

CARDIFF
UNIVERSITY

PRIFYSGOL
CAERDYDD

December 2022

Elucidating An Immune Metabolite Pathway In Sepsis

Thesis presented for the degree of Doctor of Philosophy

Linda Moet

Summary

This thesis aimed to evaluate the role of the medium-chain-fatty-acid (MCFA)-GPR84 signalling axis in sepsis. Sepsis is a life-threatening condition where the dysregulated host response to an infection can lead to death by organ failure. Several bodily systems can be affected in sepsis, including metabolism on a cellular and physiological level. GPR84 is a receptor present on innate immune myeloid cells and is upregulated in response to inflammatory stimuli. In sepsis, GPR84 expression is strongly upregulated and is an integral biomarker member of a transcriptomic signature that has been shown to accurately predict sepsis in a neonatal sepsis study. The ligands that it binds comprise MCFAs with a chain length of 10 or 12 carbons. These lipids are currently understudied and their presence in relevant amounts in the human body has been doubted in the science community. Moreover, their presence in the inflammatory context, such as during infection or in sepsis, has not been studied previously. Furthermore, their potential origin, if present, remains to be elucidated as well. I hypothesize that the MCFA-GPR84 axis is a key regulated host response in sepsis. To address this hypothesis, it is pivotal to test whether MCFAs are present in the blood and whether their concentrations change in sepsis patients and if so, to determine their possible origin. To this end, I established a lipid quantification method using liquid-chromatography tandem mass spectrometry (LC-MS/MS) that was able to reliably measure MCFAs and a range of longer chain fatty acids and acylcarnitines. Employing this method on plasma samples from a sepsis cohort led to the measurement of C10:0 and C12:0, with C10:0 being significantly increased in sepsis and C12:0 significantly decreased in some types of sepsis compared to controls. Notably, the measured lipid levels were heterogeneous and only a subset of patients showed increased C10:0 levels. Furthermore, I found using transcriptomic data that oleoyl-acp-hydrolase (*OLAH*), the gene encoding an enzyme pivotal in cellular MCFA production, is upregulated in sepsis. *OLAH*'s cellular role has previously been established with regards to MCFA production by the mammary gland during lactation, upregulation of *OLAH* in sepsis is likely caused by cortisol. I found that a synthetic glucocorticoid (dexamethasone) lead to increased *OLAH* gene expression in immune cell lines *in vitro*, in addition, the glucocorticoid receptor was found as potential transcription factor regulating *OLAH* in an *in silico* approach. Also, peripheral blood mononuclear cells were found to be able to produce C10:0 *in vitro*. Further analyses

were conducted in relation to a broader range of lipids and their levels in the plasma of sepsis patient to determine the plasma lipid profile associated with sepsis. In these investigations I found increased short-chain and medium-chain acylcarnitines and decreased poly-unsaturated fatty acids, namely arachidonic acid and eicosapentaenoic acid. An altered plasma lipid profile seems to be indicative of sepsis and was partially shown to be associated with mortality and severity. Finally, the transcriptional regulation of *GPR84* and its temporal expression pattern was examined using an *in silico* approach and *in vitro* experiments. These results indicate that *GPR84* is likely a secondary immune response gene, upregulated after around 2 hours of an inflammatory stimulus in myeloid cell types. Overall, it appears that the MCFA-GPR84 signalling axis is likely activated in a subset of sepsis patients, most likely those with increased cortisol and at a higher risk of complications and mortality. This tentatively indicates that C10:0-GPR84 signalling in sepsis is detrimental to survival. In conclusion, the data of my thesis supports the hypothesis that the MCFA-GPR84 signalling axis is a key regulated host response with the limitation that this is likely only the case in a subset of patients. In these patients though this signalling axis could likely be a mediator of the increased mortality observed in patients with strongly elevated cortisol.

Contents

Summary	i
Contents	i
Table of figures and tables	vii
Acknowledgements	xviii
Overview Lipids	xix
Abbreviation list	xxiv
1 Introduction	1
1.1 Sepsis	1
1.1.1 Incidence	2
1.1.2 Sepsis definitions and diagnosis.....	2
1.1.3 Treatment.....	5
1.1.4 The immune system	5
1.1.5 The inflammatory response to an infection	6
1.1.6 Resolution of the immune response.....	8
1.1.7 The immune response in sepsis	9
1.1.8 Metabolism in the immune response and sepsis	14
1.1.9 Metabolism on a physiological level	14
1.1.10 Cellular metabolism	15
1.1.11 Lipid metabolism in the cell	16
1.1.12 Bacterial sepsis.....	19
1.1.13 GPR43	20
1.1.14 Biomarkers	20
1.2 G-protein coupled receptors	21

1.2.1	Structure of G-protein coupled receptors and different conformational states	24
1.2.2	Ligands	28
1.2.3	G-proteins	28
1.2.4	G α_i and G α_s signalling.....	28
1.2.5	G α_q signalling.....	31
1.2.6	G $\alpha_{12/13}$ signalling.....	31
1.2.7	G $\beta\gamma$ signalling.....	33
1.2.8	Non-canonical signalling pathways	33
1.3	GPR84.....	35
1.3.1	Discovery.....	35
1.3.2	Receptor structure.....	35
1.3.3	GPR84 ligands – putative physiological ligands.....	35
1.3.4	GPR84 ligands – synthetic ligands	36
1.3.5	G-protein signalling.....	37
1.3.6	Regulation of <i>GPR84</i> expression.....	39
1.3.7	Expression in tissues and cells	39
1.3.8	Functions of GPR84 - Cytokine and chemokine secretion.....	41
1.3.9	Chemotaxis	43
1.3.10	Phagocytosis	43
1.3.11	Granule secretion	44
1.3.12	Pre-clinical studies on GPR84 antagonism as remedy for inflammation and fibrosis	44
1.4	“Omics” and other high throughput approaches to decipher biological activity.	46
1.4.1	Transcriptomics as window into cellular activity and regulation	46
1.4.2	Big data	47

Contents

1.4.3	Transcriptional regulation of gene expression	47
1.4.4	Analysis of transcriptional regulation	48
1.4.5	Measurements of metabolites using LC-MS/MS techniques	49
1.4.6	Liquid chromatography	49
1.4.7	Mass spectrometry	50
1.5	Rationale.....	54
2	Materials and methods.....	56
2.1	Optimization of a method for reliable lipid detection from blood samples	56
2.1.1	Materials	56
2.1.2	LC-MS/MS conditions.....	56
2.1.3	Derivatization optimization.....	56
2.1.4	Detection limits on the machine.....	57
2.1.5	Extraction from blood in PAXgene buffer	58
2.1.6	Whole blood extraction	58
2.2	Measurement of lipids in plasma of adult sepsis patients and healthy controls .	62
2.2.1	Calibration curve	62
2.2.2	Sample preparation and measurement.....	62
2.2.3	Data analysis	64
2.3	Lipid measurements in neonatal suspected sepsis cases and neonatal controls .	64
2.4	Mapping the transcription factor network associated with <i>GPR84</i> expression ...	65
2.5	The biosynthesis of GPR84 ligands.....	68
2.5.1	Cell culture - General remarks	68
2.5.2	General techniques	68
2.5.3	Cell maintenance.....	68
2.5.4	Freezing	69
2.5.5	Charcoal-treatment of serum	69

2.5.6	Sampling for RNA extraction	70
2.5.7	Sampling for lipid quantification	70
2.5.8	RNA sampling and extraction	70
2.5.9	HUVEC.....	72
2.5.10	Extraction of immune cells from healthy donor blood using Ficoll-density gradients.....	72
3	Optimization of a method for reliable lipid detection from blood samples.....	74
3.1	Results.....	79
3.1.1	Derivatisation optimization	79
3.1.2	Quenching optimization	81
3.1.3	LC-MS/MS conditions	85
3.1.4	Analyte recovery from whole blood.....	96
3.1.5	Efforts to decrease the background in extraction and derivatization.....	111
3.2	Discussion.....	114
4	Measurement of lipids in plasma of adult sepsis patients and healthy adult volunteers	118
4.1	Results.....	121
4.1.1	Demography of study populations	121
4.1.2	Evaluation of the reliability of lipid measurements in this cohort.....	124
4.1.3	Measurements of potential ligands for GPR84	129
4.1.4	Measurements of additional straight chain lipids.....	131
4.1.5	Measurements of acylcarnitines	135
4.1.6	Identification of lipids significantly altered in sepsis.....	138
4.1.7	Explorative, unsupervised investigation of lipids altered in sepsis.....	140
4.1.8	Lipid concentrations and survival	145
4.1.9	Disease severity and sepsis cause	150

Contents

4.1.10	Lipid concentrations in the plasma of patients with different sepsis causes	153
4.1.11	Correlation between lipids.....	156
4.1.12	Main findings.....	159
4.2	Discussion	159
5	Lipid measurements in neonatal suspected sepsis cases and neonatal controls	167
5.1	Results	167
5.1.1	Cohort.....	167
5.1.2	Quality control	168
5.1.3	Measurements of free fatty acids.....	170
5.1.4	Measurements of acylcarnitines.....	174
5.1.5	Lipid profile changes in suspected sepsis cases.....	176
5.1.6	Correlations between lipids	181
5.2	Discussion	183
6	Mapping the transcription factor network associated with <i>GPR84</i> expression in sepsis	186
6.1.1	Transcription factor binding site analysis	188
6.2	Results	190
6.2.1	Transcriptional regulation in sepsis	190
6.2.2	Functional enrichment of transcription factors in sepsis	196
6.2.3	Physical transcription factor interactions with the <i>GPR84</i> promoter according to existing Chip data	199
6.2.4	High confidence networks of transcription factors regulating <i>GPR84</i> in sepsis	201
6.2.5	The role of regulatory regions in the transcriptional regulation of <i>GPR84</i>	207
6.2.6	Experimental data on cell-specific and temporal expression patterns	219
6.3	Discussion	223

7	Biosynthesis of ligands for GPR84	234
7.1	Results	237
7.1.1	Systemic expression of OLAH in paediatric and neonatal sepsis patients. .	237
7.1.2	The promoter region of OLAH	239
7.1.3	Enhancers in the transcriptional regulation of <i>OLAH</i>	239
7.1.4	Cell specific expression of <i>OLAH</i> in sepsis	239
7.1.5	Transcriptional regulation of <i>OLAH</i> in sepsis.....	243
7.1.6	Induction of <i>OLAH</i> expression in the THP-1 cell line	251
7.1.7	MCFA production by THP-1 cells	253
7.1.8	<i>OLAH</i> expression in endothelial cells and MCFA production	254
7.1.9	<i>OLAH</i> expression blood cells and MCFA production	256
7.1.10	ACADM expression	260
7.1.11	Expression of genes involved in lipid metabolism pathways in sepsis.....	262
7.2	Discussion.....	266
8	Discussion	270
8.1.1	Main findings	270
8.1.2	The increase of C10:0 (and C8:0) in plasma of sepsis patients and the decrease of C12:0	272
8.1.3	Acylcarnitines in sepsis	274
8.1.4	Unsaturated fatty acid changes in sepsis	278
8.1.5	Urosepsis compared to sepsis OUA and pneumonia sepsis.....	279
8.1.6	Cortisol in sepsis	279
8.1.7	The role of GPR84 in sepsis	280
8.1.8	Pitfalls	283
8.1.9	Future directions.....	285
9	References	292

Table of figures and tables

10	Supplementary materials.....	314
----	------------------------------	-----

Table of figures and tables

TABLE 1	OVERVIEW OF LIPIDS MEASURED IN THE THESIS.....	XIX
TABLE 2	THE SOFA SCORE CURRENTLY IN USE PER THE SEPSIS-3 DEFINITION. SOURCE: SINGER, M. ET AL. THE THIRD INTERNATIONAL CONSENSUS DEFINITIONS FOR SEPSIS AND SEPTIC SHOCK (SEPSIS-3). JAMA - JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION 315 , 801–810 (2016).	2
FIGURE 1	A DIAGRAM DEPICTING THE IMMUNE RESPONSE IN SEPSIS. IT IS CHARACTERIZED BY A DYSREGULATED INNATE IMMUNE RESPONSE WITH PERSISTENT INFLAMMATION ACCOMPANIED BY A SUPPRESSED ADAPTIVE IMMUNE RESPONSE. AFTER THE INITIAL RESPONSE THAT CAN LEAD TO DEATH BY ORGAN FAILURE, SEPSIS ALSO LEADS TO POOR HEALTH IN THE LONG TERM IN SURVIVORS DUE TO A PRO-LONGED INNATE IMMUNE ACTIVITY AND DECREASED ADAPTIVE IMMUNE SYSTEM. SEPSIS SURVIVORS FREQUENTLY DIE IN THE MONTHS AND YEARS FOLLOWING THEIR INITIAL SEPSIS EPISODE. GRAPH TAKEN FROM DELANO AND WARD (2016)	12
FIGURE 2	THESE GRAPHS SHOW THE DIFFERENCE IN EXPRESSION AS FOLD CHANGE TO CONTROL IN CONTROLS (BLUE DOTS) AND SEPTIC NEONATES (RED DOTS) OF HIGHLY SIGNIFICANT GENES. THE RED LINES INDICATE GENES OF THE ADAPTIVE IMMUNE RESPONSE, THE GREEN LINES ARE METABOLIC GENES AND THE BLUE LINES ARE INNATE IMMUNE RESPONSE GENES. THESE RESULTS INDICATE AN INCREASED EXPRESSION OF INNATE IMMUNE GENES, METABOLISM GENES AND A DECREASE IN ADAPTIVE IMMUNE GENES. GRAPHS TAKEN FROM SMITH ET AL. (2014)	13
FIGURE 3	THE GENERAL SIGNALLING PATHWAYS THAT CAN BE INITIATED WHEN A LIGAND BINDS TO A G-PROTEIN COUPLED RECEPTOR. IN CANONICAL SIGNALLING, UPON BINDING OF THE LIGAND, THE TRANSDUCERS $G\alpha$ AND $G\beta\gamma$ DISSOCIATE AND ARE ABLE TO TRANSDUCE THE SIGNALS BY REGULATING EFFECTORS. THESE EFFECTORS CAN LEAD TO THE SYNTHESIS OR RELEASE OF SECONDARY MESSENGERS, WHICH WILL ACTIVATE DISTINCT SIGNALLING PATHWAYS AND DOWNSTREAM EFFECTS. IN NON-CANONICAL SIGNALLING, AN EFFECTOR, MOST NOTABLY ARRESTIN, CAN INHIBIT G-PROTEIN MEDIATED SIGNALLING, OFTEN VIA RECEPTOR INTERNALIZATION WHICH LEADS TO RECEPTOR RECYCLING OR DEGRADATION. THIS EFFECTOR CAN ALSO ACTIVATE ITS OWN DISTINCT SIGNALLING PATHWAYS AND DOWNSTREAM EFFECTS, SOME OF WHICH OVERLAP WITH CANONICAL SIGNALLING ACTIVATED PATHWAYS.	23
FIGURE 5	THE GENERAL STRUCTURE OF GPCR IN A TWO-DIMENSIONAL DEPICTION SHOWING THE EXTRACELLULAR N-TERMINUS AND EXTRACELLULAR LOOPS, TRANSMEMBRANE HELICES AND INTRACELLULAR LOOPS AND C-TERMINUS.....	25
FIGURE 6	THREE STRUCTURES OF GPCR, BOVINE RHODOPSIN (PURPLE), AVIAN $\beta 1AR$ (YELLOW) AND HUMAN ADENOSINE $A2A$ (GREEN) ARE IMPOSED ONTO HUMAN $\beta 2AR$ (BLUE) AND SHOWN FROM TWO DIFFERENT ANGLES. b) EACH OF THE RECEPTORS IS SHOWN IN THE INACTIVE STATE, THE LIGANDS SHOWN INHIBIT RECEPTOR SIGNALLING. THE RECEPTORS ARE SHOWN AS VIEWED FROM THE EXTRACELLULAR SURFACE, THE VIEWER HENCE LOOKS AT THE BINDING POCKET FROM THE TOP. ECL = EXTRACELLULAR LOOP, ICL = INTRACELLULAR LOOP, TM = TRANSMEMBRANE, H8 = HELIX 8. PICTURE TAKEN FROM ROSENBAUM ET AL. (2009), LICENSE FOR RE-USE OBTAINED FROM SPRINGER NATURE (LICENSE NUMBER 5602540288428).....	26

FIGURE 7 THE CONFORMATIONAL LANDSCAPE OF GPCR. R' AND R'' ARE INACTIVE-LIKE STATES, WHEREAS R* AND R** ARE ACTIVE-LIKE STATES. THE BLACK LINE (APO) INDICATES THE ABSENCE OF LIGAND, WHEREAS BLUE INDICATES THE PRESENCE OF EITHER AGONIST OR G-PROTEIN, AND GREEN INDICATES THE PRESENCE OF BOTH. ON THE Y-AXIS THE REQUIRED ACTIVATION ENERGY IS SHOWN. ADAPTED FROM WOOTTEN ET AL. (2018), LICENSE FOR RE-USE OBTAINED FROM SPRINGER NATURE (LICENSE NUMBER 5602541091077)27

FIGURE 8 AN OVERVIEW OF PATHWAYS DOWNSTREAM OF G_{A5} AND G_{Ai}- PROTEIN SIGNALLING WITH EFFECTS IN MACROPHAGES AS ELUCIDATED BY ARONOFF ET AL. G_{A5} PROTEIN STIMULATES ADENYLYL CYCLASE (AC), WHILE G_{Ai} INHIBITS IT. AC CONVERTS ATP TO CAMP, THE SECONDARY MESSENGER THAT ACTIVATES PKA AND EPAC. PKA CAN INCREASE IL-6 AND IL-10 SECRETION, INHIBIT TNF-A AND CCL3 SECRETION AND INHIBIT PHAGOCYTIC LYSIS. EPAC HAS NO EFFECT ON CYTOKINE SECRETION BUT INHIBITS BOTH FcR-MEDIATED PHAGOCYTOSIS AND THE LYSIS OF PHAGOCYTOSED BACTERIA.....30

FIGURE 9 DIAGRAM OF G_{α12/13} SIGNALLING. G_{α12} HAS BEEN SHOWN TO ACTIVATE SEVERAL TRANSCRIPTION FACTORS VIA THE RAS-MEK-ERK PATHWAY IN AN OVARIAN CANCER CELL LINE. G_{α12/13} PROTEINS ARE KNOWN TO ACTIVATE RHO AND LEAD TO CHEMOTAXIS, THIS IS POTENTIALLY MEDIATED VIA JNK WHICH IS KNOWN TO BE AFFECTED BY G_{A12/13} AND RHO AND AFFECT THE CYTOSKELETON (GREEN) BUT HAS NOT BEEN CONCLUSIVELY PROVEN. BLUE ARROWS INDICATE THE POTENTIAL INTERACTION OF G-PROTEIN AND RTK OR THE NON-CANONICAL SIGNALLING OF RTK LEADING TO CHEMOTAXIS; THIS INTERACTION REQUIRES FURTHER INVESTIGATION.32

FIGURE 10 OVERVIEW OF A LIQUID-CHROMATOGRAPHY-MASS SPECTROMETER. A) PHOTOGRAPH OF THE LC-MS MACHINE USED IN THIS STUDY, AND A SCHEMATIC REPRESENTATION OF A LC-MS SYSTEM. (TAKEN FROM WIKIMEDIA COMMONS) B) FLOW DIAGRAM OF THE STEPS INVOLVED IN QUANTIFICATION OF COMPONENTS, AND THE MACHINE PARTS AND SOFTWARE REQUIRED TO EXECUTE THESE. C) SHOWS THE MECHANISM OF LIQUID CHROMATOGRAPHY (LEFT SIDE) AND THE ELECTRO SPRAY IONIZATION (MIDDLE, FIGURE TAKEN FROM [HTTPS://NATIONALMAGLAB.ORG/USER-FACILITIES/ICR/TECHNIQUES/ESI](https://nationalmaglab.org/user-facilities/icr/techniques/esi)) AND THE MACHINE PART IN WHICH THE ELECTRO SPRAY IONIZATION IS CARRIED OUT.....53

TABLE 3 THE CONCENTRATION OF THE INTERNAL STANDARDS IN THE INTERNAL STANDARDS MIX.61

FIGURE 11 DERIVATIZATION OF FATTY ACYLS USING 3PNH AND EDC AS DESCRIBED BY HAN ET AL (2015)(A) AND DERIVATISATION OF ACYLCARNITINE AS DESCRIBED BY MEIERHOFER (2019) (B). : CATALYSED BY EDC THE CARBOXYLIC GROUP OF THE FATTY ACID AND THE ACYLCARNITINE, RESPECTIVELY, BINDS COVALENTLY TO THE NITROUS GROUP OF 3-NITROPHENYLHYDRAZINE. THIS REACTION RELEASES ONE MOLECULE OF WATER (NOT DEPICTED). PYRIDINE IS REQUIRED AS IT STABILIZES THE PH OF THE REACTION MIXTURE. HAN ET AL. (2015) FOUND BEST DERIVATISATION RESULTS BY HEATING THE SAMPLES TO 40 °C WITH 30 MINUTES OF REACTION TIME BEING SUFFICIENT, AND BOTH ACETONITRILE AND METHANOL AS SOLVENTS WERE ADEQUATE¹⁵⁷. THE FATTY ACID DEPICTED AS EXAMPLE IS OCTANOIC ACID, AND THE DEPICTED ACYLCARNITINE IS HEXANOYLARNITINE. CHEMICAL STRUCTURES WERE GENERATED USING CHEMSPIDER.78

80

FIGURE 12 DERIVATISATION OF MCFA UNDER DIFFERENT EXPERIMENTAL CONDITIONS. LC-MS/MS PEAK AREAS OF DIFFERENT COMPOUNDS DERIVATIZED USING EITHER 75% METHANOL OR 75% ACETONITRILE AS REACTION SOLVENTS. THE TESTED TIME PERIODS RANGED FROM 30 MINUTES TO 3 HOURS. DERIVATISATION WAS CARRIED OUT IN A WATERBATH SET TO 40°C. DATA ARE AVERAGES (N=3, ERROR BARS INDICATE STANDARD DEVIATION). BACKGROUND WAS MEASURED USING BOTH AN ACETONITRILE AND A METHANOL BLANK THAT WAS SUBTRACTED FROM THE SAMPLES.80

