protein profiling ([77](#_ENREF_77)). Comparative proteomics can be performed by using these databases ([77](#_ENREF_77), [78](#_ENREF_78)). Mass-spectrometry (MS) is one of the most used techniques to study peptides and proteins ([79](#_ENREF_79)) since it is an analytical technique that determines the number of ions in gas-phase and the composition of a ionised sample in mass-to-charge ratios. MS is a technique that has been in vogue during the last 20 years and has made available complete sequence databases, as well as improvement and development of computational data analysis and instrumentation. Not only MS, but also large-scale proteomics analysis, like for example stable isotope labelling techniques, have also been used for protein profiling ([79](#_ENREF_79)). These and many other analysis have been possible due to the availability of these databases, and the computational and instrumentation development ([79](#_ENREF_79)). MS is a technique with a broad range of variations, which include: Tandem MS (MS/MS), in which the first device to separate the sample is another MS; Matrix-Assisted Laser Desorption/Ionization-Time of Flight MS (MALDI-TOF), which performs a soft ionization of fragile samples, impeding their fragmentation; and Liquid chromatography-mass spectrometry (LC-MS) or Liquid chromatography-tandem mass spectrometry (LC-MS/MS), a combined technique that physically separates a sample by liquid chromatography and analyses it by MS.

There are many other techniques with different sample separation and comparison methods, and label/label-free protein profiling techniques, with its specific advantages and limitations. The broad ranging techniques include Isotope-coded affinity-tag-based protein profiling (ICAT), Stable Isotopic Labelling by Amino Acids in Cell Culture (SILAC), Isobaric tagging for relative and absolute quantification (iTRAQ), ProteomeLab PF2D (2D LC), Multi-dimensional Protein Identification Technology (MudPIT), Label Free Quantitation (LFQ), and Tandem mass tags (TMT). In order to detect PTMs, a combination of the techniques mentioned above can be used, although in the proteomic scale, PTMs can also be studied by using the advances in based proteomics quantitative MS. Table 6 summarises these techniques.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Method | Labelling | Separation | Sample comparison | Limits | Pros |
| ICAT | C12/C13 at cysteine residues | CEX, Avidin and LC MS/MS | 2-plex MS quantitation | Only detects cysteine-containing proteins | Simplifies complex mixtures; no PTMs |
| SILAC | Stable isotopes labelled peptides | CEX and LC-MS/MS | 2-plex and 3-plex MS quantitation | Incomplete production of derivates that could create side products, limiting the sensitivity of the analysis. | Labelling occurs early on in the sample preparation |
| iTRAQ (4 or 8-plex) and TMT 2, 6 or 10-plex) | Isobaric tags at N-terminus and ε-N of lysine residues | CEX and LC-MS/MS | 2, 6, 4, 8 and 10-plex MS/MS quantitation based on intensity of reported ions. | Quantification on MS/MS selected peptides only | 8, 10-plex allows greater multiplexing |
| 2D LC | None | 2D chromato-focusing and RP HPLC | Quantity pair wise RP fractions by UV | Number of samples that can be prepared | Comparison of 2 samples at >500μg |
| MudPIT | None | CEX and LC-MS/MS | 2D LC approach used to catalogue proteins from complex samples | Addition prefractioning aids in identification of proteins in the proteome | No direct quantification  Ideal for serum/plasma discoveries |
| LFQ | None | LC-MS | LC-MS based disease biomarker discovery tool | Limited or no initial protein identification. Repeat run often used for identification | Countless samples can be compared |
| HDMS | None | LC-MS | Mass of peptides and fragments are measured and the spectra of every component are aligned in time | Samples completely catalogued in an HDMS or analysis | Data are recorded without discrimination or pre-selection, hence samples are completely catalogued in a single MS or HDMS analysis |

Table 6. ****Protein Profiling Techniques. *This table summarises some of the techniques currently in use. CEX, cation exchange; RP HPLC, Reversed Phase High Performance Liquid Chromatography; RP, Reverse Phase; HDMS, High Definition MS.*****

As aforesaid, protein profiling techniques can be used to study PTMs, including acetylation. In terms of acetylation sites, Zhao *et al.*, 2010 ([11](#_ENREF_11)) did an MS analysis and found that nearly every single metabolic enzyme is found to be acetylated in nuclear, cytosolic and mitochondrial fractions of human liver tissue. They also quantified the acetylation of β-oxidation enzymes by using iTRAQ. Schlicker *et al.,* 2008 ([57](#_ENREF_57)) did an MS analysis to identify the sites where isocitrate dehydrogenase is acetylated. They were only able to find lysine residue 241 to be acetylated, and were unable to find lysine residue 75 to be acetylated, opposite to Kim *et al.,* 2006 ([80](#_ENREF_80)). In contrast, they identified a novel and additional acetylation site, the double lysine (KK) motif at lysine residues 211 and 212. Schwer *et al.,* 2009 ([81](#_ENREF_81)) identified 72 proteins that may be involved in a wide range of metabolic pathways with altered acetylation patterns (2.5-fold) due to CR via large-scale MS screening, and the acetylated peptides were subsequently analysed by LFQ and MS, finding 287 unique acetylated proteins.

Cimen *et al.,* 2010 ([59](#_ENREF_59)) mapped various acetylated lysine residues in isocitrate dehydrogenase by MS/MS and determined the role of acetylation in the activity of complex II of the electron transport chain in Sirt3 knockout mice. An MS analysis to elucidate which proteins interact with Sirt3 and to identify subunits of complex II and V of the electron transport chain was done by Finley *et al.,* 2011 ([58](#_ENREF_58)). By using MS/MS they also identified 13 acetylation sites on succinate dehydrogenase in mouse liver mitochondria, six of which are residues reported to be acetylated for the first time. Based on the lack of information about the role of HATs/KATs and HDACs/KDACs in controlling several essential functions via non-histone substrates in *Saccharomyces cerevisiae*, Lin *et al.,* 2009 ([67](#_ENREF_67)) analysed by tandem mass spectrometry yeast proteome microarrays, validating and identifying many non-chromatin substrates of the crucial NuA4 complex. They also determined two acetylation sites in phosphoenolpyruvate carboxykinase, at lysine residues 19 and 514. Lundby *et al.*, 2012 ([12](#_ENREF_12)) used MS/MS measurements to generate an atlas of lysine acetylation sites of 16 different tissues in rats. The acetylome data set they generated increased four times the existing number of lysine acetylation sites and two times the number of acetylated proteins. They were able to map in between 3000 and 6000 acetylation sites. Their study also found that the majority of the acetylated proteome is localised in either the nucleus or the cytoplasm, approximately 30% of the all the acetylated proteome, whereas the percentage of acetylated proteins in the plasma membrane and the mitochondria is approximately 15% of the total acetylated proteins. On the other hand, the Golgi apparatus, the endoplasmic reticulum as well as the extracellular space account for 5% of the acetylated proteome. Shevade *et al.,* 2013 ([82](#_ENREF_82)) used MS/MS to show PTMs in enolase. They found two acetylation sites at lysine residues 142 and 384. Their analysis also revealed the conjugation, through lysine residue 147, of three ubiquitin molecules to enolase.

Choudhary et al, 2009 ([64](#_ENREF_64)) also identified by MS 3600 acetylation sites on 1750 proteins. They also assessed *in vivo* the reproducibility and depth of the acetylome coverage by SILAC using A549 and Jurkat cells, finding that lysine acetylation is related to cellular processes such as nuclear transport, cell cycle, splicing, and chromatin remodelling. In order to determine if the acetylation patterns of metabolic enzymes were controlled by different carbon sources, Wang *et al.,* 2010 ([83](#_ENREF_83)) did an MS analysis of immunoprecipitated acetylated peptides. They identified 235 peptides that matched to 191 proteins in *Salmonella enterica*, where around 50% of the proteins that were found to be acetylated participated in multiple metabolic pathways, and around 90% of the enzymes participating in central metabolism were found to be acetylated. By SILAC quantitative analysis they determined that from the 15 enzymes that had altered acetylation patterns in response to various carbon sources, all of them showed a higher acetylation rate in the presence of glucose as a carbon source than when citrate was the carbon source. Wirth *et al.,* 2013 ([70](#_ENREF_70)) performed a SILAC to quantify differences in the acetylation of specific sites in pyruvate carboxylase. They also performed an MS/MS comparison of the non- and acetylated relative abundance of peptides revealing. New Sirt4 interacting factors were identified to be highly acetylated. 13 new acetylated sites in pyruvate carboxylase were mapped, with acetylated lysine residue 748 found to be regulating the activity of this enzyme. This analysis also revealed only minor differences in the acetylation levels of this enzyme in the presence of Sirt4.

Lv *et al.,* 2011 ([66](#_ENREF_66)) analysed immunopurified peptides from cytosolic, nuclear and mitochondrial fractions from LNCaP human prostate cell line and cancer tissues by LC-MS/MS, identifying 113 acetylated proteins, including the acetylation of pyruvate kinase on lysine residue 305. Finkemeier *et al.,* 2011 ([55](#_ENREF_55)) also used LC-MS/MC analysis to identify novel acetylated proteins in *Arabidopsis thaliana.* Kim *et al.,* 2012 ([61](#_ENREF_61)) identified several proteins showing significant quantitative changes based on an analysis of the acetylome in 3T3-L1 preadipocytes during adipogenesis by LC-MS/MS, focusing their research on malate dehydrogenase. A characterisation of enolase from pancreatic ductal adenocarcinoma cells and healthy pancreatic ductal cells was made by Zhao *et al.,* 2010 ([84](#_ENREF_84)) by LC-MS/MS, identifying 26 acetylation sites in pancreatic ductal adenocarcinoma cells, and 17 in healthy pancreatic ductal cells (see Appendix II for the acetylation sites). By using (MALDI-TOF) MS analysis Dihazi *et al.,* 2005 ([50](#_ENREF_50)) elucidated one phosphorylation site in phosphofructokinase.

Table 7 summarises all the findings made by using different protein profiling techniques and approaches.

|  |  |  |
| --- | --- | --- |
|  | Protein profiling techniques | Findings |
| Choudhary *et al.,* 2009 | MS  SILAC | Identified 3600 lysine acetylation sites on 1750 proteins.  Quantified acetylation changes in response to HDACi. |
| Cimen *et al.,* 2010 | MS/MS | Mapped various acetylated lysine residues.  Determined the role of acetylation in the activity of isocitrate dehydrogenase. |
| Dihazi *et al.,* 2005 | MALDI-TOF | Found one acetylation site on lysine residue 3 in phosphofructokinase. |
| Finkemeier *et al.,* 2011 | LC-MS/MS | Identified 91 lysine acetylation sites on 74 organellar and cytosolic proteins in *Arabidopsis thaliana*. |
| Finley *et al.,* 2011 | MS/MS | Identified 13 acetylation sites on succinate dehydrogenase, six of which are residues reported to be acetylated for the first time. |
| Kim *et al.,* 2006 | MS | Identified 388 acetylation sites in 195 proteins. Acetylation of isocitrate dehydrogenase at lysine residues 75 and 241. |
| Kim *et al.,* 2012 | LC-MS/MS | Identified several proteins showing significant quantitative changes during adipogenesis, including malate dehydrogenase. |
| Lin *et al.,* 2009 | MS/MS | Validated and identified many non-chromatin substrates of the NuA4 complex and determined two acetylation sites in phosphoenolpyruvate carboxykinase, at lysine residues 19 and 514. |
| Lundby *et al.*, 2012 | MS/MS | Provided a map of acetylation sites from rat tissues. |
| Lv *et al.,* 2011 | LC-MS/MS | Showed the acetylation of pyruvate kinase on lysine residue 305. |
| Schlicker *et al.,* 2008 | MS | Identified lysine residue 241 as the site where isocitrate dehydrogenase is deacetylated via Sirt3 |
| Schwer *et al.,* 2009 | LFQ  MS | Identified 72 proteins involved in a wide range of metabolic pathways with altered acetylation patterns due to CR. |
| Shevade *et al.,* 2013 | MS/MS | Showed acetylation of enolase on lysine residue 142 and 384. |
| Wang *et al.,* 2010 | MS  SILAC | Demonstrated the acetylation of central metabolism enzymes. |
| Wirth *et al.,* 2013 | SILAC  MS/MS | Found that the acetylation of lysine residue 748 in pyruvate carboxylase regulates its activity. |
| Zhao *et al.,* 2010 | MS  iTRAQ | Discovered that acetylation plays a major role in the regulation of metabolism. |

Table 7. Summary of findings made using protein profiling techniques.

# Modelling overview

Modelling is used to predict complex systems, processes and networks based on mathematical approaches taking into account the scope and limitations of the model(s). The modelling of these systems, processes and networks is very important because it contributes to the understanding of different phenomena. Simulations, the interpretation of the model, can be done based on biological data in order to understand these different phenomena. From all the models available, there are models able to simulate β-oxidation and the action of SCFAs on acetyl-CoA and CoA levels, like the ones developed by Bernard Corfe (BC) (unpublished). All these models have SCFAs (acetate, butyrate and propionate) defined as species, which is a key feature to simulate β-oxidation, since SCFAs are oxidized to CO2 and ketone bodies ([48](#_ENREF_48)) and metabolised in this pathway to produce acetyl-CoA. Another important aspect to be considered in these models is CoA-ligation, since this reaction could affect the levels of both CoA and acetyl-CoA. The mechanism by which this could occur is when CoA attaches to the end of butyrate, propionate and acetate to produce a transient compound, their acyl forms, butyryl-CoA and propionyl-CoA. These acyl forms will be, in turn, metabolised to acetyl-CoA in β-oxidation and then enter the TCA cycle. CoA-sink is another essential feature that should be included in the model. This process involves butyryl-CoA, propionyl-CoA, and acetyl-CoA detachment which potentially could increase CoA levels, due to the release of free CoA from these acyl forms (CoA-sink)

Models can be developed by using specialized software, like the complex pathway simulator (COPASI). This is the case of the models butyrate\_v3-1 and SCFA uptake and oxidation v4 developed by BC (unpublished). Butyrate\_v3-1 model was developed to simulate butyrate uptake and oxidation. Although this model was developed for the aforementioned SCFA, propionate, acetate, and glucose as well as acetyl-CoA and CoA were defined as species, whereas SCFAs uptake and CoA-ligation were defined as reactions in this model as well (Figure 9).

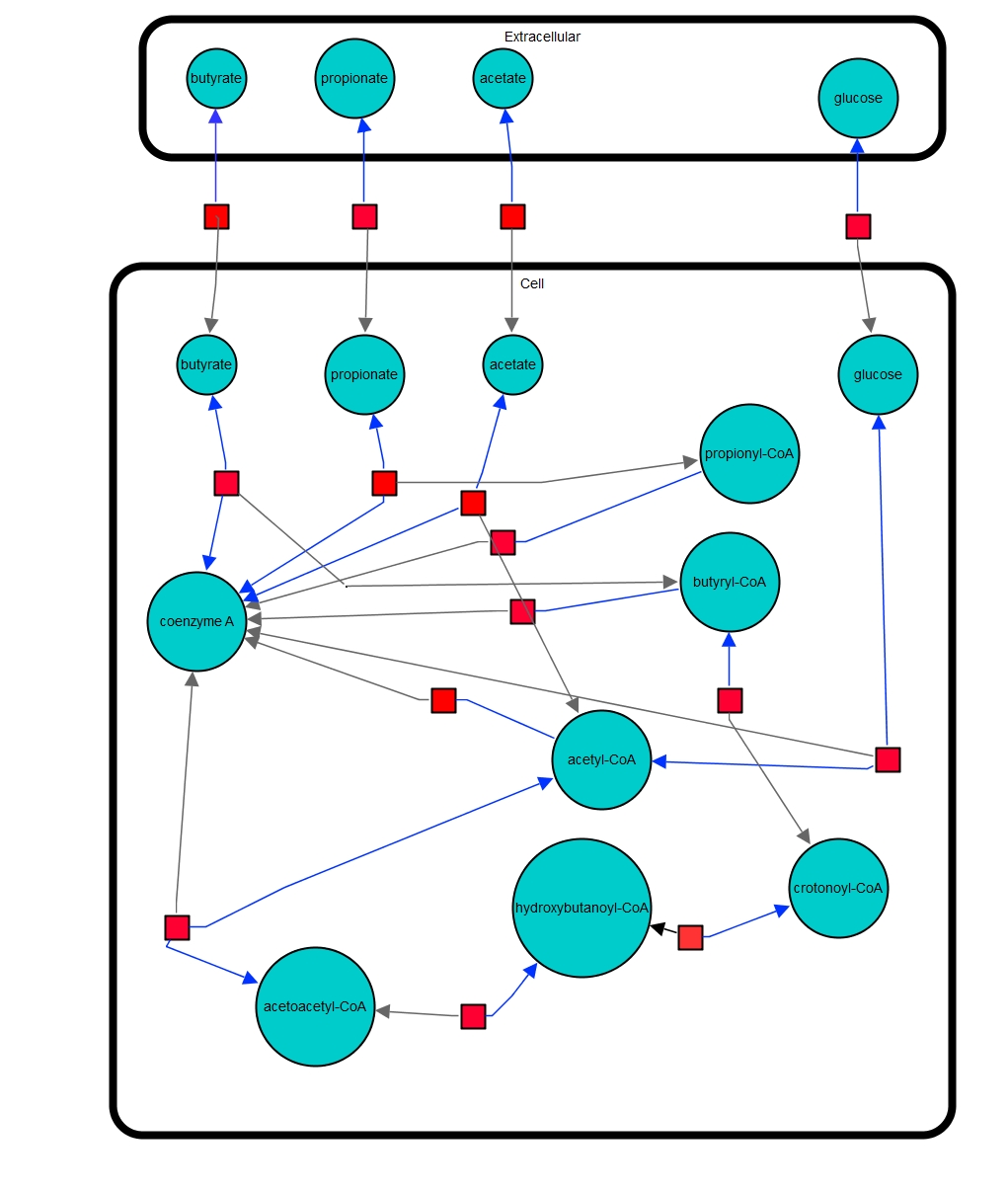


Figure 9. Butyrate\_v3-1 model. *Blue bubbles show the different species defined in the model (e.g. butyrate); the size of the bubbles does not indicate the concentration of each species. Blue arrows show two species needed to form a third species (e.g. butyrate + coenzyme A). Grey arrows show the product of two species ligation (e.g. butyrate + coenzyme A = butyryl-CoA). Red boxes show a reaction. Two compartments are defined in the model, Extracellular and Cell (intracellular).*

SCFA Uptake and Oxidation v4 was developed to simulate SCFAs uptake and oxidation, and hence has butyrate, propionate, acetate CoA and acetyl-CoA (mitochondrial) defined as species; SCFAs uptake and CoA-ligation (only in the mitochondria) were defined as reactions in this model (Figure 10).

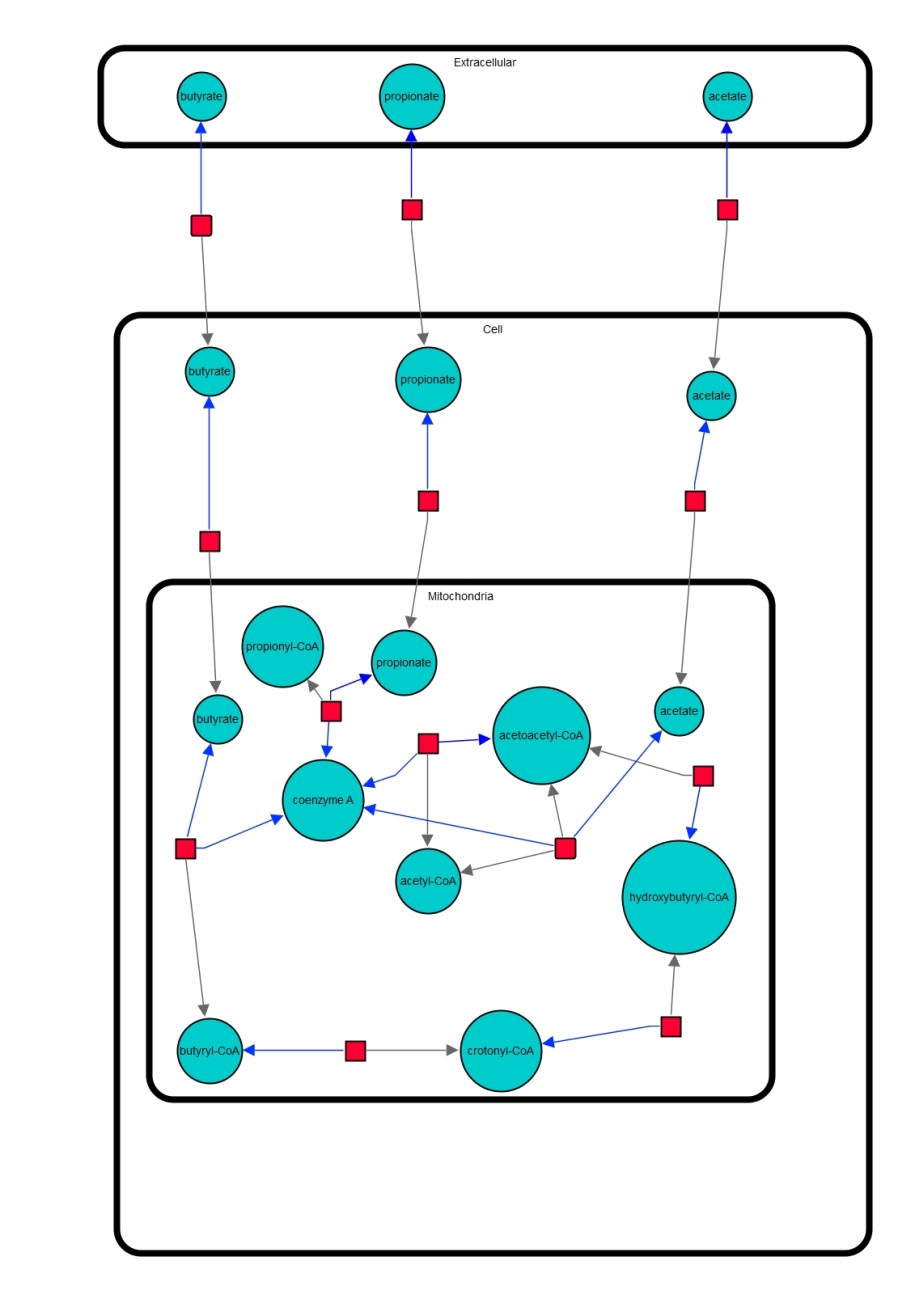


Figure 10. SCFA Uptake and Oxidation v4 model. *Blue bubbles show the different species defined in the model (e.g. butyrate); the size of the bubbles does not indicate the concentration of each species. Blue arrows show two species needed to form a third species (e.g. butyrate + coenzyme A). Grey arrows show the product of two species ligation (e.g. butyrate + coenzyme A = butyryl-CoA). Red boxes show a reaction. Three compartments are defined in the model, Extracellular, Cell (intracellular) and Mitochondria.*

Butyrate\_v3-1 and SCFA Uptake and oxidation v4 models have several advantages. They have defined the three main SCFAs, butyrate, propionate, and acetate, besides glucose (just butyrate\_v3-1 model), acetyl-CoA and CoA as species (Figure 9 and 10, Table 8). Acetate, butyrate, propionate, and glucose (just butyrate\_v3-1 model) uptake is a reaction defined in both models, as well as CoA-ligation, and CoA-sink reactions (Figure 9 and 10, Table 8). It is important that these reactions are defined in the models, since the transport of butyrate, propionate, acetate and glucose from the extracellular to the intracellular milieu is considered (SCFAs and glucose uptake), as well as the ligation of butyrate, propionate and acetate with CoA to form their acyl forms (CoA-ligation), and the release of free CoA from these acyl forms (CoA-sink). These models also consider the production of acetyl-CoA from glucose and CoA-ligation (Figure 9 and 10, Table 8), which is important since there could be a competition between SCFAs and glucose for CoA as a substrate to either produce acyl forms (SCFAs) or acetyl-CoA (glucose). The main disadvantages of these two models are that they have not been published, and therefore have not been tested and criticised by more users, and they still need to be validated. The lack of complexity and the restriction of compartments is another disadvantage of these models, since just basic aspects of SCFAs and glucose metabolism (Uptake, CoA-ligation, and CoA-sink) through β-oxidation can be simulated using these models.