Table of figures and tables

FIGURE 13 CALIBRATION CURVES OF A SELECTION OF DEUTERATED FFA AND ACYLCARNITINE STANDARDS. THE X-AXIS SHOWS THE AMOUNT INJECTED ON THE LC-MS/MS COLUMN (PICOGRAMS; CALCULATED BASED ON THE INPUT AMOUNT OF STANDARDS STOCKS BEFORE DERIVATISATION). THE Y-AXIS SHOWS THE AREAS AS MEASURED BY LC-MS/MS. THE LINES INDICATE THE INTERPOLATED CURVE BASED ON THE DATAPPOINTS, THE GOODNESS OF FIT SHOWED R^2 VALUES ABOVE 0.99 FOR ALL LIPIDS EXCEPT C18-D35 ($R^2=0.985$).....	80
FIGURE 14 THE EFFECT OF DIFFERENT QUENCHING METHODS ON CONTAMINATING FATTY ACIDS. 4 DIFFERENT QUENCHING METHODS WERE COMPARED TO THE CONTROL (NO QUENCHING) TO CHECK WHICH METHOD COULD REDUCE BACKGROUND LEVELS OF FATTY ACIDS MEASURED ON THE MACHINE. THE Y-AXIS DEPICTS THE RATIO OF THE LIPIDS MEASURED IN THE "SAMPLE" THAT UNDERWENT DERIVATISATION AND QUENCHING, TO THE LIPIDS MEASURED IN THE CONTROL – A SIMILAR SAMPLE THAT ONLY UNDERWENT DERIVATISATION BUT NO QUENCHING. ALL SAMPLES WERE DILUTED TO THE SAME VOLUME TO ASSURE COMPARABILITY. ON THE LEFT SIDE ALL LIPIDS THAT ARE MEASURED ARE CONTAMINANTS, SINCE NO NOT DEUTERATED LIPID STANDARDS WERE ADDED. ON THE RIGHT SIDE, INTERNAL DEUTERATED STANDARDS ARE DEPICTED. THESE WERE ADDED RIGHT BEFORE DERIVATISATION AND ARE USED HERE AS AN INDICATION OF DELETERIOUS EFFECTS OF THE QUENCHING METHODS ON DERIVATIZED LIPIDS.....	83
FIGURE 15 A) EXAMPLE CHROMATOGRAM OF FREE FATTY ACIDS (STANDARD MIX, 50 μ M, THE DEUTERATED STANDARDS WERE ADDED AS INTERNAL STANDARD MIX AT CONCENTRATIONS AS SET OUT IN THE METHODS CHAPTER) RUN ON THE OPTIMISED LC-MS/MS SYSTEM. FOR MOST LIPIDS ONLY THE UNLABELLED STANDARD IS INDICATED, THE DEUTERATED STANDARD IS TYPICALLY VISIBLE OVERLAPPING (OFTEN THEY DONT HAVE THE EXACT SAME RT, THERE IS SLIGHT DIFFERENCE) WITH ITS CORRESPONDING STANDARD. B) SHOWS A CLOSER LOOK AT THE LIPIDS THAT ELUTE UP TO 18 MINUTES INTO THE RUN. C) SHOWS A CLOSER LOOK AT LIPIDS ELUTING AFTER 18 MINUTES. WITH SOME EXEMPTIONS, PEAKS ARE BROADER AND FLATTER FOR SMALLER LIPIDS AND NARROWER AND HIGHER FOR LONGER LIPIDS.	87
88	
FIGURE 16 EXAMPLE CHROMATOGRAM OF ACYLCARNITINES (STANDARD MIX) RUN ON THE OPTIMISED LC-MS SYSTEM. ALL CARNITINES ARE PRESENT AT 10 μ M, AND THE DEUTERATED STANDARDS WERE ADDED AS INTERNAL STANDARD MIX WHOSE CONCENTRATIONS ARE SET OUT IN THE METHODS CHAPTER. FOR ALL LIPIDS ONLY THE NOT DEUTERATED STANDARD IS INDICATED, THE DEUTERATED STANDARD IS TYPICALLY VISIBLE OVERLAPPING WITH ITS CORRESPONDING STANDARD. AT THIS STAGE OF OPTIMIZATION C18:2 AND C18:3 STANDARDS WERE NOT USED YET.	88
TABLE 4 MS SETTINGS FOR THE DETECTION OF FATTY ACIDS AND ACYL CARNITINES. DECLUSTERING POTENTIAL (DP), COLLISION ENERGY (CE), MASS IN Q1 AND MASS IN Q3. THE RETENTION TIME (RT) IS AN INDICATION SINCE RETENTION TIMES VARY BETWEEN RUNS DUE TO VARIOUS FACTORS.....	91
FIGURE 17 CALIBRATION CURVES OF A SELECTION OF FREE FATTY ACIDS (LEFT SIDE) AND ACYLCARNITINES (RIGHT SIDE). THE X-AXIS DEPICTS THE CONCENTRATION IN SAMPLE IN μ M THAT CORRESPONDS TO THE CONCENTRATION IN A BLOOD SAMPLE WITH THE ESTABLISHED EXTRACTION METHOD CALCULATED FROM THE STOCK CONCENTRATION BEFORE DERIVATISATION. THE Y-AXIS DEPICTS THE AREAS AS MEASURED BY LC-MS/MS. ALL DEPICTED STANDARD CURVES HAVE A R^2 -GOODNESS-OF-FIT VALUE OF ABOVE 0.99.....	93
TABLE 5 THE TABLE SHOWS THE CONCENTRATION RANGE IN WHICH MEASUREMENTS BY LC-MS/MS WERE IN THE LINEAR RANGE. THE CALIBRATION CURVE WAS PREPARED BY DILUTION OF LIPID STOCKS AND SUBSEQUENT DERIVATISATION, AND I CALCULATED TO WHAT CONCENTRATIONS OF LIPIDS IN SAMPLES BEFORE EXTRACTION THAT WOULD CORRESPOND. IN	

ADDITION, THE CORRESPONDING AMOUNT ON COLUMN (IN PG) THAT WAS INJECTED PER MEASUREMENT, OF EACH COMPOUND, IS INDICATED IN THE RIGHT COLUMN.	94
FIGURE 18 OVERVIEW OF METHODS UTILIZED TO EXTRACT FATTY ACIDS AND CARNITINES FROM WHOLE BLOOD. THIS OVERVIEW WAS CREATED USING BIORENDER.COM.	98
99	
FIGURE 19 PERCENTAGE RECOVERY OF DEUTERATED FATTY ACIDS AND ACYLCARNITINES (SPIKED INTERNAL STANDARDS) BY DIFFERENT EXTRACTION METHODS FROM PBS AND BLOOD. THE EXTRACTION METHODS USED WERE THE BIPHASIC LAYER EXTRACTION USING ETHYL ACETATE OR DICHLOROMETHANOL (MODIFIED BIPHASIC) OR THE PROTEIN CRASH METHOD. THE INTERNAL STANDARDS WERE SPIKED INTO 50 µL BLOOD OR PBS SAMPLES BEFORE EXTRACTION AND COMPARED TO PEAK AREAS OF THE INTERNAL STANDARDS SPIKED INTO BLANK SAMPLES (100% METHANOL) AT THE SAME CONCENTRATION DIRECTLY BEFORE DERIVATIZATION WITHOUT EXTRACTION. RECOVERY WAS CALCULATED AS PERCENTAGE OF PEAK AREAS FROM EXTRACTED SAMPLES BY PEAK AREA OF DIRECTLY DERIVATIZED SOLVENT. ALL EXTRACTIONS WERE DONE WITH N=5. THE STATISTICAL TEST PERFORMED IS A 2-WAY ANOVA. SIGNIFICANCE LEVELS DISPLAYED ARE: * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = P ≤ 0.0001.	99
FIGURE 20 SHORT- AND MEDIUM CHAIN FATTY ACIDS EXTRACTED FROM 50 µL PBS OR 50 µL BLOOD USING THE BIPHASIC METHOD, THE MODIFIED BIPHASIC METHOD AND THE PROTEIN CRASH METHOD. THE Y-AXIS SHOWS THE AREAS AS MEASURED BY LC-MS/MS. WHEN COMPARING AREAS MEASURED, KEEP IN MIND THAT THE METHODS LEAD TO DIFFERENT FRACTION OF INITIAL SAMPLE BEING DERIVATISED. SIGNIFICANCE WAS CALCULATED USING TWO-WAY ANOVA. SIGNIFICANCE LEVELS DISPLAYED ARE: * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = P ≤ 0.0001. ALL EXTRACTIONS WERE DONE WITH N=5.	100
101	
FIGURE 21 ACYLCARNITINES EXTRACTED FROM BLOOD AND PBS USING THE BIPHASIC METHOD, MODIFIED BIPHASIC METHOD AND PROTEIN CRASH METHOD. THE Y-AXIS SHOWS THE PEAK AREA THAT WAS MEASURED BY LC-MS/MS. DUE TO THE POOR RECOVERY OF THE INTERNAL STANDARD, THE PEAK AREA WAS DEEMED A BETTER REPRESENTATION OF EXTRACTION THAN THE CALCULATED AMOUNT IN SAMPLE. ALL EXTRACTIONS WERE DONE WITH N=5.	101
FIGURE 22. THE FIGURE SHOWS THE AREAS AS MEASURED BY LC-MS/MS FOR THE EXTRACTED DEUTERATED STANDARDS (FFA, A, ACYLCARNITINES, B) FOR ALL EXTRACTIONS AS AVERAGE OF THE 5 REPLICATES WITH STANDARD DEVIATION ON THE ERROR BARS. THE X-AXIS DEPICTS THE AMOUNT OF BLOOD USED IN THE EXTRACTION, AND THE Y-AXIS THE AREA MEASURED BY LC-MS/MS.	104
FIGURE 23 LIPIDS TO INTERNAL STANDARDS RATIOS OF FFA MEASURED IN SAMPLES WITH VARYING BLOOD TO MeOH RATIO AND THE CONTROL SAMPLE CONTAINING NO BLOOD. 50, 40, 30, 20, 10 µL OR 0 µL BLOOD WERE MIXED WITH 10 µL INTERNAL STANDARD MIX AND A CORRESPONDING MeOH VOLUME TO AN END VOLUME OF 260 µL AS PART OF A PROTEIN CRASH EXTRACTION. THE RATIO OF LIPID TO IS FROM THE 0 µL BLOOD SAMPLE WAS SUBTRACTED FROM THE OTHER SAMPLES FOR BACKGROUND CORRECTION. THE Y-AXIS DEPICTS THE RATIO OF THE MEASURED LIPID TO ITS CORRESPONDING OR CLOSEST FFA AS SET OUT IN THE EXPLANATION ABOVE. THE LINE INDICATES THE AVERAGE, AND THE DOTS INDICATE REPLICATE MEASUREMENTS. THIS EXPERIMENT WAS CONDUCTED WITH N=5.	107
FIGURE 24 LIPIDS TO INTERNAL STANDARDS RATIOS OF ACYLCARNITINES MEASURED IN SAMPLES WITH VARYING BLOOD TO MeOH RATIO AND THE CONTROL SAMPLE CONTAINING NO BLOOD. 50, 40, 30, 20, 10 µL OR 0 µL BLOOD WERE MIXED WITH 10	

Table of figures and tables

<p>µL INTERNAL STANDARD MIX AND A CORRESPONDING MEOH VOLUME TO AN END VOLUME OF 260 µL AS PART OF A PROTEIN CRASH EXTRACTION. THE RATIO OF LIPID TO IS FROM THE 0 µL BLOOD SAMPLE WAS SUBTRACTED FROM THE OTHER SAMPLES FOR BACKGROUND CORRECTION. THE Y-AXIS DEPICTS THE RATIO OF THE MEASURED LIPID TO ITS CORRESPONDING OR CLOSEST FFA. C18:1 CARNITINE TO INTERNAL STANDARD RATIO WAS CALCULATED BASED ON DEUTERATED C18:0 CARNITINE. THIS EXPERIMENT WAS CONDUCTED WITH N=5.</p>	109
<p>FIGURE 25 THE PLOT SHOWS THE PERCENTAGE OF COVARIATION FOR EVERY LIPID ON THE Y-AXIS SET OUT AGAINST THE VOLUME OF BLOOD EXTRACTED FROM. THE TOTAL EXTRACTION VOLUME WAS KEPT CONSTANT. THIS ANALYSIS INDICATES HOW MUCH THE EXTRACTION MATRIX IMPACTS THE RELIABILITY OF LIPID QUANTIFICATION. FATTY ACIDS ARE SHOWN IN A) AND ACYLCARNITINES ARE SHOWN IN B).</p>	110
<p>FIGURE 26 LEVELS OF FFA IN BLANK SOLVENTS. BAR PLOTS OF BACKGROUND MEASURED WHEN EXTRACTING FROM EITHER A 10 µL METHANOL (MEOH) OR WATER (H₂O) "SAMPLE" ACCORDING TO THE ESTABLISHED PROTOCOL. THERE WERE NO SIGNIFICANT DIFFERENCES BETWEEN MEOH AND WATER EXTRACTS (2-WAY ANOVA).</p>	112
<p>FIGURE 27 THE FOLD CHANGE OF MEASURED LIPID AFTER DERIVATISATION WITH DIFFERENT DILUTIONS OF DERIVATISATION REAGENT. THE DERIVATISATION REAGENT WAS DILUTED TO 75%, 50% OR 25% OF THE CONCENTRATION IN THE ESTABLISHED PROTOCOL. THE PERCENTAGE OF DERIVATISATION REAGENT USED IS PLOTTED AGAINST THE AREAS MEASURED PER DEUTERATED LIPID. THE PERCENTAGE DERIVATISATION REAGENT RELATES TO THE ESTABLISHED PROTOCOL (100%).</p>	113
<p>FIGURE 28 THE OPTIMIZED WORKFLOW AS ESTABLISHED IN THIS CHAPTER FOR FREE FATTY ACID AND ACYLCARNITINE QUANTIFICATION IN BLOOD SAMPLES.</p>	117
<p>TABLE 6 DEMOGRAPHY OF SEPSIS PATIENTS AND HEALTHY CONTROLS INCLUDED IN THIS STUDY</p>	122
<p>FIGURE 29 MEASUREMENT OF LIPIDS IN THE QUALITY CONTROL. A) MEASUREMENT OF FREE FATTY ACIDS ACROSS 13 MEASUREMENTS (UNDILUTED SAMPLES) OR 9 (DILUTED SAMPLES) OF THE QUALITY CONTROL SAMPLE. C) AND D) TABLES SUMMARIZING THE MEAN, %CV AND N OF QUALITY CONTROL MEASUREMENTS FOR FFA AND ACYLCARNITINES. B) MEASUREMENTS OF ACYLCARNITINES ACROSS 11 QUALITY CONTROL MEASUREMENTS. QC1 IS THE FIRST QUALITY CONTROL MEASUREMENT AND QC13, 9 AND 11 ARE THE LAST QUALITY CONTROL MEASUREMENT FOR FFA DILUTED AND UNDILUTED AND ACYLCARNITINES, RESPECTIVELY. THE DIFFERENCE IN HOW MANY QUALITY CONTROL SAMPLE MEASUREMENTS WERE PERFORMED IS CAUSED BY DIFFERENCES IN HOW MANY MEASUREMENTS WERE EXECUTED PER RUN, I.E., WHEN SAMPLE MEASUREMENTS WERE REPEATED DUE TO TECHNICAL ISSUES SUCH AS LEAKS IN THE LIQUID CHROMATOGRAPHY OR MOVING PEAKS.</p>	127
<p>FIGURE 30 BOXPLOTS OF THE CONCENTRATIONS IN µM OF POTENTIAL GPR84 LIGANDS IN THE PLASMA OF SEPSIS PATIENTS (N=52) AND OF HEALTHY CONTROLS (N=18); A) C10:0 OR DECANOIC ACID, B) C12:0 OR DODECANOIC ACID, C) ALL SAMPLES WITH C10:0 LEVELS BELOW 10 µM. THE DASHED LINE INDICATES THE EC50 CONCENTRATION REQUIRED TO ACTIVATE GPR84 AS EVALUATED BY WANG ET AL. USING A cAMP ASSAY⁹⁶(C10:0 4.5 µM, C12:0 8.8 µM). THE STATISTICAL TEST PERFORMED WAS A MANN-WHITNEY-WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05. THE LOWER HALF OF EACH BOX IN THE BOXPLOT DEPICTS THE 2ND QUANTILE OF VALUES, WITH THE UPPER HALF DEPICTING THE THIRD QUANTILE. THE EXTENDED LINES DEPICT THE FIRST AND LAST QUANTILE OF VALUES, WITH ANY POINTS REACHING BELOW OR ABOVE THESE LINES BEING STATISTICAL OUTLIERS.</p>	130
<p>FIGURE 31 CONCENTRATIONS IN µM IN THE PLASMA OF SEPSIS PATIENTS AND HEALTHY CONTROLS; A) C3-OH OR LACTIC ACID, B) C8:0 OR OCTANOIC ACID, C) C14:0 OR MYRISTIC ACID, D) C16:0 OR PALMITIC ACID, E) C18:0 OR STEARIC ACID, F) C8:0</p>	

IN ALL SAMPLES WITH CONCENTRATIONS UP TO 15 μ M. THE STATISTICAL TEST PERFORMED IS A MANN-WHITNEY-WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05.	133
FIGURE 32 CONCENTRATIONS IN μ M IN THE PLASMA OF SEPSIS PATIENTS AND HEALTHY CONTROLS; A) C18:1 OR OLEIC ACID, B) C18:2 OR LINOLEIC ACID, C) C18:3 OR α -LINOLENIC ACID, D) C20:4 OR ARACHIDONIC ACID, E) C20:5 OR EICOSAPENTAENOIC ACID, F) C22:6 OR DOCOSAHEXAENOIC ACID. THE STATISTICAL TEST PERFORMED IS A MANN-WHITNEY-WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05.	134
FIGURE 33 BOXPLOTS DEPICTING THE ACYLCARNITINE CONCENTRATIONS MEASURED IN PLASMA SAMPLES OF SEPSIS PATIENTS AND HEALTHY CONTROLS. SHOWN ARE A) ACETYLCARNITINE (C2:0-CARNITINE), B) PROPIONYLCARNITINE (C3:0-CARNITINE), C) BUTYRYLCARNITINE (C4:0-CARNITINE), D) HEXANOYLCARNITINE (C6:0-CARNITINE), E) OCTANOYLCARNITINE (C8:0-CARNITINE), F) DECANOYL-CARNITINE (C10:0-CARNITINE) G) DODECANOYLCARNITINE (C12:0-CARNITINE), H) MYRISTOYLCARNITINE (C14:0-CARNITINE), I) PALMITOYLCARNITINE (C16:0-CARNITINE), J) STEAROYLCARNITINE (C18:0-CARNITINE) AND K) OLEOYLCARNITINE (C18:1-CARNITINE). THE STATISTICAL TEST PERFORMED IS A MANN-WHITNEY-WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05.	137
FIGURE 34 VOLCANO PLOT DEPICTING LIPID CONCENTRATIONS ALTERED IN THE PLASMA OF SEPSIS PATIENTS COMPARED TO HEALTHY CONTROLS. THE Y-AXIS DEPICTS THE NEGATIVE LOG ₁₀ OF THE P-VALUE, HENCE THE HIGHER THE VALUE THE HIGHER THE SIGNIFICANCE. THE X-AXIS DEPICTS THE LOG ₂ OF THE FOLD CHANGE CALCULATED BASED ON NOT-NORMALIZED LIPID CONCENTRATIONS. THE DASHED LINES AT LOG ₂ (FC) 1 INDICATE A 2-FOLD CHANGE ON RAW DATA. SIGNIFICANCE WAS CALCULATED USING THE KRUSKAL-WALLIS TEST ON LOG ₁₀ -TRANSFORMED LIPID CONCENTRATIONS AND IS ADJUSTED FOR MULTIPLE TESTING. DASHED LINES INDICATE A LOG ₂ (FC) OF 1 AND A -LOG ₁₀ P-VALUE THAT CORRESPONDS TO P=0.1 AFTER MULTIPLE TESTING CORRECTION.	139
FIGURE 35 A) SHOWS THE HEATMAP OF ALL SAMPLES AND ALL LIPIDS MEASURED (LOG-NORMALIZED AND CENTRE-SCALED). BOTH LIPIDS AND SAMPLES ARE CLUSTERED USING D. WARD CLUSTERING, DISTANCES ARE EUCLIDEAN. THE COLOUR SCALE IN THE HEATMAP GOES FROM 4 (INCREASED) TO -4 (DECREASED) AND THE COLOUR INDICATES THE Z-SCORE OF THE LIPID MEASUREMENT WITHIN THE ROW FOR EACH SAMPLE. B) SHOWS A 3D-PLOT OF A PCA ON THE DATA WITH THE COLOURS INDICATING CONTROLS, SURVIVORS AND NON-SURVIVORS. C), D) AND F) ARE ALSO PCA PLOTS OF THE EXACT SAME DATA BUT SHOWN AS BIPLLOTS. IN C) THE DATAPOINTS ARE COLOURED BASED ON THEIR BELONGING INTO EITHER THE CONTROL, SURVIVOR OR NON-SURVIVOR GROUP, WITH PC1 BEING THE X-AXIS AND PC2 THE Y-AXIS. D) SHOWS THE BIPLLOT OF CONTROLS, SURVIVORS AND NON-SURVIVORS WITH PC2 AS THE X-AXIS AND PC3 AS THE Y-AXIS. F) SHOWS THE SAME PCA AS C), BUT THE DATAPOINTS ARE COLOURED BASED ON THEIR BELONGINGS INTO EITHER THE CONTROL, PNEUMONIA, SEPSIS OR UROSEPSIS GROUP. THE COLOURED "HALOS" INDICATE THE 95% CONFIDENCE REGIONS OF THE GROUPS. E) ALSO SHOWS A BIPLLOT WITH PC2 ON THE X-AXIS AND PC3 ON THE Y-AXIS, AND IT SHOWS BOTH THE SPREAD OF SAMPLES, INDICATED BY THEIR NUMBERS, AND THE LOADINGS OF THE PRINCIPAL COMPONENTS AS INDICATED BY THE RED ARROWS.	143
FIGURE 36 A) SHOWS THE 2D-PLOT OF THE PLS-DA PERFORMED ON ALL DATA WITH THE GROUPS OF CONTROLS, SURVIVORS AND NON-SURVIVORS. THE PLS-DA REDUCES ALL DATA PROVIDED INTO PRINCIPAL COMPONENTS THAT BEST SEPARATE THE GROUPS BY COMBINING AND WEIGHING THE TRANSFORMED FEATURES, I.E., LIPID MEASUREMENTS. B) SHOWS THE LOADINGS PLOT OF THE PLS-DA, INDICATING THE FEATURES, I.E., LIPIDS, THAT SHOWED THE STRONGEST EFFECT ON THE CLASSIFICATION OF SAMPLES INTO THEIR GROUPS. THE X-AXIS INDICATES THE VIP SCORE, VIP = VARIABLE IMPORTANCE IN	

Table of figures and tables

PROJECTION, I.E., HOW MUCH THE VARIABLE CONTRIBUTES TO THE PROJECTION OF THE SAMPLES IN THE MODEL. THE PANEL NEXT TO IT INDICATES THE AVERAGE CONCENTRATION IN THE GROUP IN RELATION TO THE OTHER GROUPS, WITH BLUE INDICATING A RELATIVELY LOW AVERAGE CONCENTRATION AND RED INDICATING A RELATIVELY HIGH AVERAGE CONCENTRATION. C) SHOWS A TABLE OF THE LIPIDS WITH THE HIGHEST FOLD-CHANGES BETWEEN SURVIVORS AND NON-SURVIVORS AS CALCULATED FROM AVERAGE CONCENTRATION IN THE NON-SURVIVOR POPULATION DIVIDED BY THE AVERAGE CONCENTRATION IN THE SURVIVOR POPULATION. THE P-VALUES ARE OBTAINED FROM WILCOXON TESTS, AND THE RIGHT COLUMN SHOWS THE FALSE DISCOVERY RATE (FDR). D) SHOWS THE RESULT FROM THE CROSS-VALIDATION OF THE MODEL WHICH DETERMINES WHICH NUMBER OF COMPONENTS BEST PROJECTS THE MODEL. ACCURACY, R2 AND Q2 EXAMINE THE MODEL'S PERFORMANCE. HIGHER VALUES INDICATE BETTER PERFORMANCE WITH 1 BEING THE MAXIMUM.....	148
149	
FIGURE 37 A) SHOWS THE TABLE OF AUC VALUE, SENSITIVITY AND SPECIFICITY OF THE MODEL IN TRAINING AND 10-FOLD CROSS-VALIDATION. B) SHOWS THE ROC OF THE MODEL.	149
TABLE 7 A AND B SHOW SEVERAL INDIRECT INDICATORS OF SEVERITY AND THE VALUES ASSOCIATED WITH THESE FOR EITHER A) DIFFERENT SEPSIS CAUSES OR B) SURVIVORS AND NON-SURVIVORS. AP2 STANDS FOR APACHE II SCORE, A SCORING SYSTEM FOR PREDICTING MORTALITY IN THE HOSPITAL.....	152
FIGURE 38 DOTPLOTS SHOWING LIPID CONCENTRATIONS IN THE SEPARATE SEPSIS CAUSES (A-H) AND (I) THE C10:0 (DECANOIC ACID) CONCENTRATIONS IN SEPSIS CAUSED BY DIFFERENT ORGANISMS AND HEALTHY CONTROLS (HC). THE LINES INDICATE THE MEAN VALUE OF THE GROUP, AND THE Y-AXIS DEPICTS THE CONCENTRATION IN THE PLASMA SAMPLE IN μM . THE STATISTICAL TEST PERFORMED IS A REPEAT KRUSKAL-WALLIS-TEST COMPARING ALL GROUPS USING DUNN'S MULTIPLE TESTING CORRECTION YIELDING AN ADJUSTED P-VALUE WITH SIGNIFICANCE LEVEL $P=0.05$. NOT SIGNIFICANT TEST RESULTS ARE NOT SHOWN. THE SIGNIFICANCE LEVELS DISPLAYED ARE: * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$	154
FIGURE 39 CORRELATION HEATMAP OF ALL LIPIDS MEASURED IN THIS ANALYSIS. THE LIPIDS ARE LOG-NORMALIZED AND CENTRE-SCALED, THE COLOUR INDICATES THE PEARSON'S R CORRELATION VALUE RANGING FROM -1 TO 1.....	158
TABLE 8 OVERVIEW OF THE EC50 (AND K_i) VALUES OBTAINED FOR C10 (DECANOIC ACID) AND C12:0 (DODECANOIC ACID) IN DIFFERENT LIGAND BINDING ASSAYS FOR GPR84. ADAPTED FROM CHEN ET AL. (2020) ¹³¹	162
TABLE 9 THE COEFFICIENT OF VARIATION IN QUALITY CONTROL 1 AND 2. QC1 WAS MEASURED 7 TIMES AND QC2 WAS MEASURED 6 TIMES. C8:0 AND C3:0-OH WERE NOT MEASURED DUE TO TECHNICAL DIFFICULTIES.....	169
FIGURE 40 BOXPLOTS OF SHORT CHAIN- AND MEDIUM CHAIN FATTY ACIDS MEASURED IN WHOLE BLOOD SAMPLES OF A COHORT OF SUSPECTED SEPSIS CASES AND A CONTROL COHORT. FOR C10:0 TWO BOXPLOTS ARE SHOWN (D AND E) IN ORDER TO BE ABLE TO SHOW THE FULL RANGE (E) AND THE LOWER RANGE (D) CONTAINING THE MAJORITY OF MEASUREMENTS. THE STATISTICAL TEST PERFORMED WAS A WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05. THE LOWER HALF OF EACH BOX IN THE BOXPLOT DEPICTS THE 2 ND QUANTILE OF VALUES, WITH THE UPPER HALF DEPICTING THE THIRD QUANTILE. THE EXTENDED LINES DEPICT THE FIRST AND LAST QUANTILE OF VALUES, WITH ANY POINTS REACHING BELOW OR ABOVE THESE LINES BEING STATISTICAL OUTLIERS.	171
FIGURE 41 BOXPLOTS OF LONG CHAIN FATTY ACIDS MEASURED IN WHOLE BLOOD SAMPLES OF A COHORT OF SUSPECTED SEPSIS CASES AND A CONTROL COHORT. FOR C18:2 (D, E) AND C18:3 (F,G) TWO BOXPLOTS ARE SHOWN IN ORDER TO SHOW THE FULL RANGE AND THE LOWER RANGE CONTAINING THE MAJORITY OF VALUES. THE STATISTICAL TEST PERFORMED WAS A	

WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05. THE LOWER HALF OF EACH BOX IN THE BOXPLOT DEPICTS THE 2ND QUARTILE OF VALUES, WITH THE UPPER HALF DEPICTING THE THIRD QUARTILE. THE EXTENDED LINES DEPICT THE FIRST AND LAST QUARTILE OF VALUES, WITH ANY POINTS REACHING BELOW OR ABOVE THESE LINES BEING STATISTICAL OUTLIERS....172

FIGURE 42 BOXPLOTS OF ACYLCARNITINES MEASURED IN WHOLE BLOOD SAMPLES OF A COHORT OF SUSPECTED SEPSIS CASES AND A CONTROL COHORT. THE STATISTICAL TEST PERFORMED WAS A WILCOXON TEST WITH A SIGNIFICANCE LEVEL OF 0.05. THE LOWER HALF OF EACH BOX IN THE BOXPLOT DEPICTS THE 2ND QUARTILE OF VALUES, WITH THE UPPER HALF DEPICTING THE THIRD QUARTILE. THE EXTENDED LINES DEPICT THE FIRST AND LAST QUARTILE OF VALUES, WITH ANY POINTS REACHING BELOW OR ABOVE THESE LINES BEING STATISTICAL OUTLIERS.175

FIGURE 43 VOLCANO PLOT OF LIPIDS ALTERED IN SEPSIS. ALL LABELLED LIPIDS ARE STATISTICALLY SIGNIFICANTLY ALTERED. THE Y-AXIS DEPICTS THE NEGATIVE LOG₁₀ OF THE P-VALUE, HENCE THE HIGHER THE VALUE THE HIGHER THE SIGNIFICANCE. THE X-AXIS DEPICTS THE LOG₂ OF THE FOLD CHANGE CALCULATED BASED ON NOT-NORMALIZED LIPID CONCENTRATIONS. THE DASHED LINES AT LOG₂(FC) 1 INDICATE A 2-FOLD CHANGE ON RAW DATA. SIGNIFICANCE WAS CALCULATED USING THE KRUSKAL-WALLIS TEST ON LOG₁₀-TRANSFORMED LIPID CONCENTRATIONS AND IS ADJUSTED FOR MULTIPLE TESTING. DASHED LINES INDICATE A LOG₂(FC) OF 1 AND A -LOG₁₀ P-VALUE THAT CORRESPONDS TO P=0.1 AFTER MULTIPLE TESTING CORRECTION.178

FIGURE 44 A) HEATMAP OF ALL SAMPLES AND LIPIDS. PRIOR TO ANALYSES THE LIPIDS WERE LOG-NORMALIZED AND CENTRE-SCALED. BOTH LIPIDS AND SAMPLES ARE CLUSTERED USING D. WARD CLUSTERING, DISTANCES ARE EUCLIDEAN. THE COLOUR SCALE IN THE HEATMAP GOES FROM 4 (INCREASED) TO -4 (DECREASED) AND THE COLOUR INDICATES THE Z-SCORE OF THE LIPID MEASUREMENT WITHIN THE ROW FOR EACH SAMPLE. B) 3D PCA PLOT OF ALL SAMPLES. C) 2D PCA PLOT WITH RED ARROWS INDICATING THE LOADINGS. INSTEAD OF DOTS SAMPLE NAMES ARE SHOWN. D, E, F SHOW THE 2D PCA PLOTS WITH RESPECTIVELY PC1 AND PC2, PC2 AND PC3 AND PC1 AND PC3 AS AXES. THE COLOURED "HALOS" INDICATE THE 95% CONFIDENCE REGIONS OF THE GROUPS.179

FIGURE 45 HEATMAP OF CORRELATIONS BETWEEN LIPIDS MEASURED IN THIS COHORT. LIPIDS WERE LOG-NORMALIZED AND CENTRE-SCALED BEFORE CORRELATIONS WERE COMPUTED. THE CORRELATIONS WERE CALCULATED USING PEARSON'S CORRELATION, THE COLOUR INDICATES THE R-VALUE AND RANGES FROM -1 TO +1.....182

FIGURE 46 A FLOW DIAGRAM OF THE ANALYSIS STEPS PERFORMED IN THIS CHAPTER AND THE RESULTING INTERPRETATION OF GPR84'S TRANSCRIPTIONAL REGULATION THEY CONTRIBUTED TO.189

FIGURE 47 VOLCANO PLOT OF TRANSCRIPTION FACTOR EXPRESSION IN A NEONATAL SEPSIS DATASET SHOWING THE -LOG₁₀ OF THE SIGNIFICANCE ON THE Y-AXIS AND THE LOG₂ FOLD CHANGE COMPARING SEPSIS CASES TO HEALTHY CONTROLS. THE DASHED LINES INDICATE A LOG₂ FOLD CHANGE OF 1 AND A P-VALUE OF 0.05 AS -LOG₁₀ VALUE, RESPECTIVELY.....192

FIGURE 48 VOLCANO PLOT OF TRANSCRIPTION FACTOR EXPRESSION IN A DATASET OF PAEDIATRIC SEPSIS SHOWING THE -LOG₁₀ OF THE SIGNIFICANCE ON THE Y-AXIS AND THE LOG₂ FOLD CHANGE COMPARING SEPSIS CASES TO HEALTHY CONTROLS. THE DASHED LINES INDICATE A LOG₂ FOLD CHANGE OF 1 AND A P-VALUE OF 0.05 AS -LOG₁₀ VALUE, RESPECTIVELY.192

TABLE 10 THIS TABLE SHOWS THE SIGNIFICANTLY UP- AND DOWNREGULATED TRANSCRIPTION FACTOR GENES IDENTIFIED IN THE PAEDIATRIC SEPSIS TRANSCRIPTOMICS DATA AND THE NEONATAL SEPSIS TRANSCRIPTOMICS DATA.193

FIGURE 49 DEPICTION OF TRANSCRIPTION FACTORS AND THEIR ASSOCIATION WITH TISSUES (A) AND GO PATHWAYS (B) BASED ON GTE:X - TF DATA. FIGURE GENERATED BY CHEA3 BASED ON TFs PREDICTED TO REGULATE UPREGULATED GENES OF A TRANSCRIPTOMICS DATASET OF PAEDIATRIC SEPSIS CASES. EACH DOT REPRESENTS A TF, AND THE COLOUR OF THE DOT

Table of figures and tables

INDICATES THE TISSUE (A) OR GO PATHWAY (B). THE TOP 50 RANKING TFs ENRICHED IN SEPSIS ARE SURROUNDED BY A GREY CIRCLE.....	198
FIGURE 50 VENN DIAGRAM OF CHIP HITS FOUND IN ENCODE AND THE SIGNALLING PATHWAYS PROJECT (SPP) DATABASES USING ONLINE TOOLS HARMONIZOME AND OMINER. VENN CREATED USING BioVENN ²⁰⁶	200
203	
FIGURE 51 VENN DIAGRAM SHOWING THE OVERLAP BETWEEN THE RESULTS FROM TWO ANALYSIS APPROACHES.	203
TABLE 12 HIGH CONFIDENCE AND MEDIUM CONFIDENCE TFs DRIVING GPR84 EXPRESSION IN SEPSIS	203
FIGURE 52 STRING NETWORK OF TRANSCRIPTION FACTORS PREDICTED TO REGULATE GPR84 IN SEPSIS WITH HIGH AND MEDIUM CONFIDENCE. THE EDGES INDICATE THE EVIDENCE FOR INTERACTION, CONFIDENCE OF INTERACTION IS MINIMUM 0.6. TURQUOISE EDGES INDICATE INTERACTIONS FROM CURATED DATABASES, PURPLE INDICATES INTERACTIONS THAT WERE EXPERIMENTALLY DETERMINED. THE NODES WERE CLUSTERED USING MCL CLUSTERING WITH AN INFLATION PARAMETER OF 3. THE EDGES INDICATE THAT PROTEINS HAVE A SPECIFIC JOINT SHARED FUNCTION, INDEPENDENT OF IF THEY PHYSICALLY INTERACT.	205
FIGURE 53 STRING NETWORK OF PHYSICAL ASSOCIATIONS BETWEEN HIGH CONFIDENCE AND MEDIUM CONFIDENCE TFs. EDGES INDICATE CONFIDENCE, MINIMUM CONFIDENCE BEING 0.6. EDGES STAND FOR PHYSICAL INTERACTION BETWEEN THE PROTEINS. NODES WERE CLUSTERED USING MCL CLUSTERING WITH INFLATION PARAMETER OF 3.....	206
FIGURE 54 ANNOTATION OF THE GENOMIC REGION AROUND GPR84 FROM UCSC GENOME BROWSER ²¹¹ . TRACKS RELATING TO GENE EXPRESSION REGULATION WERE CHOSEN INCLUDING PREDICTED PROMOTER SITES AND ENHANCERS FROM GENEHANCER, ENCODE AND OREGANNO.	209
FIGURE 55 THE GENOMIC REGION UPSTREAM OF GPR84 VISUALIZED IN THE UCSC GENOME BROWSER WITH TRACKS RELATING TO TRANSCRIPTIONAL REGULATION INCLUDING DNase I HYPERSENSITIVITY, SEQUENCE CONSERVATION, CpG ISLANDS AND TF CHIPSEQ CLUSTERS.	210
TABLE 13 SHOWS THE TFBS FOUND IN THE HUMAN GPR84 PROMOTER, THE TFBS FOUND IN THE MAJORITY OF GENES CORRELATING WITH GPR84 IN SEPSIS, AND THE TFBS FOUND IN THE PROMOTER SEQUENCE THAT IS CONSERVED IN HUMANS AND MICE.	213
FIGURE 56 OVERLAP IN TFBS BETWEEN THE HUMAN GPR84 PROMOTER SEQUENCE (B), A CONSERVED SUBPART OF THE HUMAN GPR84 PROMOTER (C), AND TFBS FOUND IN AT LEAST 75% OF GENES CORRELATING WITH GPR84 (A).....	215
FIGURE 57 CELL TYPES PREDICTED TO EXPRESS Gpr84 BASED ON THREE OPEN CHROMATIN REGIONS (Gpr84_1, Gpr84_2, Gpr84_3) ASSOCIATED WITH Gpr84 EXPRESSION IN MICE, AS ESTABLISHED BY THE ENHANCERCONTROL TOOL BASED ON DATA FROM YOSHIDA ET AL. ¹³⁶ , LEGEND ADDED IN INKSCAPE.	218
FIGURE 58 GPR84 EXPRESSION IN A) THP-1 CELLS, B) NEUTROPHILS AND C) PBMCs TREATED WITH EITHER DEXAMETHASONE (DEX), LPS AND DEXAMETHASONE (LD), LPS OR UNTREATED (UNTR) AS FOLD CHANGE TO TIME-POINT 0 WITH THE TIME IN HOURS ON THE X-AXIS. DATA FROM THP-1 CELLS WAS OBTAINED IN 3 SEPARATE EXPERIMENTAL REPEATS (N=3), WHEREAS NEUTROPHIL AND PBMC DATA ORIGINATED FROM THE CELLS FROM 4 HEALTHY VOLUNTEERS(N=4).	222
FIGURE 59 OLAH EXPRESSION IN NEONATAL SEPSIS (A) AND PAEDIATRIC SEPSIS (B). THE GENE EXPRESSION WAS MEASURED BY MICROARRAYS, DATA IS RMA-NORMALIZED. SIGNIFICANCE WAS CALCULATED USING THE KRUSKAL-WALLIS TEST.	238
FIGURE 60 THE GENOMIC REGION AROUND OLAH AS SHOWN IN THE UCSC GENOME BROWSER. ON THE LEFT THE NAMES OF THE DISPLAYED TRACKS ARE SHOWN. EVERY TRACK INDICATES A DIFFERENT GENOMIC FEATURE. THE REGION CODING FOR THE	

OLAH GENE IS INDICATED IN THE TOP TRACK CALLED GENCODE V41. THE ARROWS INDICATE THE READING DIRECTION OF THE GENE. EPDNEW V006 IS A DATABASE FOR EXPERIMENTALLY VALIDATED PROMOTERS BASED ON TRANSCRIPTION START SITE MAPPING FROM EXPERIMENTAL HIGH-THROUGHPUT DATA. RAMPAGE INDICATES THE TRANSCRIPTION START SITE BASED ON CDNA SEQUENCING DATA. FANTOM5 ALSO PREDICTS, IN THIS CASE, THE TRANSCRIPTION START SITE.241

TABLE 14 EXPRESSION OF THE FIRST ENHANCERS FOUND TO BE LIKELY REGULATING OLAH ACCORDING TO FANTOM5. THE TABLE SHOWS THE EXPRESSION OF THE ENHANCERS IN DIFFERENT CELL TYPES, BOTH AS PERCENTAGE OF TOTAL EXPRESSION OF THE ENHANCERS AND AS TAGS PER MILLION.....242

TABLE 15 EXPRESSION OF THE SECOND ENHANCER FOUND TO BE LIKELY REGULATING OLAH ACCORDING TO FANTOM5. THE TABLE SHOWS THE EXPRESSION OF THE ENHANCERS IN DIFFERENT CELL TYPES, BOTH AS PERCENTAGE OF TOTAL EXPRESSION OF THE ENHANCERS AND AS TAGS PER MILLION.....242

FIGURE 61 TF NETWORK LIKELY REGULATING OLAH EXPRESSION IN SEPSIS, GENERATED USING STRING. THE CONFIDENCE OF INTERACTION BETWEEN PROTEINS WAS SET TO A MINIMUM OF 0.6, THE COLOURS OF THE CONNECTING LINES INDICATE THE TYPE OF EVIDENCE FOR INTERACTION. TURQUOISE INDICATES THE INTERACTION STEMS FROM CURATED DATABASES; PURPLE CONNECTIONS INDICATE EXPERIMENTALLY DETERMINED INTERACTIONS. THE INTERACTION INDICATES A SHARED FUNCTION RATHER THAN A PHYSICAL INTERACTION. TRANSCRIPTION FACTORS WERE CLUSTERED USING A MCL INFLATION PARAMETER OF 3.246

TABLE 16 SHOWS THE 25 (EXPERIMENTAL) CONDITIONS IN WHICH OLAH EXPRESSION IS THE MOST AFFECTED. THE DATA IS OBTAINED FROM THE EBI GENE EXPRESSION ATLAS, ACCESSED 24/10/2022.247

FIGURE 62 THE OLAH EXPRESSION IN HEALTHY CONTROLS (CTRL), CORTICOSTEROID TREATED SEPSIS PATIENTS (YES) AND PATIENTS NOT TREATED WITH CORTICOSTEROID TREATMENTS (NO). FOR A SUBSET OF PATIENTS, IT WAS UNKNOWN IF THEY WERE TREATED WITH CORTICOSTEROIDS (UNKNOWN). THE GENE EXPRESSION IS RMA-NORMALIZED AND LOG-TRANSFORMED. THE LOWER HALF OF EACH BOX IN THE BOXPLOT DEPICTS THE 2ND QUANTILE OF VALUES, WITH THE UPPER HALF DEPICTING THE THIRD QUANTILE. THE EXTENDED LINES DEPICT THE FIRST AND LAST QUANTILE OF VALUES, WITH ANY POINTS REACHING BELOW OR ABOVE THESE LINES BEING STATISTICAL OUTLIERS. THE STATISTICAL TEST PERFORMED WAS A KRUSKAL-WALLIS-TEST.247

FIGURE 63 THE EXPRESSION OF OLAH IN PBMCs TREATED WITH DIFFERENT CYTOKINES. THE DATA WAS OBTAINED FROM A PUBLIC REPOSITORY AND STEMS FROM AN EXPERIMENT ON BLOOD DERIVED PBMCs FROM HEALTHY VOLUNTEERS²⁴⁴. THE ERROR BARS REPRESENT THE STANDARD DEVIATION. APART FROM TIME-POINT 0, THERE WAS ONLY ONE REPLICATE PER TIMEPOINT AND TREATMENT. GENE EXPRESSION IS RMA-NORMALIZED AND LOG-TRANSFORMED.....248

FIGURE 64 THE EXPRESSION OF OLAH IN DIFFERENT CELL TYPES (B-CELL, CD4+ T-CELL, CD8+ T-CELL, MONOCYTE, NK CELL) IN A TIME COURSE WITH EITHER MOCK TREATMENT OR 0.6 pM IFN γ . THE CELLS WERE OBTAINED FROM HEALTHY VOLUNTEER BLOOD²⁴³. THE ERROR BARS REPRESENT THE STANDARD DEVIATION. APART FROM TIME-POINT 0, THERE WAS ONLY ONE REPLICATE PER TIMEPOINT AND TREATMENT. GENE EXPRESSION IS RMA-NORMALIZED AND LOG-TRANSFORMED.....249

FIGURE 65 THE EXPRESSION OF OLAH IN DIFFERENT CELL TYPES AT 0h, 24h AND 96h AFTER TREATMENT WITH DIFFERENT ACTIVATING COMPOUNDS²⁴⁴. CELLS WERE OBTAINED FROM HEALTHY VOLUNTEER BLOOD. THE ERROR BARS REPRESENT THE STANDARD DEVIATION. GENE EXPRESSION IS RMA-NORMALIZED AND LOG-TRANSFORMED.250

FIGURE 66 OLAH EXPRESSION IN THP-1 CELLS AFTER TREATMENT WITH DIFFERENT STIMULI: DEXAMETHASONE (Dex), LPS AND DEXAMETHASONE (LD), LPS AND NO STIMULI/UNTREATED (UNTR). DEPICTED IS THE TIME COURSE OF EXPRESSION WITH

Table of figures and tables

TIME IN HOURS ON THE X-AXIS AND THE FOLD CHANGE COMPARED TO THE BASELINE EXPRESSION AT T=0 ON THE Y-AXIS. THE EXPRESSION DATA WAS COLLECTED IN THREE SEPARATE EXPERIMENTAL REPEATS. THE ERROR BARS DEPICT THE SEM.	252
FIGURE 67 THE OLAH EXPRESSION IN ENDOTHELIAL CELLS SHOWS AS THE FOLD CHANGE COMPARED TO THE EXPRESSION AT TIMEPOINT 0. THE X-AXIS DEPICTS THE TIME IN HOURS. THIS DATA COMES FROM 2 EXPERIMENTAL REPEATS.	255
FIGURE 68 EXPRESSION OF OLAH AS FOLD CHANGE COMPARED TO T=0 IN A) NEUTROPHILS AND B) PBMCs EXTRACTED FROM HEALTHY VOLUNTEER BLOOD (N=5) IN A TIME COURSE OF A) 6 HOURS AND B) 24 HOURS. MCFA WERE EXTRACTED FROM SUPERNATANT SAMPLES TAKEN SIMULTANEOUSLY AND QUANTIFIED BY PREVIOUS ESTABLISHED METHOD YIELDING CONCENTRATION IN μM AS RESULT. RESULTS SHOWN SHOW C) MCFA IN PBMCs AND D) MCFA IN NEUTROPHIL TIME COURSE EXPERIMENTS. ERROR BARS DEPICT SEM IN ALL PLOTS. DEX = DEXAMETHASONE, LD = LPS + DEXAMETHASONE, UNTR = UNTREATED.	258
FIGURE 69 ACADM EXPRESSION IN A) THP-1 CELLS (N=3), B) NEUTROPHILS (N=2) AND C) PBMCs (N=3). THE TIME IS DEPICTED IN HOURS ON THE X-AXIS, AND THE Y-AXIS SHOWS THE FOLD CHANGE OF ACADM EXPRESSION TO T=0. THE ERROR BARS DEPICT THE SEM.	261
FIGURE 70 GENES INVOLVED IN LIPID METABOLISM SIGNIFICANTLY UP- OR DOWNREGULATED IN SEPSIS. THE Y-AXIS SHOWS THE LOG ₂ FOLD CHANGE COMPARED TO HEALTHY CONTROL NEONATES.	265
TABLE 17 SHOWS THE MAIN FINDINGS OF MY THESIS.	282
FIGURE 71 OVERVIEW PICTURE OF AN IMMUNE CELL DEPICTING SEVERAL OF MY MAIN RESEARCH FINDINGS FROM THE THESIS. SPEECH BUBBLES WITH ORANGE BACKGROUND INDICATE BIOLOGIC PROCESSES, GREEN OUTLINED, ROUNDED FIGURES CONTAIN GENES, BLUE LINED SQUARES INDICATE LIPIDS.	289
FIGURE 72 SECOND OVERVIEW PICTURE WITH A MORE IN-DEPTH VIEW OF PROCESSES WITHIN THE CELL THAT MIGHT LEAD TO THE OBSERVED LIPID CONCENTRATIONS. BLUE ARROWS INDICATE POTENTIAL FLUXES OF LIPIDS WITHIN THE CELL. TWO CIRCLE ARROWS INDICATE A SYNTHESIS OR CATABOLIC CYCLE. LIPID MEDIATORS AND OTHER ARROWS ARE MARKED WITH DASHED LINES TO INDICATE A LEVEL OF UNCERTAINTY, AS THE EXPORT MECHANISMS FOR THESE LIPIDS ARE NOT QUITE SURE, AND THE LIPID BIOSYNTHESIS IN ENDOPLASMATIC RETICULUM CAN LEAD TO MEMBRANE SYNTHESIS AND LIPID MEDIATORS. IN THIS CONTEXT LIPID MEDIATORS APPEAR MORE LIKELY.	290
SUPPLEMENTARY TABLE 1 DEPICTS THE CHEA3 TF ENRICHMENT HITS FOR PAEDIATRIC SEPSIS AND NEONATAL SEPSIS FILTERED FOR HITS WITH A HIT FOR GPR84 IN ONE OF THE DATABASES. THE RIGHT COLUMN INDICATES THE OVERLAP IN TF HITS BETWEEN THE TWO DATASETS.	314
SUPPLEMENTARY TABLE 2 TFs BINDING TO THE GPR84 REGULATORY REGION ACCORDING TO CHIP HITS FROM THE ENCODE AND SIGNALLING PATHWAYS PROJECT (SPP) DATABASES.	316
SUPPLEMENTARY TABLE 3 LIST OF ALL TFs IDENTIFIED IN AT LEAST ONE ANALYSIS AND COLUMNS INDICATING IF A TF WAS FOUND IN THIS ANALYSIS OR NOT.	321
SUPPLEMENTARY TABLE 4" KNOWLEDGE" ON GPR84 EXPRESSION ACCORDING TO ENRICH.	331
SUPPLEMENTAL FIGURE 1 GPR84 EXPRESSION IN BMDM INFECTED WITH MCMV OR B) TREATED WITH IFN- γ . THE Y-AXIS DEPICTS THE LOG ₂ FOLD CHANGE IN EXPRESSION RELATIVE TO EXPRESSION AT T=0, THE X-AXIS THE HOURS AFTER INFECTION OF THE CELLS WITH THE VIRUS (N=1) OR STIMULATED BY IFN- γ	334

Acknowledgements

First of all, I would like to thank my supervisors for their support during my PhD. Peter, thank you for always having an open door, for your enthusiasm and for your insightful guidance. Val and Matthias, thank you for your feedback, your honest advice and your help throughout my PhD.