# AIMS

Based on the findings that

1. Acetylation patterns of metabolic enzymes can be controlled by acetyl-CoA levels and SCFAs
2. A shift in the acetylated state of metabolic enzymes can be produced by a product inhibition of deacetylation
3. High levels of butyrate and pyruvate could be related to enhanced acetylation patterns

The aim of this project is to model the action of SCFAs using a mitochondrial proteomic dataset as a start point. An existing mitochondrial proteomic dataset generated by BC and Caroline Evans (CE) in collaboration with Joanne Connolly (JC) of Waters Corporation (Manchester, United Kingdom, UK) will be used to model SCFAs action. The data will be mined in terms of which detected proteins are acetylated by quantitatively examining the data.

# 

# OBJECTIVES

* Determine which enzymes for major metabolic pathways involving acetyl-CoA as a component (glycolysis, the TCA cycle, pyruvate metabolism and β-oxidation) are substrates of acetylation.
* Compare the enrichment scores of different functional annotations and their respective gene ontology (GO) terms.
* Establish the value of the percentage coverage of pathways.
* Analyse different pathways by developing a matching pathway representation.
* Map position of acetylations.
* Model feedback loop

# METHODS

# Mitochondrially enriched fractions

A mitochondrial proteomic dataset was generated by BC and CE in collaboration with JC of Waters Corporation (Manchester, UK) after HCT116 colon cancer cells were stimulated for 6 hours with 10mM propionate and/or butyrate in a 2x2 factorial design with two independent repeats. The cells were lysate and the mitochondrially enriched fractions were extracted and then analysed by using a label-free workflow with a nanoACQUITY and SYNAPT G2S system solution to perform HDMSE (elevated HDMS) experiments. For further details and full methodology please see Patel *et al,* 2009 ([85](#_ENREF_85)). The mitochondrial proteomic dataset was divided in two subsets, global and acetylated proteome. The global proteome dataset consisted of 3963 hits (accession numbers), this is the whole proteome extracted from the mitochondrial fractions of HCT116 cells. On the other hand, the acetylated proteome dataset consisted of 686 hits (accession numbers), gathering all the acetylated proteins extracted from the mitochondrially enriched fractions after the treatment with propionate and/or butyrate. The existing mitochondrial proteomic dataset generated, as explained above, by BC and CE in collaboration with JC of Waters Corporation (Manchester, UK) was mined in terms of which proteins are detected to be acetylated by qualitatively examining the data. By qualitatively examining the data, the metabolic pathways acetylation (Database for Annotation, Visualization and Integrated Discovery, DAVID), functional annotations (DAVID), pathways coverage (Kyoto Encyclopaedia of Genes and Genome, KEGG) and matching pathways representation (Reactome) were determined. Specific quantifications of the mitochondrial proteomic dataset were not possible due to discrepancies in the magnitude of the units between the global and the acetylated proteome datasets.

# Acetylation of metabolic pathways

An existing proteomic dataset generated by BC and CE in collaboration with JC of Waters Corporation (Manchester, UK) was used to determine the acetylation of metabolic pathways. Representation analysis of the existing proteomic dataset (acetylated and global proteome) from mitochondrially enriched fractions was undertaken DAVID (<http://david.abcc.ncifcrf.gov/>), and Reactome (<http://www.reactome.org/>) an open-source curated peer reviewed pathway database. Enzymes for major metabolic pathways involving acetyl-CoA as a component (glycolysis, the TCA cycle, pyruvate metabolism and β-oxidation) were searched for by the Universal Protein Resource (Uniprot) identifier (<http://www.uniprot.org/>), a catalogue of information on proteins, and a map of these metabolic pathways was made. The list of enzymes from the total proteome data set obtained from Uniprot identifier was then compared with the global and acetylated proteome dataset to find which enzymes were acetylated. This comparison was made in order to elucidate the acetylation patterns in enzymes for these major metabolic pathways involving acetyl-CoA as a component. The enzymes found in the acetylated data set were then compared with the manually curated database, PhosphoSitePlus *(*[www.phosphosite.org](http://www.phosphosite.org)) in order to have an accurate representation of the acetylated proteome. Once these enzymes were compared, a map of the acetylated proteome was made.

# Functional annotations

Functional annotations are the attachment of biological data to elements related to the genome. The functional annotations of the global and acetylated proteome datasets, generated by BC and CE in collaboration with JC of Waters Corporation (Manchester, UK), were taken from DAVID (<http://david.abcc.ncifcrf.gov/>). The GO terms, a standard set of terms to resolve all information of different names in a conserved vocabulary, were selected according to the aims of this work. These GO terms were Biological Process (BP), extensive biological fields accomplished by a systemic gathering of various molecular functions; Cellular Compartment (CC), different subcellular macromolecular complexes, structures and locations; and Molecular Function (MF), tasks made by individual gene products. The fifth level of GO of the functional annotations was chosen over the other levels (1-4), since the first levels of GO represent the most general functional annotations, lacking specificity, whereas the latest levels of GO cover more specific functional annotations. The fifth level of GO was chosen because a wider range of specific functional annotations fulfils better the aims of this project. Besides, this level of GO has the right balance in between general functional annotations and specific functional annotations. The enrichment scores of the top 10 general functional annotations and the top 10 metabolism-related functional annotations of the fifth level of GO were ranked and plotted. The enrichments scores of every GO functional annotation showed by DAVID are a rank of biological significance of a group of proteins evidencing a change in the expression of a specific pathway protein set revealing the degree to which a protein set is overrepresented in an expression dataset indicating the percentage of proteins contributing to the enrichment score in every functional annotation of each Go term. These enrichment scores are calculated by clustering an overall EASE score (Fisher Exact modified *p-*values) of all the enrichment annotation terms, where the lower the *p­*-value is, the higher the enrichment (see Appendix 3 for the *p*-values of every functional annotation in each GO term for the global and acetylated proteome datasets). The enrichments scores show multiple proteins evidencing a shift in the expression of a specific pathway protein set, where the main result of the protein analysis reveals the degree to which a protein set is overrepresented at the top or bottom of a ranked list of proteins in the expression dataset. The results of this analysis indicate the percentage of proteins contributing to the enrichment score in every functional annotation of each GO term. To determine if the distributions of the data are not significantly different from a normal distribution, a Shapiro-Wilk test was done in R. If the *p*-value in this test is >0.05, the distribution does not significantly differ from the normal distribution at a five percent significance level, so the Student’s test (Welch Two Sample t-test) was used; if the *p*-value is <0.05, the distribution significantly differs from the normal distribution at a five percent significance level, so Wilcoxon ran sum test was used.

# Pathways coverage

The total number of proteins in each pathway (All pathway proteins), as well as the total number of proteins in the global and acetylated proteome datasets (All list proteins) generated by BC and CE in collaboration with JC of Waters Corporation (Manchester, UK) were taken from KEGG (<http://www.genome.jp/kegg/>), a database for the understanding of the functions and utilities of biological systems, through DAVID (<http://david.abcc.ncifcrf.gov/>, accessed on 24/05/14). The value of the percentage coverage of pathways was taken dividing the total number of enzymes (All pathway enzymes) in each pathway by the number of proteins found in the global or acetylated proteome (All list enzymes) in order to establish diferences in the number of enzymes participating in each metabolic pathway in the global and acetylated proteome dataets.

# Matching pathways representation

The open access software Reactome (<http://www.reactome.org/>, accessed on 06/08/14), a precise manually curated and peer-reviewed pathway database that denotes human biological processes as inter-connected molecular events or reactions, was used to do an overrepresentation analysis of the global and the acetylated proteome datasets generated by BC and CE in collaboration with JC of Waters Corporation (Manchester, UK). By using Reactome, different pathways matching with the global and acetylated proteome datasets were found. From all the different pathways analysed and found to match the global proteome dataset, the most representative for this project, Metabolism, was selected. This same pathway selected for the global proteome dataset was analysed and searched for the acetylated proteome dataset through Reactome.

# Modelling

COPASI, a software application designed to simulate and analyse the dynamics of biochemical networks, was used to simulate the action of butyrate, propionate, acetate and glucose increasing concentration on CoA and acetyl-CoA levels. From several models explored to simulate the oxidation of SCFAs and after describing them, as well as analysing their merits and limitations (see Modelling overview sub-heading, where all these aspects are discussed), butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models were selected to run simulations on the action of increasing concentrations of SCFAs and glucose on acetyl-CoA and CoA and to simulate the effect of physiological ranges of butyrate, propionate and acetate on a time course. These models were chosen because they both meet the requirements searched in a model, which are: having acetate, butyrate, propionate, acetyl-CoA and CoA defined as species; acetate, butyrate, and propionate uptake as well as CoA-ligation and CoA-sink defined as reactions (Figure 9 and 10, Table 8). Burytate\_v3-1 and SCFA Uptake and Oxidation v4 models are able to interrogate and simulate the oxidation of butyrate, propionate, acetate and glucose (just butyrate\_v3-1 model). As aforesaid, these models have defined the three main SCFAs, butyrate, propionate, and acetate, besides glucose (just butyrate\_v3-1 model), and acetyl-CoA and CoA as species (Figure 9 and 10, Table 8). They also have butyrate, propionate, acetate and glucose (just butyrate\_v3-1 model) uptake defined as reactions, as well as CoA-ligation, and CoA-sink (Figure 9 and 10, Table 8). These models also consider the production of acetyl-CoA from glucose and CoA-ligation (Figure 9 and 10, Table 8).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Species in the model | SCFAs uptake | CoA-ligation | CoA-sink |
| Butyrate\_v3-1 | Butyrate, propionate, acetate, and glucose | YES | YES | YES |
| SCFA Uptake and Oxidation v4 | Butyrate, propionate and acetate | YES**+** | YES**‡** | NO |

Table 8. Models developed by BC. *This table shows the species defined in each model, and the reactions defined (uptake, CoA-ligation and CoA-sink of the species defined in the model). + indicates that glucose uptake is not defined in this model, since glucose is not a defined species within it; ‡ indicates that CoA-ligation is only defined as a reaction in the mitochondria in this model.*

The underlying kinetics that control the behaviour of the models are various. In the butyrate\_v3-1 model acetate, butyrate and propionate initial concentrations in the cell are set at 0.1mmol/l, whereas acetyl-CoA concentration is set at 60mmol/l, CoA concentration set at 40mmol/l, and glucose concentration set at 0.5mmol/l (Table 9). All these species type is defined as reactions in this model, which means that the kinetic laws of the reactions modify the amount/concentration of the species. The extracellular initial concentrations of acetate and butyrate are defined at 0mmol/l, propionate is defined at 10mmol/l, and glucose at 0.5mmol/l (Table 9). All these species type is defined as fixed, which means that the species have a constant concentration/amount value corresponding to their given initial concentration. In terms of SCFAs and glucose uptake, this model assumes that acetate, butyrate and propionate uptake is via MCT-1 (monocarboxylate transporter-1) and that all the uptake reactions are reversible. Acetate uptake reaction has acetate in the cell defined as a product, extracellular acetate as substrate, and extracellular and in the cell and as well as extracellular propionate and in the cell as modifiers, with a flux of -0.02mmol/s. Butyrate uptake reaction has butyrate in the cell defined as a product, extracellular butyrate as substrate, and extracellular acetate and in the cell and as well as extracellular propionate and in the cell as modifiers, with a flux of -0.02mmol/s. Propionate uptake reaction has propionate in the cell defined as a product, extracellular propionate as substrate, and extracellular acetate and in the cell as well as extracellular butyrate and in the cell as modifiers, with a flux of 2.75mmol/s. Glucose uptake reaction has glucose in the cell defined as a product and extracellular glucose as substrate, with a flux of 0mmol/s. As uptake reactions, all CoA-ligation reactions are also defined as reversible. Acetate - CoA-ligation has acetyl-CoA in the cell defined as product, with acetate and CoA in the cell defined as substrates, with a flux of 0.35mmol/s. Butyrate - CoA-ligation has butyryl-CoA in the cell defined as product, with butyrate and CoA in the cell defined as substrates, with a flux of 2.43mmol/s. Propionate - CoA-ligation has propionyl-CoA in the cell defined as product, with propionate and CoA in the cell defined as substrates, with a flux of 2.43mmol/s. Glucose – CoA-ligation has acetyl-CoA in the cell defined as product, with glucose and CoA in the cell defined as substrates, with a flux of 2.82mmol/s.

In the SCFA Uptake and Oxidation v4 model acetate, butyrate and propionate initial concentrations in the cell are set at 1mmol/l, whereas CoA concentration is set at 40mmol/l. Acetyl-CoA concentration is only set in the mitochondria at 60mmol/l (Table 9). As in butyrate\_v3-1 model, all these species type is defined as reactions in this model, which means that the kinetic laws of the reactions modify the amount/concentration of the species. The extracellular initial concentration of SCFAs is set as follows: acetate at 25mmol/l, butyrate at 5mmol/l, propionate at 7mmol/l (Table 9). All these species type is defined as fixed, which means that the species have a constant concentration/amount value corresponding to their given initial concentration. SCFA Uptake and Oxidation v4 model has also several species defined in the mitochondria: acetate, butyrate and propionate defined at 1mmol/l. All these species type is defined as reactions in this model, which means that the kinetic laws of the reactions modify the amount/concentration of the species. Opposite to butyrate\_v3-1 model, SCFA Uptake and oxidation v4 model assumes that acetate, butyrate and propionate uptake is simpler, with no modifiers defined, just the extracellular SCFA defined as product, and the SCFA in the cell as the product. As in butyrate\_v3-1 model, SCFA Uptake and Oxidation v4 model also defines the SCFA uptake reactions as reversible. CoA-ligation reactions are also defined as reversible, but opposite to butyrate\_v3-1 model, these reactions are only defined in the mitochondria. Acetate – CoA-ligation has acetyl-CoA as product with acetate and CoA as substrates at a flux of 0mmol/min. Butyrate – CoA-ligation has butyryl-CoA defiend as a product with butyrate and CoA defined as substrates at a flux of 0mmol/min. Propionate – CoA-ligation has propionyl-CoA defined as a product with propionate and CoA defiend as substrates at a flux of 0mmol/min.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Initial concentration in the cell  (mmol/l) | | | | | | Extracellular initial concentration (mmol/l) | | | | | |
|  | Ace | But | Prop | Glu | AcCoA | CoA | Ace | But | Prop | Glu | AcCoA | CoA |
| Butyrate\_v3-1 | 0.1 | 0.1 | 0.1 | 0.5 | 40 | 60 | 0 | 0 | 10 | 0.5 | **\*** | **\*** |
| SCFA Uptake and Oxidation v4 | 1 | 1 | 1 | **×** | 40 | 60 | 25 | 5 | 7 | **×** | **\*** | **\*** |

Table 9. Initial concentrations of the substrate species defined in the models. *Ace = acetate, But= butyrate, Prop = propionate, Glu = glucose, AcCoA = acetyl-CoA. × indicates this species is not defined in the model. \* Indicates no extracellular concentrations are defined for these species.*

The simulations with butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models were run in COPASI using the parameter scan task at increasing concentrations of butyrate (1-20mM), propionate (1-20mM), acetate (1-52mM) and glucose (just butyrate\_v3-1) (0.5-2mM) at physiological ranges. Each of the aforementioned species in the models was independently set as parameters to scan in the ranges described above at 1000 intervals. The initial extracellular concentrations of these species were set to be scanned and run the simulations in response to the levels of either acetyl-CoA or CoA by the increasing concentrations of butyrate, propionate, acetate or glucose.

In order to simulate the effect of time on the acyl forms as well as on acetyl-CoA and CoA levels the same models were used. These models were chosen following the same rationale as before (Figure 9 and 10, Table 8). The simulations were run in COPASI using the time course task, at 0.1, 100, 1000 and 10000 seconds at 100 intervals. The concentrations of butyryl-CoA, propionyl-CoA, and acetyl-CoA as well as CoA were simulated at each time point. The changes in these concentrations between the first time point of 0.1 and 100 seconds were measured, as well as the difference between 100 and 1000 seconds and 1000 and 10000 seconds. As there was no difference between 1000 and 10000 seconds in both models, 1000s was chosen as the maximum time to run the simulations at a 1000 intervals.

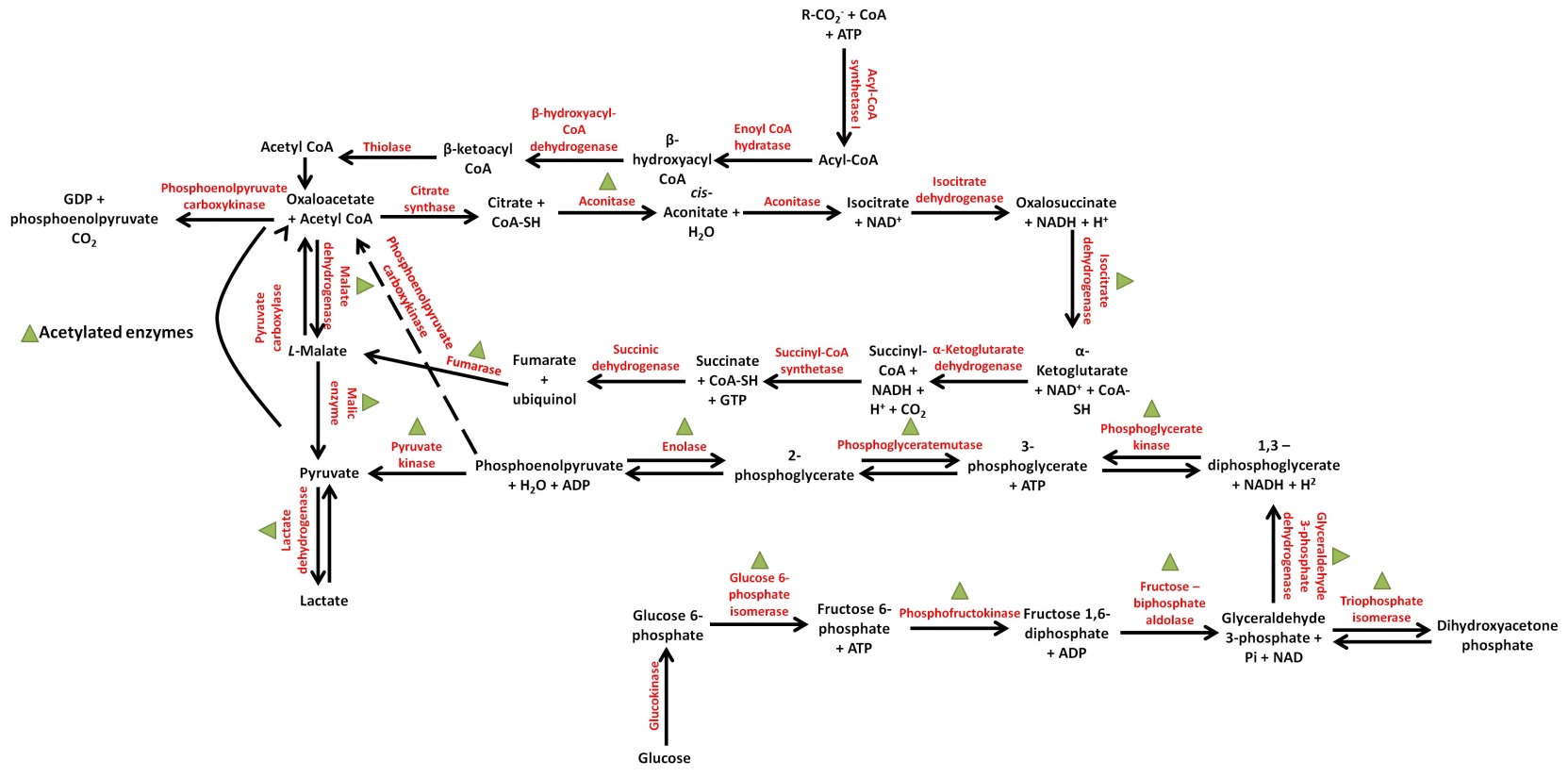
# Validation of the models

Andy Nichols and Marilena Epitropou, two MSc students under my supervision, culturedHCT116 human colorectal carcinoma cells (Health Protection Agency Culture Collections, Porton Down, UK) in Dulbecco’s Modified Eagle Medium (DMEM) supplemented with 500 units/500ml of streptomycin (Lonza, Cambridge, UK) and 500 units/ml of penicillin (Lonza, Cambridge, UK) and 10% foetal bovine serum (FBS) (Biosera, Uckfield, UK) between 37 and 62 passages. Cells were seeded in 6 well plates at sub confluence and incubated for 24 hours at 37°C 4% CO2 prior to a 6 hour time course treatment with increasing concentrations of SCFAs: sodium acetate (0-50mM) (Sigma Aldrich, Poole, UK), sodium butyrate (0-20mM) and sodium propionate (0-20mM) (Sigma Aldrich, Poole, UK). CoA levels were determined using a commercially available CoA detection kit (Abcam, Cambridge, UK) at Ex/Em = 490/520 nm using a Spectramax plate reader (Molecular Devices LLC, Reading, UK). These results allowed future comparison to the models previously described (Modelling overview sub-heading, and Modelling in Methods section).

# RESULTS

# Representation and acetylation of metabolic pathways

The accession numbers of all the hits in the global proteome dataset were searched for by Uniprot identifier ([www.uniprot.org/](http://www.uniprot.org/)) to find the function and the name of each protein and select the enzymes taking part in major metabolic pathways with acetyl-CoA as a component: glycolysis, the TCA cycle, pyruvate metabolism and β-oxidation. A map of the metabolic pathways was produced with these enzymes, by using data taken from KEGG ([www.genome.jp/kegg](http://www.genome.jp/kegg)). Once the metabolic pathway map for the global proteome dataset was made, acetylated enzymes were searched for by Uniprot ([www.uniprot.org/](http://www.uniprot.org/)) identifier to find which metabolic enzymes participating in the aforementioned pathways were acetylated; these enzymes were mapped (Figure 11). In this map the 21 metabolic enzymes that were found in the acetylated proteome dataset are highlighted with a green legend (Figure 11). Nine glycolytic enzymes were found to be acetylated namely glucose 6-phosphate isomerase, phosphofructokinase, fructose-biphosphate aldolase, glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutase, enolase, pyruvate kinase, and triophosphate isomerase (Figure 11, bottom right). In the TCA cycle the 5 enzymes that were found to be acetylated are aconitase, isocitrate dehydrogenase, succinate dehydrogenase, fumarase and malate dehydrogenase (Figure 11, centre). Regarding β-oxidation, the 3 enzymes that were found to be acetylated are acyl-CoA synthetase, enoyl-CoA hydratase, and thiolase (Figure 11, top). In terms of pyruvate metabolism, the 4 enzymes that were found to be acetylated in this metabolic pathway are lactate dehydrogenase, malic enzyme, malate dehydrogenase, and pyruvate kinase (Figure 11, bottom left).



**Figure 11. Representation and acetylation of major metabolic pathways involving acetyl-CoA as a component. *The metabolic pathways depicted in this map are β-oxidation (top, centre), the TCA cycle (centre), glycolysis (bottom, right) and pyruvate metabolism (bottom, left). Metabolic enzymes in these pathways are shown in red bold font; metabolic products in these pathways are shown in black bold font. The enzymes found in the acetylated proteome dataset are highlighted with a green legend.***

All of the aforementioned metabolic enzymes that were observed in the acetylated proteome have previously been reported to be acetylated (see Appendix 2 for the complete list of acetylated lysine residues and their implications).