I would also like to thank the funders for making this PhD possible: Cardiff University and the Welsh government together with the EU that provided the Sêr Cymru grant.

I would also like to express a heartfelt thanks to everyone in the Project Sepsis team for all the help I received and the laughs we shared. For the helping hand, the open ear, the coffee break when needed. I very much enjoyed learning not just about receptors and metabolism, but also about chickens, U.K. politics, the intricacies of a British hen-do, live in the U.K., British food, and many more subjects of cultural relevancy. A special shout-out to Dan White for all the patient help and explanations with the LC-MS and your guidance on the method development. Without you this method would not have been established. In the cell culture any experiments would not have been possible without the help of first Widad (and Kostas) and later Luke. Thank you for all the explanations, discussions, and the helping hand or supervision when needed.

I have been very lucky and grateful to make great friends during my stay in Cardiff, who made this time something I look back at with fond memories and a smile on my face. I would like to thank Ruth and Patricia for all the running, sewing and walks in the park. Furthermore, Oliwia, Agata and Mikolaj, for including me in your friend group and all the fun game sessions we had! These always provided a much-needed refuge. Thank you Natasa and Agata for consecutively being great flatmates. Thank you, Ross, for the advice on anything code/statistics related and the great “discussions” about PhD life. Mariama, ich hoffe wir sehen uns mal wieder für einen Brunch oder einen Cocktail!

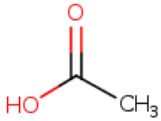
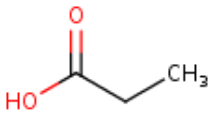
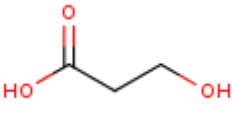
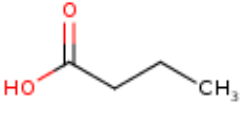
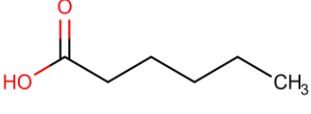
Finally, I had a great “safety net” – friendships from before and, of course and foremost, my parents. Thank you for always being there and having an open ear, even if it usually was on the other side of a computer or phone screen. Gio, Emelie, Sigrun – you’re the best.

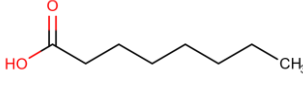
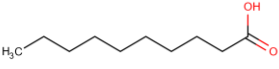
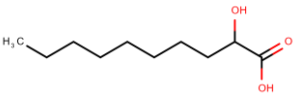
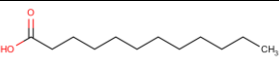
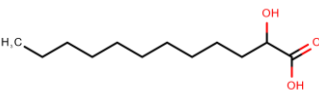
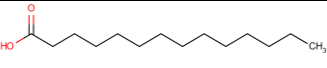
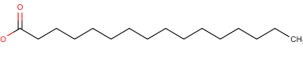
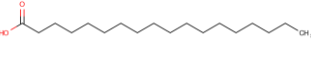
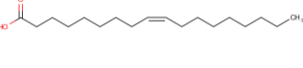
Tack Jonas.

Overview Lipids

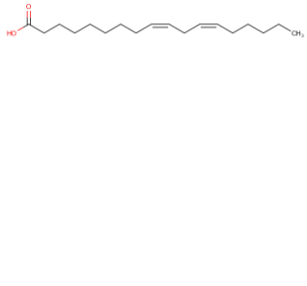
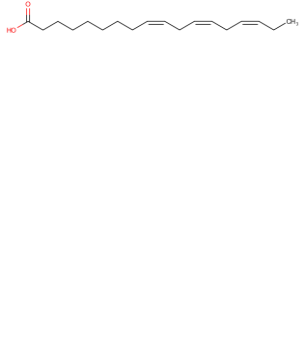
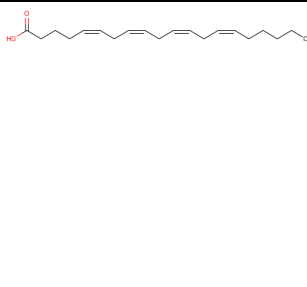


Note: I generated this overview of the lipids as a hopefully helpful “cheat sheet” to anyone reading my thesis. Names in **bold** are the names that will be used throughout my thesis in combination with the abbreviation provided in the first column. For lipids where the abbreviation could indicate several lipids, only the name for the lipid actually measured in my thesis and the associated structure are shown (e.g., C18:3 can be γ -linolenic acid or α -linolenic acid, but only α -linolenic acid is shown as I did not measure γ -linolenic acid). Molecular structures obtained from Chemspider (A. Chemist, *ChemSpider SyntheticPages*, 2001, <http://cssp.chemspider.com/123>)

Table 1 Overview of lipids measured in the thesis

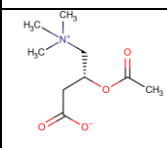
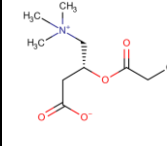
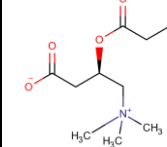
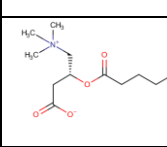
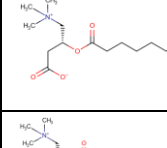
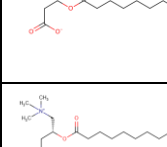
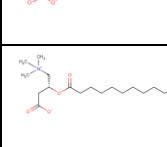
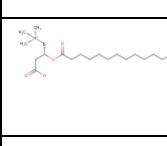
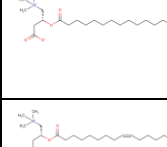
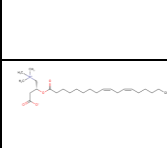
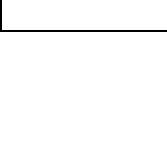

Abbreviation	Traditional/alternative name	IUPAC name	Molecular structure	Chemical formula	Comment
C2:0	Ethanoic acid, Vinegar acid	Acetic acid		C ₂ H ₄ O ₂	SCFA, saturated
C3:0	Propionic acid	propanoic acid		C ₃ H ₆ O ₂	SCFA, saturated
C3-OH	Lactic acid	2-hydroxypropanoic acid		C ₃ H ₆ O ₃	SCFA, saturated
C4:0	Butyric acid	Butanoic acid		C ₄ H ₈ O ₂	SCFA, saturated
C6:0	Caproic acid	Hexanoic acid		C ₆ H ₁₂ O ₂	MCFA, saturated

C8:0	Caprylic acid	Octanoic acid		C ₈ H ₁₆ O ₂	MCFA, saturated
C10:0	Capric acid	Decanoic acid		C ₁₀ H ₂₀ O ₂	MCFA, saturated
C10-2OH	2-hydroxy-capric acid	2-hydroxy-decanoic acid		C ₁₀ H ₂₀ O ₃	MCFA, saturated, hydroxylated
C12:0	Lauric acid	Dodecanoic acid		C ₁₂ H ₂₄ O ₂	MCFA, saturated
C12-2OH	2-hydroxy-lauric acid	2-hydroxy-dodecanoic acid		C ₁₂ H ₂₄ O ₃	MCFA, saturated, hydroxylated
C14:0	Myristic acid	Tetradecanoic acid		C ₁₄ H ₂₈ O ₂	LCFA, saturated
C16:0	Palmitic acid	Hexadecanoic acid		C ₁₆ H ₃₂ O ₂	LCFA, saturated
C18:0	Stearic acid	Octadecanoic acid		C ₁₈ H ₃₆ O ₂	LCFA, saturated
C18:1	Oleic acid	(Z)-octadec-9-enoic acid		C ₁₈ H ₃₄ O ₂	PUFA, omega-9

Overview Lipids

C18:2	Linoleic acid	(9Z,12Z)-octadeca-9,12-dienoic acid		C ₁₈ H ₃₂ O ₂	PUFA, omega-6
C18:3	α-linolenic acid	(9Z,12Z,15Z)-octadeca-9,12,15-trienoic acid		C ₁₈ H ₃₀ O ₂	PUFA, omega-3
C20:4	Arachidonic acid	(5Z,8Z,11Z,14Z)-icosa-5,8,11,14-tetraenoic acid		C ₂₀ H ₃₂ O ₂	PUFA, omega-6
C20:5	Eicosapentaenoic acid	(5Z,8Z,11Z,14Z,17Z)-icosa-5,8,11,14,17-pentaenoic acid		C ₂₀ H ₃₀ O ₂	PUFA, omega-3
C22:6	Docosahexaenoic acid	(4Z,7Z,10Z,13Z,16Z,19Z)-docosa-4,7,10,13,16,19-hexaenoic acid		C ₂₂ H ₃₂ O ₂	PUFA, omega-3

Overview Lipids

Abbreviation	Name	Alternative name	Molecular structure	Chemical formula
C2:0-carnitine	Acetylcarnitine			$C_9H_{17}NO_4$
C3:0-carnitine	Propionylcarnitine	Propanoylcarnitine,		$C_{10}H_{19}NO_4$
C4:0-carnitine	Butyrylcarnitine	Butanoylcarnitine		$C_{11}H_{21}NO_4$
C6:0-carnitine	Hexanoylcarnitine	Caproylcarnitine		$C_{13}H_{26}NO_4$
C8:0-carnitine	Octanoylcarnitine			$C_{15}H_{29}NO_4$
C10:0-carnitine	Decanoylcarnitine			$C_{17}H_{33}NO_4$
C12:0-carnitine	Dodecanoylcarnitine	Lauroylcarnitine		$C_{19}H_{38}NO_4$
C14:0-carnitine	Myristoylcarnitine	Tetradecanoylcarnitine		$C_{21}H_{42}NO_4$
C16:0-carnitine	Palmitoylcarnitine	Hexadecanoylcarnitine		$C_{23}H_{46}NO_4$
C18:0-carnitine	Stearoylcarnitine	Octadecanoylcarnitine		$C_{25}H_{50}NO_4$
C18:1-carnitine	Oleoylcarnitine	(9Z)-Octadecenoyl-L-carnitine		$C_{25}H_{48}NO_4$
C18:2-carnitine	Linoleoylcarnitine	Octadecadienyl-L-carnitine		$C_{25}H_{46}NO_4$

Abbreviation list

3NPH	3-nitrophenylhydrazine
"-OH"	hydroxgroup
3T3-L1	murine adipocytic cell line
6-OAU	6-n-octylaminouracil
AC	adenylylcyclases
acetyl-CoA	acetyl-Coenzyme A
adj.	adjusted
APACHE II score	clinical score for mortality
ATAC-seq	assay for transposase-accessible chromatin using sequencing
ATP	adenosine-triphosphate
AUC	area under curve
BMDM	bone marrow derived macrophages
C	Carbon
Ca ²⁺	Calcium
cAMP	cyclic adenosine-monophosphate
CCCL	(C-C)motif ligand
CHO	Chinese Hamster Ovary cells
COX-2	cyclic oxidase 2
CREB	cAMP Response-Element Binding protein
CRP	C-reactive protein
CXCL	(C-X-C motif) ligand
DAG	diacylglycerol
DHA	docosahexaenoic acid
DIM	diindolylmethane
DMSO	dimethylsulfoxide
DP	declustering potential
DSS	dextran sodium sulphate
EC50	half maximal effective concentration
ECL	extracellular loop

Abbreviation list

	<i>N</i> -Ethyl- <i>N'</i> -(3-dimethylaminopropyl)carbodiimide
EDC	hydrochloride
ELISA	enzyme-linked immunoassay
EPAC	exchange protein directly activated by cAMP
EST	expressed sequence tags
FC	Fold change
FCS	foetal calf serum
FFAR	free fatty acid receptor
GDP	guanosine diphosphate
GPCR	G-protein coupled receptors
GR	glucocorticoid receptor
GTP	guanosine triphosphate
H ⁺	hydrogen ion
HEK293	human embryonic kidney cell line
HE-NECA	2-Hexynyl-adenosine-5'-N-ethyluronamide
HLA-DR	human leucocyte antigen DR subtype
HMDM	human monocyte derived macrophages
HUVEC	human umbilical vein endothelial cells
ICL	intracellular loop
ICU	intensive care unit
IFN	interferon
IL	interleukin
ILTIS	the Innate-like T-cells in Sepsis
IP3	inositol triphosphate
IV	intravascular
JNK	Jun N-terminal kinase
KD	knock-down
kDa	kiloDalton
KO	knock-out
LC-MS/MS	liquid-chromatography tandem mass spectrometry
LCFA	long chain fatty acids

LC-MS	liquid-chromatography mass spectrometry
LPS	lipopolysaccharide
MAPK	Mitogen Associated Protein Kinase
MCFA	medium chain fatty acid
MCMV	murine cytomegalovirus
MeOH	methanol
mRNA	messenger RNA
NAFLD	non-alcoholic fatty liver disease
NAS2	nitric oxidase 2
NET	Neutrophil extracellular trap
NGAL	neutrophil gelatinase associated lipocalin
NICU	neonatal intensive care unit
NO	nitrous oxide
nSep	neonatal sepsis (study)
OCR	open chromatin region
OUA	of unknown aetiology
PAMP	pathogen associated molecular patterns
PBMC	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PCA	principal component analysis
PGE2	prostaglandin e2
PI3K/Akt	phosphatidylinositol 3-kinase
PIP2	phosphatidylinositol4,5-bisphosphate
PIP2	phosphatidylinositol4,5-bisphosphate
PKA	phosphokinase A
PLS-DA	Partial least squares-discriminant analysis
PMN	polymorphonuclear cells
Ptges2	prostaglandin synthase E2

Abbreviation list

PUFA	polyunsaturated fatty acid
QC	quality control
qPCR	quantitative polymerase chain reaction
qSOFA	quick SOFA
ROS	reactive oxygen species
RPM	
RPMI	Roswell Park Memorial Institute
RQ	respiratory quotient
RTK	receptor tyrosine kinase
SCFA	short chain fatty acid
SIRS	systemic inflammatory response syndrome
SOFA score	sequential organ failure assessment
TAG	triacylglycerol
TCA	tricarboxylic acid cycle
TE	thioesterase
TF	transcription factor
TFBS	Transcription Factor binding site(s)
THP-1	human monocytic cell line
TLR	toll-like receptor
TM	transmembrane
TNF- α	tumour necrosis factor alpha
Treg	regulatory T-cell
TSS	transcription start site
WT	wild type
yrs	years

1 Introduction

1.1 Sepsis

Sepsis is a life-threatening health condition that has been around for centuries but whose pathophysiology we are only starting to understand in the past decades. The word sepsis is thought to originate from Hippocrates (460–370 BC) and was used by him to describe a process that would let tissues rot, generate foul odours in swamps and lead to wounds festering¹. Since then, our understanding of infections and the immune system has grown exponentially; not only are we aware of pathogenic microorganisms, e.g. bacteria, viruses and fungi, as the cause of infections, but also of the immune system as driver of inflammatory processes. The inflammatory reaction, despite commonly being uncomfortable and potentially dangerous to the host, is understood to be a vital part of our response to pathogenic microorganisms in the body. The World Health Organization (WHO) describes sepsis in a recent report as “[...]the clinical deterioration of common and preventable infections such as those of the respiratory, gastrointestinal and urinary tract, or of wounds and skin”². The official definition from 2016 states that sepsis is a “life-threatening organ dysfunction caused by a dysregulated host response to infection”³. A key feature of this is that an invading microorganism can trigger an immune response that adversely affects several bodily systems and is potentially fatal. To the patient this might initially feel like a strong flu, gastroenteritis or chest infection, including symptoms like muscle pain, shivering, confusion, lethargy and breathlessness⁴. As sepsis progresses, patients can lose consciousness, become delirious and require extensive medical attention including medication, mechanical ventilation and enteral feeding⁵. The disease course is heterogeneous; some patients die within hours after onset of symptoms, while others remain in a critical condition for weeks before resolution or eventual death. Sepsis survivors often suffer from physical weakness due to loss of muscle, cognitive decline and fatigue for a prolonged time. Especially elderly sepsis survivors suffer from a loss in physical performance without recovery to baseline functioning, while younger survivors can recover to their pre-sepsis physical performance⁶. In addition, sepsis survivors have an increased mortality in the follow-up period to their septic episode, assessed in two studies to be 21% and 17.6% in the 1st year, due to infections, cardiovascular disease and cancer^{7,8}.

1.1.1 Incidence

Sepsis occurs in low-, middle- and high-income countries; it can affect all age groups and is a danger to both healthy and pre-disposed persons. Reports of its incidence vary and are scarce for low- and middle-income countries despite these carrying the main burden; a meta-analysis by Fleischmann et al. (2016) however estimates the global annual burden to be 50.9 million sepsis cases, with an estimated death toll of 5.3 million deaths⁹. This estimation is based on data from high-income countries and hence thought to be probably underestimating the real burden of sepsis. The Global Burden of Disease Consortium analysed sepsis incidence and disease using a model based on hospital data and death certificates leading to an estimation of 48.9 million cases of sepsis in 2017, among which 20.3 million occurred in children below the age of 5, leading to death in 11 million cases in the overall population and 2.9 million in children below the age of 5. This means that sepsis was the cause of death in 19.7% of deaths in 2017¹⁰. Among the vulnerable populations, highlighted in this report and a report by the WHO are neonates, infants, pregnant women and elderly people^{2,11}. Sepsis is the second most common cause of deaths in neonates globally¹². In the UK, sepsis incidence is estimated at ~245 000 cases in 2017 with related deaths at 47 860¹⁰.

1.1.2 Sepsis definitions and diagnosis

The high mortality of sepsis has been attributed in part to the often-late diagnosis of the condition. There is no fast and unequivocal diagnosis method available, yet time is of considerable importance in treating sepsis effectively^{13,3}. An explanation for this lack is the complexity and heterogeneity of sepsis, which we are only starting to understand. This is also reflected in the evolving definitions of sepsis that are directly linked to its diagnosis. Its first official definition was given only in 1991; this was also the time when our understanding of sepsis changed from being caused by a toxin in the blood to an infection with detrimental immune response. Initially sepsis was described as a condition with four degradations: systemic inflammatory response syndrome (SIRS), sepsis, severe sepsis and septic shock. Each degradation was characterized by clinical manifestations: presence of clinical symptoms including increased heart rate, changes in body temperature and blood cell count (SIRS), SIRS and proven infection (sepsis), sepsis and organ dysfunction (severe sepsis), sepsis and persistent hypotension despite fluid resuscitation (septic shock). These degradations were kept in the 2001 consensus definition while signs and symptoms for

1 Introduction

diagnosis were added. The 2013 consensus definition decreased the number of degradations to sepsis and septic shock and provided for both a general definition and clinical criteria. The organ dysfunction described to define sepsis would be diagnosed using the sequential organ failure assessment (SOFA) score, a scoring system already used in critical care. The SOFA score consists of a point system in which a score between 1 and 4 is given to the severity of dysfunction of each of six components: respiratory, cardiovascular, hepatic, coagulation, renal and neurological systems (Table 2). An acute change in score of at least two points is required for the diagnosis of sepsis. This point system requires some laboratory measurements, which can slow down diagnosis. To alleviate this, the alternative scoring system quick SOFA (qSOFA) allows for quick diagnosis at the bedside based on high respiratory rate, an altered mental state and low blood pressure. Septic shock was defined as a subset of sepsis presenting with circulatory, cellular and/or metabolic abnormalities that considerably increase mortality. A patient is assessed to have septic shock when a persistent hypotension requires prolonged treatment with vasopressors and the serum lactate concentration exceeds 2 mmol/l³. Blood cultures are taken in patients with suspected sepsis, however their diagnostic value is limited, with one study stating only 53% of cases being blood culture positive¹⁴.

*Table 2 The SOFA score currently in use per the Sepsis-3 definition. Source: Singer, M. et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA - Journal of the American Medical Association **315**, 801–810 (2016) (removed since no copyright license was obtained in time).*

1 Introduction

1.1.3 Treatment

As mentioned before, time is of considerable importance in treating sepsis. That is why upon sepsis diagnosis, the patient will be treated immediately with broad-spectrum antibiotics. A delay in antibiotic treatment is thought to lead to an increased mortality and increased progression to septic shock per hour delay¹⁵⁻¹⁷, though contested recently in a meta-review that claims no benefit to mortality by earlier antibiotic treatment¹⁸ and a randomized controlled trial where earlier antibiotics did not improve survival¹⁹. Confirmation of bacterial presence will be done via blood cultures, potentially leading to refinement of the antibiotic therapy. Further interventions to be implemented within the first 1 to 3 hours after septic shock diagnosis are fluid resuscitation and vasopressors to reduce hypotension and subsequent hypoperfusion¹³. Mechanical ventilation or oxygen masks are used to improve oxygenation of the blood, and sometimes blood transfusions are used to improve haemoglobin levels. Patients in an advanced stage might require regulation of their blood glucose levels, enteral feeding and dialysis. Sometimes corticosteroids are used, despite their use being disputed⁵. Previous studies on corticosteroids show conflicting results on survival benefits, and a recent large-scale randomized Australian study with 3800 sepsis patients showed no survival benefit but an improvement in some secondary outcomes including time till resolution of shock²⁰. As evident from the definition of sepsis, the dysregulated immune response in sepsis affects several bodily and organ systems and may lead eventually to death by multiple organ failure; patients who survive the initial septic response often die due to opportunistic infections during the following phase of immunosuppression. Notably, only corticosteroids target the immune response, while other treatments, apart from antibiotics, aim at restoring the homeostatic state of different bodily functions. Despite a high research interest into blockage of certain immune responses and several clinical trials, none of these immunomodulatory treatments have been approved for sepsis treatment yet. There is hence still a dire need for improved treatment, and due to the heterogeneity of sepsis, personalized treatment is thought to be the key²¹.