# Percentage coverage of pathways

The total number of proteins encoded in each pathway was taken from KEGG (<http://www.genome.jp/kegg/>), through DAVID (<http://david.abcc.ncifcrf.gov/>). The total number of proteins was divided by the number of proteins found in the total and acetylated proteome datasets and plotted as a percentage of the coverage of pathways (Figure 12). In accordance with the enrichment scores, there was a shift between the number of proteins covered in each pathway when the global and the acetylated proteome datasets were compared. The global proteome dataset had the highest coverage of pathways. The ranges of the percentage coverage of pathways were between 81 and 55% in this dataset (Figure 12). In the global proteome dataset, the TCA cycle was the pathway with the highest percentage coverage of pathways with 81%, followed by glycolysis and pyruvate metabolism both with 65%, whereas β-oxidation was the pathway with the lowest percentage coverage of pathways with 55% (Figure 12). Regarding the acetylated proteome dataset, the pathway with the highest coverage of pathways was glycolysis with 23%, followed by pyruvate metabolism with 20%, the TCA cycle with 19%, and finally and in accordance to what was observed in the global proteome dataset, β-oxidation was the pathway with the least percentage coverage of pathways with 18%.

A lower percentage cover of pathways is expected in the acetylated proteome dataset, since not all the proteins found in each pathway are substrates of acetylation or instead, this decrease in the percentage cover of pathways in the acetylated proteome dataset could also be due to stoichiometry since not all enzymes have the same sites. Interestingly, and even if there was a 3-fold decrease, from 65 to 25%, in its percentage coverage of pathways (when compared to the global proteome dataset), glycolysis was the pathway with the highest percentage coverage of pathways with 23% in the acetylated proteome dataset, suggesting that glycolytic enzymes are more likely to be substrates of acetylation or that these enzymes have more sites than the other enzymes in the TCA cycle, β-oxidation and pyruvate metabolism. On the other hand, and as in the global proteome dataset the pathway with the lowest percentage coverage of pathways in the acetylated proteome dataset was β-oxidation with 18% coverage of pathways (Figure 12), this suggests that the enzymes in this pathway are less likely to be substrates of acetylation or have less sites than the enzymes in glycolysis, the TCA cycle, or pyruvate metabolism.

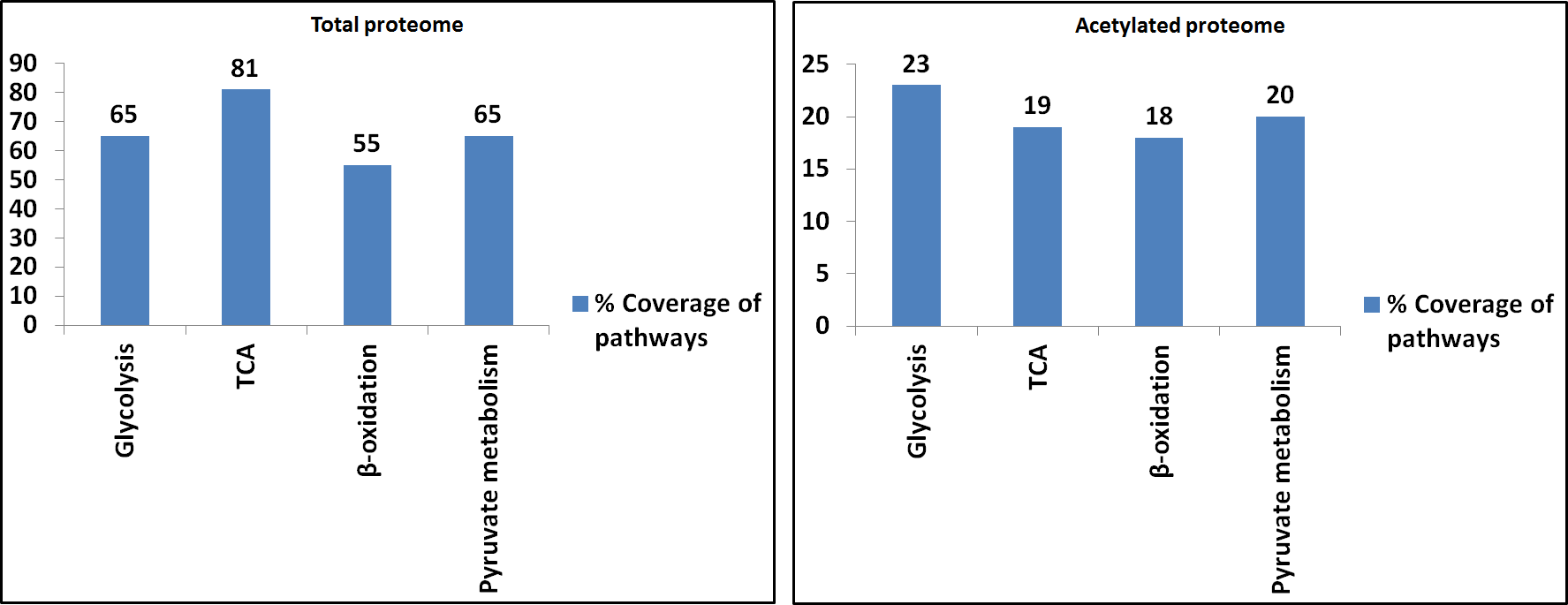


Figure 12. Percentage Coverage of pathways. *Representation of the percentage of proteins in the global and acetylated proteome datasets to be found to match the total number of proteins taking part in each metabolic pathway.*

When every pathway is analysed separately, it can be noticed that for glycolysis, 60 proteins were found to take part in this metabolic pathway and 39 were found to have a match in the global proteome dataset (65% coverage of pathway), whereas just 14 were found to have a match in the acetylated proteome dataset (23% coverage of pathway) (Table 10). In the case of the TCA cycle, 31 proteins are involved in this metabolic pathway, with 25 having a match with the global proteome dataset (81% coverage cover of pathway), and scarcely 6 having a match according to the acetylated proteome dataset (19% coverage of pathway). In β-oxidation 40 proteins are implicated, with 22 found in the global proteome dataset (55% coverage of pathway) and merely 7 found in the acetylated proteome dataset (18% coverage of pathway) (Table 10). Regarding pyruvate metabolism, 40 proteins are associated to this metabolic pathway. Of those 40 proteins, 26 were found in the total proteome dataset (65% coverage of pathway), whilst only 8 were found to match the acetylated proteome dataset (20% coverage of pathway) (Table 10).

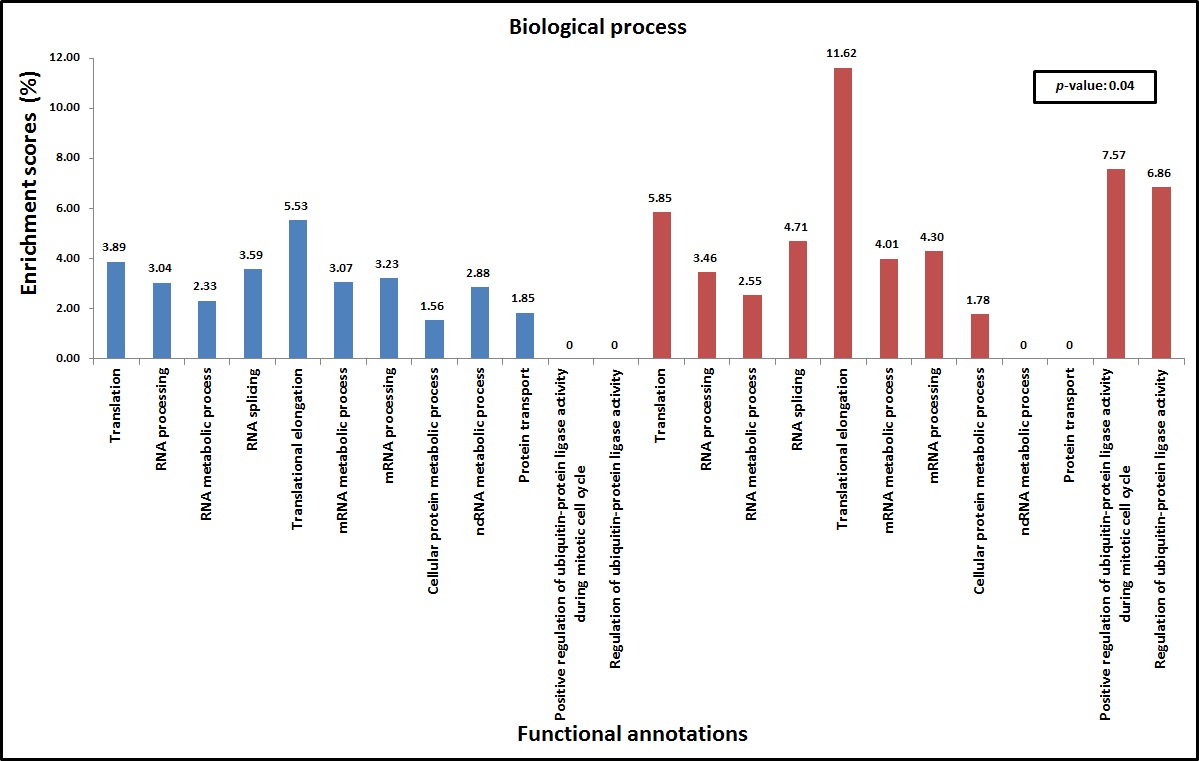
|  |  |  |
| --- | --- | --- |
| Pathway | All pathway proteins | All list proteins  (global proteome/acetylated proteome) |
| Glycolysis | 60 | 39/14 |
| The TCA cycle | 31 | 25/6 |
| β-oxidation | 40 | 22/7 |
| Pyruvate metabolism | 40 | 26/8 |

Table 10. Coverage ratio. *Number of proteins in the global and acetylated proteome datasets found to match the total number of proteins taking part in glycolysis, the TCA cycle, β-oxidation, and pyruvate metabolism pathways.* *All pathway proteins column shows the total number of enzymes in each metabolic pathway; All list proteins column shows the number of enzymes found in the global (left) and acetylated (right) proteome datasets.*

# Functional annotations

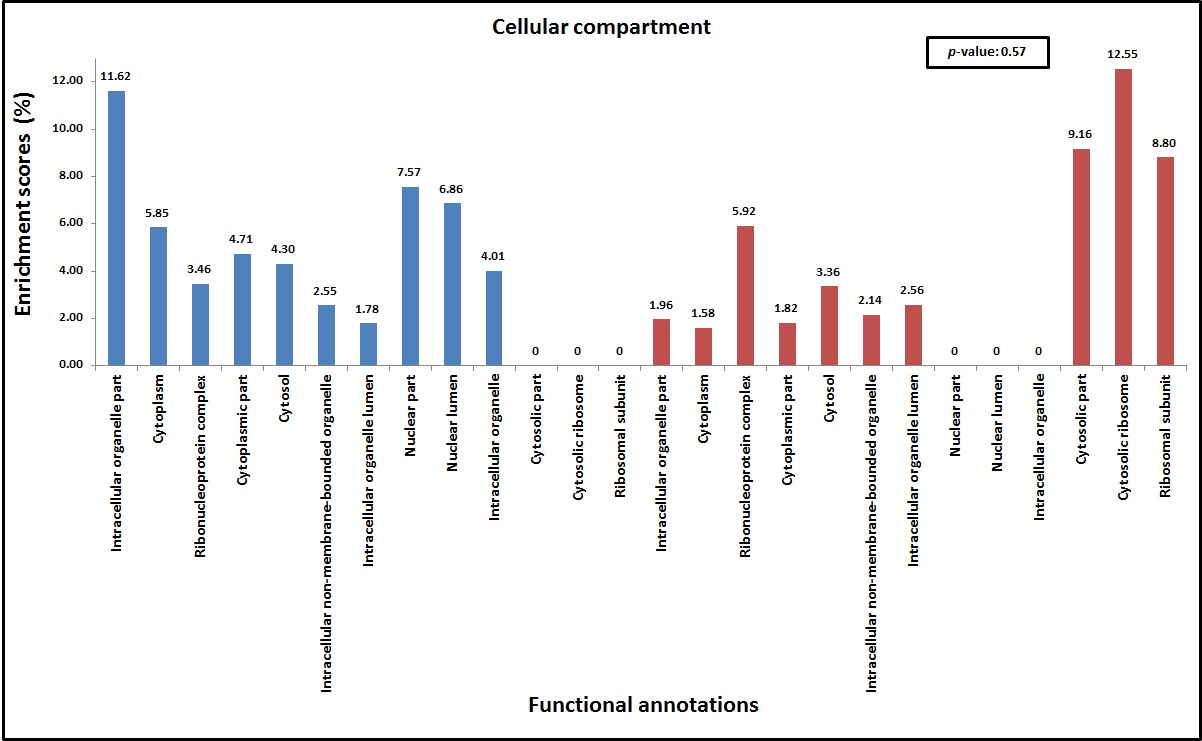
Protein representation in the mitochondrial fractions were taken from DAVID (<http://david.abcc.ncifcrf.gov/>). The enrichment scores of the functional annotations in the three different GO terms chosen according to the aims of this project were analysed. The three different GO terms were BP, which includes extensive biological fields, like translation and RNA processing, all of them accomplished by a systematic gathering of various molecular functions ([86](#_ENREF_86)); CC, which gathers different subcellular macromolecular complexes, structures and locations such as intracellular organelle part and cytoplasm ([86](#_ENREF_86)); and MF, which includes tasks made by individual gene products, for example pryophosphatase activity and adenyl rinonucleotide binding ([86](#_ENREF_86)). The top 10 functional annotations in the aforementioned GO terms of the fifth level of ontology were plotted and ranked (Figures 13, 14, and 15). These graphs show the enrichment scores of every functional annotation in each GO term. Enrichment scores are a biological cluster of a group of proteins, based on overall *p*-values of all the enriched functional annotation terms (see Appendix 3 for the *p*-values of every functional annotation in each GO term for the global and acetylated proteome datasets).

From the three different GO terms chosen, the one that shared the most functional annotations, when comparing the global and the acetylated proteome datasets was BP. This analysis could indicate that the activation of the biological fields and various molecular functions in BP GO term, at least in 8 out of 10 of the top 10 functional annotations, do not depend on acetylation. The functional annotations shared between the global and the acetylated proteome datasets in the BP GO term were Translation, RNA processing, RNA metabolic process, RNA splicing, Translational elongation, mRNA metabolic process, mRNA processing, and Cellular protein metabolic process (Figure 13). Interestingly, and despite 8 out of 10 functional annotations, aforementioned, were shared in the BP GO term between the global and the acetylated proteome datasets (Figure 13), the enrichment scores were different. An increase in the enrichment scores might be related to an up-regulation associated with acetylation of a group of proteins in a specific functional annotation in any GO term, since this increase is expressed as a higher percentage of proteins contributing to that specific functional annotation. For example, in the BP GO term, the functional annotation Translational Elongation has an enrichment score of 5.53 in the global proteome dataset and 11.62 in the acetylated proteome dataset (Figure 13). This shift in the enrichment scores of these functional annotations represents a 2-fold increase. Even a decrease in the enrichment scores in the acetylated proteome dataset is expected, since not all proteins are substrate of acetylation, this observation was not made in BP GO term, since the enrichment scores of every functional annotation in the acetylated proteome dataset are higher than in the global proteome dataset (Figure 13).



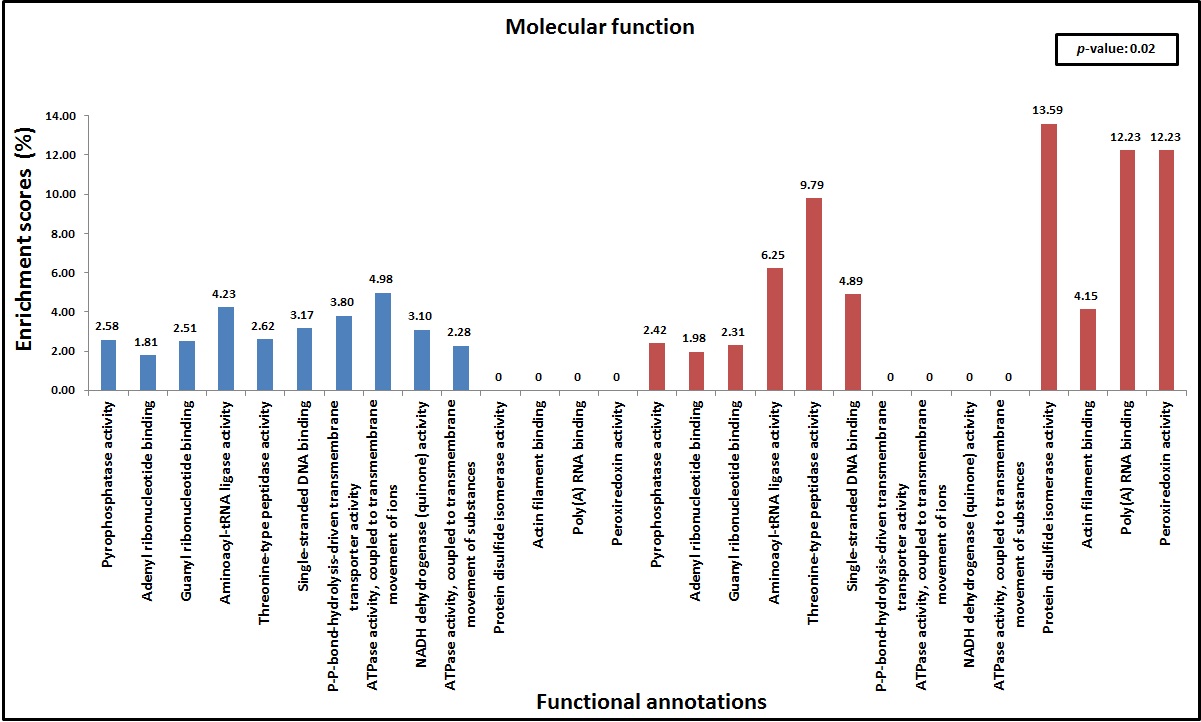
**Figure 13. BP enrichment scores. *Enrichment scores of the global proteome (left, blue) and the acetylated proteome (right, red) datasets of the top ten functional annotations of the fifth level of ontology for BP GO term. Functional annotations that were not found in both global and acetylated proteome datasets were scored as 0, since no comparison was feasible to be made amongst these functional annotations. Differences in the enrichment scores were observed in the global and acetylated proteome datasets. For example, Translational elongation functional annotation has a higher enrichment score in the acetylated proteome dataset (11.62) when compared to the global proteome dataset (5.53). The p*-*value of 0.04 indicates that there are statistically significant differences in the changes observed in the enrichment scores between the global and the acetylated proteome datasets. The Welch Two Sample t-test was used to compare the global and the acetylated proteome datasets (Shapiro-Wilk test p-value: 0.51 for the global proteome and 0.37 for the acetylated proteome).***

For the CC GO term, 7 out of 10 functional annotations were shared (Figure 14). These functional annotations were Intracellular organelle part, Cytoplasm, Ribonucleoprotein complex, Cytoplasmic part, Cytosol, Intracellular non-membrane-bounded organelle, Intracellular organelle lumen, and Intracellular organelle part (Figure 14). As in the BP GO term, a shift in the enrichment scores was observed between the global and the acetylated proteome datasets. The functional annotation Ribonucleoprotein complex was the one with the highest increase in its enrichment score in the CC GO term, since there was a shift from 3.46 in the global proteome dataset to 5.92 observed in the acetylated proteome dataset (Figure 14). The functional annotation Intracellular organelle part was the one with the highest decrease in its enrichment score in the CC GO term, since there was a shift from 11.62 in the global proteome dataset to 1.96 observed in the acetylated proteome dataset. This decrease in the enrichment scores indicates less proteins contributing to a specific functional annotation, which as aforesaid is expected, since not all proteins are substrate of acetylation.



**Figure 14. CC enrichment scores. *Enrichment scores of the global proteome (left, blue) and the acetylated proteome (right, red) datasets of the top ten functional annotations of the fifth level of ontology for CC GO term. Functional annotations that were not found in both global and acetylated proteome datasets were scored as 0, since no comparison was feasible to be made amongst these functional annotations. Differences in the enrichment scores were observed in the global and acetylated proteome datasets. For example the functional annotation Intracellular organelle part has a higher enrichment score in the global proteome dataset (11.62) when compared to the acetylated proteome dataset (1.96); on the other hand, Ribonucleoprotein complex functional annotation has a higher enrichment score in the acetylated proteome dataset (5.92) when compared to the global proteome dataset (3.46%). The p*-*value of 0.57 indicates that there are no statistically significant differences in the changes observed in the enrichment scores between the global and the acetylated proteome datasets. Wilcoxon rank sum test* *was used to compare the global and the acetylated proteome datasets (Shapiro-Wilk test p-value: 0.37 for the global proteome and 0.03 for the acetylated proteome).***

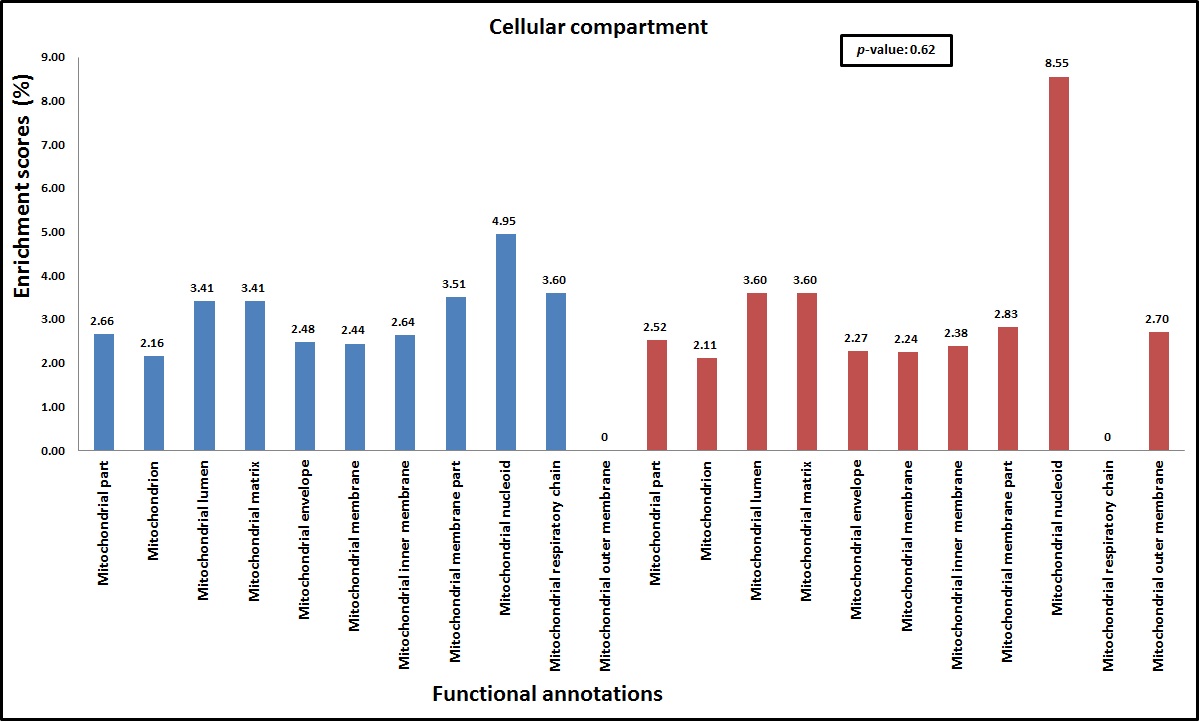
The results for the MF GO term followed the same trend, although the number of functional annotations shared in between the global and the acetylated proteome datasets was the lowest, with just 6 of these functional annotations shared (Figure 16). The functional annotations shared in the MF GO term in between the global and the acetylated proteome dataset were Pyrophosphatase activity, Adenyl ribonucleotide binding, Guanyl ribonucleotide binding, Aminoacyl-tRNA ligase activity, Threonine-type peptidase activity, and Single-stranded DNA binding (Figure 15). As in the BP and the CC GO terms the enrichments scores in the MF GO term also changed between the global and acetylated proteome datasets. For example, Threonine-type peptidase activity was the functional annotation with the largest change in its enrichment score in the MF GO term. In this functional annotation, a shift from 2.62 in the global proteome dataset to 9.79 in the acetylated proteome dataset was observed (Figure 15), indicating an increase in the percentage of proteins contributing to this functional annotation. On the other hand, the functional annotation with the highest decrease in its enrichment score in the MF GO term was Guanyl ribonucleotide binding, since a shift from 2.51 in the total proteome dataset to 2.31 in the acetylated proteome dataset was observed.