1.1.4 The immune system

The immune system consists of several cells and acellular components that work together to protect our body from outside threats. The immune system has two arms: the innate

and the adaptive arm. The innate immune system is generally quick to respond, non-specific and evolutionary conserved. Its cells are monocytes, macrophages, neutrophils, eosinophils, basophils, mast cells, dendritic cells, NK cells and megakaryocytes. The adaptive immune system is thought to be acquired, slower to respond and specific to the infection. Its cells are T- and B-cells²². The arms of the immune system complement each other and depend on cross-talk to elicit an effective response to infection without excessive harm to the body. Therefore, the immune system must regulate itself using pro- and anti-inflammatory signals, including cytokines and lipid mediators. Binding of a cytokine to its receptor on an immune cell will influence the behaviour of the immune cell and in some cases, its phenotype. During the inflammatory response, many pro-inflammatory cytokines will be released, leading to increased recruitment and activity of immune cells. The inflammatory response needs to be resolved after successful elimination of the threat. This is mediated by pro-resolving processes, which actively affect the immune response towards resolution, and is seen as different from anti-inflammatory signalling despite overlap in functions and effect. Resolution and inflammation do not occur in distinct, separate phases, but rather co-occur partially, with elements of resolution increasing while inflammation decreases as the immune response progresses²³.

1.1.5 The inflammatory response to an infection

Our environment is crowded with microorganisms, including viruses, fungi, bacteria and parasites. Many of these microorganisms can invade the human body, and several have developed sophisticated techniques in order to do so. The first hurdle a pathogen faces are the outer defensive layers of our body: the skin barrier, the mucus layer and the epithelial layer in our respiratory, gastro-intestinal and genito-urinary system. A pathogen who has overcome this first obstruction and entered the body will face the cellular and acellular defence mechanisms of the innate and adaptive arms of the immune system.

A pathogen entering the body will be quickly recognized by cells of the innate immune system. These can recognize pathogens based on common molecular signatures, called pathogen associated molecular patterns (PAMPs). These PAMPs are distinct from human molecular patterns, hence allowing the body to differentiate between invaders and its own tissues. By utilizing receptors recognizing PAMPs, the innate immune system can fight pathogens it has not previously encountered. An important class of these receptors are toll-like receptors (TLR) able to recognize e.g., bacterial lipopolysaccharides (LPS) on the

1 Introduction

outer membrane and viral RNA intracellularly²⁴. Recognition of a pathogen by a TLR will lead to activation of a common and of alternative signalling pathways. The common pathway will lead to release of pro-inflammatory cytokines by NF- κ B activation, including interleukin-6 (IL-6), IL-1 β , IL-12 and tumour necrosis factor α (TNF- α). Activation of the immune system also leads to the production of prostaglandins from arachidonic acid, e.g., PGE₂, which will be part of the inflammatory response²⁵. The alternative signalling pathway depends on the type of TLR and the associated pathogen, including antiviral type I interferons (IFN) after activation of TLR3, TLR4, TLR7, TLR8 and TLR9²⁴. Cytokines and chemokines released will alert and attract further immune cells²². These can move towards the infection by sensing the gradient of chemokines, e.g., IL-8, in a process called chemotaxis. The majority of cells present initially are neutrophils, followed by monocytes and dendritic cells. The release of pro-inflammatory cytokines and mediators leads to a pro-inflammatory environment at the site of infection, which poises the immune cells for their pathogen-killing effector functions²³. Especially monocytes, macrophages and neutrophils will kill by eating pathogens: phagocytosis. Upon physically encountering the invading microorganism, these cells engulf it into a vesicle called the phagosome. The phagosome will unite itself with a second vesicle inside the cell to form the phagolysosome and kill the contained microorganism utilizing reactive oxygen species (ROS), diverse hydrolytic enzymes and a low pH²⁶. Neutrophils often die after phagocytosis. Especially bigger microorganisms will trigger the extracellular release of high quantities of ROS and lytic enzymes from granules by neutrophils and nitrous oxide (NO) by monocytes and macrophages. Prolonged stimulation by pathogens will lead to neutrophil death by apoptosis, necroptosis or neutrophil extracellular trap (NET) formation. NET formation is a still controversial topic with many uncertainties, however commonly agreed on are the presence of cell components including nuclear or mitochondrial DNA, histones and granules in NETs. These NETs either just immobilize bacteria or kill them, and probably kill the neutrophil as well, which is disputed²⁷. Monocytes take part in the inflammatory response by migrating to the site of infection, releasing high amounts of cytokines, e.g., IL-12 and IFN- γ , and phagocytosing pathogens. At the site of infection monocytes can differentiate into macrophages due to being exposed to high amounts of cytokines and other inflammatory signals. If these are pro-inflammatory, the macrophage will be programmed to be more pro-inflammatory, and vice versa for anti-inflammatory signals. *In*

vitro macrophages are divided into 3 types (M0, M1, M2) based on their differentiation, *in vivo* it is likely more accurate to describe macrophages on a functional spectrum rather than with distinct types and associated functions²⁷. The innate and adaptive immune system is connected by dendritic cells. These cells take up pathogenic particles and present them as antigens to T-cells in order to activate them.

The response of the adaptive immune cells occurs later than the innate immune response and includes the generation of antibodies by B-cells and the maturation of T-cells to different subtypes depending on cytokines and other signalling molecules in the environment. Activation of T-cells requires the presentation of an antigen by either a dendritic cell, a B-cell or a macrophage and a co-stimulatory signal. The different T-cell subtypes will regulate the immune response differently by releasing different mediators. This includes mediators important in monocyte differentiation. T-cells also play an important role in the resolution of inflammation²³.

The interaction of the immune system with pathogens does not happen in a vacuum separated from other bodily systems, but rather requires and involves several of them. An inflamed wound for example presents itself with swelling, redness, localized heat, pain and a loss of function. The heat, swelling and redness are mediated by the involvement of the cardiovascular system: an increase in blood flow and capillary dilatation facilitating increased immune cell invasion. Cardiovascular dilatation is mediated by hormonal signalling involving the pituitary gland – adrenal axis. A fever that might be developed during an infection is caused by cytokines instigating brain cells to release lipid mediators that mediate hormone release from the nervous system leading to an upregulation of the body temperature. The body temperature itself will affect the pathogens, but also affects the immune response²⁸. Coagulation is an important mechanism initiated in case of vessel damage. It leads to the forming of a blood clot, which can occlude blood flow and hence cease bleeding; this is required for the damage to be repaired. Coagulation will be activated in response to a pro-inflammatory environment, including DAMPs, NETs, and an activated complement system, and in response to vessel damage.

1.1.6 Resolution of the immune response

Resolution is required to protect the body from damage by excessive inflammation. Similar to inflammation, it is a highly regulated process involving a multitude of mediators. These mediators regulate each other in positive feedback loops and will also increase anti-

1 Introduction

inflammatory signalling, including IL-10. Resolution stops the inflammation and leads to tissue repair and homeostasis, preventing scarring and fibrosis. It is important to note that several resolution pathways are elicited as part of the pro-inflammatory response and that some pro-inflammatory mediators can also signal pro-resolution. A selective inhibition of pro-inflammatory signalling can negatively impact the resolution of the inflammation; for example the inhibition of COX-2, the enzyme necessary to produce pro-inflammatory prostaglandin (PGE₂) is detrimental to resolution since COX-2 is also required for pro-resolution lipoxin synthesis²³.

Important hallmarks of resolution are the inhibition of further recruitment of neutrophils, promotion of the apoptosis and programmed cell death of neutrophils, clearance of apoptotic cells by macrophages and the re-programming of monocytes to pro-resolving monocytes. Histologically, absence of neutrophils is a sign of successful resolution. Monocytes and macrophages play an important role in the resolution and will switch from a pro- to an anti-inflammatory polarization, also called classical and alternative, as the immune response progresses to resolution. Signalling by T-cells is pivotal for this polarization. T-cells will be activated by antigen presentation and move to the site of infection where they regulate the response. Alternative macrophages will release anti-inflammatory cytokines and remove dead neutrophils²³.

1.1.7 The immune response in sepsis

Our understanding of the immune response in sepsis is improving and various aspects have been characterized. What was initially thought to be a hyperinflammatory response and later on understood to be followed by an immunosuppressive period is now often interpreted as a simultaneous occurring of hyper innate immune activation and adaptive immune suppression^{29,30}(Figure 1). It is important to bear in mind that the septic response is heterogeneous with differences due to age groups³¹, the causing infection and further underlying factors³⁰. What causes these immune features to arise, how to address them and many of the interactions between different factors still remain poorly understood. The immune response in sepsis is described as dysregulated, with an increased activation of the innate immune system and suppression of the adaptive immune system^{3,21,32}. Cell counts of immune cells in sepsis show increased monocyte, neutrophil and myeloid suppressor cell populations. In addition, myeloid suppressor cells and a proportion of neutrophils are immature cells contributing to functional deficiencies of the response. Immature

neutrophils have deficits in oxidative burst, cell migration, complement activation and bacterial clearance, while myeloid suppressor cells suppress adaptive immune responses²⁹. Our understanding of the immune response in sepsis has been driven forward by transcriptomic studies on patient populations and healthy controls and the systems biology analyses of these, elucidating underlying signalling mechanisms. A transcriptomic approach by Smith et al. (2014) in neonatal bacterial sepsis found an upregulation of innate immune and metabolic genes and downregulation of adaptive immune response genes (Figure 2)³². An in-depth analysis of the data elucidated macrophage activation markers, signatures of increased chemokine signalling and FcγR-mediated phagocytosis on the one hand, on the other hand signs of T-cell suppression including signatures of impaired T-cell proliferation, cytokine production and T-cell receptors. In addition, decreased priming of T-cells by antigen presentation and co-stimulatory signals was deemed likely due to decreases in T-cell receptor gene expression, HLA-DR genes, which are required for antigen presentation, and decreases in expression of co-stimulatory genes. Simultaneously, inhibitory signalling pathways affecting the innate immune system, including negative regulators of IL-1, TLR-signalling and NF-κB were upregulated as well³². A transcriptomic analysis by Wynn et al. (2011) of differences between paediatric age groups found a significantly lower white blood cell count, neutrophil count and percentage of neutrophils in neonates (aged <28 days), compared to infants (aged 1-12 months), toddlers (2-5 years) and school-age children (6-10 years)³¹. In contrast, lymphocytes were increased in neonates and the percentage of monocytes was the same as in other age groups. The global gene expression pattern differed among age groups, with neonates and school—age children differing the most and only minor differences occurring between infants, toddlers and school-age children. In neonates a higher proportion of downregulated gene probes was found, especially within adaptive immunity and the antigen presentation pathway³¹. The downregulation of adaptive immunity and the antigen presentation pathway is in accordance with findings by Smith et al. (2014).

As mentioned before, the immune response does not occur in a vacuum but rather interacts with several bodily systems. This interaction is thought to be part of sepsis pathology and lead to organ dysfunction. One affected organ system is the vasculature. The endothelial lining of the blood vessels is highly responsive to mediators released by the immune response. Increased release of pro-inflammatory cytokines, including TNF-α, IL-1,

1 Introduction

IL-6 and interferon, strong activation of the complement system and increased release of NETs are thought to be part of a vicious cycle with a destructive endothelial response^{33,34}. This destructive endothelial response leads to microvascular disruption by increased coagulopathy and hence presence of thrombi in the vasculature³³. In addition, arterial vasodilation can lead to loss of vascular tone which affects both blood flow and cardiac function. Vascular dysfunction hence likely leads to or at least contributes to organ dysfunction by decreasing and disrupting blood flow to the organs. Pro-inflammatory signalling and vascular damage can also affect the lungs, both in patients with and without an infection localized to the lungs. Increased alveolar permeability leads to the development of oedema, and damages to the alveoli can affect the breathing capability of the patient. The aforementioned vascular dysfunction is also linked to cardiac dysfunction. The disrupted systemic blood flow affects the proper functioning of the heart. A further avenue by which the immune response in sepsis might lead to organ dysfunction is by affecting metabolism on a cellular and physiological level, further described below.

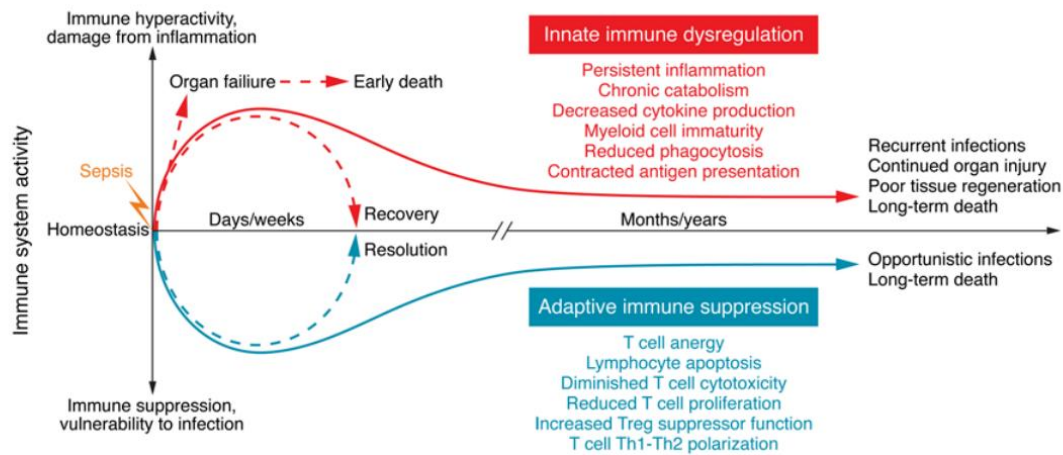


Figure 1 A diagram depicting the immune response in sepsis. It is characterized by a dysregulated innate immune response with persistent inflammation accompanied by a suppressed adaptive immune response. After the initial response that can lead to death by organ failure, sepsis also leads to poor health in the long term in survivors due to a prolonged innate immune activity and decreased adaptive immune system. Sepsis survivors frequently die in the months and years following their initial sepsis episode. Graph taken from Delano and Ward (2016) with permission from JCL (license number 1383176-1)

1 Introduction

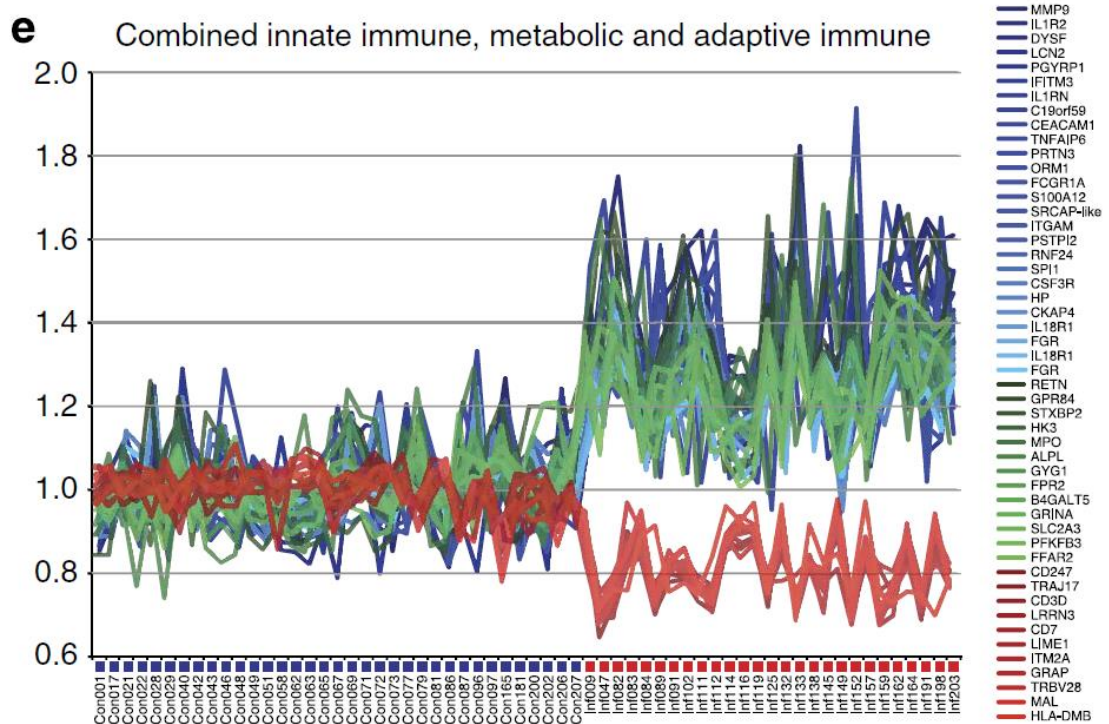


Figure 2 These graphs show the difference in expression as fold change to control in controls (blue dots) and septic neonates (red dots) of highly significant genes. The red lines indicate genes of the adaptive immune response, the green lines are metabolic genes and the blue lines are innate immune response genes. These results indicate an increased expression of innate immune genes, metabolism genes and a decrease in adaptive immune genes. Graphs taken from Smith et al. (2014) with author permission.

1.1.8 Metabolism in the immune response and sepsis

Metabolism is defined as the chemical reactions occurring within a living organism³⁵. A key function of metabolism is to make energy available to the body. Metabolism consists of two processes: catabolism, the breaking down of bigger molecules to smaller ones to gain energy, and anabolism, the energy-requiring process of building bigger molecules that are needed for cellular processes. Also, the catabolism of bigger molecules can free up smaller building blocks required for the synthesis of cellular components. Catabolism hence has a dual function in providing energy and building blocks. In the body, metabolism occurs both on a cellular level and physiologically, i.e., several organs contribute different functions to the overall metabolism of the body.

1.1.9 Metabolism on a physiological level

The metabolism of our body is tightly regulated and aimed at achieving homeostasis³⁵. In other words, the metabolic system is regulated to meet the body's energy demands without exceeding it. Excess energy is stored to be released when needed, i.e., in times of high energy demand or in between meals. A key organ in energy homeostasis is the liver. The liver receives the majority of amino acids, carbohydrates and lipids from the digestive system via the portal vein³⁶, with the remaining amino acids, carbohydrates and lipids entering the systemic circulation directly. In the liver glucose is stored as glycogen and lipids as triacylglycerol (TAG). The liver releases these throughout the day to provide a steady energy supply meeting the body's energy demand^{35,37}. The main energy storage organ however is the fat tissue, which stores lipids. Importantly, glucose and lipids do not just originate from the diet but can also be synthesized in the body, for example in the liver or kidney³⁷. The distribution of nutrients within our body depends on our blood circulation. When needed and hence regulated to do so, the liver, kidney or fat tissue release these molecules into the systemic circulation. Changes in circulation, i.e., widening and tightening of different blood vessels, can direct the blood flow to increase or decrease blood supply, and hence the supply of nutrients, to specific tissues and organs. Both metabolism and blood flow are regulated by the sympathetic nervous system³⁷. Hormones also play an important role in the immediate and more long-term regulation of metabolism³⁷. Insulin, glucagon and the catecholamines adrenaline and noradrenaline have an immediate effect on the body's metabolism. An example is that a blood glucose spike, which should occur after a meal, is followed within minutes by a spike in insulin. The hormone insulin allows

1 Introduction

for cells to take up glucose and is hence pivotal for the body's utilization of this nutrient and for the regulation of glucose levels in the blood. Adrenaline and noradrenaline decrease insulin secretion and stimulate glucose production, leading to increased glucose levels. Hormones like glucagon, growth hormone and cortisol have a less immediate effect and can affect the body's metabolism on a scale of hours or days, for example by altering insulin sensitivity. In order to achieve homeostasis of the metabolic system, the free fatty acid and glucose levels are monitored by the body and their levels influence hormone secretion and hence metabolism as well.

1.1.10 Cellular metabolism

Within the cell, energy can be harvested from lipids by oxidative phosphorylation, from glucose by glycolysis and from various nutrients, including amino acids, using the tricarboxylic acid cycle³⁸. These are catabolic pathways breaking down molecules into smaller metabolites or "building blocks". Other metabolic pathways exist but are not further explained here for sake of simplicity. The metabolism of a cell will correspond to situation-specific requirements; in other words, cells change their metabolism when their behaviour changes to satisfy the need for building blocks and changed energy requirements. As a result, cellular metabolism does not just passively provide the required energy or building blocks, but also regulates the immune response. The field of immunometabolism focuses on this intersection of metabolism and immunity. Within this field especially the role of glycolysis has received much attention. A switch in cellular metabolism to glycolysis is required for an effective immune response due to its role in effector functions of several immune cells, including phagocytosis and inflammatory cytokine production in macrophages, antigen presentation and the release of effector cytokines by T-cells³⁸. Firstly, an induction of glycolysis and simultaneous inhibition of oxidative phosphorylation leads to an increase in acetyl-CoA moieties, ATP and lipids³⁸. The presence of aforementioned acetyl-CoA moieties is required for histone modifications underlying altered transcriptional activity in the immune response³⁹. Glycolysis will also fuel protein synthesis and secretion. The metabolic switch can be initiated by e.g. TLR signalling that activates hexokinase, a rate-limiting enzyme in glycolysis, a key metabolic biomarker gene (Figure 2)^{32,40}. In line with this, an immunotolerogenic state of the cell is described for cells utilizing fatty acid oxidation, whereas LPS-activated effector cells shift to glycolysis and fatty acid synthesis⁴⁰.

1.1.11 Lipid metabolism in the cell

Since lipids are the main focus of my studies, I will introduce these molecules in more detail. Fatty acids can be divided into saturated fatty acids and mono- or poly-unsaturated fatty acids. Saturated fatty acids contain only single bonds between their carbons, whereas mono- or poly-unsaturated fatty acids contain one or several double bonds in their carbon chain, respectively. Based on the position of the first double bond, poly-unsaturated fatty acids (PUFAs) are divided into omega-3 or omega-6 lipids. The straight chain saturated fatty acids C14:0, C16:0 and C18:0 are dietary lipids, and C16:0 and C18:0 are among the most common lipids in our diet⁴¹. In addition, these lipids can be synthesized in our bodies. In this process, C16:0 is the most commonly synthesized lipid by the process of lipid synthesis by fatty acid synthetase in the cytoplasm⁴¹. C16:0 can be further elongated to C18:0 in a separate process. To store lipids they will be esterified, and it is likely that C18:0 will be desaturated to C18:1, oleic acid, before esterification⁴¹. Esterification with glycerol leads to the formation of triacylglycerols. C18:1 can also originate from the diet. C18:2 is the most common PUFA in the diet⁴². Further PUFAs in our diet are C18:3, C20:4, C20:5 and C22:6. Apart from C18:2 and C18:3, all of these lipids can be synthesized in the body as well⁴³. Furthermore, these lipids can be oxidized in mitochondria to generate energy. In order to enter the mitochondria, most lipids require the use of carnitine for molecular shuttling across the membrane⁴⁴. This depends partially on the chain length of the lipid, where shorter lipids can potentially diffuse into the mitochondria as well as being shuttled⁴⁵. However, the transport of carnitine-bound fatty acids is more efficient and seems to be the preferred method even for lipids able to diffuse⁴⁶. This can be seen from a set of experiments which firstly prove that medium-chain fatty acids can diffuse into mitochondria and secondly, that inhibiting the carnitine shuttle led to a strongly decreased uptake of medium-chain fatty acids into mitochondria⁴⁶. After entering the mitochondria, the fatty acids will release the carnitine and bind acetyl-CoA instead before being either oxidized or elongated to longer fatty acids⁴⁷.

Since in sepsis the immune response is activated, it is conceivable that metabolism should be changed accordingly. Indeed, metabolism in general, and lipid metabolism as well, has been shown to change in sepsis, as has been acknowledged in the latest paper setting out the consensus sepsis-3 definition³. This is evidenced in previous studies by both transcriptomic and proteomic changes showing alterations in enzymes and other

1 Introduction

molecules impacting (lipid) metabolism and availability^{48,49}, and by changes to lipids and other metabolites that indicate metabolic alterations^{50,51}. Smith et al. found transcriptomic signatures of changes to metabolism in sepsis (Figure 2). Increased expression of glucose transporters and hexokinase indicated an increase in glycolytic flux³². They also found significant increases in expression of lipid-binding receptors GPR43 and GPR84, which poses the question if their lipid ligands are altered in sepsis. Sharma et al. analysed the differences in the proteome between healthy volunteers and sepsis patients at admission and 7 days after admission⁴⁸. Amongst the proteins whose levels they found most changed were proteins involved in fatty acid metabolism (decreased), lipid transport (decreased) and lipid uptake (increased).

How do the changes in lipid metabolism relate to lipid functions in the body? Firstly, an important function of lipids is to provide energy to the body's cells. That is also why the human body stores lipids, for example in fat and liver tissue, which can be liberated from storage when needed^{52,53}. Any immune reaction, but especially the aberrant immune reaction in sepsis, requires a high amount of energy to sustain itself⁵². A shift towards fatty acid metabolism has been argued to be the result of a depletion of other energy sources, as seen in a meta-study by Golucci et al.(2018)⁵⁴ and a mouse study by Irahara et al.(2018)⁵⁵. Golucci et al. (2018) performed a meta-study concatenating the information from 29 studies on lipid profiles in sepsis. They conclude that the increased energy demand depletes the available carbohydrates within 12-24 hours in sepsis, leading to a shift towards fatty acid metabolism evidenced by an increase in triacylglycerols in the circulation. The TAG are likely released from fat storage, and are accompanied by a decrease in cholesterol, low-density lipoproteins and high density lipoproteins, the latter decrease being associated with sepsis severity⁵⁴. Irahara et al.(2018) conducted a study inducing sepsis in mice using either a low (1mg/kg body weight) or high (5mg/ kg body weight) dose of *E. coli* LPS⁵⁵. Their results suggest a shift towards a higher fatty acid utilization in the acute phase of inflammation, i.e., the first 48 hours after LPS exposure, but saw a decrease in free fatty acid (FFA) and TAG in the circulation in contrast to the Golucci et al. findings. Interestingly, Irahara et al. did find increased FFA and TAG levels in the liver of septic mice and hypothesized a transport of FFA to the liver for energy storage, while Golucci et al. hypothesized that impaired circulation to the liver in human sepsis patients leads to reduced clearance of various substances, including fats, from the blood. The observed

differences between the sepsis mouse model and studies in human patients might hence be caused by differences in blood circulation to the liver. Since hormones and the sympathetic system regulate the metabolic system, it would have been interesting to know the hormone levels in these studies. Furthermore, the human study would have been impacted by both parental or enteral feed and intravascular fluids (IV) given to the patients, while in the mouse study mice had access to standard chow. Parenteral or enteral feed administered to the patients should hence contribute to their energy needs being met. This feed should contain all required nutrients, i.e., lipids, carbohydrates, protein and micronutrients, though it is worth noting that patients receive fewer calories than the daily reference intake for adults^{56,57}. Patients with sepsis often have increased levels of adrenaline and noradrenaline in their blood, which might lead to hyperglycaemia due to attenuated insulin release. In fact, in the clinical setting glucose levels of sepsis patients are monitored with the aim to avoid hyperglycaemia by administering insulin in the IV when blood glucose levels reach 181 mg/l, i.e. 1 mmol/l,⁵⁶ which is an actually relatively high glucose level. In a study by Finfer et al., 69% of sepsis patients treated according to this rule required insulin administration⁵⁶.

In addition to the role of lipids in metabolism and immunometabolism discussed above, lipid mediators play a direct role in the regulation of inflammation⁴². Lipid mediators constitute of the so-called eicosanoids, e.g., prostaglandins, leukotrienes, and specialised pro-resolving mediators (SPM) including lipoxins⁴², however the presence and function of SPM is subject to discussion at the moment⁵⁸⁻⁶⁰. These lipid mediators originate from C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid), and C22:6 (docosahexaenoic acid). C20:5 and C22:6 are omega-3 PUFAs and can originate from the omega-3 PUFA C18:3, whereas C20:4 is an omega-6 PUFA that can be metabolized from the omega-6 PUFA C18:2. Omega-3-PUFAs cannot be converted into omega-6-PUFAs or vice versa. Omega-3 PUFAs are often classified as anti-inflammatory and omega-6-PUFAs as pro-inflammatory, based on the signalling molecules they give rise to. Arachidonic acid (C20:4) gives rise to prostaglandins, leukotrienes and lipoxins⁴². Arachidonic acid becomes available for conversion into lipid mediators after TLR stimulation and will initially lead to an increase in prostaglandins and leukotrienes as they play a role in the initiation of inflammation, at a later stage lipoxins are generated that play a role in resolving inflammation. It is hence important to note that omega-6-PUFA metabolites are not exclusively pro-inflammatory. C20:5 and C22:6 can give

1 Introduction

rise to resolving lipids. In addition, omega-3-PUFAs can impede omega-6-PUFA metabolism in the cell by acting as a competitive substrate. Furthermore, mediatory functions of so called octadecanoids, originating from C18:0 and C18:1, have been identified but are less researched so far^{61,62}.

Since metabolism and the immune system are tightly interconnected, it has been attempted by several research groups to measure metabolites that would correlate with survival in sepsis. Liu et al. recruited patients in an early stage of sepsis and measured metabolites to find indicators of survival⁵⁰. They found 43 metabolites with significantly different levels. Increased in non-survivors were TCA cycle metabolites, lactic acid, and metabolites indicating liver injury (bile acids), muscle tissue damage and -catabolism (creatine, creatinine, branched chain amino acids) and disruption of intestinal environment and kidney damage (indoxylactate, indoxysulfate). Decreased levels were found of long-chain and short chain fatty acids and C2:0-carnitine in non-survivors which the authors hypothesized indicated an increased β -oxidation of lipids. The study authors hypothesized that increased lactate production and enhanced catabolism indicate an energy scarcity in combination with hypoxia and energy generation inefficiency⁵⁰. An analysis conducted by Langley et al. (2014) integrating lipidomics with proteomics lead the authors to conclude that carnitine shuttling in sepsis non-survivors is decreased due to a decrease in fatty acid transport proteins⁵¹. They found differences for 6 acylcarnitines between sepsis survivors and healthy controls, and, more strikingly, that short-chain and medium-chain-acylcarnitines and branched-chain acylcarnitines differed between sepsis survivors and non-survivors. The hypothesized mitochondrial dysfunction would severely impact any lipid oxidation and the metabolic functioning of the cell in general^{51,63,64}.

1.1.12 Bacterial sepsis

Bacteria, fungi and viruses can cause sepsis. It is, however, often referred to as caused by bacteria and much research focusses on bacterial sepsis. This is in part because blood cultures are routinely taken to confirm presence of bacteria in the blood and monitor the effectiveness of antibiotic treatment. When a blood culture is negative, it leaves uncertainty about the type of infection and the presence of infection. Bacteria can also play a role in immunity as part of the microbiota. The microbiota is defined as all microorganisms colonizing the human body, foremost the skin and intestines. It influences

our health in various ways, including providing colonization resistance to other bacteria, crosstalk with our immune system and the digestion of food in the intestines. The composition of the microbiota determines its effect, and sometimes bacteria are part of our microbiota that are detrimental to our health. In addition, sometimes the microbiota is “incomplete” leading to vulnerabilities that can be exploited by pathogens. It has been hypothesized that sepsis causing pathogens can enter the body via the gut, but evidence is weak: Stewart et al. found that in 6 out of 7 preterm babies within a cohort that developed late-onset sepsis, the causing pathogen (identified by blood culture) was the most prevalent or second most prevalent intestinal colonizer before the onset of sepsis⁶⁵.

The intestinal microbiota digest food components leading to the release of a multitude of metabolites, some of which will nourish our intestinal lining and other tissues in our body after entering the blood stream. These metabolites are also important for the crosstalk between our microbiota and immune system. A well-researched example are the short chain fatty acids (SCFA) butyrate, acetate and propionate, which are the cognate ligands for the receptor GPR43.

1.1.13 GPR43

The G-protein coupled receptor GPR43 binds mainly acetate and propionate, and with weaker activation, butyrate⁶⁶. These are produced when the intestinal microbiota breaks down fibre, with the concentration ratio in the colon being roughly 60:25:15. In blood acetate will be the most prevalent SCFA⁶⁷. A high-fibre diet will increase SCFA production and is thought to be associated with reduced systemic inflammation⁶⁸. The receptor is expressed on neutrophils, monocytes, macrophages, dendritic cells, eosinophils and regulatory T-cells (Tregs). Its expression is upregulated by inflammatory mediators. Activation leads to downregulation of IL-6 and IL-1 β , which are inflammatory cytokines and to chemotaxis of neutrophils. Furthermore, GPR43 is known to inhibit NF-kB signalling, IL-6 and IL-1 β ^{68,69}.

1.1.14 Biomarkers

Biomarkers in sepsis have been researched for both diagnostic and treatment use. A biomarker is a compound that can be measured in order to diagnose a disease or pathological condition, e.g., the measurement of glucose in urine to diagnose diabetes mellitus. The measurement of CO₂, bilirubin and lactate are current biomarkers used to

1 Introduction

identify underlying dysfunctions present in sepsis. The need for better diagnosis requires biomarkers to be measured fast, give a reliable indication of sepsis and at best stratifies the patient cohort into subtypes indicating which medical attention is required²¹. Biomarkers that have been considered for sepsis diagnosis are molecules whose concentration are altered in the blood, e.g., C-reactive protein (CRP) and procalcitonin, and host gene signatures. CRP and procalcitonin, despite being significantly higher in sepsis cases than controls, were found to be not sufficient to diagnose sepsis as singular markers¹⁴. The use of transcriptomic analysis of whole blood taken within 24 hours of admission led Smith et al. to a 52-gene classifier differentiating robustly between bacterial sepsis and controls. These host signatures include gene expression of the innate and adaptive immune system and of metabolism, including GPR43 and GPR84. This classifier only recognises bacterial and not viral sepsis cases³². Wong et al. have developed a prognostic marker that stratifies paediatric septic shock patients into risk groups and in addition has shown to have predictive value, too, which could inform treatment choices⁷⁰. PERSEVERE is based on transcriptomic data identifying gene signatures measured in blood taken within 24 hours of admission significantly associated with mortality in paediatric septic shock patients. A subset of genes was chosen based on their biological function and the possibility to measure the gene-product (the protein encoded by the gene) in the blood. This led to a classifier of 5 proteins called PERSEVERE: C-C chemokine ligand 3, IL-8, heat shock protein 70 kDa 1B, granzyme B and matrix metalloproteinase 8. The inclusion of more proteins and mRNA lead to the improved model PERSEVERE-XP which was deemed to have a negative likelihood ratio in the clinically useful range and a positive likelihood ratio in the modestly useful range⁷¹. The relevance of stratifying patient cohorts lies both in the option to improve treatment for specific patient cohorts and in enabling studies to test effect on mortality with enough power by selecting high-risk cohorts.

1.2 G-protein coupled receptors

A vital component of the regulation of every process in our body, both on an organismal and a cellular level, is the communication between tissues and between cells. This communication is mediated by membrane receptors that bind molecules present in their direct vicinity. Not just our own body produces these ligands; our diet, our microbiota and pathogenic organisms all contain and/or release molecules known to interact with human

receptors. Medication frequently exerts its desired effect on the human body by receptor interaction or affecting molecules that bind to receptors. The largest class of receptors in humans is called G-protein coupled receptors (GPCR)⁷². Every cell in our body harbours numerous GPCRs that transmit signals from the extracellular compartment to the intracellular compartment. Between them, GPCRs can bind a pleiotropy of different compounds, ranging from ions, e.g. Ca^{2+} and H^+ , to complex large molecules like hormones, e.g. dopamine, histamine and adrenaline. GPCRs are involved in a broad variety of processes, including reception of taste, mood regulation and the immune response⁷³. A recent analysis indicates that 25-36% of all registered drugs target GPCRs, hence attesting to the pivotal role of GPCR in several conditions and diseases⁷².

GPCRs are coupled to G-proteins. Upon binding of a ligand, the receptor will commonly signal through G-proteins, which transduce the signal to cytosolic or membrane-bound effectors. These effectors activate a signalling cascade that will result in the release of secondary messengers, e.g., cAMP or Ca^{2+} , which in turn lead to cellular effects including the activation of other signalling cascades, changes in gene expression or chemotaxis. This general principle is illustrated in Figure 3. Apart from the name-giving G-proteins, the receptors also signal through other proteins, most notably arrestins. Signalling through G-proteins is also called canonical signalling, whereas G-protein independent signalling is called non-canonical⁷⁴. There is a variety of G-proteins, effectors and secondary messengers and one receptor can often activate several of these. Two receptors can have different effects despite overlapping signalling mechanisms, and different signalling mechanisms can lead to similar downstream effects⁷⁵. This principle is impressively shown when looking at the regulation of systemic functions by GPCRs; despite several receptors transducing signals through the same G-protein type (further explained below) outcomes are divergent; simultaneously, different transducers can lead to similar effects.

1 Introduction

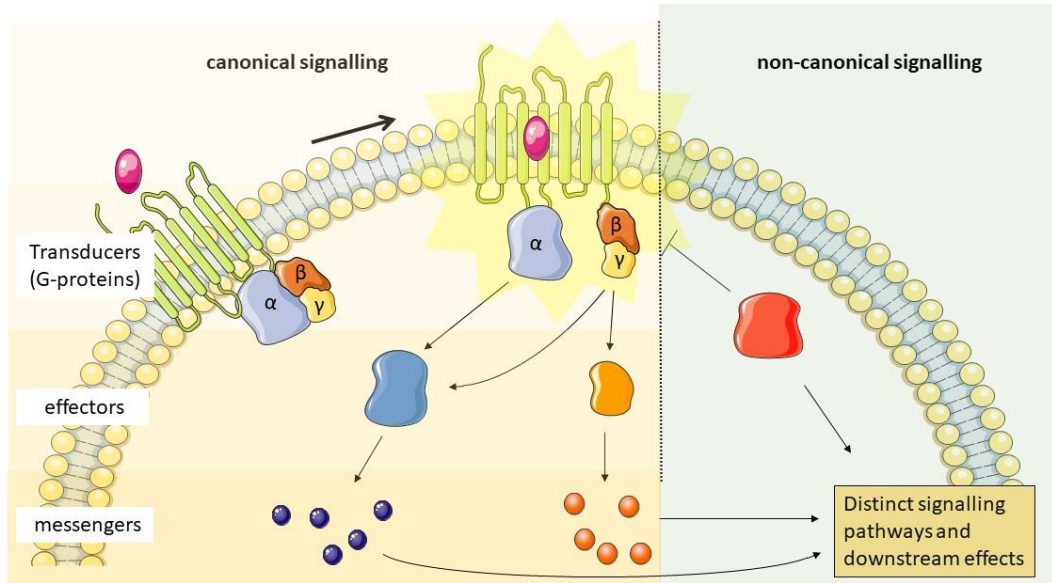


Figure 3 The general signalling pathways that can be initiated when a ligand binds to a G-protein coupled receptor. In canonical signalling, upon binding of the ligand, the transducers $G\alpha$ and $G\beta\gamma$ dissociate and are able to transduce the signals by regulating effectors. These effectors can lead to the synthesis or release of secondary messengers, which will activate distinct signalling pathways and downstream effects. In non-canonical signalling, an effector, most notably arrestin, can inhibit G-protein mediated signalling, often via receptor internalization which leads to receptor recycling or degradation. This effector can also activate its own distinct signalling pathways and downstream effects, some of which overlap with canonical signalling activated pathways. Parts of the figure were drawn by using pictures from Servier Medical Art. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

1.2.1 Structure of G-protein coupled receptors and different conformational states

G-protein coupled receptors (GPCR) are situated in the cell membrane and consist of 7 transmembrane helices connected by alternating intra- and extracellular loops (Figure 4). The N-terminus and C-terminus stick out into the extra- and intracellular compartment, respectively (Figure 5 a, right side). The binding pocket for ligands is formed by the transmembrane helices, while the extracellular loops form the 'entrance' to the pocket and, in case of some GPCR, restrict access⁷⁶. This is called the orthosteric binding site. Numerous GPCRs appear to have a second or more additional binding pockets, called the allosteric binding sites. The binding pocket binding the typical ligand is called the orthosteric binding pocket, while the allosteric binding pocket binds distinct ligands⁷⁴. This general structure is common to all GPCR, with distinct structural features of binding pockets, intra- and extracellular loops and the N- and C-terminus further determining receptor-specific functional features including which ligands it binds, which downstream signalling molecules it interacts with and its signalling activity.

The structure of a GPCR is not rigid, but rather thought to change between different states or conformations. A change in conformation will occur in presence of a ligand, but can also occur in its absence, and is required for receptor activation. The associated energy cost of changing conformation makes changes unlikely to occur by chance, some GPCRs are however known to signal constitutively. The conformations harbour different degrees of being active or inactive and dictate the energy required to activate receptor signalling⁷⁴. This means that receptor signalling is not a simple switch between 'active' and 'inactive' operated by a ligand. Figure 6 shows an illustration of the 'landscape' of conformation states of GPCRs. Mechanistically, upon binding of an activating ligand, the receptor will contract its extracellular compartment, which also affects the intracellular conformation and its interaction with any transducers, e.g. G-proteins. Furthermore, both binding to the orthosteric and the allosteric pocket can affect receptor conformation, and binding of an allosteric ligand might enhance the effect of ligand binding at the orthosteric site⁷⁷. Binding of both a ligand and a transducer, e.g., a G-protein, might be necessary for a complete shift in conformation to a fully activated state.

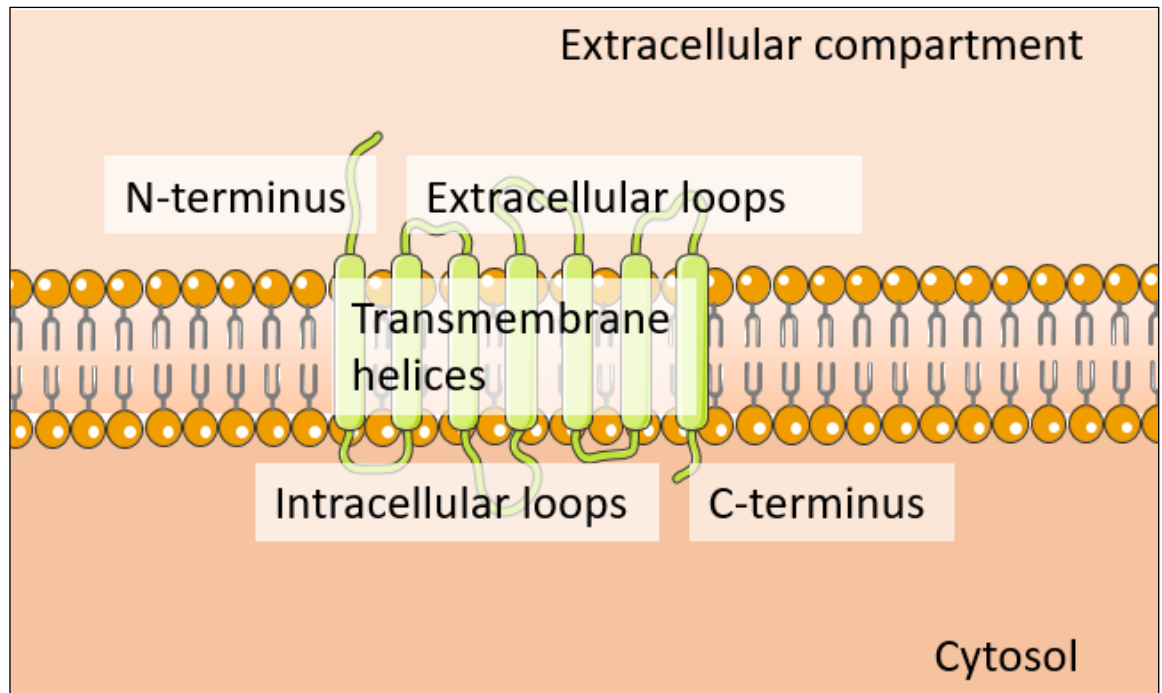


Figure 4 The general structure of GPCR in a two-dimensional depiction showing the extracellular N-terminus and extracellular loops, transmembrane helices and intracellular loops and C-terminus. Parts of the figure were drawn by using pictures from Servier Medical Art. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

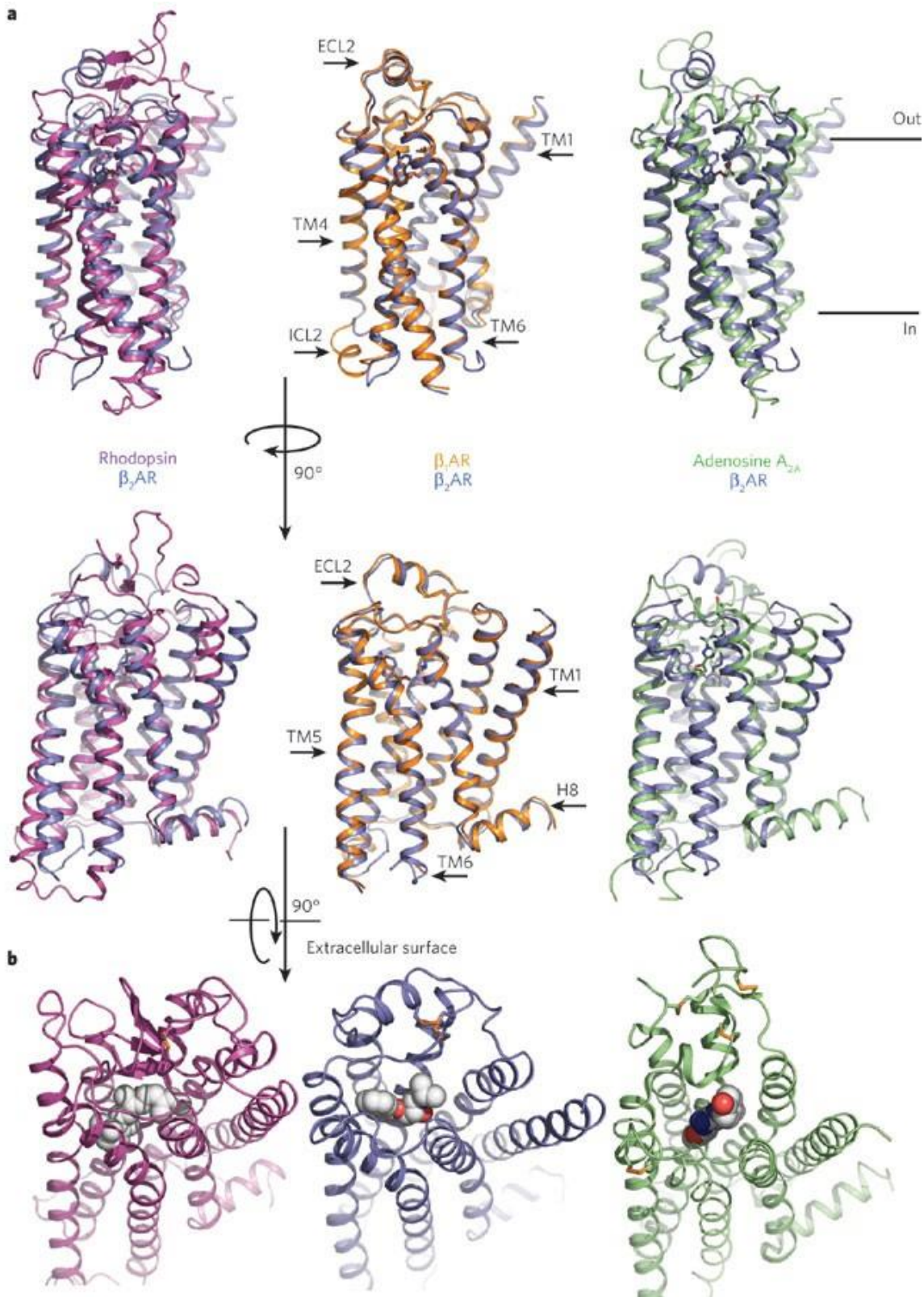


Figure 5 Three structures of GPCR, bovine rhodopsin (purple), avian $b1AR$ (yellow) and human Adenosine A_{2A} (green) are imposed onto human $b2AR$ (blue) and shown from two different angles. b) Each of the receptors is shown in the inactive state, the ligands shown inhibit receptor signalling. The receptors are shown as viewed from the extracellular

1 Introduction

surface, the viewer hence looks at the binding pocket from the top. ECL = Extracellular loop, ICL = intracellular loop, TM = transmembrane, H8 = Helix 8. Picture taken from Rosenbaum et al. (2009), license for re-use obtained from Springer Nature (license number 5602540288428).