****

**Figure 15. MF enrichment scores*. Enrichment scores of the global proteome (left, blue) and the acetylated proteome (right, red) datasets of the top ten functional annotations of the fifth level of ontology for MF GO term. Functional annotations that were not found in both global and acetylated proteome datasets were scored as 0, since no comparison was feasible to be made amongst these functional annotations. Differences in the enrichment scores were observed in the global and acetylated proteome datasets. For example the functional annotation Guanyl ribonucleotide binding has a higher enrichment score in the global proteome dataset (2.51) when compared to the acetylated proteome dataset (2.31); on the other hand, Threonine-type peptidase activity functional annotation has a higher enrichment score in the acetylated proteome dataset (9.79) when compared to the global proteome dataset (2.62). The p-value of 0.02 indicates that there are statistically significant differences in the changes observed in the enrichment scores between the global and the acetylated proteome datasets. The Welch Two Sample t-test was used to compare the global and the acetylated proteome datasets (Shapiro-Wilk test p-value: 0.57 for the global proteome and 0.1 for the acetylated proteome).***

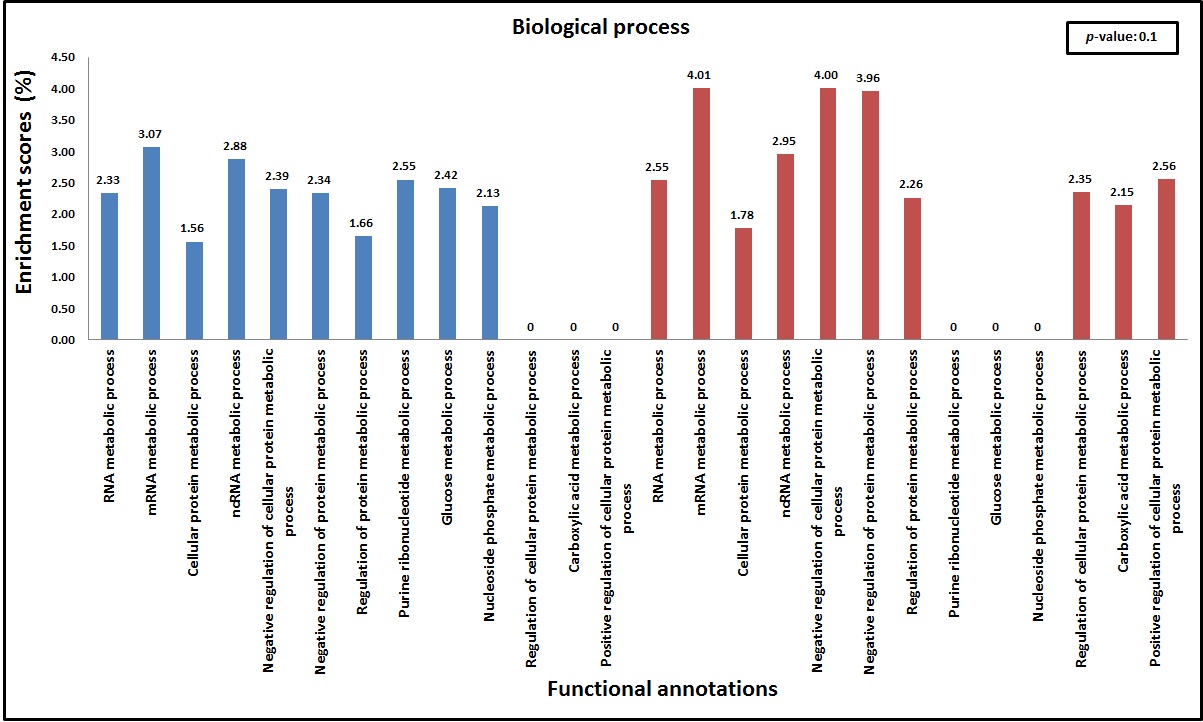
The differences observed in the enrichment scores of the top ten functional annotations in the GO terms chosen, BP, CC, and MF, give a global overview of the results, but since the aims of this project are based on changes in metabolic enzymes, the top ten functional annotations related to extensive biological fields accomplished by a systemic gathering of various molecular functions in metabolic processes were chosen for BP GO term. Different subcellular macromolecular complexes, structures and locations in which metabolism takes place were chosen for CC GO term and tasks made by individual gene products in metabolism were chosen for MF GO term.

From the three different GO terms chosen, the one that shared the most top 10 metabolism-related functional annotations, when comparing the global and the acetylated proteome datasets was CC, contrasting with the top 10 global overview in which BP shared the most functional annotations. This analysis could also indicate that the activation of subcellular macromolecular complexes, structures and locations in CC GO term, at least in 9 out of 10 of the top 10 metabolism-related functional annotations, do not depend on acetylation, in accordance to BP GO term in the global overview. The functional annotations shared between the global and the acetylated proteome datasets in the CC GO term were Mitochondrial part, Mitochondrion, Mitochondrial lumen, Mitochondrial matrix, Mitochondrial envelope, Mitochondrial membrane, Mitochondrial inner membrane, Mitochondrial membrane part, Mitochondrial nucleoid, and Mitochondrial respiratory chain (Figure 16). Contrasting the observations made in the global overview, and despite 9 out of 10 metabolism-related functional annotations, aforementioned, were shared in the CC GO term between the global and the acetylated proteome datasets, the enrichment scores did not show major differences except Mitochondrial nucleoid which decreased from 8.55 in the acetylated proteome to 4.95 in the global proteome (Figure 16). Even the differences observed were not very evident; the functional annotation Mitochondrial membrane part was the one with the highest increase, from 3.51 in the global proteome dataset to 2.83 in the acetylated proteome dataset. This apparent lack of changes in the enrichments scores is supported by statistical analysis which showed no statistically significant difference between the global and acetylated proteome datasets (Figure 16).



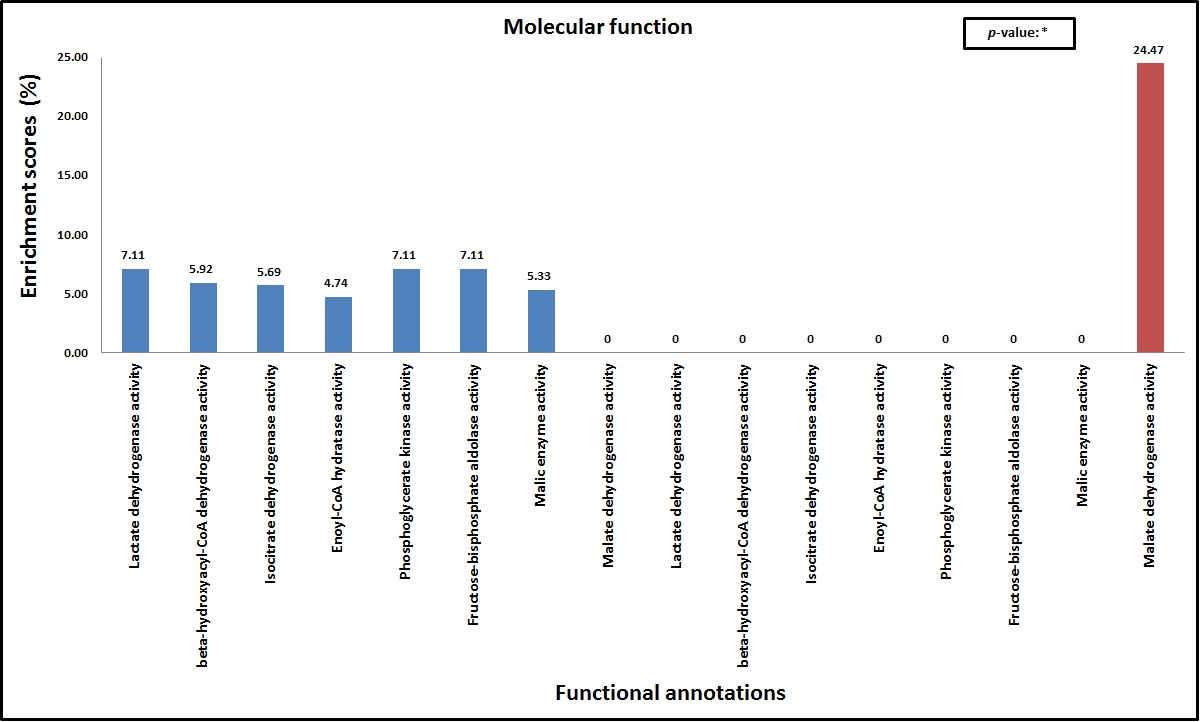
**Figure 16. CC metabolism-related enrichment scores*. Enrichment scores of the global proteome (left, blue) and the acetylated proteome (right, red) datasets of the top ten metabolism-related functional annotations of the fifth level of ontology for CC GO term. Functional annotations that were not found in both global and acetylated proteome datasets were scored as 0, since no comparison was feasible to be made amongst these functional annotations. The only functional annotation that showed major differences after comparing the global and acetylated proteome datasets was Mitochondrial nucleoid, which enrichment score changed from 8.55 in the acetylated proteome to 4.95 in the global proteome. This lack of evident differences in the enrichment scores is supported by the p-value of 0.62, indicating that there are no statistically significant differences in the changes observed in the enrichment scores between the global and the acetylated proteome datasets. The Wilcoxon rank sum test with continuity correction* *was used to compare the global and the acetylated proteome datasets (Shapiro-Wilk test p-value: 0.14 for the global proteome and 6.167e-05 for the acetylated proteome).***

For the BP GO term, 7 out of 10 metabolism-related functional annotations were shared (Figure 17). These functional annotations were RNA metabolic process, mRNA metabolic process, Cellular protein metabolic process, ncRNA metabolic process, Negative regulation of cellular protein metabolic process, Negative regulation of protein metabolic process, and Regulation of protein metabolic process (Figure 17). Opposed to the observations made in CC GO term, in BP GO term a more evident shift in the enrichment scores was observed between the global and the acetylated proteome datasets. This shift was mainly in Negative regulation of cellular protein metabolic process functional annotation, showing higher enrichment scores in the acetylated proteome dataset (3.96) when compared to the global proteome dataset (2.34) (Figure 17). Even the other functional annotations did not show great differences between the global and the acetylated proteome datasets, all these functional annotations showed higher enrichment scores in the acetylated proteome dataset (Figure 17).



**Figure 17.BP metabolism-related enrichment scores*. Enrichment scores of the global proteome (left, blue) and the acetylated proteome (right, red) datasets of the top ten metabolism-related functional annotations of the fifth level of ontology for BP GO term. Functional annotations that were not found in both global and acetylated proteome datasets were scored as 0 since no comparison was feasible to be made amongst these functional annotations. Only Negative regulation of cellular protein metabolic process functional annotation showed evident differences in the enrichment scores*** ***after comparing the global and acetylated proteome datasets. The lack of evident differences in the enrichment scores of the functional annotations is supported by the p-value of 0.1 indicating that there are no statistically significant differences in the changes observed in the enrichment scores between the global and the acetylated proteome datasets. The Welch Two Sample t-test was used to compare the global and the acetylated proteome datasets (Shapiro-Wilk test p-value: 0.58 for the global proteome and 0.06 for the acetylated proteome).***

The results for MF GO term showed that any of the metabolism-related functional annotations were shared in between the global and acetylated proteome datasets (Figure 18). Only seven metabolism-related functional annotations were found in the fifth level of ontology in the global proteome, whereas just one metabolism-related functional annotation was found in the acetylated proteome dataset (Figure 18). The functional annotations found in the global proteome datasets are Lactate dehydrogenase activity, β-hydroxyacyl-CoA dehydrogenase activity, Isocitrate dehydrogenase activity, Enoyl-CoA hydratase activity, Phosphoglycerate kinase activity, Fructose-bisphosphate aldolase activity and Malic enzyme activity. On the other hand, the functional annotation found in the acetylated proteome dataset was Malate dehydrogenase activity. As expected and due to the aims of this project all the functional annotations depicted in this analysis belong to metabolic pathways (Figure 18). Lactate dehydrogenase, malic enzyme and malate dehydrogenase take part in pyruvate metabolism; β-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase are enzymes of the β-oxidation pathway; phosphoglycerate kinase and fructose bisphosphate aldolase are glycolytic enzymes; isocitrate dehydrogenase takes part in the TCA cycle.



**Figure 18*.* MF metabolism-related enrichment scores*. Enrichment scores of the global proteome (left, blue) and the acetylated proteome (right, red) datasets of the metabolism-related functional annotations of the fifth level of ontology for MF GO term. Unlike all the previously described functional annotations in each GO term, the metabolism-related functional annotations in MF GO term were not shared in between the global and acetylated proteome datasets. Additionally, only seven functional annotations were found at this level of ontology in the global proteome dataset, and just one was found in the acetylated proteome dataset. \*No comparison was made in between the global and acetylated proteome datasets, since there were not enough observations. The only statistical analysis that could be made is the Shapiro-Wilk test for the metabolism-related functional annotations found in the global proteome dataset (p-value: 0.16).***

# Matching pathways

The analysis of the different pathways was made through Reactome (<http://www.reactome.org/>), manually curated and peer-reviewed open access software. Twenty four different pathways were found to match the global proteome, from which twenty three were found to match the acetylated proteome dataset, since no entities for Reproduction pathway were found in the acetylated proteome dataset (Table 11). Furthermore to Reproduction pathway not being found in the acetylated proteome dataset, there was a change in the number of entities found in the global and acetylated proteome datasets in each pathway, as well as the coverage in every pathway (Table 11). The global proteome showed more entities in every pathway, and an 11-46% higher coverage of pathways when compared to the acetylated proteome. This finding can also be related to stoichiometry and that not every enzyme is substrate for acetylation.

From all the pathways matching the global and the acetylated proteome datasets, Metabolism pathway was chosen, according to the aims of this project, in order to interrogate which metabolic enzymes in the global and in the acetylated proteome datasets take part in this pathway. For the global proteome, 10 glycolytic enzymes were found to take part in this pathway, whereas 3 TCA cycle enzymes were found to take part (Table 12). In terms of β-oxidation, 1 enzyme was found in Metabolism pathway, whilst 3 enzymes in pyruvate metabolism were also found to take part in this pathway (Table 12). For the acetylated proteome, 9 glycolytic enzymes were found to take part in Metabolism pathway, whilst 3 TCA cycle enzymes were found (Table 12). Regarding β-oxidation there were no enzymes found in this pathway, whereas 2 enzymes in pyruvate metabolism were found to match this pathway (Table 12).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Global proteome | | | Acetylated proteome | | |
| Pathway name | **Entities total** | **Entities found** | **Coverage (%)** | ***p*-value** | **Entities found** | **Coverage (%)** | ***p*-value** |
| DNA Replication | 104 | 70 | 67.3 | 8.88E-05 | 29 | 27.8 | 1.81E-07 |
| Apoptosis | 161 | 80 | 49.6 | 0.001 | 37 | 22.9 | 4.39E-07 |
| Cell Cycle | 532 | 205 | 38.5 | 0.068 | 69 | 12.9 | 0.001 |
| Metabolism of proteins | 685 | 307 | 44.8 | 0.105 | 114 | 16.6 | 2.57E-10 |
| Membrane Trafficking | 206 | 103 | 50 | 0.401 | 28 | 13.5 | 0.024 |
| Muscle contraction | 51 | 26 | 50.9 | 0.411 | 8 | 15.6 | 0.096 |
| Gene Expression | 1141 | 464 | 40.6 | 0.643 | 173 | 15.1 | 3.70E-12 |
| Cellular responses to stress | 236 | 108 | 45.7 | 0.882 | 27 | 11.4 | 0.132 |
| Chromatin organization | 183 | 85 | 46.4 | 0.902 | 10 | 5.4 | 0.969 |
| DNA Repair | 114 | 45 | 39.4 | 0.906 | 12 | 10.5 | 0.340 |
| Cell-Cell communication | 138 | 35 | 25.3 | 0.979 | 11 | 7.9 | 0.705 |
| Organelle biogenesis and maintenance | 55 | 20 | 36.3 | 0.983 | 3 | 5.4 | 0.874 |
| Circadian Clock | 55 | 14 | 25.4 | 0.996 | 4 | 7.2 | 0.733 |
| Hemostasis | 490 | 183 | 37.3 | 0.999 | 41 | 8.3 | 0.722 |
| Reproduction | 26 | 3 | 11.5 | 0.999 | 0 | 0 |  |
| Binding and Uptake of Ligands by Scavenger Receptors | 195 | 11 | 5.6 | 1 | 5 | 2.5 | 0.999 |
| Developmental Biology | 509 | 173 | 33.9 | 1 | 34 | 6.6 | 0.976 |
| Disease | 1309 | 481 | 36.7 | 1 | 144 | 11 | 0.007 |
| Extracellular matrix organization | 268 | 43 | 16 | 1 | 9 | 3.3 | 0.999 |
| Immune System | 1427 | 445 | 31.1 | 1 | 89 | 6.2 | 0.999 |
| Metabolism | 1587 | 551 | 34.7 | 1 | 129 | 8.1 | 0.925 |
| Neuronal System | 280 | 58 | 20.7 | 1 | 9 | 3.2 | 0.999 |
| Signal Transduction | 1994 | 357 | 17.9 | 1 | 70 | 3.5 | 1 |
| Transmembrane transport of small molecules | 549 | 100 | 18.2 | 1 | 14 | 2.5 | 1 |

Table 11. Matching pathways. *Different* p*athways found in the analysis, and the number of proteins (Entities found) matching the total number of proteins (Entities total) in each of these pathways for the global and the acetylated proteome datasets. Coverage, expressed as %, is the ratio between the Entities found and the Entities total.*

The glycolytic enzyme that was not found to match Metabolism in the acetylated proteome dataset is hexokinase. The three same TCA cycle enzymes were found to match Metabolism pathway in both the global and the acetylated proteome datasets. Phosphoenolpyruvate carboxykinase is the pyruvate metabolism enzyme that was not found to match Metabolism pathway in the acetylated proteome dataset. Enoyl-CoA hydratase, the β-oxidation enzyme that was found to match this pathway in the global proteome, was not found in the acetylated proteome dataset. That the enzymes mentioned above could not be found to match Metabolism pathway in the acetylated proteome dataset might be due to these enzymes having few lysine residues amenable to acetylation (see Appendix 2 for the complete list of acetylated lysine residues), compared to the enzymes that were found to match this pathway.

|  |  |  |
| --- | --- | --- |
| Pathway name | Global proteome | Acetylated proteome |
| Metabolism | Enolase (Glycolysis)  Enoyl-CoA hydratase (β-oxidation)  Fructose-bisphosphate aldolase (Glycolysis)  Glucose-6-phosphate isomerase (Glycolysis)  Glyceraldehyde-3-phosphate dehydrogenase (Glycolysis)  Hexokinase (Glycolysis)  Isocitrate dehydrogenase (the TCA cycle)  Lactate dehydrogenase (Pyruvate metabolism)  Malate dehydrogenase (the TCA cycle)  Malic enzyme (Pyruvate metabolism)  Phosphoenolpyruvate carboxykinase (Pyruvate metabolism)  Phosphofructokinase (Glycolysis)  Phosphoglycerate kinase (Glycolysis)  Phosphoglycerate mutase (Glycolysis)  Pyruvate kinase (Glycolysis/Pyruvate metabolism)  Succinate dehydrogenase (the TCA cycle)  Triosephosphate isomerase (Glycolysis) | Enolase (Glycolysis)  Fructose-bisphosphate aldolase (Glycolysis)  Glucose-6-phosphate isomerase (Glycolysis)  Glyceraldehyde-3-phosphate dehydrogenase (Glycolysis)  Isocitrate dehydrogenase (the TCA cycle)  Lactate dehydrogenase (Pyruvate metabolism)  Malate dehydrogenase (the TCA cycle)  Phosphofructokinase (Glycolysis)  Phosphoglycerate kinase (Glycolysis)  Phosphoglycerate mutase (Glycolysis)  Pyruvate kinase (Glycolysis/Pyruvate metabolism)  Succinate dehydrogenase (the TCA cycle)  Triosephosphate isomerase (Glycolysis) |

Table 12. Enzymes matching Metabolism pathway. *From the total number of entities found to match the global and acetylated proteome datasets in Metabolism pathway (Table 11), just the metabolic enzymes found to match this pathway are shown. These enzymes have been selected upon the interest in metabolism. In parentheses, the metabolic pathway in which each enzyme takes part; blank, enzyme not found in the acetylated proteome dataset.*

# Action of SCFAs on acetyl-CoA and CoA

As it has been shown by quantitatively examining the data through bioinformatics approaches, not all metabolic enzymes are suitable of being acetylated (Representation and acetylation of metabolic pathways), there are differences in the enrichment scores of the global and the acetylated proteome datasets (Functional annotations), the number of enzymes matching the metabolic pathways analysed is higher in the global proteome when compared to the acetylated proteome (Percentage coverage of pathways) and more enzymes in the Metabolism pathway were found to match the global proteome compared to the acetylated proteome dataset (Matching pathways).

Although functional proteomics and bioinformatics are a valuable tool to elucidate the role(s) of SCFAs, there are some other approaches. One of these approaches is to model the action of these SCFAs in a time course simulation and to model the effect of increasing concentrations of SCFAs on acetyl-CoA and CoA levels.

A comparison of butyrate\_v3-1 and SCFA Uptake and Oxidation v4 was made to model the effect of acetate, butyrate and propionate on butyryl-CoA, propionyl-CoA, acetyl-CoA and in CoA levels during a time course simulation. The simulation of the action of increasing concentrations of acetate, butyrate and propionate on acetyl-CoA and CoA was also compared using these two models.

The action of acetate, butyrate and propionate at physiological ranges (1-52mM in the case of acetate and 1-20mM in the case of both butyrate and propionate) on their acyl forms, butyryl-CoA and propionyl-CoA as well as in acetyl-CoA and CoA at different time points was simulated by using the two models aforementioned. The simulations were run at 1000 intervals, with different time points ranging from 10 to 10000 seconds. When using the butyrate\_v3-1 model, the effect of SCFAs in the concentrations of butyryl-CoA, propionyl-CoA and acetyl-CoA as well as in CoA became evident after the first time point of 0.1 seconds (Figure 19). There were no changes in the concentrations after 100 seconds time point, since at 1000 and 10000 seconds time points the concentrations remained the same as in the 100 seconds time point (Figure 19).