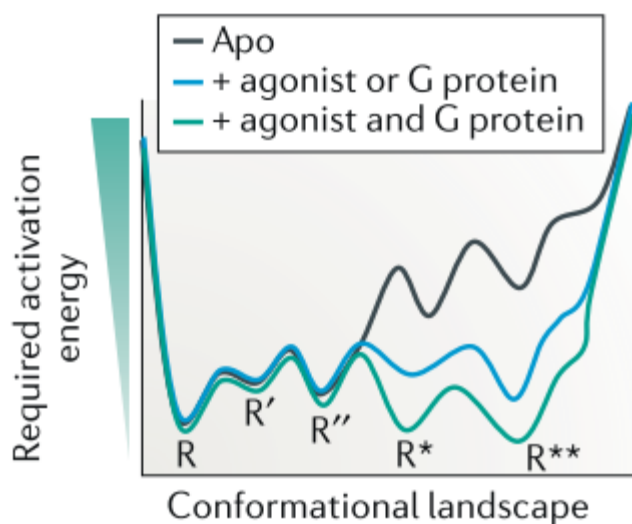


Figure 6 The conformational landscape of GPCR. R' and R'' are inactive-like states, whereas R^* and R^{**} are active-like states. The black line (Apo) indicates the absence of ligand, whereas blue indicates the presence of either agonist or G-protein, and green indicates the presence of both. On the y-axis the required activation energy is shown. Adapted from Wootten et al. (2018), license for re-use obtained from Springer Nature (license number 5602541091077)

1.2.2 Ligands

There are different types of ligands: agonists, inverse agonists and antagonists. Agonists nudge the conformation of the receptor towards a more active state and can lead to receptor signalling. Inverse agonists, on the other hand, stabilize the receptor in an inactive conformation, inhibiting further signalling. Antagonists will have no effect on the state of the receptor, but prevent the binding of other ligands, hence inhibiting activation of the receptor. Ligands also differ in their affinity – how easily they bind to the receptor – and their efficacy – the quantitative downstream effect of ligand binding relating to time elapsed. A ligand with high efficacy would e.g. lead to a comparably fast GDP/GTP exchange on the $G\alpha$ protein leading to quantitatively more downstream signalling than a ligand with lower efficacy in the same time-frame^{74,76}.

1.2.3 G-proteins

G-proteins are heterotrimeric guanine nucleotide (GTP/GDP) binding regulatory proteins; in other words they consist of three distinct units that together bind guanosine triphosphate (GTP) and guanosine diphosphate (GDP). The subunits are called α -, β - and γ -subunit. In the inactive state they will be bound as one complex at the intracellular site of the receptor. When the receptor is activated, the β - and γ -subunits ($G\beta\gamma$) will remain tightly together, whereas the α subunit ($G\alpha$) dissociates when its bound GDP is replaced by GTP. GDP to GTP exchange is required for G-protein signalling. Both the $G\alpha$ and $G\beta\gamma$ proteins can transduce signals. There are 16 different $G\alpha$, 5 $G\beta$ and 13 $G\gamma$ genes in humans. The different G-unit proteins can combine into a multitude of different combinations. The $G\alpha$ proteins are subdivided into 4 types, and these types are often used to classify G-protein coupled receptors^{74,76}. G-proteins exert their signalling effects by coupling to effectors that increase secondary messengers. The secondary messengers are cyclic adenosine monophosphate (cAMP), Ca^{2+} , inositol triphosphate (IP3), diacylglycerol (DAG) and arachidonic acid^{76,78}. The types of G-protein subunits determine which effectors are activated.

1.2.4 $G\alpha_i$ and $G\alpha_s$ signalling

The most studied $G\alpha$ protein signalling is the activation and inhibition of cAMP signalling by $G\alpha_s$ and $G\alpha_i$, respectively. $G\alpha_s$ and $G\alpha_i$ have opposing effects on signalling by regulating adenylylcylases (AC), the enzymes leading to cAMP release from ATP. Despite $G\alpha_i$ often

1 Introduction

being described as pro-inflammatory and $G\alpha_s$ as anti-inflammatory, the effect of their signalling is not that clear-cut. cAMP signalling will activate phosphokinase A (PKA), which phosphorylates several metabolic enzymes, Mitogen Associated Protein Kinase (MAPK), the transcription factor cAMP Response-Element Binding protein (CREB) and inhibits phospholipase C β 2⁷⁹. Genes that harbour a binding site for CREB include Interleukin 2 (*IL-2*), *IL-6*, *IL-10* and tumor necrosis factor α (*TNF- α*)⁸⁰. Logically, $G\alpha_s$ signalling would lead to increased expression of these CREB related genes, whereas $G\alpha_i$ would inhibit it. Furthermore, CREB is thought to influence NF- κ B signalling, though it is unclear, or at least situation-dependant, if it enhances or inhibits it⁸⁰. Apart from PKA, cAMP has also been found to activate the Exchange Protein directly Activated by cAMP (EPAC1 and EPAC2). EPAC1 is ubiquitously expressed, whereas EPAC2 is restricted to certain tissues including brain and steroidogenic tissues^{81,82}. Albeit less studied than PKA, EPAC1 has been found to regulate phagocytosis and intracellular killing in alveolar macrophages⁸¹. The effects of cAMP signalling elucidated by Aronoff et al. using PKA and EPAC specific agonists and inhibitors in rodent alveolar and peritoneal macrophages is displayed in (Figure 7). Aronoff et al. found that $G\alpha_s$ signalling inhibits FcR-mediated phagocytosis, and that this was mediated via EPAC in alveolar macrophages. A $G\alpha_s$ induced reduction in *TNF- α* in alveolar macrophages was mediated only via PKA, whereas the phagocytes' decreased ability to kill previous ingested bacteria could be caused via both PKA and EPAC⁸¹. Interestingly, they further found that cAMP signalling is different in dendritic cells and macrophages. While a PKA agonist induced *IL-6* and *IL-10* secretion in alveolar macrophages, it suppressed *IL-6* in dendritic cells. The secretion of *TNF- α* and chemokine (C-C motif) ligand 3 (*CCL3*) was decreased in both cell types. In addition, EPAC had no effect on cytokine secretion in macrophage cells, but was able to suppress *CCL3* in dendritic cells⁸³. This illustrates that the effects of a G-protein and its effectors are also cell-dependant, adding another layer of complexity to our understanding of canonical GPCR signalling. In addition, signalling by $G\alpha_s$ in this context was neither distinctly pro- nor anti-inflammatory. Another point worth raising is that despite CREB activation and inhibition playing a major role in the downstream effects of $G\alpha_s$ and $G\alpha_i$ signalling, respectively, CREB can also be activated by pathways independent of cAMP⁸⁰.

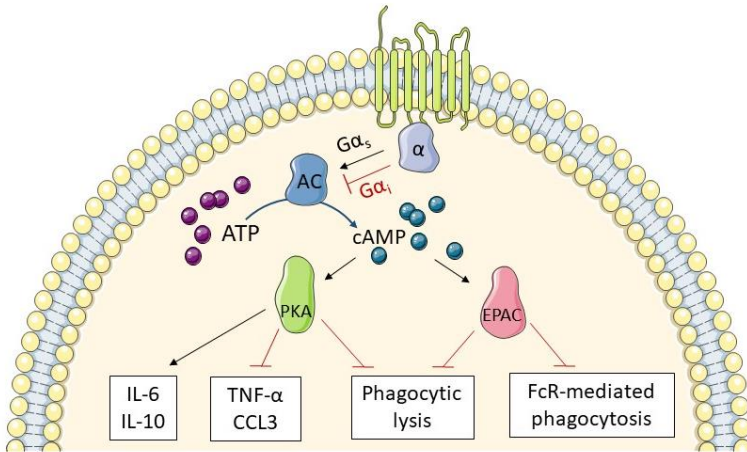


Figure 7 An overview of pathways downstream of $G\alpha_s$ and $G\alpha_i$ protein signalling with effects in macrophages as elucidated by Aronoff et al. $G\alpha_s$ protein stimulates adenylyl cyclase (AC), while $G\alpha_i$ inhibits it. AC converts ATP to cAMP, the secondary messenger that activates PKA and EPAC. PKA can increase IL-6 and IL-10 secretion, inhibit TNF- α and CCL3 secretion and inhibit phagocytic lysis. EPAC has no effect on cytokine secretion but inhibits both FcR-mediated phagocytosis and the lysis of phagocytosed bacteria. Parts of the figure were drawn by using pictures from Servier Medical Art. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

1 Introduction

1.2.5 $G\alpha_q$ signalling

The $G\alpha_q$ protein canonically activates phospholipase β , which cleaves phosphatidylinositol 4,5-bisphosphate (PIP₂) into inositol triphosphate (IP₃) and diacylglycerol (DAG). IP₃ releases Ca²⁺ from the endoplasmic reticulum, while DAG activates protein kinase C (PKC). Other phospholipases activated by $G\alpha_q$ will lead to the release of arachidonic acid. Ca²⁺ and PKC can affect several downstream pathways, including inhibition of the phosphatidylinositol 3-kinase (PI3K)/Akt pathway and activation of the MAPK pathway. These pathways play important roles in cell survival and proliferation^{78,84}. $G\alpha_q$'s main signalling mechanism is, however, deemed to be Ca²⁺ mobilization, which has ubiquitous effects by binding to and activating intracellular proteins⁸⁵. G_q -coupled receptors are frequently hormone receptors or chemokine receptors on immune cells⁷⁸.

1.2.6 $G\alpha_{12/13}$ signalling

$G\alpha_{12/13}$ proteins are thought to play a role in cellular motility, cell growth and differentiation. Among the effectors activated by both $G\alpha_{12}$ and $G\alpha_{13}$ are members of the Rho family of small GTPases, which in turn affect the actin cytoskeleton⁸⁶. The Jun N-terminal kinase (JNK) pathway has also been found to be regulated by $G\alpha_{12/13}$ ⁸⁶ and Rho GTPases, which could provide a feasible pathway since JNK can affect the cytoskeleton⁸⁷(Figure 8).

$G\alpha_{13}$ is thought to increase motility and chemotaxis of immune cells via the cytoskeleton in response to growth factors activating receptor tyrosine kinase (RTK). It remains to be studied further if the activation of $G\alpha_{13}$ in this context is mediated by GPCR transactivation by RTK or non-canonical signalling of RTK⁷⁶. Hee Ha et al. found several transcription factors to be differently activated in ligand stimulated $G\alpha_{12}^{-/-}$ and WT cells, one of the highest hits being CREB. They then established CREB activation via the Ras/Mek/Erk pathway by $G\alpha_{12}$ in an ovarian cancer cell line in ligand treated cells⁸⁸.

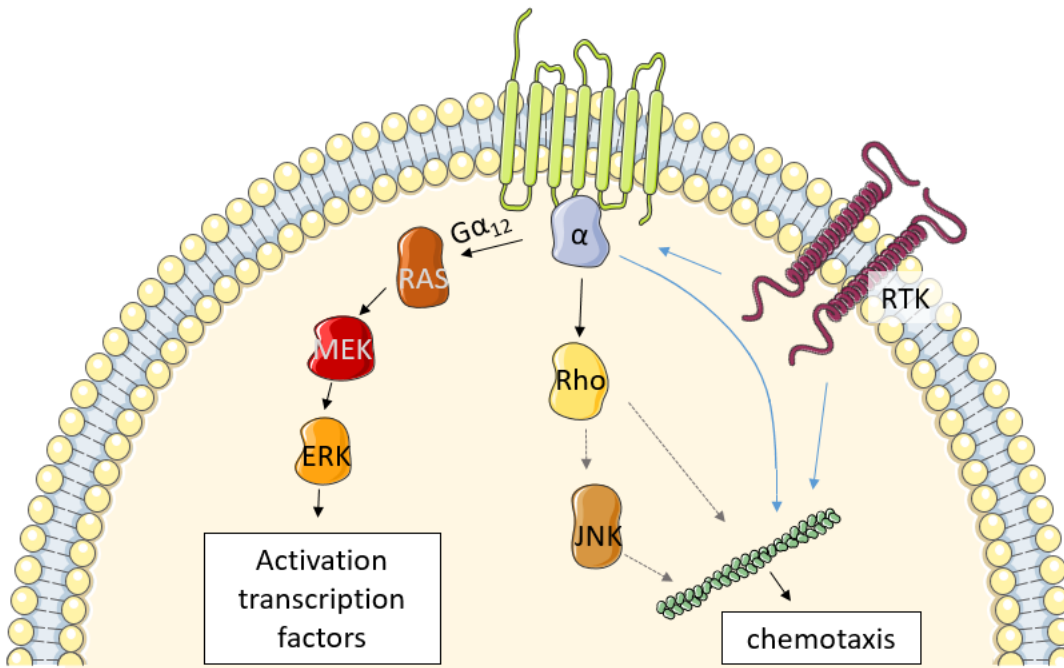


Figure 8 Diagram of $G\alpha_{12/13}$ signalling. $G\alpha_{12}$ has been shown to activate several transcription factors via the RAS-MEK-ERK pathway in an ovarian cancer cell line. $G\alpha_{12/13}$ proteins are known to activate Rho and lead to chemotaxis, this is potentially mediated via JNK which is known to be affected by $G\alpha_{12/13}$ and Rho and affect the cytoskeleton (green) but has not been conclusively proven. Blue arrows indicate the potential interaction of G-protein and RTK or the non-canonical signalling of RTK leading to chemotaxis; this interaction requires further investigation. Parts of the figure were drawn by using pictures from Servier Medical Art. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

1 Introduction

1.2.7 G $\beta\gamma$ signalling

Not all GPCR elicit G $\beta\gamma$ signalling^{89,90}. Most G $\beta\gamma$ signalling is mediated from G α_i -uncoupling. This is due to its high abundance and a relatively high threshold required for G $\beta\gamma$ signalling to be effective. G $\beta\gamma$ signalling can interact with G α signalling from both the same receptor and other receptors. G $\beta\gamma$ affects G α_i and G α_s signalling by inhibiting adenylyl cyclase or enhancing its effect. This will depend on both the subtypes of G β and G γ as well as the adenylyl cyclase isoform.⁹¹ G $\beta\gamma$ subtypes enhancing G α_s stimulated adenylyl cyclase activity usually originate from a G α_i -G $\beta\gamma$ heterotrimer coupled to a different receptor, hence leading to the contradictory occurrence of G α_i -coupled receptor signalling increasing cAMP signalling^{91,90}. Adenylyl cyclase subtypes and their prevalence differ among cell types, and G α and G $\beta\gamma$ proteins might interact differently with different subtypes⁹². As a result, signalling by these G-proteins might differ cell-type dependently, leading to contradicting outcomes of one receptor in different cell types. Pertussis toxin, which is known to inhibit G α_i proteins, inhibits G $\beta\gamma$ as well when it is bound to G α_i ⁹⁰. G $\beta\gamma$ also regulates phospholipase C β , K⁺ channels and voltage gated Ca²⁺ channels⁷⁴ and the PI3K/Akt pathway⁹⁰ and can hence also affect G α_q signalling.

1.2.8 Non-canonical signalling pathways

Among non-canonical signalling pathways, arrestin signalling is the most studied. Arrestins can block G-proteins and lead to GPCR internalization, hence abrogating GPCR signalling, or on the contrary enable additional signalling pathways by providing a scaffold for MAPK when bound to the receptor⁷⁴. Similar to GPCRs' ability to bind different types of G α , receptors are able to switch between G-proteins and other effector proteins. Which pathway will be activated depends strongly on the ligand bound to the receptor, since different ligands can induce different conformational changes. The conformational changes affect scaffolding proteins and will affect which molecules can be bound on the intracellular site of the receptor. If a ligand typically leads to increased activation of one pathway over other pathways, this is called biased agonism^{74,93}. The ability of different ligands to elicit different responses while binding to the same receptor has to be considered in discussing a receptor's function.

To conclude, GPCR signalling is a complex process whose outcome is dictated by its various participants, notably the ligand, the GPCR structure, the transducers and potentially other

receptors the GPCR or its transducers may interact with. Each participant in this interaction adds a layer of complexity. In addition, the cellular context plays an important role, leading to limitations of transferring insights from one cell type to another with regards to the function of a receptor.

1 Introduction

1.3 GPR84

1.3.1 Discovery

GPR84 was discovered independently in 2001 by two research groups. Wittenberger et al. (2001) employed a data mining strategy of conserved GPCR sequences against an expressed sequence tags (ESTs) database and found the gene sequence of GPR84⁹⁴. Yousefi et al. (2001) conducted a qPCR scan of human peripheral blood neutrophil-derived RNA converted to cDNA using degenerate primers aimed at sequences conserved in chemokine receptors and discovered GPR84⁹⁵.

1.3.2 Receptor structure

The GPR84 protein sequence has been analysed and despite similarities to known receptors, differences in consensus sequences lead Wittenberger et al. (2001) to conclude that it cannot be classified into any of the known GPCR families⁹⁴. Yousefi et al. (2001) compared its sequence to several other chemotactic receptors and found it to be only distantly related⁹⁵. GPR84 has been described to have an orthosteric and two allosteric binding pockets. The binding pockets and binding to it was investigated by Mahmud et al. (2017), who found that apart from MCFA more complex compounds could bind to the orthosteric pocket as well. Antagonists and a different class of agonists couple to different allosteric binding sites⁷⁷. When an agonist was bound at the allosteric binding site, efficacy of orthosteric binding was increased. Binding of diindolylmethane (DIM) to the allosteric pocket lead to inhibition of forskolin induced cAMP release similar to the inhibition caused by binding of embelin to the orthosteric site⁷⁷.

1.3.3 GPR84 ligands – putative physiological ligands

Wang et al. (2006) screened for the ligands of GPR84 using transfected cells and chimeric $G\alpha$ proteins in a Ca^{2+} mobilization assay and found MCFA of chain length C9-C14 to be the most potent agonists. Chimeric G-proteins are hybrids of different types of $G\alpha$ -proteins, for example a $G\alpha_q/G\alpha_i$ chimeric protein can be activated via its $G\alpha_i$ portion but elicit $G\alpha_q$ signalling, e.g. an increase in Ca^{2+} . The MCFA decanoic acid (C10:0), undecanoic acid (C11:0) and dodecanoic acid (C12:0) had the strongest effects. These results were further confirmed by $GTP\gamma S$ incorporation into the membrane upon stimulation, and inhibition of forskolin-induced cAMP increases that could be reversed by pertussis toxin⁹⁶. These MCFA occur naturally in coconut oil, palm kernel oil, dairy products and breast milk, and are

thought not to be produced by the human body itself, with exception of the mammary gland during lactation⁴⁷. Later on, Suzuki et al. (2013) claimed that hydroxylated MCFA (2-hydroxy decanoic acid, 3-hydroxy decanoic acid, 2-hydroxy dodecanoic acid and 3-hydroxy dodecanoic acid) were ligands based on GTP γ S incorporation into the membrane upon stimulation and an IP assay. A study into the binding of hydroxylated dodecanoic acid confirmed the binding of 2-hydroxy and 3-hydroxy dodecanoic acid to GPR84, and found 4-hydroxy and 12-hydroxy dodecanoic acid to be ligands as well⁹⁷. The latter observation is in contradiction with results obtained by Suzuki et al. (2013). Hydroxy-MCFA can be components of LPS or occur when MCFA are hydroxylated as part of β -oxidation in mitochondria^{44,98}. Some hydroxy-MCFA might also occur as quorum sensing molecules from bacterial strains. Schulze et al.(2022) examined 2-hydroxy-C10:0 in the cis and trans enantiomer and concluded that the trans-enantiomer had a higher potency, while both could activate the receptor (cAMP inhibition assay, EC50 2-cis-hydroxy decanoic acid 12 μ M, EC50 2-trans-hydroxy decanoic acid 4 μ M)⁹⁹. There is no consensus yet if MCFA are the physiological ligands. This is due to the relatively low concentrations at which MCFA are found in the blood of healthy humans, and the relatively high doses of MCFA required for receptor activation and downstream effects^{69,100}. The feasibility of (dietary) MCFA as physiological GPR84 ligands has been evaluated sceptically and thoroughly by Luscombe et al. in 2020. Apart from the already mentioned doubts due to concentrations, a relatively low affinity and efficacy of the receptor to these ligands are cited and a lack of spatiotemporal overlap in presence of GPR84 and dietary MCFA¹⁰¹. Most papers employ alternative ligands, both natural and synthetic, to investigate GPR84 signalling and function.

1.3.4 GPR84 ligands – synthetic ligands

There are only few other natural agonists documented for GPR84: embelin, diindolylmethane (DIM) and a fatty acid produced by a marine bacterium. Embelin is a metabolite originally isolated from the plant *Embelia Ribes* and used as a traditional Chinese medicine. It was discovered as an efficacious ligand by Gaidarov et al. (2018) and able to mediate both cellular signalling and cellular functions, as discussed in the sections on GPR84 signalling and GPR84 expression and function⁹². In short, embelin can both affect cAMP signalling and β -arrestin recruitment^{92,102}. Even though embelin is a natural ligand,

1 Introduction

its physiological occurrence in the human body seems unlikely. Diindolylmethane is a brassica derivative and could hence occur in the body after ingestion and digestion of different types of cabbage. The fatty acid produced by a marine bacterium contains a cyclopropane ring and had no effect on cAMP signalling. It did lead to β -arrestin recruitment 53- and 25-fold more potent than decanoic and dodecanoic acid, respectively, and had 54% potency compared to embelin. The synthesis of cyclopropane containing MCFA remains to be further elucidated and its occurrence in a wider population of bacteria is unsure¹⁰². Synthetic ligands for GPR84 are numerous and consist of all different types of ligands, including biased ligands. Structurally several of the natural and synthetic ligands show an aliphatic carbon chain similar to MCFA, including ZQ16, embelin and 6-n-octylaminouracil (6-OAU)^{69,92}. Experiments on embelin, which consists of a hydrophilic head group and the aforementioned carbon chain with a chain length of 11, explored the importance of the carbon chain by testing alternative versions with different carbon chain lengths. A U-shaped relationship between carbon chain length and receptor activation as measured by cAMP inhibition was found. In other words, embelin variants with a very short (<C6) or very long alkyl chain (C13-C15) showed the least activation, and alkyl chains of medium length the highest (C7-C11). Embelin variants with extreme chain length (C3 and C15) showed no detectable activity⁹². In contrast to this finding, other ligands including DIM and several antagonists are composed entirely of indole and benzene rings. These differences might be explained by the binding pockets they bind to, but need further elucidation⁷⁷. A variety of derivatives of DIM were also found to activate GPR84, with the changes to molecular structure impacting the downstream activated signalling pathway¹⁰³.