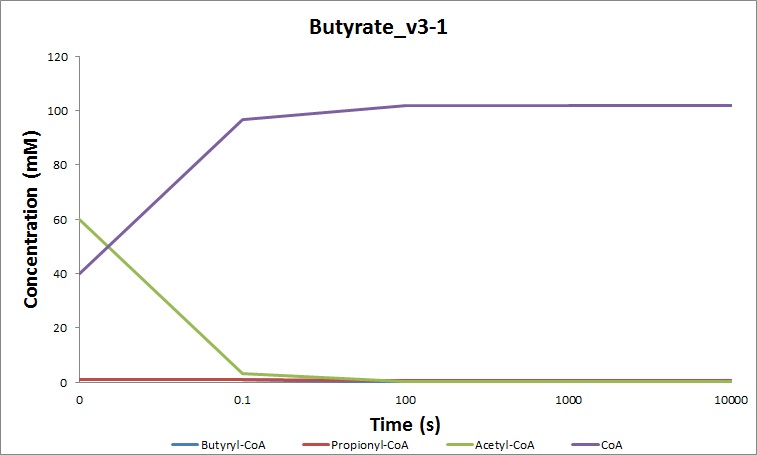
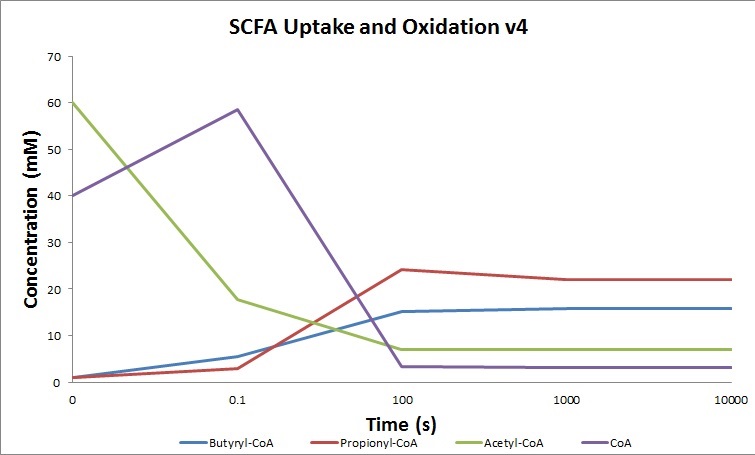


Figure 19. Changes in the concentration of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA at different time points using the butyrate\_v3-1 model. *The initial concentrations at 0s were as follows: butyryl-CoA 1mM, propionyl-CoA 1mM, acetyl-CoA 60mM, CoA 40mM; the concentrations at 0.1s were as follows: butyryl-CoA 0.88mM, propionyl-CoA 0.82mM, acetyl-CoA 3.35mM, CoA 96.60mM; the concentrations at 100s were as follows: butyryl-CoA 1.57x10-6mM, propionyl-CoA 0.61mM, acetyl-CoA 0.05mM, CoA 101.64mM. No further changes in the concentrations were observed at 1000 and 10000 seconds.*

A decrease in the concentration of butyryl-CoA, propionyl-CoA and acetyl-CoA was observed after the first time point of 0.1 seconds when compared to the initial concentration (Figure 19). The largest decrease in concentration at this time point was observed in acetyl-CoA (60-3.35mM), followed by propionyl-CoA (1-0.82mM) and butyryl-CoA (1-0.88mM). The acyl form with the highest concentration after 0.1 seconds time point was acetyl-CoA (3.35mM), followed by butyryl-CoA (0.88mM) and propionyl-CoA (0.82). Opposite to what was observed in the acyl forms, the concentration of CoA was increased (40-96.60mM) after 0.1 seconds time point. The concentration of CoA (96.60mM) was the highest in this time point.

The reductions observed at 100 seconds time point were even higher than the observed at 0.1 seconds time point. The highest decrease in concentration at this time point was observed again as in 0.1 seconds time point in acetyl-CoA (60-0.05mM), followed by butyryl-CoA (1-1.57x10-6mM) and propionyl-CoA (1-0.61mM) when compared to the initial concentration. Opposite to what was observed at 0.1 seconds, the acyl form with the highest concentration after 100 seconds time point was propionyl-CoA (0.61mM), followed by acetyl-CoA (0.05mM) and butyryl-CoA (1.57x10-6mM). As observed in the 0.1 seconds time point, an increase in the concentration of CoA (40-101.64mM) was also observed in 100 seconds time point when compared to the initial concentration.

On the other hand, when using the SCFA Uptake and Oxidation v4 model, the effect of SCFAs in the concentrations of butyryl-CoA, propionyl-CoA and acetyl-CoA as well as in CoA became evident after the first time point of 0.1 seconds (Figure 20). These changes in the concentrations were even higher after 100 seconds time point. A little difference was observed after the 1000 time point when compared to the 100 seconds time point. Alternatively, 10000 seconds time point showed no difference in the concentrations when compared to 1000 seconds time point (Figure 20).

**Figure 20. Changes in the concentration of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA at different time points using the SCFA Uptake and Oxidation v4 model. *The initial concentrations at 0s were as follows: butyryl-CoA 1mM, propionyl-CoA 1mM, acetyl-CoA 60mM, CoA 40mM; the concentrations at 0.1s were as follows: butyryl-CoA 5.55mM, propionyl-CoA 2.88mM, acetyl-CoA 17.79mM, CoA 58.56mM; the concentrations at 100s were as follows: butyryl-CoA 15.14mM, propionyl-CoA 24.19mM, acetyl-CoA 7.08mM, CoA 3.40mM; the concentrations at 1000s were as follows: butyryl-CoA 15.72mM, propionyl-CoA 21.95mM, acetyl-CoA 6.91mM, CoA 3.12mM. No further changes in the concentrations were observed at 10000 seconds.***

Interestingly and opposite to what was observed in the simulations run using the butyrate\_v3-1 model, the simulations run using the SCFA Uptake and Oxidation v4 model, an increase in the concentration of butyryl-CoA, propionyl-CoA, and CoA was observed after 0.1 seconds time point. The highest increase was observed in CoA (40-58.56mM), followed by butyryl-CoA (1-5.55mM) and propionyl-CoA (1-2.88mM). Contrasting with the observations made when using the butyrate\_v3-1 model, a reduction in acetyl-CoA concentration (60-17.79mM) was observed in the SCFA Uptake and Oxidation v4 model when compared to the initial concentration. The highest concentration in this time point was observed in CoA (58.56mM), followed by acetyl-CoA (17.79mM), butyryl-CoA (5.55mM) and propionyl-CoA (2.88mM) (Figure 20).

After the 100 seconds time point just butyryl-CoA and propionyl-CoA had an increase in their concentration (1-15.14mM and 1-24.19mM respectively) when compared to their initial concentrations. On the other hand, a reduction in the concentrations of acetyl-CoA and CoA (60-7.08mM and 40-3.40mM) was observed in the 100 seconds time point. In this time point, the highest concentration was observed in propionyl-CoA (24.19mM), followed by butyryl-CoA (15.14mM), acetyl-CoA (7.08mM), and CoA (3.40mM).

At 1000 seconds time point, butyryl-CoA and propionyl-CoA showed an increase in their concentration again (1-15.72mM and 1-21.95mM respectively) as in the 100 seconds time point when compared to the initial concentration. In terms of comparing the 100 and 1000 seconds time points, just butyryl-CoA observed an increase in its concentration (15.14-15.72mM), whereas a decrease in the concentration of propionyl-CoA (24.19-21.95mM) was observed in these terms. As observed in 100 seconds time point, a decrease in the concentration of acetyl-CoA and CoA (60-6.91mM and 40-3.12mM) was also observed when compared to the control. The highest concentration in this time point was observed in propionyl-CoA (21.95mM), followed by butyryl-CoA (15.72mM), acetyl-CoA (6.91mM) and CoA (3.12mM).

The action of increasing concentration of acetate, butyrate and propionate at physiological ranges (1-52mM in the case of acetate and 1-20mM in the case of both butyrate and propionate) on acetyl-CoA and CoA was simulated using butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models at 1000 intervals. The action of increasing concentrations of glucose at physiological ranges (0.5-2mM) at 1000 intervals was only simulated using butyrate\_v3-1 model, since SCFA Uptake and Oxidation v4 model does not have glucose defined as a specie.

By using butyrate\_v3-1 model, the increasing concentration of acetate produced an increase of 66.58mM (60-126.58mM) in the concentration of acetyl-CoA (Figure 21, left panel). Interestingly, by using SCFA Uptake and Oxidation v4 model, acetate had a scarce effect on acetyl-CoA levels. Even though the effect was scarce, it was the opposite of the effect observed when the simulations were run with butyrate\_v3-1 models, since a decrease of 0.06mM (60-59.94mM) was observed (Figure 21, right panel).

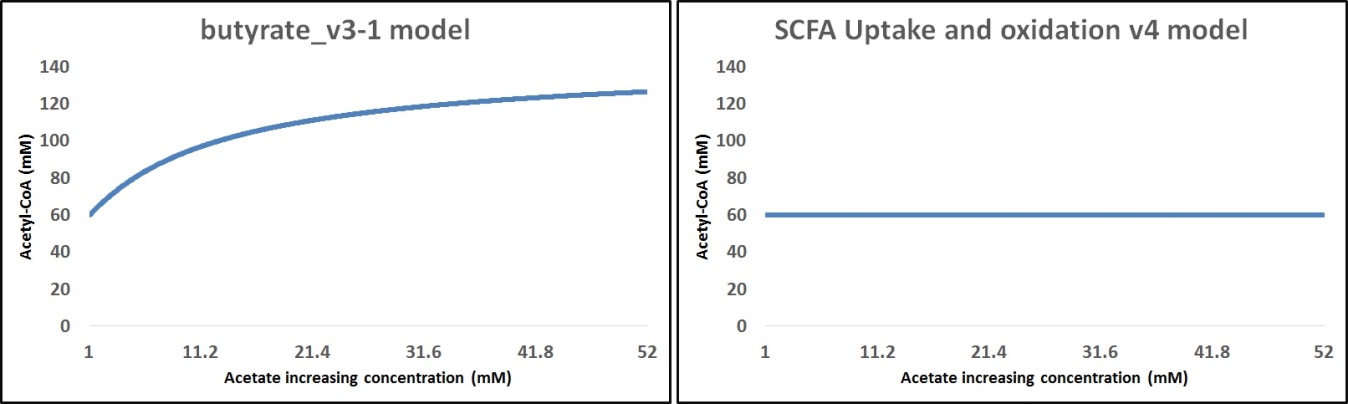


Figure 21. Effect of acetate on acetyl-CoA levels. *The left panel shows the effect of this SCFA on acetyl-CoA levels when simulations were run with butyrate\_v3-1 model. The right panel shows the effect of acetate on acetyl-CoA levels when simulations were run with SCFA Uptake and Oxidation v4 model.*

As with acetate, models showed an opposite effect on the action of increasing butyrate concentrations at physiological ranges (1-20mM) on acetyl-CoA levels. The simulations run with butyrate\_v3-1 model showed an increase of 87.23mM (60-147.23mM) on acetyl-CoA levels, whereas the simulations run with SCFA Uptake and oxidation v4 model showed a decrease of 34.62 (60-34.67mM) on acetyl-CoA levels (Figure 22, left panel byturate\_v3-1 model; right panel SCFA Uptake and Oxidation v4 model).

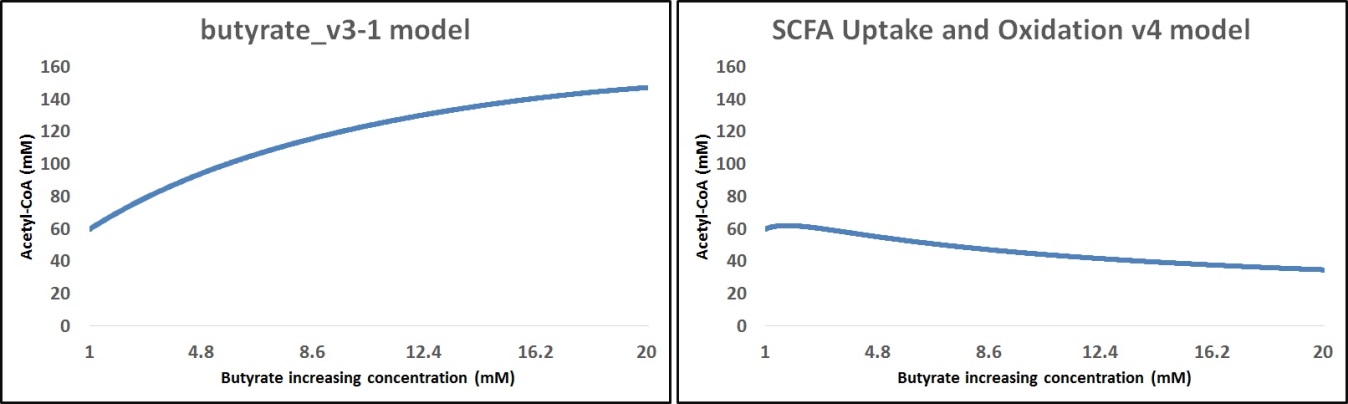


Figure 22. Effect of butyrate on acetyl-CoA levels. *The left panel shows the effect of this SCFA on acetyl-CoA levels when simulations were run with butyrate\_v3-1 model. The right panel shows the effect of acetate on acetyl-CoA levels when simulations were run with SCFA Uptake and Oxidation v4 model.*

Even the increasing concentration of propionate is not expected to produce changes in acetyl-CoA levels, since it is not able to be metabolised in β-oxidation to produce acetyl-CoA, simulations with both models, butyrate\_v3-1 and SCFA Uptake and Oxidation v4 were run to compare these two models. When the simulations were run with butyrate\_v3-1 model, the increasing concentration of propionate at physiological ranges (1-20mM) had virtually no effect on acetyl-CoA levels, since an increase of just 1.5μM (60-60.0015mM) was observed (Figure 23, left panel). In accordance to the observations made with acetate and butyrate and opposite effect in acetyl-CoA levels were observed when simulations were run with SCFA Uptake and Oxidation v4 model. A decrease of 24.52mM (60-35.47mM) was observed by using this model (Figure 23, right panel). This reduction in acetyl-CoA levels could be due to the production of acetate and propionyl-CoA from acetyl-CoA and propionate ([87](#_ENREF_87)).

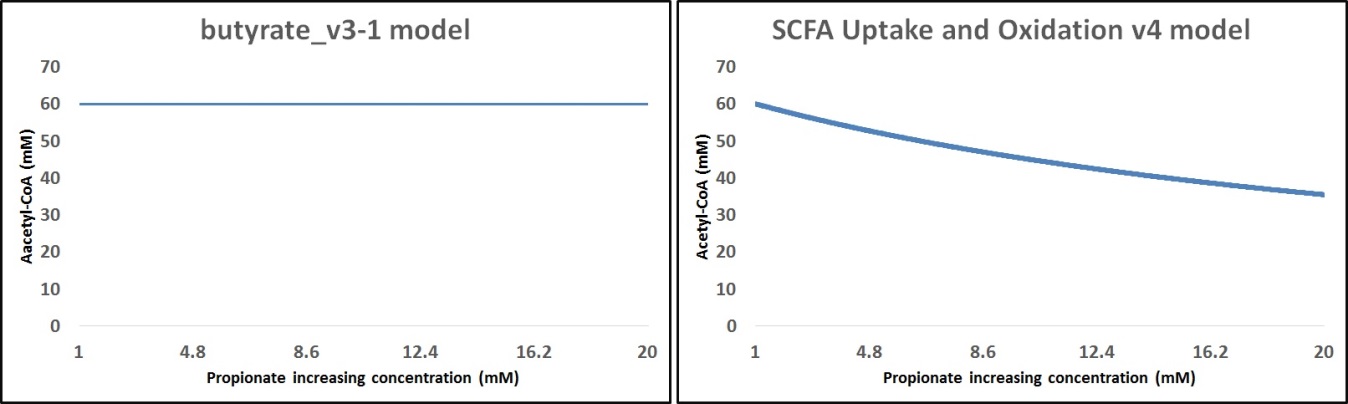


Figure 23. Effect of propionate on acetyl-CoA levels. *The left panel shows the effect of this SCFA on acetyl-CoA levels when simulations were run with butyrate\_v3-1 model. The right panel shows the effect of acetate on acetyl-CoA levels when simulations were run with SCFA Uptake and Oxidation v4 model.*

Since glucose ligation with CoA is able to produce acetyl-CoA as well, the action of this molecule on acetyl-CoA levels was also simulated by using butyrate\_v3-1 model, but not SCFA Uptake and Oxidation v4 since glucose is not defined as a specie in this model. At physiological ranges, 0.5-2mM, glucose increasing concentration produced a slight decrease of 0.04mM (60-59.96mM) on the acetyl-CoA levels (Figure 24).

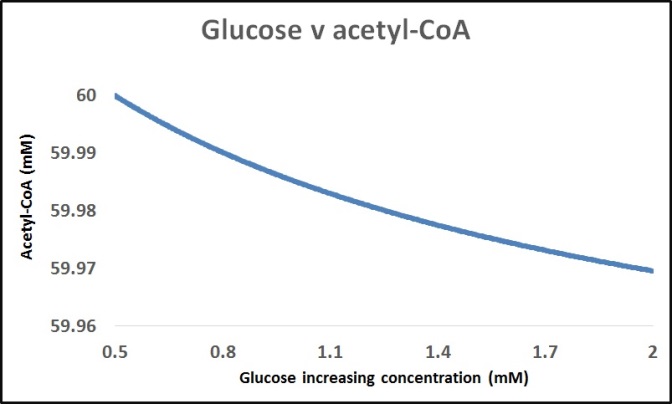


Figure 24. Effect of glucose on acetyl-CoA levels. *Effect of glucose on acetyl-CoA levels when simulations were run with butyrate\_v3-1 model.*

With regard of the action of acetate, butyrate and propionate on CoA, butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models the simulations with these models were also compared. As in respect with acetyl-CoA levels, there were mixed results in the levels of CoA, depending on the SCFA. When the simulations were run with butyrate\_v3-1 model, the increasing concentration of acetate at physiological ranges (1-52mM) showed an increase of 0.14mM (40-40.14mM) in CoA levels (Figure 25, left panel). On the other hand, when the simulations were run using the SCFA Uptake and Oxidation v4 model, a decrease of 0.12mM (40-39.88mM) was observed (Figure 25, right panel).

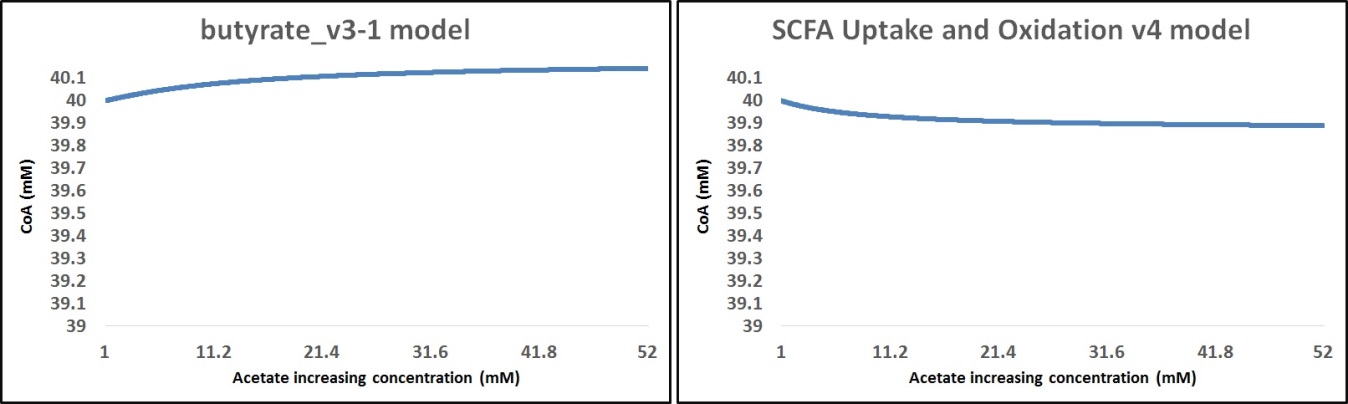


Figure 25. Effect of acetate on CoA levels. *The left panel shows the effect of this SCFA on CoA levels when simulations were run with butyrate\_v3-1 model. The right panel shows the effect of acetate on CoA levels when simulations were run with SCFA Uptake and Oxidation v4 model.*

The increasing concentration of butyrate produced a decrease in the CoA levels of 12.17mM (40-27.83mM) at physiological ranges (1-20mM) when simulations were run with butyrate\_v3-1 model (Figure 26, left panel). SCFA Uptake and Oxidation v4 simulations also showed a decrease in CoA levels (Figure 26, right panel). This decrease was higher than the one observed when simulations were run in butyrate\_v3-1, since the levels of CoA decreased by 34.82mM (40-5.18mM).

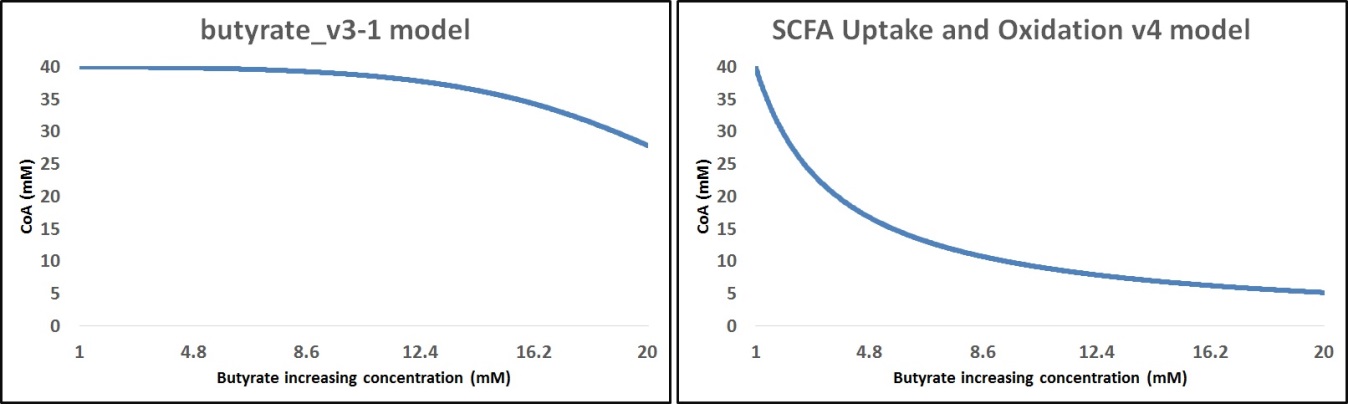


Figure 26. Effect of butyrate on CoA levels. *The left panel shows the effect of this SCFA on CoA levels when simulations were run with butyrate\_v3-1 model. The right panel shows the effect of butyrate on CoA levels when simulations were run with SCFA Uptake and Oxidation v4 model.*

In the simulations run with butyrate\_v3-1 model, the increasing concentrations of propionate produced a slight decrease of 0.15mM (40-39.85mM) in CoA levels (Figure 27, left panel). When using the SCFA Uptake and Oxidation v4 model, a decrease in CoA levels was also observed (Figure 27, right panel) with increasing concentrations of propionate. This decrease was higher, 16.35mM (40-23.65mM), than the decrease in the simulations run with butyrate\_v3-1 model (0.15mM; 40-39.85mM), and it was in accordance with the observations made for butyrate (Figure 27, right panel).

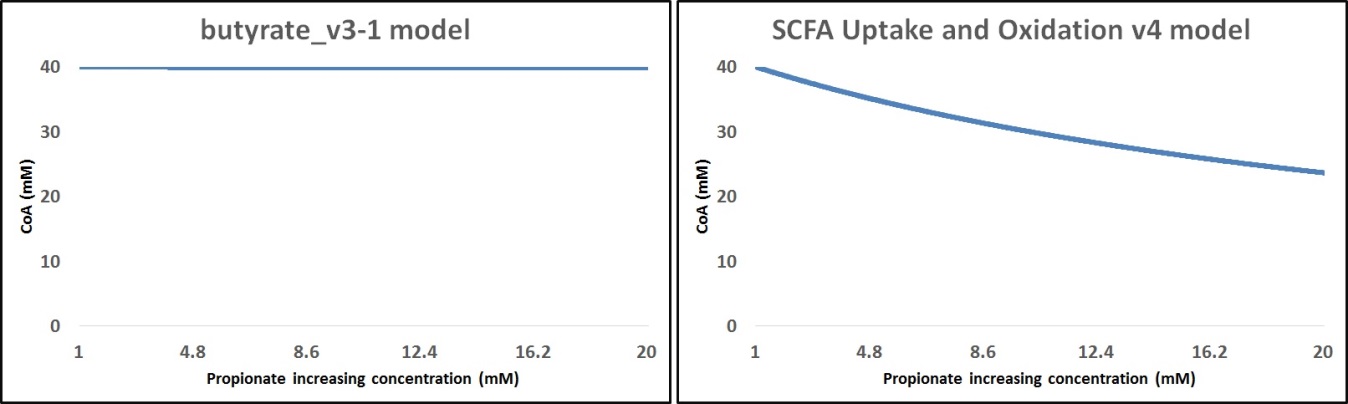


Figure 27. Effect of propionate on CoA levels. *The left panel shows the effect of this SCFA on CoA levels when simulations were run with butyrate\_v3-1 model. The right panel shows the effect of propionate on CoA levels when simulations were run with SCFA Uptake and Oxidation v4 model.*

The same rationale as with acetyl-CoA production and glucose – CoA-ligation was applied to run simulations on increasing glucose concentration on CoA levels. As aforesaid, glucose and CoA-ligation is another fashion in which acetyl-CoA con be produced. Glucose can be competing with SCFAs for CoA as a substrate, and that is why it is important to simulate the effect of glucose on CoA levels. Simulations, as in with acetyl-CoA levels, were only run with butyrate\_v3-1 model since SCFA Uptake and Oxidation v4 model has not glucose defined as specie. When used at physiological ranges (0.5-2mM) glucose produced a scarce decrease of 0.03mM (40-39.97mM) (Figure 28).

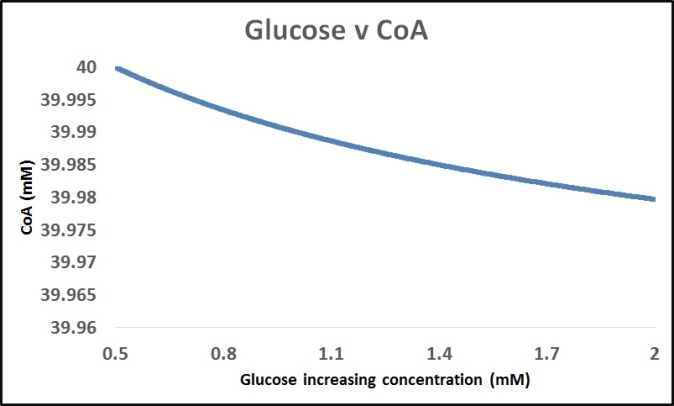


Figure 28. Effect of glucose on CoA levels. *Effect of glucose on CoA levels when simulations were run with butyrate\_v3-1 model.*

Even these results must be validated by experimental data they can also be validated based on previous findings. There is evidence that the increase in acetyl-CoA levels by the action of butyrate, as simulated using butyrate\_v3-1 model (Figure 22, left panel) could be related to its regulatory role on protein acetylation ([10](#_ENREF_10)) and its role on histone acetylation ([4](#_ENREF_4)). The production of butyryl-CoA as simulated using SCFA Uptake and Oxidation v4 model (Figure 20) could be related to the inhibitory mechanism of butyrate ([40](#_ENREF_40)). Furthermore, the increase in the production of acyl forms (Figure 20) could simulate cells to be driven to the acylated state by the action of SCFAs ([40](#_ENREF_40)). Moreover, a decrease in acetyl-CoA levels (Figure 20) could be due to the production of acetate and propionyl-CoA ([87](#_ENREF_87)).

# Validation of the models

In order to validate the simulations run using butyrate\_v3-1 and SCFA Uptake and Oxidation, HCT116 cells were cultured by Andrew Nichols and Marilena Epitropou, two MSc students under my supervision, with increasing concentrations of acetate, butyrate and propionate at physiological ranges (1-50mM in the case of acetate, and 1-20mM in the case of butyrate and propionate). The treatment with acetate produced an increase of 0.34μM (3.5x10μM - 3.84x10μM) in CoA levels (Figure 29, upper left panel). On the other hand, the treatment with butyrate produced an increase of 0.1μM (3.5x10μM - 3.6x10μM) on CoA levels (Figure 29, upper right panel), whereas the treatment with propionate produced an increase of 0.96μM (3.5x10μM - 4.46x10μM) on CoA levels (Figure 29, bottom panel).

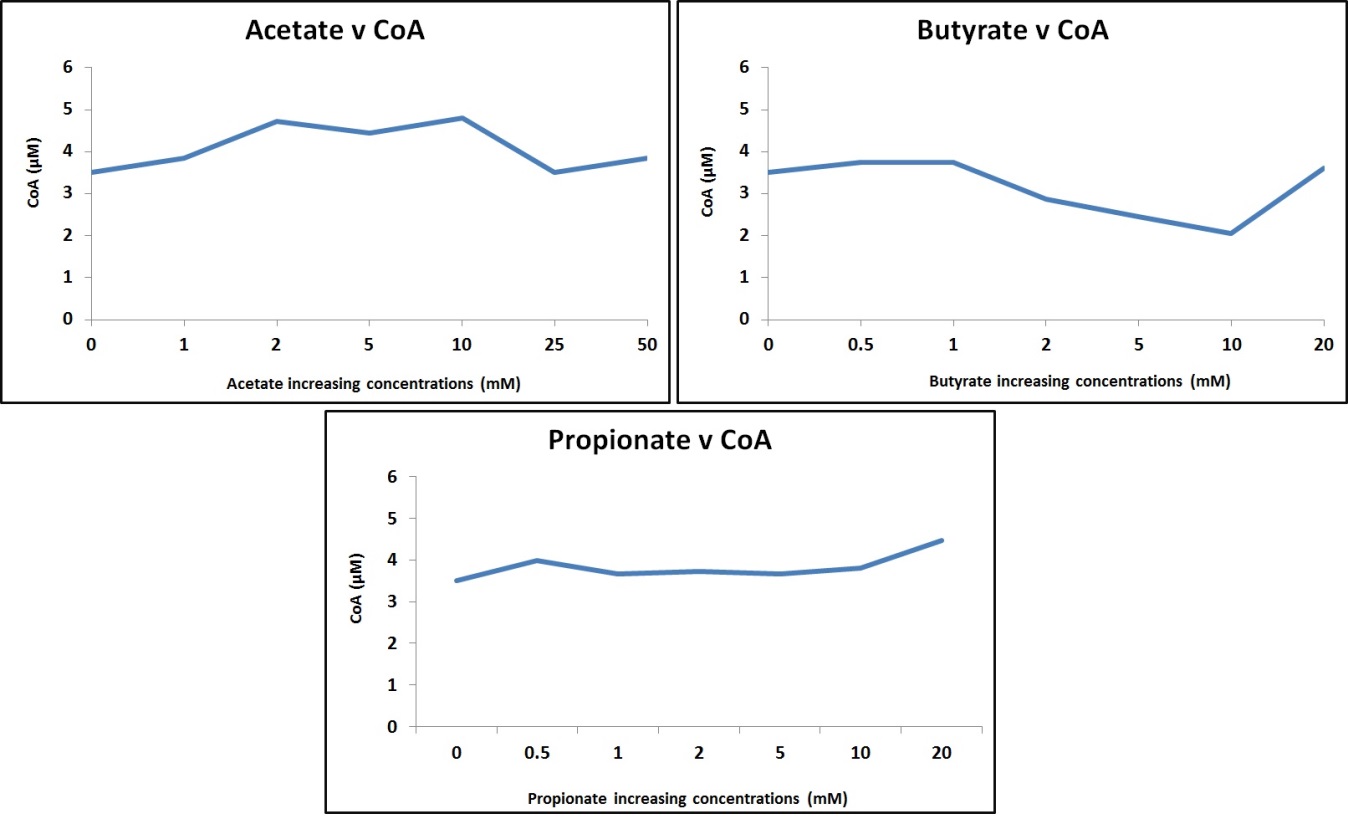
****

Figure 29. Effect of increasing concentrations of SCFAs on CoA levels. *Effect of increasing concentrations, at physiological ranges, of acetate (1-50mM, upper left panel), butyrate (1-20mM, upper right panel), and propionate (1-20mM, bottom panel) on CoA levels in HCT116 colon cancer cells treated with these SCFAs.*

The changes in CoA levels were not as steady as in the simulations and the general outcome, increase in CoA levels, was only in accordance to the simulations run with butyrate\_v3-1, in the case of acetate. In the case of butyrate and propionate the effect observed when HCT116 cells were treated with these SCFAs is opposite to the effect observed and expected, decrease in CoA levels, when simulations were run with both butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models. Another important fact to address is that the magnitude of these changes is higher in the simulations run with butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models, in the range of millimolar (mM), whereas it is lower in the case of HCT116 cells treated with acetate, butyrate and propionate, in the range of micro molar (μM).

# DISCUSSION

Recent findings are providing new evidence of a novel role of SCFAs related to control metabolic enzymes as well as acetylation ([4](#_ENREF_4)). By controlling the acetylation patterns of metabolic enzymes, SCFAs could be modulating metabolism. The modulation of metabolism via acetylation, and an up-regulation of proteins has previously been described by Wellen and Thompson, 2012 ([88](#_ENREF_88)). It has also been suggested that acetylation regulates proteins/enzymes and cellular metabolism ([89](#_ENREF_89)) by preserving the equilibrium between carbohydrate and metabolism of fatty acid synthesis ([61](#_ENREF_61)) as well as controlling metabolic enzymes transcription levels ([9](#_ENREF_9)). A reduction in the acetylation levels of metabolic enzymes could be linked to the KDAC inhibitory role of SCFAs ([41](#_ENREF_41), [90](#_ENREF_90)), altering the equilibrium between the opposing effects of KATs and KDACs.

Changes in the acetylation patterns of metabolic enzymes are very important in terms of their expression and function. For example in the TCA cycle, the acetylation of malate dehydrogenase ([60](#_ENREF_60), [61](#_ENREF_61)) and isocitrate dehydrogenase enhances their activity ([56](#_ENREF_56), [57](#_ENREF_57)), whereas in glycolysis, phosphoglycerate mutase deacetylation via Sirt1 inhibits its activity using the same mechanism and under the same conditions as Sirt3 to deacetylate and inhibit isocitrate dehydrogenase ([54](#_ENREF_54)). Therefore, it is reasonable to propose that Sirt1 and Sirt3 could have a key role in regulating metabolic enzymes, tumour proliferation and survival. This regulation is tightly linked to energy metabolism, metabolic syndromes and cancer (Warburg effect) ([54](#_ENREF_54)). On the other hand, some of the effects of the change in the acetylation patterns of metabolic enzymes are related to enzymatic activity. For example, the glycolytic enzymes fructose-biphosphate aldolase and glyceraldehyde 3-phosphate dehydrogenase function as cellular instruments to deactivate enzymatic action when they are acetylated ([12](#_ENREF_12)). This function could be related to fructose-biphosphate aldolase specific pattern of acetylation, depending on their localisation and also that glyceraldehyde 3-phosphate dehydrogenase is acetylated in a specific binding site ([12](#_ENREF_12)). Due to this specificity in terms of the acetylation patterns of the different isoforms of fructose-biphosphate aldolase, which depends on their localisation, a tissue-specific KATs mechanism has been proposed by Lundby *et al.,* 2012 ([12](#_ENREF_12)).

Deacetylation of metabolic enzymes is not the only feature associated with glucose deprivation, since it has been shown that in these conditions, the Warburg effect, the metabolic shift from oxidative metabolism to aerobic glycolysis, is prevented ([4](#_ENREF_4)). Coupled with metabolic plasticity, there is also a common thought that excessively high rates of glycolysis in cancer are in response to the loss of ATP which could enhance the Warburg effect, determined primarily by glucose and aerobic glycolysis requirements. The shift from oxidative metabolism to glucose and aerobic glycolysis is a hallmark of cancer cells metabolism ([4](#_ENREF_4)). The Warburg effect is also related to acetylation patterns, since Sirt1 reversible regulation of acetylation of glycolytic enzymes could play a key role in energy metabolism, metabolic syndromes and in enhancing the Warburg effect ([54](#_ENREF_54)).

Another enzyme for which acetylation patterns are important is pyruvate kinase, since it is involved in glycolysis and pyruvate metabolism. The importance of this enzyme relies on opposing effects, since the up-regulation of this enzyme produces a decrease in the glycolytic flux ([91](#_ENREF_91)) and the ablation of glycolysis could be due to a failure in the acetylation patterns of this enzyme ([66](#_ENREF_66)). This is possibly one of the key roles of pyruvate kinase in cancer and tumour growth in neoplastic colon tissue ([92](#_ENREF_92)). Moreover, pyruvate kinase is also related to independent caspase apoptosis in tumour cells ([93](#_ENREF_93)), and transcriptional activity in cancer ([94](#_ENREF_94)).

These changes in the acetylation patterns of metabolic enzymes can also be reflected in terms of the representation and analysis of metabolic pathways, since only 19 out of 26 enzymes were detected to be acetylated in the metabolic pathways analysed. There were 9 out of 10 acetylated enzymes in glycolysis, 5 out of 8 in the TCA cycle, 3 out of 4 in β-oxidation, and 4 out of 6 in pyruvate metabolism. As some metabolic enzymes were found to be taking part in more than one metabolic pathway, an overlap of glycolysis, the TCA cycle, β-oxidation and pyruvate metabolism indicates these pathways to be closely related. The enzymes and molecules taking part in more than one metabolic pathway are oxaloacetate, malate, pyruvate and acetyl-CoA. Acetyl-CoA produced in β-oxidation enters the TCA cycle where it is used along with oxaloacetate by malate dehydrogenase as substrates to produce citrate and malate, and becomes involved in β-oxidation and the TCA cycle. Malate produced from oxaloacetate (metabolised by malate dehydrogenase) or fumarate (metabolised by fumarase) is used as a substrate by malic enzyme to produce pyruvate, linking the TCA cycle and pyruvate metabolism. Glycolysis is linked to TCA cycle in two ways. The first one involves the production of pyruvate, which is the last step in glycolysis, and is used as a substrate by pyruvate carboxylase and metabolised into oxaloacetate. The other way involves oxaloacetate production since phosphoenolpyruvate is used as a substrate by phosphoenolpyruvate carboxykinase to produce oxaloacetate.

The percentage coverage of pathways is another way in which the metabolic enzymes are affected by acetylation. In the global proteome dataset the TCA cycle was the pathway with the highest percentage coverage of pathways with 81%, followed by glycolysis and pyruvate metabolism with 65%, and β-oxidation with 55%. In the acetylated proteome dataset, the pathway with the highest coverage was glycolysis with 23%, followed by pyruvate metabolism with 20%, the TCA cycle with 19%, and β-oxidation with 18%. Matching pathways, the number of pathways found to match the global and acetylated proteome datasets, showed differences as well. Twenty four different pathways were found to match the global proteome, from which twenty three were found to match the acetylated proteome dataset. The global proteome showed more entities in each pathway, and an 11-46% higher coverage of pathways when compared to the acetylated proteome. The specific analysis of Metabolism pathway showed that 10 glycolytic enzymes, 3 TCA cycle enzymes, 3 enzymes in pyruvate metabolism and 1 β-oxidation enzyme were found to take part in this pathway in the global proteome, whereas 9 glycolytic enzymes, 3 TCA cycle enzymes, 2 pyruvate metabolism and none β-oxidation enzymes were found to take part in this pathway in the acetylated proteome.

Changes related to acetylation of the proteome were also observed in terms of the enrichment scores of the functional annotations in the GO terms chosen. The higher enrichment scores in some functional annotations in the acetylated proteome dataset show an up-regulation associated with acetylation of a group of proteins in a specific GO term, since these enrichments scores are a rank of biological significance of a group of proteins evidencing a shift in the expression of a specific pathway protein set. This analysis reveals the degree to which a protein set is overrepresented in an expression dataset, indicating the percentage of proteins contributing to the enrichment score in every functional annotation of each GO term. Differences in the number of functional annotations shared between the global and acetylated proteome datasets were also observed. BP GO term shared the most functional annotations between the global and the acetylated proteome datasets, 8 out of 10, followed by CC GO term 7 out of 10 and MF GO term 6 out of 10. In the metabolism-related functional annotations, CC GO term shared the most metabolism-related functional annotations between the global and acetylated proteome datasets, 9 out of 10, followed by BP GO term, 7 out of 10, and MF GO term with none metabolism-related functional annotations shared between the global and acetylated proteome datasets. Finding functional annotations being shared in both proteome datasets could indicate that there are similar proteins contributing to a specific functional annotation, but the percentage of proteins contributing to each enrichment score in every functional annotation of every GO term will be different, inducing a change in the enrichment scores. For example, a decrease in the enrichment scores in the acetylated proteome dataset indicates less proteins contributing to a specific functional annotation, which as aforesaid is expected, since not all proteins are substrate of acetylation. Even a decrease in the enrichment scores in the acetylated proteome is expected, the metabolism-related functional annotation Malate dehydrogenase activity in MF GO term was only found in the acetylated proteome dataset with an enrichment score of 24.47 (Figure 18), indicating that the activation of this TCA cycle enzyme is related to acetylation, as suggested by Zhao *et al*., 2010 ([11](#_ENREF_11)) and showed by Finkemeier *et al.,* 2011 ([55](#_ENREF_55)) and Kim *et al.,* 2012 and 2013 ([60](#_ENREF_60), [61](#_ENREF_61)).

The increase in acetyl-CoA levels by butyrate, as simulated by using butyrate\_v3-1 model (Figure 22, left panel), could be related to its regulatory role on protein acetylation patterns by both its concentration levels and its role in inhibiting KDACs ([10](#_ENREF_10)). This increase in acetyl-CoA levels could also be related to the role of butyrate in histone acetylation. It has been shown that this role is more complex than previously accepted. Donohoe *et al.,* 2012 ([4](#_ENREF_4)) found that butyrate is metabolised to acetyl-CoA, and depending on their concentration it determines if butyrate inhibits KDACs or stimulates KATs, leading to the regulation of the expression of different target genes. Moreover, it has also been proposed by Corfe B, 2012 ([40](#_ENREF_40)) that the inhibitory mechanism of butyrate is related to the inhibition of deacetylation by the production of butyryl-CoA, which is a substrate for protein acylation and is the product of deacetylation. Abundance of this and other acyl forms will change the acylation and deacetylation equilibrium to the acylated state. As can be seen in Figure 20 there was an increase in the production of all of these acyl forms when SCFA Uptake and Oxidation v4 model was used to run the time course simulations, suggesting that cells might be driven to the acylated state by the action of butyrate and other SCFAs. Regarding CoA levels, the reduction observed in the simulations by the action of butyrate could be related to the formation of butyryl-CoA. The increase in acetyl-CoA levels by the action of butyrate is also expected since two molecules of acetyl-CoA are needed to form butyrate yielding acetoacetyl-CoA, which is then converted, via the intermediates L(+)-β-hydroxybutyryl CoA and crotonyl-CoA, to butyryl-CoA. Then, butyrate metabolism (breakdown) produces acetyl-CoA. This pathway links butyrate directly to acetyl-CoA production, mainly through CoA-ligation and the formation of butyryl-CoA. It is also interesting to note that when the effect of increasing concentrations of propionate were simulated using SCFA Uptake and oxidation v4 model, a reduction in the levels of acetyl-CoA (60-35.47mM, Figure 23 right panel) was observed, whereas when time course simulations were run with this same model, an increase in propionyl-CoA (1-21.95mM, Figure 20) was observed. This reduction in acetyl-CoA levels and an increase in propionyl-CoA levels could be due to the production of acetate and propionyl-CoA from acetyl-CoA and propionate via the enzymatic activity of propionate-CoA transferase ([87](#_ENREF_87)).

The effect of glucose on acetyl-CoA was also simulated to evaluate the participation of glucose on the production of acetyl-CoA. This is relevant, since glucose and CoA-ligation is needed to produce acetyl-CoA, so glucose could be competing with SCFAs for CoA as substrate, since SCFAs also need CoA-ligation to form their acyl forms. The increase in acetyl-CoA by the increasing concentration of glucose, even it was slight, could indicate a competition between glucose and SCFAs for CoA as a substrate to whether form acetyl-CoA, in the case of glucose, or either form their respective acyl form, in the case of SCFAs. The ligation of CoA with increasing concentrations of glucose was also simulated. The slight decrease in CoA levels observed in the simulations suggests glucose and CoA-ligation, showing again a possible competition between glucose and SCFAs for CoA as a substrate. The slight increase in acetyl-CoA and the scarce decrease in CoA levels observed in the simulation support, at least in part, the idea of SCFAs and glucose competition for CoA as a substrate to either produce acyl forms or acetyl-CoA.

The simulations run at different time points at physiological ranges of butyrate, propionate and acetate (1-20, 1-20, and 1-52mM respectively) with SCFA Uptake and Oxidation v4 model were more heterogeneous, making it the best model to simulate the effect of a time course on the levels of butyryl-CoA, propionyl-CoA, acetyl-CoA and CoA. According to the simulations run, butyrate\_v3-1 model is the best model to simulate the action of increasing concentrations of butyrate, propionate, acetate, and glucose on CoA and acetyl-CoA levels.

The validation of the models with HCT166 cells treated with increasing concentrations of acetate, butyrate and propionate showed that changes in CoA levels were not as steady as in the simulations and the general outcome of the validation of the models on HCT116 cells, increase in CoA levels, was only in accordance to the time course simulations run with butyrate\_v3-1 model (Figure 19), and the simulations run to predict the effect of acetate on CoA levels using this model (Figure 25, left panel). In the case of butyrate and propionate, the effect observed in the validation of the models on HCT116 cells when treated with these SCFAs is opposite to the effect observed and expected, a decrease in CoA levels, when simulations were run with both butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models. The effect of butyrate on CoA levels, according to the simulations run, showed a decrease in CoA levels using butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models (Figure 26). On the other hand, the effect of propionate on CoA levels, showed no effect in the simulations run with butyrate\_v3-1 model, and a decrease in CoA levels in the simulations run with SCFA uptake and oxidation v4 model (Figure 27). Another important fact to address is that the magnitude of these changes is higher in the simulations run with butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models, in the range of mM, whereas it is lower in the case of HCT116 cells treated with acetate, butyrate and propionate, since it is in the range of micro molar μM. According to the results of the action of SCFAs on the levels of CoA in HCT116 cells, it is not possible to conclude if these models are able to simulate the effect of increasing concentrations of SCFAs in the levels of CoA. Further *in vitro* and *in vivo* studies are needed to validate these models. Despite the fact that further experiments are needed to validate these models, the simulations run with butyrate\_v3-1 and SCFA Uptake and Oxidation v4 models are able to predict various features of the action of SCFAs (28, 12, 34, 86).

# CONCLUSIONS

* Functional proteomics can contribute to the study of the shift in acetylation/deacetylation patterns and the control of enzymes involved in metabolic enzymes, which are molecular mechanism of SCFAs that could be studied and benefited by functional proteomics. Based on these acetylation/deacetylation patterns, the biological function of metabolic enzymes could be elucidated by using the emerging research field related to functional proteomics. In metabolic pathways, the cellular mechanisms of acetylation and deacetylation could be described by functional proteomics as well, making it a valuable tool.
* As it has been shown by quantitatively examining the data through bioinformatics approaches, not all metabolic enzymes are suitable of being acetylated (Representation and acetylation of metabolic pathways), there are differences in the enrichment scores of the global and the acetylated proteome datasets (Functional annotations), the number of enzymes matching the metabolic pathways analysed is higher in the global proteome when compared to the acetylated proteome (Percentage coverage of pathways) and more enzymes in the Metabolism pathway were found to match the global proteome compared to the acetylated proteome dataset (Matching pathways). A shift in the enrichment scores in the functional annotations might be related to an up- or down-regulation associated with acetylation of a group of proteins in a specific functional annotation in any GO term, since these enrichments scores are a rank of biological significance of a group of proteins evidencing a shift in the expression of a specific pathway protein set. This analysis reveals the degree to which a protein set is overrepresented in an expression dataset, indicating the percentage of proteins contributing to the enrichment score in every functional annotation of each GO term.

The elucidation of the acetylation status of metabolic enzymes and to provide more information about the functional proteome can be done by interrogating SCFAs oxidation models of metabolic pathways and by systematically interogating a proteomic dataset of mitochondrially enriched fractions. The modelling of the action of SCFAs and glucose increasing concentrations on acetyl-CoA and CoA levels and the effect of physiological ranges of SCFAs on a time course *in vitro*  and *in vivo* is essential to elucidate their role in controlling the enzymes in metabolic pathways. There is evidence that: 1) the increase in acetyl-CoA levels by the action of butyrate, as simulated using butyrate\_v3-1 model could be related to its regulatory role on protein acetylation and its role on histone acetylation, 2) the production of butyryl-CoA as simulated using SCFA Uptake and Oxidation v4 model could be related to the inhibitory mechanism of butyrate, 3) the increase in the production of acyl forms could simulate cells been driven to the acylated state by the action of SCFAs, 4) a decrease in acetyl-CoA levels could be due to the production of acetate and propionyl-CoA. No validation of the models in regard of increasing concentrations of acetate, butyrate and propionate on acetyl-CoA levels or the effect of physiological ranges of SCFAs on a time course were done, increasing the need for further studies to validate the models used in order to elucidate the action of SCFAs and glucose increasing concentrations on acetyl-CoA and CoA levels and the effect of physiological ranges of SCFAs on a time course *in vitro* and *in vivo*.

* Acetate, butyrate, and propionate play a role in controling metabolic enzymes (which are amenable to proteomic analysis), at least with regard to its acetylation patterns.

# REFERENCES

**1. Medina V, Edmonds B, Young G, James R, Appleton S, Zalewski P. Induction of caspase-3 protease activity and apoptosis by butyrate and trichostatin A (inhibitors of histones deacetylase): dependence on protein synthesis and synergy with a mitochondrial/cytochrome c-dependet pathway. Cancer Res. 1997;57(17):3697-707.**

**2. Jahns F, Wilhelm A, Jablonowski N, Mothes H, Radeva M, Wölfert A, et al. Butyrate suppresses mRNA increase of osteopontin and cyclookygenase-2 in human colon tumor tissue. Carcinogenesis. 2011;32(6):913-20.**

**3. Sun S, Li W, Zhang H, Zha L, Xue Y, Wu X, et al. Requirement for store-operated calcium entry in sodium butyrate-induced apoptosis in human colon cancer cells. Biosci Rep. 2012;32(1):83-90.**

**4. Donohoe D, Collins L, Wali A, Bigler R, Sun W, Bultman S. The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. J Lipid Res. 2012;53(9):1864-76.**

**5. Burgess D. Metabolism: Warburg behind the butyrate paradox? Nat Rev Cancer. 2012;12(12):798.**

**6. Thangaraju M, Gopal E, Martin P, Ananth S, Smith S, Prasad P, et al. SLC5A8 Triggers Tumor Cell Apoptosis through Pyruvate-Dependent Inhibition of Histone Deacetylases. Cancer Res. 2006;66(24):11560-4.**

**7. Matthews G, Howarth G, Butler R. Short-chain fatty acids induce apoptosis in colon cancer cells associated with changes to intracellular redox state and glucose metabolism. Chemotherapy. 2012;58(2):102-9.**

**8. Kim G, Gocevski G, Wu C, Yang X. Dietary, metabolic, and potentially environmental modulation of the lysine acetylation machinery. Int J Cell Biol. 2010;2010:632739 doi: 10.1155/2010/632739.**

**9. Grillon J, Johnson K, Kotlo K, Danziger R. Non-histone lysine acetylated proteins in heart failure. Biochimica et Biophysica Acta. 2012;1822(4):607–14.**

**10. Corfe B, Williams E, Bury J, Riley S, Croucher L, Lai D, et al. A study protocol to investigate the relationship between dietary fibre intake and fermentation, colon cell turnover, global protein acetylation and early carcinogenesis: the FACT study. BMC Cancer. 2009;9(332):doi: 10.1186/471-2407-9-332.**

**11. Zhao S, Xu W, Jiang W, Yu W, Lin Y, Zhang T, et al. Regulation of cellular metabolism by protein lysine acetylation. Science. 2010;327(5968):1000-4.**

**12. Lundby A, Lage K, Weinert B, Bekker-Jensen D, Secher A, Skovgaard T, et al. Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep. 2012;2(2):419-31.**

**13. Salyers A. Energy sources of major fermentative anaerobes. Am J Clin Nutr. 1979;32(1):158-63.**

**14. Salyers A, Vercellotti J, West S, Wilkins T. Fermentation of mucin and plant polysaccharides bv strains of Bacteroides from the human colon. Appl Environ Microbiol 1977;33(2):319-22.**

**15. McNeil N, Cummings J, James W. Short chain fatty acid absorption by the human large intestine. Gut. 1978;19(9):819-22.**

**16. Cummings J, Hill M, Jenkins D, Pearson J, Wiggins H. Changes in fecal composition and colonic function due to cereal fiber. Am J Clin Nutr. 1976;29(12):1468-73.**

**17. Thornton J. High colonic pH promotes colorectal cancer. Lancet. 1981;1(8229):1081-3.**

**18. Isfort R, Cody D, Asqtlilh T, Ridder G, Stuard S, LeBoeuf R. Induction of protein phosphorylation, protein synthesis, immediate early gene expression and cellular proliferation by intercellular pH modification. Implications of hydrogen ions in signal transduction. Eur J Biochem. 1993;213(1):349-57.**

**19. Hinnebusch B, Meng S, Wu J, Archer S, Hodin R. The effects of short-chain fatty acids on human colon cancer cell phenotype are associated with histone hyperacetylation. J Nutr. 2002;132:1012-17.**

**20. Waldecker M, Kautenburger T, Daumann H, Busch C, Schrenk D. Inhibition of histone-deacetylase activity by short-chain fatty acids and some polyphenol metabolites formed in the colon. Journal of Nutritional Biochemistry. 2008;19:587-93.**

**21. Vinoloa M, Rodrigues A, E H, Satoa F, Sampaio S, Curia R. Suppressive effect of short-chain fatty acids on production of proinflammatory mediators by neutrophils. Journal of Nutritional Biochemistry. 2011;22:849-55.**

**22. Davie J. Inhibition of histone deacetylase activity by butyrate. J Nutr. 2003;133(7 Suppl):2485S-93S.**

**23. Toscani A, Soprano D, Soprano K. Molecular analysis of sodium butyrate-induced growth arrest. Oncogene Res. 1988;3(3):223-38.**

**24. Yu D, Waby J, Chirakkal H, Staton C, Corfe B. Butyrate suppresses expression of neuropilin I in colorectal cell lines through inhibition of Sp1 transactivation. Mol Cancer. 2010;9(276):doi: 10.1186/476-4598-9-276.**

**25. Chirakkal H, Leech S, Brookes K, Prais A, Waby J, Corfe B. Upregulation of Bak by butyrate in the colon is associated with increased Sp3 binding. Oncogene. 2006;25(54):7192-200.**

**26. Hague A, Manning A, Huschlscha L, Hart D, Paraskeva C. Sodium butyrate induces apoptosis in human colonic tumour cell lines in a p53-independent pathway: implications for the possible role of dietary fibre in the prevention of large bowel cancer. Int J Cancer. 1993;55(3):498-505.**

**27. Velázquez O, Lederer H, Rombeau J. Butyrate and the colonocyte - Implications for neoplasia. Dig Dis Sci. 1996;41(4):727-39.**

**28. Heerdt B, Houston M, Augenlicht L. Potentiation by specific short chain fatty acids of differentiation and apoptosis in human colonic carcinoma cell lines. Cancer Res. 1994;54(12):3288–94.**

**29. Le Blay G, Blottiere H, Ferrier L, Le Foll E, Bonnet C, Galmiche J, et al. Short-chain fatty acids induce cytoskeletal and extracellular protein modifications associated with modulation of proliferation in primary culture of rat intestinal smooth muscle cells. Dig Dis Sci. 2000;45(8):1623-30.**

**30. Csordas A. Butyrate, aspirin and colorectal cancer. Eur J Cancer Prev. 1996;5(4):221-31.**

**31. Freeman H. Effects of differing concentrations of sodium butyrate on 1, 2-dimethylhydrazine-induced rat intestinal neoplasia. Gastroenterology. 1986;91(3):596-602.**

**32. Zoran D, Turner N, Taddeo S, Chapkin R, Lupton J. Wheat bran diet reduces tumor incidence in a rat model of colon cancer independent of effects on distal luminal butyrate concentrations. J Nutr. 1997;127(11):2217-25.**

**33. Deschner E, Ruperto J, Lupton J. Dietary butyrate (tributyrin) does not enhance AOM-induced colon tumorigenesis. Cancer Lett. 1990;52(1):79-82.**

**34. Caderni G, Luceri C, De Filippo C, Salvadori M, Giannini A, Tessitore L, et al. Slow-release pellets of sodium butyrate do not modify azoxymethane (AOM)-induced intestinal carcinogenesis in F344 rats. Carcinogenesis. 2001;22(3):525-7.**

**35. Riester D, Wegener D, Hildmann C, Schwienhorst A. Members of the histone deacetylase superfamily differ in substrate specificity towards small synthetic substrates. Biochemical and Biophysical Research Communications. 2004;324:1116-23.**

**36. Soliman M, Rosenberger T. Acetate supplementation increases brain histone acetylation and inhibits histone deacetylase activity and expression. Mol Cell Biochem. 2011;352:173-80.**

**37. Roediger W. Utilization of nutrients by osilated epithelial cells of the rat colon. Gastroenterology. 1982;83(2):424-9.**

**38. Gonçalves P, Martel F. Butyrate and colorectal cancer: the role of butyrate transport. Curr Drug Metab. 2013;14(9):994-1008.**

**39. Lupton J. Microbial degradation products influence colon cancer risk: the butyrate controversy. Journal of Nutrition. 2004;134(2):479-82.**

**40. Corfe B. Hypothesis: butyrate is not an HDAC inhibitor, but a product inhibitor of deacetylation. Mol BioSyst. 2012;8(6):1609-12.**

**41. Candido E, Reeves R, Davies J. Sodium butyrate inhibits histone deacetylation in culture cells. Cell. 1978;14(1):105-13.**

**42. Matthews G, Howarth G, Butler R. Short-chain fatty acids induce apoptosis in colon cancer cells associated with changes to intracellular redox state and glucose metabolism. Chemotherapy. 2012;58(2):102-9.**

**43. Kilner J, Waby J, Chowdry J, Khan A, Noirel J, Wright P, et al. A proteomic analysis of differential cellular responses to the short-chain fatty acids butyrate, valerate and propionate in colon epithelial cancer cells. Mol BioSyst. 2012;8:1146-56.**

**44. Kilner J, Corfe B, Wilkinson S. Modelling modifications affecting the tubulin code as a route to study the chemopreventive and chemotherapeutic actions of SCFAs as HDAC**

**inhibitors. Mol BioSyst. 2011;7:975-83.**

**45. Knight G, Hager L, Palmiter R. Butyrate and Related Inhibitors of Histone Deacetylation Block the Induction of Egg White Genes by Steroid Hormones. Cell. 1980;22:469-77.**

**46. Liu Q, Wang C, Huang Y, Dong K, Yang W, Zhang S, et al. Effects of isovalerate on ruminal fermentation, urinary excretion of purine derivatives and digestibility in steers. Journal of Animal Physiology and Animal Nutrition. 2009;93:716-25.**

**47. Zambell K, Fitch M, Fleming S. Acetate and butyrate are the major substrates for de novo lipogenesis in rat colonic epithelial cells. J Nutr. 2003;133(11):3509-15.**

**48. Ardawi M, Newsholme E. Fuel utilization in colonocytes of the rat. Biochem J. 1985;231(3):713-9.**

**49. Anderson J, Bridges S. Short-chain fatty acid fermentation products of plant fiber affect glucose metabolism of isolated rat hepatocytes. Proc Soc Exp Biol Med. 1984;177 (2):372-6.**

**50. Dihazi H, Kessler R, Müller G, Eschrich K. Lysine 3 acetylation regulates the phosphorylation of yeast 6-phosphofructo-2-kinase under hypo-osmotic stress. Biol Chem. 2005;386(9):895-900.**

**51. Ventura M, Mateo F, Serratosa J, Salaet I, Carujo S, Bachs O, et al. Nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase is regulated by acetylation The International Journal of Biochemistry & Cell Biology. 2010;42:1672–80.**

**52. Sen N, Hara M, Kornberg M, Cascio M, Bae B, Shahani N, et al. Nitric oxide-induced nuclear GAPDH activates p300/CBP and mediates apoptosis. Nat Cell Biol. 2008;10(7):866-73.**

**53. Li T, Liu M, Feng X, Wang Z, Das I, Xu Y, et al. Glyceraldehyde-3-phosphate Dehydrogenase Is Activated by Lysine 254 Acetylation in Response to Glucose Signal. J Biol Chem. 2014;289(6):3775-85.**

**54. Hallows W, Yu W, Denu J. Regulation of glycolytic enzyme phosphoglycerate mutase-1 by Sirt1 protein-mediated deacetylation. J Biol Chem. 2012;287(6):3850-8.**

**55. Finkemeier I, Laxa M, Miguet L, Howden A, Sweetlove L. Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. Plant Physiol. 2011;155(4):1779-90.**

**56. Yu W, Dittenhafer-Reed K, Denu J. SIRT3 protein deacetylates isocitrate dehydrogenase 2 (IDH2) and regulates mitochondrial redox status. J Biol Chem. 2012;287(17):14078-86.**

**57. Schlicker C, Gertz M, Papatheodorou P, Kachholz B, Becker C, Steegborn C. Substrates and Regulation Mechanisms for the Human Mitochondrial Sirtuins Sirt3 and Sirt5. J Mol Biol. 2008;382:790-801.**

**58. Finley L, Haas W, Desquiret-Dumas V, Wallace D, Procaccio V, Gygi S, et al. Succinate dehydrogenase is a direct target of sirtuin 3 deacetylase activity. PLoS One. 2011;6(8):doi: 10.1371/journal.pone.0023295.**

**59. Cimen H, Han M, Yang Y, Tong Q, Koc H, Koc E. Regulation of succinate dehydrogenase activity by SIRT3 in mammalian mitochondria. Biochemistry. 2010;49(2):304-11.**

**60. Kim E, Han B, Kim W, Lee S, Bae K. Acceleration of adipogenic differentiation via acetylation of malate dehydrogenase 2. Biochem Biophys Res Commun. 2013;441(1):77-82.**

**61. Kim E, Kim W, Kang H, Kim J, Chung S, Seo Y, et al. Acetylation of malate dehydrogenase 1 promotes adipogenic differentiation via activating its enzymatic activity. J Lipid Res. 2012;53(9):1864-76.**

**62. Zhao X, León I, Bak S, Mogensen M, Wrzesinski K, Højlund K, et al. Phosphoproteome analysis of functional mitochondria isolated from resting human muscle reveals extensivephosphorylation of inner membrane protein complexes and enzymes. Mol Cell Proteomics. 2011;10(1):doi: 10.1074/mcp.M110.000299.**

**63. Frahm J, Li L, Grevengoed T, Coleman R. Phosphorylation and Acetylation of Acyl-CoA Synthetase- I. J Proteomics Bioinform. 2011;4(7):129-37.**

**64. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science. 2009;325(5942):834-40.**

**65. Crosby H, Escalante-Semerena J. The acetylation motif in AMP-forming acyl-CoA synthetases contains residues critical for acetylation and recognition by the protein acetyltransferase Pat of Rhodopseudomonas palustris. J Bacteriol. 2014:1496-504.**

**66. Lv L, Li D, Zhao D, Lin R, Chu Y, Zhang H, et al. Acetylation targets the M2 isoform of pyruvate kinase for degradation through chaperone-mediated autophagy and promotes tumor growth. Mol Cell. 2011;42(6):719-30.**

**67. Lin Y, Lu J, Zhang J, Walter W, Dang W, Wan J, et al. Protein acetylation microarray reveals that NuA4 controls key metabolic target regulating gluconeogenesis. Cell. 2009;136(6):1073-84.**

**68. Jiang W, Wang S, Xiao M, Lin Y, Zhou L, Lei Q, et al. Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase. Mol Cell. 2011;43(1):33-44.**

**69. Zhao D, Zou S, Liu Y, Zhou X, Mo Y, Wang P, et al. Lysine-5 acetylation negatively regulates lactate dehydrogenase A and is decreased in pancreatic cancer. Cancer Cell. 2013;23(4):464-76.**

**70. Wirth M, Karaca S, Wenzel D, Ho L, Tishkoff D, Lombard D, et al. Mitochondrial SIRT4-type proteins in Caenorhabditis elegans and mammals interact with pyruvate carboxylase and other acetylated biotin-dependent carboxylases. Mitochondrion. 2013;13(6):705-20.**

**71. Monti M, Orrù S, Pagnozzi D, Pucci P. Functional proteomics. Clin Chim Acta. 2005;357(2):140-50.**

**72. Neubauer G, Gottschalk A, Fabrizio P, Seraphin B, Lurhmann R, Mann M. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc Natl Acad Sci U S A. 1997;94(2):385-90.**

**73. Shevchenko A, Keller P, Scheiffele P, Mann M, Simons K. Identification of components of trans-Golgi network-derived transport vesicles and detergent-insoluble complexes by nanoelectrospray tandem mass spectrometry. Electrophoresis. 1997;18(14):2591-600.**

**74. Pandey A, Fernandez M, Steen H, Blagoev B, Nielsen M, Roche S, et al. Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways. J Biol Chem. 2000;275(49):38633-9.**

**75. Hinsby A, Olsen J, Bennett K, Mann M. Signaling initiated by overexpression of the fibroblast growth factor receptor-1 investigated by mass spectrometry. Mol Cell Proteomics. 2003;2(1):29-36.**

**76. Godovac-Zimmermann J, Brown L. Perspectives for mass spectrometry and functional proteomics. Mass Spectrom Rev. 2001;20(1):1-57.**

**77. Chen W, Kang S, Lubec G. Protein profiling by the combination of two independent mass spectrometry techniques. Nat Protoc. 2006;1(3):1446-52.**

**78. Smolka M, Zhou H, Aebersold R. Quantitative Protein Profiling Using Two-dimensional Gel Electrophoresis, Isotope-coded Affinity Tag Labeling, and Mass Spectrometry Molecular & Cellular Proteomics. 2002;1(1):19-29.**

**79. Lengqvist J, Sandberg A. Stable isotope labeling methods in protein profiling. Methods Mol Biol. 2013;1023:21-51.**

**80. Kim S, Sprung R, Chen Y, Xu Y, Ball H, Pei J, et al. Substrate and Functional Diversity Resource of Lysine Acetylation Revealed by a Proteomics Survey Molecular Cell. 2006;23(607-618).**

**81. Schwer B, Eckersdorff M, Li Y, Silva J, Fermin D, Kurtev M, et al. Calorie restriction alters mitochondrial protein acetylation. Aging Cell. 2009;8(5):604-6.**

**82. Shevade S, Jindal N, Dutta S, Jarori G. Food vacuole associated enolase in *Plasmodium* undergoes multiple post-translational modifications: evidence for atypical ubiquitination. PLoS One. 2013;8(8):doi: 10.1371/journal.pone.0072687.**

**83. Wang Y, Tsai C, Hong T, Tsou C, Lin P, Pan S, et al. An informatics-assisted label-free quantitation strategy that depicts phosphoproteomic profiles in lung cancer cell invasion. J Proteome Res. 2010;9(11):doi: 10.1021/pr100394u.**

**84. Zhou W, Capello M, Fredolini C, Piemonti L, Liotta L, Novelli F, et al. Mass Spectrometry Analysis of the Post-Translational Modiﬁcations of alpha-Enolase from Pancreatic Ductal Adenocarcinoma Cells  J Proteome Res. 2010;9:2929-36**

**85. Patel V, Thalassinos K, Slade S, Connolly J, Crombie A, Murrell J, et al. A Comparison of Labeling and Label-Free Mass Spectrometry-Based Proteomics Approaches Journal of Proteome Research. 2009;8:3752–9.**

**86. Dennis G, Sherman B, Hosack D, Yang J, Gao W, Lane C, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;4(5):P3.**

**87. Stadtman E. Functional group of coenzyme A and its metabolic relations, especially in the fatty acid cycle: Discussion. Fed Proc. 1953;12(3):692-3.**

**88. Wellen K, Thompson C. A two-way street: reciprocal regulation of metabolism and signaling. Nat Rev Mol Cell Biol. 2012;13(4):270-6.**

**89. Xiong Y, Lei Q, Zhao S, Guan K. Regulation of glycolysis and gluconeogenesis by acetylation of PKM and PEPCK. Cold Spring Harb Symp Quant Biol. 2011;76(285-289).**

**90. Lea M, Randolph V. Induction of reporter gene expression by inhibitors of histone deacytilase. Anticancer Res. 1998;18(4A):2717-22.**

**91. Kung C, Hixon J, Choe S, Marks K, Gross S, Murphy E, et al. Small Molecule Activation of PKM2 in Cancer Cells Induces Serine Auxotrophy. Chem Biol. 2012;19(9):1187-98.**

**92. Jahns F, Wilhelm A, Greulich K, Mothes H, Radeva M, Wölfert A, et al. Impact of butyrate on PKM2 and HSP90β expression in human colon tissues of different transformation stages: a comparison fo gene and protein data. Genes Nutr. 2012;7(2):235-46.**

**93. Steták A, Veress R, Ovádi J, Csermely P, Kéri G, Ullrich A. Nuclear translocation of the tumor marker pyruvate kinase M2 induces programmed cell death Cancer Res. 2007;67(4):1602-8.**

**94. Luo W, Semenza G. Emerging roles of PKM2 in cell metabolism and cancer progression. Trends Endocrinol Metab. 2012;23(11):560-6.**

**95. Amoêdo N, Rodrigues M, Pezzuto P, Galina A, da Costa R, de Almeida F, et al. Energy metabolism in H460 lung cancer cells: effects of histone deacetylase inhibitors. PLoS One. 2011;6((7):e22264):doi: 10.1371/journal.pone.0022264.**

**96. Tiedge M, Lenzen S. Effects of sodium butyrate on glucose transporter and glucose-phosphorylating enzyme gene expression in RINm5F insulinoma cells. J Mol Endocrinol. 1996;17(1):19-26.**

**97. Fernández-Mejía C, Davidson M. Effect of sodium butyrate on glucose transport and glucose-phosphorylating enzymes in RIN-m5F cells. Pancres. 1993;8(5):589-96.**

**98. Wang Q, Zhang Y, Yang C, Xiong H, Lin Y, Yao J, et al. Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. Science. 2010;327(5968):1004-7.**

**99. Van Dyk D, Misztal D, Wilkins M, Mackintosh J, Poljak A, Varnai J, et al. Identification of cellular changes associated with increased production of human growth hormone in a recombinant Chinese hamster ovary cell line. Proteomics. 2003;3(2):147-56.**

**100. Ngo J, Chute H, Sanders D, Orlando R, Ibsen K. Regulation of pyruvate kinase expression and growth in mastocytoma cells. I. Initial observations. Exp Cell Res. 1983;149(2):565-75.**

**101. Shimada N, Yamada K, Tanaka T, Kawata H, Mizutani T, Miyamoto K, et al. Alterations of gene expression in endoderm differentiation of F9 teratocarcinoma cells. Mol Reprod Dev. 2001;60(2):165-71.**

**102. Valvassori S, Calixto K, Budni J, Resende W, Varela R, de Freitas K, et al. Sodium butyrate reverses the inhibition of Krebs cycle enzymes induced by amphetamine in the rat brain. J Neural Transm. 2013;120(12):1737-42.**

**103. Shahriari A, Dawson N, Bell R, Storey K. Stable Suppression of Lactate Dehydrogenase Activity during Anoxia in the Foot Muscle of Littorina littorea and the Potential Role of Acetylation as a Novel Posttranslational Regulatory Mechanism. Enzyme Res. 2013;2013(461374):doi: 10.1155/2013/461374.**

# APPENDIX I. Abbreviations

Adenosine triphosphate (ATP)

Azoxymethane (AOM)

Bcl antagonist killer (Bak)

Bernard Corfe (BC)

Biological Process (BP)

Calorie restriction (CR)

Caroline Evans (CE)

**Cation exchange (CEX)**

Cellular Compartment (CC)

Complex pathway simulator (COPASI)

Database for Annotation, Visualization and Integrated Discovery (DAVID)

Deoxyribonucleic acid (DNA)

Dulbecco’s Modified Eagle Medium (DMEM)

Elevated high definition mass spectrometry (HDMSE)

Enzyme-linked immunosorbent assay (ELISA)

Foetal bovine serum (FBS)

Gene ontology (GO)

**High Definition MS (HDMS)**

Histone deacetylases (HDACs)

Histone-3 (H3)

Histone-4 (H4)

Human embryonic kidney (HEK)

Isobaric tagging for relative and absolute quantification (iTRAQ)

Isotope-coded affinity-tag-based protein profiling (ICAT)

Joanne Connolly (JC)

KDAC6-like enzyme from *Bordetella/Alcaligenes* strain (FB188FB188 HDAH)

Kyoto Encyclopaedia of Genes and Genome (KEGG)

Label Free Quantitation (LFQ)

Liquid chromatography-mass spectrometry (LC-MS)

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

LNCaP (Lymph node prostate cancer cell line)

Lysine acetyltransferases (KATs)

Lysine deacetylase inhibitors (KDACi)

Lysine deacetylase-8 (HDAC8)

Lysine deacetylases (KDACs)

Lysine deactylase-6 (HDAC6)

Mass-spectrometry (MS)

Matrix-Assisted Laser Desorption/Ionization-Time of Flight MS (MALDI-TOF)

Messenger ribonucleic acid (mRNA)

Michigan Cancer Foundation-7 (MCF-7)

Micro molar (μM)

Millimolar (mM)

Molecular Function (MF)

Mouse embryonic fibroblasts (MEF)

Multi-dimensional Protein Identification Technology (MudPIT)}

NAD dependant deacetylase sirtuin-1 (Sirt1)

NAD dependant deacetylase sirtuin-2 (Sirt2)

NAD dependant deacetylase sirtuin-3 (Sirt3)

NAD dependant deacetylase sirtuin-4 (Sirt4)

NAD dependant deacetylase sirtuin-4 (Sirt5)

Nicotinamide adenine dinucleotide (NADH)

Nicotinamide adenine dinucleotide phosphate (NADPH)

Post-translational modifications (PTMs)

ProteomeLab PF2D (2D LC)

Reactive oxygen species (ROS)

**Reverse Phase (Rp)**

**Reversed Phase High Performance Liquid Chromatography (RP HPLC)**

Short chain fatty acid (SCFA)

Short chain fatty acids (SCFAs)

Specific protein-1 (Sp1)

Specific protein-3 (Sp3)

Stable Isotopic Labelling by Amino Acids in Cell Culture (SILAC)

Tandem mass tags (TMT)

Tandem MS (MS/MS)

Tricarboxylic acid (TCA)

Universal Protein Resource (Uniprot)

# APPENDIX II. Implications of acetylated lysine residues in metabolic enzymes/Variations in response to butyrate in metabolic enzymes

|  |  |  |  |
| --- | --- | --- | --- |
| Glycolytic enzyme | Acetylated Lysine residues | Variation in response to butyrate | Impact of Lys acetylation on the enzyme |
| Hexokinase | 21, 24, 62, 77, 176, 187, 191, 333, 344, 346, 353, 418, 429, 481, 544 and 763. | Increased mitochondrial bound activity in H460 lung cancer cells ([95](#_ENREF_95)) and increased expression and activity in RINm5F insulinoma cells ([96](#_ENREF_96), [97](#_ENREF_97)). | No effect reported |
| Glucose 6-phosphate isomerase | 12 ([64](#_ENREF_64)), 142 ([64](#_ENREF_64)), 211, 252, 362, and 466. | No effect reported | No effect reported |
| Phosphoglycerate kinase | 6, 11 ([64](#_ENREF_64)), 30 ([64](#_ENREF_64)), 48 ([64](#_ENREF_64)), 56, 75 ([64](#_ENREF_64)), 86 ([64](#_ENREF_64)), 91 ([64](#_ENREF_64)), 97 ([64](#_ENREF_64)), 131 ([64](#_ENREF_64)), 133, 141, 146 ([64](#_ENREF_64)), 149 ([55](#_ENREF_55)), 169 ([55](#_ENREF_55)), 191, 199 ([64](#_ENREF_64)), 216, 220 ([11](#_ENREF_11)), 267 ([64](#_ENREF_64)), 272, 275, 291 ([64](#_ENREF_64)), 322, 323 ([64](#_ENREF_64)), 361, and 406 ([64](#_ENREF_64)). | No effect reported | Inhibition in *Arabidopsis thaliana* ([55](#_ENREF_55)) |
| Fructose- bisphosphate aldolase | 7 , 13, 14, 42, 108 ([64](#_ENREF_64)), 139, 147 ([12](#_ENREF_12), [64](#_ENREF_64)), 153, 200, 208, 230, 294, 312, 330 ([64](#_ENREF_64)), 342, 395 . | No effect reported | Inhibition *in vivo* in rats ([12](#_ENREF_12)).  Impact on protein-protein interactions in *Arabidopsis thaliana* ([55](#_ENREF_55)) |
| Triophosphate isomerase | 43, 51 ([64](#_ENREF_64)), 96, 106, 168, 193, 225 ([64](#_ENREF_64), [80](#_ENREF_80)), 275 ([11](#_ENREF_11), [64](#_ENREF_64)), and 285. | No effect reported | No effect reported |
| Glyceraldehyde 3-phosphate dehydrogenase | 55, 61 ([64](#_ENREF_64)), 107, 117 ([51](#_ENREF_51)), 120 , 139 ([64](#_ENREF_64)), 160 ([52](#_ENREF_52)), 162, 186([64](#_ENREF_64)), 194 ([64](#_ENREF_64)), 215, 219 ([64](#_ENREF_64)), 227 ([51](#_ENREF_51), [64](#_ENREF_64)), 251 ([51](#_ENREF_51)), 254 ([11](#_ENREF_11), [53](#_ENREF_53), [64](#_ENREF_64)), 259, 260, 314 and 334. | No effect reported | Impact on protein-protein interactions in *Arabidopsis thaliana* ([55](#_ENREF_55))  Increase activity in HEK 293T, A549, and stable pool cells at Lys 254 residue ([53](#_ENREF_53)).  Regulation of nuclear translocation in NIH3T3 and HCT116 cells lines at Lys residues 117, 227 and 251 ([51](#_ENREF_51)).  Increase of glycolytic activity and inhibition of gluconeogenetic activity in *Salmonella* ([98](#_ENREF_98)). |
| Phosphofructokinase | 3 , 395 ([64](#_ENREF_64)), 480, 486 ([64](#_ENREF_64)), 574, 688 ([64](#_ENREF_64)), 700 ([11](#_ENREF_11)), 736, and 747 ([11](#_ENREF_11)). | No effect reported | Regulation of phosphorylation in yeast ([50](#_ENREF_50)). |
| Phosphoglycerate mutase | 5 ([64](#_ENREF_64)), 39, 61, 100 ([64](#_ENREF_64)), 106 , 113 ([64](#_ENREF_64)), 125, 157, 251 ([54](#_ENREF_54)), 253 ([54](#_ENREF_54))and 254 ([54](#_ENREF_54)). | No effect reported | Down-regulation in HEK293 cells ([54](#_ENREF_54)). |
| Enolase | 5 ([64](#_ENREF_64)), 28 ([84](#_ENREF_84)), 60 ([64](#_ENREF_64)), 64 ([64](#_ENREF_64), [84](#_ENREF_84)), 71 ([64](#_ENREF_64), [84](#_ENREF_84)), 80 ([64](#_ENREF_64), [84](#_ENREF_84)), 81 ([84](#_ENREF_84)), 89 ([64](#_ENREF_64), [84](#_ENREF_84)), 92 ([11](#_ENREF_11), [84](#_ENREF_84)), 103 ([84](#_ENREF_84)), 105 ([84](#_ENREF_84)), 126 ([64](#_ENREF_64), [84](#_ENREF_84)), 142 ([82](#_ENREF_82)), 193 ([64](#_ENREF_64), [84](#_ENREF_84)), 197, 199 ([64](#_ENREF_64)), 202 ([84](#_ENREF_84)), 221 ([64](#_ENREF_64), [84](#_ENREF_84)), 228 ([64](#_ENREF_64), [84](#_ENREF_84)), 233 ([84](#_ENREF_84)), 239 ([84](#_ENREF_84)), 256 ([64](#_ENREF_64), [84](#_ENREF_84)), 262 ([84](#_ENREF_84)), 281 ([64](#_ENREF_64), [84](#_ENREF_84)), 285 ([64](#_ENREF_64), [84](#_ENREF_84)), 306, 326, 330 ([84](#_ENREF_84)), 335 ([84](#_ENREF_84)), 343 ([84](#_ENREF_84)), 358 ([84](#_ENREF_84)), 384 , 406 ([64](#_ENREF_64), [84](#_ENREF_84)), 420 ([64](#_ENREF_64), [84](#_ENREF_84)) and 434. | Induction of expression in CHO cells ([99](#_ENREF_99)). | No effect reported |
| Pyruvate kinase | 3, 62 ([64](#_ENREF_64)), 89 ([64](#_ENREF_64)), 135 ([64](#_ENREF_64)), 151, 162, 166 ([64](#_ENREF_64)), 186, 206, 207, 230, 261, 266 ([64](#_ENREF_64)) 305 ([11](#_ENREF_11), [64](#_ENREF_64), [66](#_ENREF_66)), 422, 433 ([64](#_ENREF_64)) and 498. | Enhanced expression in P-815 mastocytoma cells ([100](#_ENREF_100)).  Decrease in the amount of transcripts in human colon tissues ([92](#_ENREF_92)).  Decrease in gene expression in F9 terotocarcinoma cells ([101](#_ENREF_101)). | Regulation of activity and stability at Lys residue 305 in 293T cells ([66](#_ENREF_66)). |

Table 13. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on glycolytic enzymes. *Some of* t*he acetylated lysine residues were found in other species, but match the reported in humans.*

|  |  |  |  |
| --- | --- | --- | --- |
| TCA cycle enzyme | Acetylated Lys residues | Variation in response to butyrate | Impact of Lys acetylation on enzyme function |
| Citrate synthase | None reported | Reversion of D-amphetamine (mania inducer in rats) inhibitory effect on this enzyme ([102](#_ENREF_102)). | No effect reported |
| Aconitase | 50 ([64](#_ENREF_64)), 233, 520, 521, 549 ([11](#_ENREF_11)), 573 ([64](#_ENREF_64)), 577, 605 ([64](#_ENREF_64)), 689, 700, and 743. | No effect reported | No effect reported |
| Isocitrate dehydrogenase | 81, 93 ([64](#_ENREF_64)), 126, 224 ([11](#_ENREF_11), [64](#_ENREF_64)), 211 ([57](#_ENREF_57)), 212 ([57](#_ENREF_57)), 233, 321 ([64](#_ENREF_64)), and 413 ([56](#_ENREF_56)). | No effect reported | Deacetylation at Lys 413 stimulates its activity in HEK293 and MEFs cells ([56](#_ENREF_56)).  Deacetylation enhances activity in Rosetta2 cells ([57](#_ENREF_57)). |
| α-Ketoglutarate dehydrogenase | None reported | No effect reported | No effect reported |
| Succninyl-CoA synthetase | None reported | No effect reported | No effect reported |
| Succinate dehydrogenase | 179 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 182 ([58](#_ENREF_58)), 250 ([58](#_ENREF_58)), 335 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 423 ([11](#_ENREF_11), [64](#_ENREF_64), [80](#_ENREF_80)), 480 ([58](#_ENREF_58)), 485 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 498 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 517 ([11](#_ENREF_11), [64](#_ENREF_64), [80](#_ENREF_80)), 538 ([11](#_ENREF_11), [64](#_ENREF_64), [80](#_ENREF_80)), 541 ([11](#_ENREF_11), [64](#_ENREF_64), [80](#_ENREF_80)), 547 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 550 ([58](#_ENREF_58)), 598 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 608 ([11](#_ENREF_11), [58](#_ENREF_58), [64](#_ENREF_64), [80](#_ENREF_80)), 624 ([58](#_ENREF_58)) and 633 ([58](#_ENREF_58)). | Increased activity in H640 lung cancer cells ([95](#_ENREF_95)).  Reversion of D-amphetamine (mania inducer in rats) inhibitory effect on this enzyme ([102](#_ENREF_102)). | Regulation of its activity and the activity of complex II electron transport chain. Conformational change in the structure of the enzyme. Acetylated lysine residues 179, 485, 498 and 538 in mouse are human homologues ([59](#_ENREF_59)). |
| Fumarase | 61, 66 ([64](#_ENREF_64)), 80 ([64](#_ENREF_64)), 94 ([64](#_ENREF_64)), 115, 172, 223, 256 ([64](#_ENREF_64)), 263, 292 ([64](#_ENREF_64)), 467, 470, 473, and 477. | No effect reported | No effect reported |
| Malate dehydrogenase | 74, 78, 91, 118 ([61](#_ENREF_61)), 121 ([61](#_ENREF_61)) 156, 157, 165 ([64](#_ENREF_64)), 185 ([11](#_ENREF_11), [64](#_ENREF_64)), 203, 215, 249 , 296, 297, 298 ([61](#_ENREF_61)), 301 ([11](#_ENREF_11), [64](#_ENREF_64)), 307 ([11](#_ENREF_11)), 314 ([11](#_ENREF_11), [64](#_ENREF_64)), 324, 328, 329 ([64](#_ENREF_64)), and 335 ([64](#_ENREF_64)). | Reversion of D-amphetamine (mania inducer in rats) inhibitory effect on this enzyme ([102](#_ENREF_102)). | Dramatically increased in the late stage of adipogenesis ([60](#_ENREF_60), [61](#_ENREF_61)).  Enhances activity in nuclear, cytosolic and mitochondrial fractions of human liver tissue ([11](#_ENREF_11)).  Inhibition in *Arabidopsis thaliana* ([55](#_ENREF_55)). |

Table 14. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on the TCA cycle enzymes. *Some of* t*he acetylated lysine residues were found in other species, but match the reported in humans.*

|  |  |  |  |
| --- | --- | --- | --- |
| β-oxidation enzyme | Acetylated Lys residues | Variation in response to butyrate | Impact of Lys acetylation on enzyme function |
| Acyl-CoA synthetase | 49 ([63](#_ENREF_63)), 52 ([63](#_ENREF_63)), 217 ([63](#_ENREF_63)), 223 ([63](#_ENREF_63)), 241 ([63](#_ENREF_63)), 250 ([63](#_ENREF_63)), 341 ([63](#_ENREF_63)), 387 ([63](#_ENREF_63)), 418 ([64](#_ENREF_64)), 428 ([63](#_ENREF_63)), 504 ([63](#_ENREF_63)), 533 ([63](#_ENREF_63)), 534 , 544 ([63](#_ENREF_63)), 562 ([63](#_ENREF_63)), 641([63](#_ENREF_63)), and 649 ([63](#_ENREF_63)). | No effect reported | Inactivation in *Rhodopseudomonas palustris* ([65](#_ENREF_65)).  Inhibition by mutations at Lys 676 *in vitro* and *in* vivo. Decrease in activity by mutations at Lys 285 *in vitro* and *in vivo* ([63](#_ENREF_63)). |
| Enoyl-CoA hydratase | 40, 118 ([64](#_ENREF_64)), 165 ([11](#_ENREF_11)), 171 ([11](#_ENREF_11)), 346 ([11](#_ENREF_11)), and 584 ([11](#_ENREF_11)). | No effect reported | Enhances activity in nuclear, cytosolic and mitochondrial fractions of human liver tissue ([11](#_ENREF_11)). |
| β-hydroxyacyl CoA dehydrogenase | None reported | No effect reported | Enhances activity in nuclear, cytosolic and mitochondrial fractions of human liver tissue ([11](#_ENREF_11)). |
| Thiolase | 83 ([64](#_ENREF_64)), 124, 131, 174 ([64](#_ENREF_64)), 181 ([64](#_ENREF_64)), 190, 202, 203, 251 ([64](#_ENREF_64)), 257, 263 ([64](#_ENREF_64)), and 302. | No effect reported | No effect reported |

Table 15. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on β-oxidation enzymes. *Some of* t*he acetylated lysine residues were found in other species, but match the reported in humans.*

|  |  |  |  |
| --- | --- | --- | --- |
| Pryruvate metabolism enzyme | Acetylated Lys residues | Variation in response to butyrate | Impact of Lys acetylation on enzyme function |
| Lactate dehydrogenase | 5 ([64](#_ENREF_64), [69](#_ENREF_69)), 14 ([11](#_ENREF_11), [64](#_ENREF_64)), 22, 57 ([64](#_ENREF_64)), 81 ([64](#_ENREF_64)), 118 ([64](#_ENREF_64)), 126 ([64](#_ENREF_64)), 132, 222 ([64](#_ENREF_64)), 228, 278 ([64](#_ENREF_64)), 318 ([64](#_ENREF_64)), and 328. | No effect reported | Inhibition *in Littorina littorea*([103](#_ENREF_103))**.**  Down-regulation at Lys 5 in pancreatic cancer ([69](#_ENREF_69)). |
| Malic enzyme | 24, 94 ([64](#_ENREF_64)), 156 ([64](#_ENREF_64)), 224 ([64](#_ENREF_64)), 240 ([64](#_ENREF_64)), 272 ([64](#_ENREF_64)), 339, and 346 ([64](#_ENREF_64)). | No effect reported | No effect reported |
| Malate dehydrogenase | 74, 78, 91, 156, 157, 165 ([64](#_ENREF_64)), 185 ([64](#_ENREF_64)), 203, 215, 296, 297, 301 ([64](#_ENREF_64)), 307, 314 ([11](#_ENREF_11), [64](#_ENREF_64)), 324, 328, 329 ([64](#_ENREF_64)), and 335 ([64](#_ENREF_64)). | Reversion of D-amphetamine (mania inducer in rats) inhibitory effect on this enzyme ([102](#_ENREF_102)). | Dramatically increased in the late stage of adipogenesis ([60](#_ENREF_60), [61](#_ENREF_61)). |
| Pyruvate carboxylase | 35 ([70](#_ENREF_70)), 39 ([70](#_ENREF_70)), 79 ([70](#_ENREF_70)), 148 ([70](#_ENREF_70)), 152 ([70](#_ENREF_70)), 237 ([70](#_ENREF_70)), 241 ([70](#_ENREF_70)), 316 ([70](#_ENREF_70)), 434 ([70](#_ENREF_70)), 589 ([70](#_ENREF_70)), 717 ([70](#_ENREF_70)), 748 ([70](#_ENREF_70)), 892 ([70](#_ENREF_70)), 969 ([70](#_ENREF_70)), 992 ([11](#_ENREF_11), [70](#_ENREF_70)), and 1090 ([64](#_ENREF_64)), 1909 ([70](#_ENREF_70)). | No effect reported | Activity regulation at Lys 748 in *C. elegans,* mouse, HEK293 and HEK293T cells ([70](#_ENREF_70)). |
| Phosphoenolpyruvate carboxykinase | 19 , 70 ([11](#_ENREF_11)), 71 ([11](#_ENREF_11), [68](#_ENREF_68)), 72 ([68](#_ENREF_68)), 514 , 594 ([11](#_ENREF_11)) and 624 ([11](#_ENREF_11)). | No effect reported | Regulation of stability ([11](#_ENREF_11), [68](#_ENREF_68)).  Acetylation at Lys 514 crucial for enzymatic activity in *Saccharomyces cerevisiae* ([67](#_ENREF_67)). |
| Pyruvate kinase | 3, 62 ([64](#_ENREF_64), [66](#_ENREF_66)), 89 ([64](#_ENREF_64)), 135 ([64](#_ENREF_64)), 151, 162, 166 ([64](#_ENREF_64)), 186, 206, 207, 230, 261, 266 ([64](#_ENREF_64)) 305 ([11](#_ENREF_11), [64](#_ENREF_64), [66](#_ENREF_66), [89](#_ENREF_89)), 422, 433 ([64](#_ENREF_64)) and 498. | Enhanced expression in P-815 mastocytoma cells ([100](#_ENREF_100)).  Decrease in the amount of transcripts in human colon tissues ([92](#_ENREF_92)).  Decrease in gene expression in F9 terotocarcinoma cells ([101](#_ENREF_101)). | Regulation of activity and stability at Lys residue 305 in 293T cells ([66](#_ENREF_66)). |

Table 16. Acetylation of lysine residues, variation in response to butyrate and impact of lysine residues acetylation on Pyruvate metabolism enzymes. *Some of* t*he acetylated lysine residues were found in other species, but match the reported in humans.*

# Appendix III. *p*-values in every functional annotation of each GO term chosen in the global and acetylated proteome datasets.

|  |  |  |
| --- | --- | --- |
| Biological process | | |
| Functional annotation | ***p*-value** | |
|  | **Global proteome dataset** | **Acetylated proteome dataset** |
|  |  |
| RNA metabolic process | 1.25E-63 | 1.62E-21 |
| mRNA metabolic process | 7.06E-53 | 2.50E-24 |
| Cellular protein metabolic process | 4.92E-38 | 4.46E-19 |
| ncRNA metabolic process | 1.03E-27 | 6.72E-08 |
| Nucleoside phosphate metabolic process | 8.44E-15 | \* |
| Negative regulation of cellular protein metabolic process | 3.51E-13 | 6.04E-12 |
| Negative regulation of protein metabolic process | 9.44E-13 | 3.86E-12 |
| Purine ribonucleotide metabolic process | 3.03E-12 | \* |
| Glucose metabolic process | 1.13E-11 | \* |
| Regulation of protein metabolic process | 9.37E-11 | 4.78E-09 |
| Regulation of cellular protein metabolic process | \* | 9.49E-09 |
| Carboxylic acid metabolic process | \* | 5.75E-08 |
| Positive regulation of cellular protein metabolic process | \* | 1.17E-06 |

Table 17. BP GO term *p-*values. *This table shows the p-values of the functional annotations in BP GO term for the global and acetylated proteome datasets. \* indicates this functional annotation was not found in the dataset.*

|  |  |  |
| --- | --- | --- |
| Cellular compartment | | |
| Functional annotation | ***p*-value** | |
|  | **Global proteome dataset** | **Acetylated proteome dataset** |
|  | |
| Mitochondrial part | 1.23E-48 | 3.50E-12 |
| Mitochondrion | 3.26E-48 | 2.59E-13 |
| Mitochondrial lumen | 4.03E-35 | 4.05E-11 |
| Mitochondrial matrix | 4.03E-35 | 4.05E-11 |
| Mitochondrial envelope | 5.56E-28 | 1.10E-06 |
| Mitochondrial membrane | 3.23E-25 | 3.63E-06 |
| Mitochondrial inner membrane | 1.47E-24 | 8.87E-06 |
| Mitochondrial membrane part | 5.08E-21 | 5.13E-04 |
| Mitochondrial respiratory chain | 1.35E-11 | \* |
| Mitochondrial nucleoid | 1.82E-11 | 5.47E-08 |
| Mitochondrial outer membrane | \* | 0.00731247 |

Table 18. CC GO term *p-*values. *This table shows the p-values of the functional annotations in CC GO term for the global and acetylated proteome datasets. \* indicates this functional annotation was not found in the dataset.*

|  |  |  |
| --- | --- | --- |
| Molecular function | | |
| Functional annotation | ***p*-value** | |
|  | **Global proteome dataset** | **Acetylated proteome dataset** |
|  | |
| Lactate dehydrogenase activity | 0.00172786 | \* |
| beta-hydroxyacyl-CoA dehydrogenase activity | 0.00460503 | \* |
| Isocitrate dehydrogenase activity | 0.02222638 | \* |
| Enoyl-CoA hydratase activity | 0.03984772 | \* |
| Phosphoglycerate kinase activity | 0.05371314 | \* |
| Fructose-bisphosphate aldolase activity | 0.05371314 | \* |
| Malic enzyme activity | 0.09749902 | \* |
| Malate dehydrogenase activity | \* | 0.07988411 |

Table 19. MF GO term *p-*values. *This table shows the p-values of the functional annotations in MF GO term for the global and acetylated proteome datasets. \* indicates this functional annotation was not found in the dataset.*