1.3.5 G-protein signalling

GPR84 has been found to couple to $G\alpha_i$ proteins according to several studies. To start with, Wang et al. (2006) found GPR84 mediated Ca^{2+} mobilization in transfected CHO cells occurred in presence of the chimeric G-protein G_{q19} but not in presence of G_{q55} indicating the requirement for the $G\alpha_i$ subpart. In absence of chimeric proteins, agonists did not induce Ca^{2+} mobilization, indicating as well that it is not a $G\alpha_q$ coupled receptor. MCFA-GPR84 signalling inhibited a forskolin-induced cAMP increase, which could be abrogated by pertussis toxin⁹⁶. In a separate study, *Gpr84*^{-/-} bone marrow derived macrophages (BMDM) and biogel-elicited peritoneal macrophages showed significantly higher cAMP levels after

forskolin treatment, both in absence and presence of LPS, than their *WT* counterparts, showing that GPR84 signalling suppresses cAMP¹⁰⁴. In another set of experiments it was examined if GPR84 could couple to $G\alpha_{12/13}$ proteins using chimeric G-proteins consisting of a $G\alpha_s$ protein hybridized with either $G\alpha_{12}$ or $G\alpha_{13}$. HEK293 cells, an embryonic kidney cell line, transfected with both GPR84 and either $G\alpha_s$ - $G\alpha_{12}$ or $G\alpha_s$ - $G\alpha_{13}$ and treated with embelin, showed modest dose-dependant increases of cAMP. These increases were potentiated by treatment of pertussis toxin, which silenced $G\alpha_i$ -mediated inhibition of cAMP. That the observed increase in cAMP was GPR84 mediated was confirmed by repeating the experiment on parental HEK293 cells using the same chimeras. In this instance, no cAMP increase was observed⁹².

Interestingly, human monocyte derived macrophages (HMDM) showed dose-dependant increases in cAMP upon agonist treatment in presence of either PGE₂ or HE-NECA, which are G_s-protein coupled receptor agonists. This effect was abrogated by pertussis toxin, indicating its dependance on $G\alpha_i$ signalling. Hence a crosstalk between $G\alpha_s$ receptors and GPR84 elicited $G\alpha_i$ signalling seems likely to explain these results. To explore this further, assays to investigate the downstream effectors of GPCR signalling were investigated. Embelin was found to induce both Erk and Akt phosphorylation dose-dependantly in IFN- γ primed macrophages, and this process was pertussis toxin sensitive. In addition, using the PI3K inhibitor wortmannin lead to inhibition of Akt phosphorylation. This indicates that the PI3K/Akt pathway was activated by embelin in a $G\alpha_i$ -dependant manner, probably by $G\beta\gamma$ signalling which couples to PI3K. Even though the authors did not have a negative control like *GPR84*^{-/-} cells or a GPR84 specific antagonist, they corroborate this finding by demonstrating that another GPR84 agonist had the same effect⁹². Both agonists were demonstrated previously to be highly specific for GPR84. A separate study conducted in transfected CHO cells also found ERK signalling induced by GPR84 activation, using both decanoic acid and hydroxydecanoic acid as ligands¹⁰⁵. In contrast to these findings, GPR84 was found to be able to signal through $G\alpha_{15}$ in macrophages and neutrophils in a more recent study utilizing a GPR84 and GA15 knockdowns¹⁰⁶. In this study, GA15 knockdown led to reduced ERK phosphorylation and Ca²⁺ signalling in response to 6-OAU. In addition to G-protein signalling, GPR84 has also been found to be able to activate the β -arrestin pathway. This non-canonical signalling pathway leads to some overlapping functions with the

1 Introduction

canonical G-protein signalling. Some GPR84 ligands are able to activate both pathways, whereas others are biased towards one pathway. Pillaiyar et al.(2017) studied the effects of different diindolylmethane derivatives and found that while DIM potently activates both pathways, several derivatives had a strong bias to either the canonical or non-canonical pathway¹⁰³.

1.3.6 Regulation of *GPR84* expression

It is important to note that *GPR84* expression at a basal level is thought to be low, but significantly upregulated in presence of various inflammatory stimuli, including pro-inflammatory cytokines TNF- α , IL-1 β , IFN γ , presence of viral or bacterial agents, or PAMPs including LPS and zymosan^{100,107,108}. It is not upregulated in response to IL-6 or PGE₂¹⁰⁷. The upregulation in response to LPS was found to be partially TNF- α and IL-1 dependant, since genetic ablation of these pro-inflammatory cytokines significantly reduced *Gpr84* expression in a microglial model¹⁰⁷. The variation in PAMPs and cytokines able to elicit expression indicates that *GPR84* expression is regulated by pro-inflammatory pathways activated by TLRs. Previous promoter analysis and experiments utilizing NF- κ B inhibitors indicate a TNF- α – NF κ B axis leading to GPR84 upregulation in murine 3T3-L1 adipocytes. In this experiment, *GPR84* expression could be induced by TNF- α , but was abrogated when I κ B, a NF- κ B inhibitor, was administered¹⁰⁹. In a similar experiment, both TNF- α and LPS could upregulate GPR84 expression in a microglial cell line despite the presence of dexamethasone, a glucocorticoid. Dexamethasone can disrupt NF- κ B signalling by abrogating the interaction of its p65 unit with interferon regulatory factor 3 (Irf3), which affects roughly 50% of all NF- κ B regulated genes¹⁰⁷. Taken together, these findings indicate a role for GPR84 as part of the inflammatory response.

1.3.7 Expression in tissues and cells

The first study aiming to characterize the receptor found it to be predominantly expressed in immune cells, more specifically neutrophils, eosinophils, CD4+ T-Cells, CD8+ T-cells, B cells, RAW264.7 cells, which is a murine cell line with similarities to macrophages, and resident peritoneal macrophages. Mice ablated from the *GPR84* gene did not show any changes to the overall number and distribution of these different cell subtypes compared to their wild type (WT) peers and *ex vivo* no difference in T- and B-cell proliferation in response to stimulants was found¹¹⁰. The expression of *Gpr84* in different tissues was also

examined by Recio et al. in mice injected with LPS and sacrificed after either 2 or 8h. In both time points, significantly higher *Gpr84* expression was found in adipose tissue, bone marrow, brain and lungs compared to mice injected with sterile salt solution. In addition, after 8h expression in kidney and intestines was significantly higher, too. The strongest upregulation was found in bone marrow, hence indicating a strong increase in expression of *Gpr84* on myeloid cells¹⁰⁰. This is in concordance with the previous finding by Venkataraman et al. (2005) that GPR84 is expressed on neutrophils, eosinophils and macrophages¹¹⁰. Also, an experiment with THP-1 cells, which are a human transformed cell line thought to mimic monocytes and macrophages, showed a stronger elevation of *GPR84* expression in response to LPS in adherent, differentiated THP-1 cells than in non-adherent THP-1 cells. This indicates expression in both monocytes and macrophages, albeit higher in adherent macrophages than monocytes¹⁰⁷. In addition, GPR84 expression in brain and fat tissue was confirmed later by separate studies. Microglia, which are a subtype of macrophages present exclusively in the brain, exhibited upregulated *Gpr84* expression in both LPS-treated mice and an experimental autoimmune encephalomyelitis (EAE) model¹⁰⁷. Nagasaki et al. (2012) showed the expression of *Gpr84* in response to inflammatory stimuli in 3T3-L1 cells, a murine adipocytic cell line¹⁰⁹. In a mouse model of lung injury, *Gpr84* was significantly upregulated in alveolar macrophages, specifically in a subpopulation found to secrete increased amounts of pro-inflammatory cytokines and have enhanced phagocytic ability¹¹¹. Another study found GPR84 to be present in the epithelial lining of the oesophageal mucosa in reflux esophagitis in an animal model and in patient tissues¹¹². Further studies indicate a role of GPR84 in pain, in brain-related diseases, in kidney injury and fibrosis, in skeletal muscle and in the adipose tissue^{100,104,107,109,113,114}. Upregulation of GPR84 expression is also often found in infectious diseases or infection models, including patients suffering from COVID-19^{115,116}, sepsis^{32,117}, *Brucella* infection of a murine cell line¹⁰⁸, treatment of whole blood with superantigens. In a study of cancer and immunotherapy, GPR84 expression was found to be increased in several types of cancer tissue compared to normal tissue, its expression correlated with the expression of inflammatory genes in these tissues¹¹⁸. At least two studies identified GPR84 to be part of gene classifiers for sepsis^{32,117}, one study found it as part of a gene identifier for Alzheimer's¹¹⁹ and another study found that GPR84 is part of a prognostic gene classifier indicating worse outcome and higher immune activation in glioblastoma¹²⁰. Despite GPR84

1 Introduction

expression being widespread, it seems to be consistently associated with inflammation. One exemption is the increased presence of GPR84 in the epithelial cells of the gastric mucosa in breastfed mouse pups that disappears upon weaning; in this context GPR84 expression appears to be diet-dependant¹²¹.

1.3.8 Functions of GPR84 - Cytokine and chemokine secretion

GPR84 affects cytokine and chemokine secretion upon activation as indicated by a variety of mostly *in vitro* experiments. The first to research this was Venkatamaran et al. (2005) who extracted T cells from WT and *Gpr84*^{-/-} mice and stimulated them *ex vivo* with lipopolysaccharide (LPS) or antibodies leading to an increase in IL-4 secreted from *Gpr84*^{-/-} T cells, but no differences in IFN- γ , IL-2 secretion or NFAT, STAT6 or NF κ B as detected by ELISA and western blot, respectively¹¹⁰. This indicates a very limited role of GPR84 in inflammatory immunoregulation of the adaptive arm by regulation of IL-4 secretion. GPR84's role in immunoregulation appears to be more prevalent in the innate arm of immunity. It is important to note that inflammatory stimulation is required and will alter the outcome of GPR84 activation. *In vitro* experiments utilizing PMNs and U937 macrophages showed that in presence of LPS, GPR84 ligands mediated significantly increased IL-8 (PMNs), a chemokine that attracts neutrophils, or TNF- α production (U937), as measured by ELISA. The same experiments without LPS showed significantly less production of cytokines. As a control the experiment was repeated using a siRNA knock-down technique which decreased *GPR84* expression by 90% in U937 cells; these siRNA treated cells showed significantly less production of TNF- α in presence of ligand compared to control cells. There was no significant difference in TNF- α production between control and knock-down cells in absence of ligand. It can thus be concluded that ligand-GPR84 signalling enhances LPS induced TNF- α secretion of these cells¹²². Müller et al. (2017) also investigated the modulation of *TNF- α* expression by GPR84. To start with, they confirmed the significant increase of GPR84 on the cell surface 24 and 48 hours after LPS stimulation in a glycoproteomic analysis of both THP-1 cells and monocytes obtained from healthy volunteer blood. Furthermore, they added MCFA to monocytes and THP-1 cells in presence and absence of a pre-stimulation with LPS and a second addition of it. When adding a high dose of MCFA to pre-stimulated monocytes this significantly increased *TNF- α* expression, both in presence and absence of a second dose of LPS. In absence of MCFA pre-stimulated

monocytes and THP-1 cells' *TNF- α* expression in response to LPS was significantly lower than in cells which had not been stimulated before. Unfortunately, no control was included in which GPR84 signalling would be inhibited or ablated, and hence it cannot be concluded for sure that the effect was mediated by GPR84. The authors describe this experiment as a model for endotoxin tolerance in monocytes¹²³. Other cytokines have been found to be altered by GPR84 signalling as well. After stimulation with LPS and treatment with MCFA, RAW264.7 cells, a murine macrophage cell line, secrete IL12p40 in a MCFA dose-dependant manner⁹⁶. A gene expression microarray performed on murine *Gpr84*^{-/-} and *WT* peritoneal macrophages stimulated with LPS showed significantly higher expression of *Tnf- α* , *Il-6*, chemokine (C-C motif) ligand 2 (*Ccl2*), *Ccl3* and nitric oxidase 2 (*Nos2*) in wild type macrophages. Even though not significant, after induction with LPS expression of prostaglandin synthase E2 (*Ptges2*) and *Il12b* was 1.5 times higher in WT cells, and expression of *Il10* was 1.5 times lower¹⁰⁴. In the murine cell line MH-S, DIM stimulation of LPS-treated cells led to significantly higher expression levels of several inflammation and chemotaxis related genes, including *Il6*, *Il12b*, and *Cxcl2*. Similar results were found in murine BMDMs: stimulation of *WT* BMDMs with 6-OAU and LPS led to significantly higher expression of *Tnf- α* , *Il6*, *Ccl2*, *Il12b*, *Ccl5*, *Cxcl1* and Fc γ receptor 1 (*Fc γ 1*) compared to vehicle. This effect was not observed in *Gpr84*^{-/-} cells. These results were confirmed on a protein level for TNF- α , IL6 and CCL2 by ELISA. In a separate study, peritoneal exudates of *Gpr84*^{-/-} and *WT* mice were challenged with LPS, *WT* immune cells secreted significantly more IL6, CXCL1, CXCL2 and neutrophil gelatinase associated lipocalin (NGAL)⁹². In contrast, experiments on LPS-stimulated primary murine microglia in cell culture found several pro-inflammatory cytokines, including *Il-6*, *Tnf- α* , *Il-1 β* and *Cxcl1*, to be unaffected or even significantly decreased when GPR84 ligands were added. Addition of agonist also lead to less translocation of p65 into the nucleus, hence indicating that NF- κ B signalling was decreased by GPR84 signalling in microglia¹¹³. This difference between microglia and macrophages might be due to cell-specific functions. The cytokine and chemokine profile indicates an important role for GPR84 in the onset of inflammation including in the chemotaxis of cells and the secretion of inflammatory cytokines by monocytes, macrophages and neutrophils.

1 Introduction

1.3.9 Chemotaxis

Suzuki et al. (2013) found *GPR84* expression to be high in human polymorphonuclear cells (PMN), in other words neutrophils, eosinophils and basophils, and macrophages. They hypothesized that GPR84 would have a chemotactic function in these cells and confirmed this using *in vitro* transwell migration assays and *in vivo* using the rat air pouch model. In the transwell migration assay PMN prepared from human blood, Chinese hamster ovary cells (CHO) transfected with *GPR84* and differentiated U937 cells, a macrophage-like cell line, exhibited migration towards both a synthetic ligand for GPR84 and MCFA (2-OH-C12:0, 3-OH-C12:0). CHO cells without the *GPR84* insert did not show any migration. When inoculating the synthetic ligand into an air pouch in the skin of a rat, PMNs and macrophages showed significantly increased numbers in the pouch after 4 hours compared to a pouch inoculated with vehicle. Furthermore, injection of the synthetic ligand into the jugular vein increased chemokine (C-X-C motif) ligand 1(*Cxcl1*) expression, the rodent homologue to IL-8, but no increased expression of various tested pro- and anti-inflammatory cytokines. It is important to note that both migration and all animal experiments were conducted in absence of an inflammatory stimulus¹²². Chemotaxis of neutrophils in response to a natural ligand (embelin) was also demonstrated by Gaidarov et al (2018) in a transwell migration assay⁹² and by Sundqvist et al. to a synthetic GPR84 ligand (ZQ16)⁶⁹. Puengel et al. examined migration of human neutrophils, murine neutrophils and murine monocytes in response to embelin and MCFA and found significant chemotactic activity of all cells in response to embelin, and of human neutrophils to MCFA¹²⁴. In contrast to these findings no increased migration of monocytes or macrophages to nerve injury occurred in *WT* mice compared to *Gpr84*^{-/-} mice¹⁰⁴.

1.3.10 Phagocytosis

Phagocytosis is an effector function which is important in the eradication of pathogenic microorganisms. Phagocytosis of fluorescent bacteria by BMDMs was increased in presence of 6-OAU in wild type cells, but not in *Gpr84*^{-/-} cells¹⁰⁰. Also, transfection of RAW 264.7 cells to overexpress zebrafish GPR84 lead to significantly increased phagocytosis of fluorescent bacteria¹²⁵.

1.3.11 Granule secretion

Another important effector function is the release of ROS and enzymes from granules, especially in neutrophils. Sundqvist et al. researched this using human blood-derived neutrophils and the synthetic GPR84 ligand ZQ16. They found a minor though significant effect of ZQ16 on granule mobilization in unstimulated cells. TNF- α primed neutrophils secreted significantly more ROS than naïve neutrophils when treated with ZQ16 or the positive control. Mediation of this effect through GPR84 was confirmed by demonstrating the inhibition of ROS secretion when adding a GPR84 antagonist, which had no effect on the positive control⁶⁹. Another study on human blood derived neutrophils also found ROS production in response to ZQ16 and, albeit weaker, another synthetic ligand. This study found that both the canonical, G-protein coupled, and the non-canonical β -arrestin mediated pathway could lead to ROS production mediated by GPR84 in neutrophils¹²⁶. A similar experiment in human primary neutrophils showed no ROS secretion in response to GPR84 agonist embelin alone, but addition of embelin and known inducers of ROS lead to enhanced ROS secretion compared to ROS induction without embelin⁹². Peters et al. also found ROS release in inflammatory THP-1 cells in response to decanoic acid and hydroxylated decanoic acid, confirmed to be GPR84-mediated using a knockdown¹⁰⁶. In response to the same ligands, peripheral mononuclear cells showed increased NET formation as evidenced by loss of nuclei and staining of DNA outside of cells.

1.3.12 Pre-clinical studies on GPR84 antagonism as remedy for inflammation and fibrosis

The antagonistic inhibition of GPR84 has been explored in various studies as a mean to reduce fibrosis and/or inflammation and alleviate their consequences. Fibrosis is the scarring of inflamed or damaged tissue when extracellular tissue accumulates leading to stiffening and thickening of the tissue¹¹⁴. Fibrosis can occur as a result of different inflammatory or metabolic diseases and commonly impacts the functioning of the affected tissue. In a heart failure rat model, rats were treated with PBI-4050, a synthetic antagonist that has been shown to inhibit GPR84 activity and its expression¹²⁷. This led to improved heart function and a reduction in restrictive respiratory lung syndrome and lung scarring compared to vehicle treated controls. The same compound, PBI-4050, has also been tested in murine fibrosis and found to alleviate fibrosis in several different murine fibrosis models, including chronic kidney disease models, liver injury, pulmonary fibrosis^{114,128}. The

1 Introduction

causative effect of GPR84 on fibrosis was confirmed using KO mice, since wild type mice showed more severe fibrosis than KO mice in the fibrosis model. That the alleviation of fibrosis by the antagonist was mediated via GPR84 was also confirmed using KO mice, since fibrosis was not alleviated by the antagonist in the KO mice. In some fibrosis models though a different receptor, GPR40, was found to play a role as well. Both GPR40 and GPR84 can bind PBI-4050. A study with a similar antagonist, PBI-4547, investigated the effect of this antagonist on metabolic disruptions and non-alcoholic fatty liver disease (NAFLD)¹²⁹. In this study, mice were fed a high-fat-diet in a placebo and an antagonist treated group. Mice treated with the antagonist had better insulin sensitivity, lower body fat and lower fasting glucose levels than controls on the same diet. In addition, the treated mice exhibited few indicators of NAFLD, including significantly reduced steatosis in the liver and reduced liver triglyceride levels. Since PBI-4547 can bind other receptors apart from GPR84, *Gpr84*^{-/-} mice were used to establish if the effects were mediated by GPR84. PBI-4547 had no effect in KO mice compared to vehicle treated KO mice. Another study into NAFLD used yet two other antagonists, CpdB and CpdA¹²⁴. In their fibrosis mouse model, fibrosis was also ameliorated by their GPR84 antagonists, similarly to the other study. Antagonist treated mice had significantly less infiltration of monocytes and neutrophils into their livers upon induced injury and when initiating NAFLD using a diet model. In addition, histologically, less ballooning and less fibrosis was present in antagonist treated mice. In this study no knock-out was used as negative control, however specificity of the ligand to GPR84 was examined and confirmed. Further studies investigated the use of antagonists to treat colitis. A later study sought to identify the applicability of an antagonist on inflammatory bowel disease¹³⁰. The compound under study proceeded to phase II trials after successfully reducing neutrophil infiltration and inflammation scores in IBD mouse models. In phase II trials in humans the antagonist did not reach the desired efficacy, however it is currently being tested in phase II trials for idiopathic pulmonary fibrosis¹³¹. In a separate study, *Gpr84*^{-/-} mice developed significantly less severe colitis in a dextran sulphite sodium (DSS) model than their WT controls¹³². When treating WT mice with an antagonist in the same DSS colitis model, inflammation and severity were significantly lower than in vehicle treated mice. The severity was determined by weight loss, immune cell infiltration and mucosal damage, all of which were less severe in both KO and antagonist-treated mice.

To sum up the findings on GPR84's role in immune function, there are strong indications that it is a chemotactic, pro-inflammatory receptor that only exerts its function in inflammation primed cells.

1.4 “Omics” and other high throughput approaches to decipher biological activity

1.4.1 Transcriptomics as window into cellular activity and regulation

There are several molecular systems entities one can analyse to understand the biology of the body better: the genome, the transcriptome, the proteome and the lipidome. While the body's genetic code is relatively stable, the transcriptome, proteome and lipidome are acutely dynamical systems, that are under stringent control, responding to different physiological and cell specific situations. Great advances have been made in the analysis of these different “-omics” and “multi-omics” are popular as comprehensive approaches to analysing cellular and tissue-specific activity.

In transcriptomics, all mRNA transcripts are quantified using methods such as microarrays or RNAseq technologies. Analysing the mRNAs present will provide a snapshot of a cell's or tissues' transcriptional activity. In addition to this untargeted approach, a targeted approach to analyse a fraction of the transcriptome is possible by utilizing qPCR for target genes only. By comparing the transcriptome of, for example, sepsis patients with non-sepsis controls, the gene expression significantly altered in sepsis can be revealed. The gene expression is an indication but not proof of the presence of proteins and hence provides a prediction of cellular activities at play. Gene expression is tightly regulated on several levels. First of all, transcriptional regulation by transcription factors will activate or repress the transcription of genes in a context-specific manner, secondly microRNA can inhibit gene translation on a post-transcriptional level as microRNA can capture or alter mRNA decreasing or abolishing its function. Thirdly, on a post-translational level proteins are inactivated or activated by post-translational modification or broken down at differing rates, and these rates and modifications depend on the physiological situation, e.g. post-translational modifications to a protein can alter its stability. Thus, quantity of mRNA transcripts of a gene is not strictly proportionate to presence of protein and in fact no direct indication of protein activity. Nevertheless, transcriptomic studies in combination with statistical analyses can yield important insights into cellular and physiological processes and further our understanding of complex processes.

1 Introduction

1.4.2 Big data

The availability of publicly available transcriptomic data in the 21st century in science is a proud example of collaboration within the scientific community. Many transcriptomic datasets are shared online either as raw or processed datasets or as part of curated databases. These datasets and other datasets generated from high-throughput studies enable researchers to address diverse questions without entering the lab. By combining several datasets or approaches, very powerful analyses are possible. This however requires both technical know-how and computing power. However, the collaboration within big data analysis does not only consist in the sharing of data, but also in the making available and freely sharing of tools. Many research groups and institutions generate and provide sophisticated tools that make complex analyses possible without requiring programming knowledge or -power from the individual researcher using their tools. Instead, curated databases can be queried using online interfaces or free-to-use software packages, and valuable hypothesis generating insights are only the click of a mouse away.

1.4.3 Transcriptional regulation of gene expression

The regulation of gene expression by transcription factors is an important regulatory mechanism that allows our body to tailor the function of its cells to their specific context. Transcription factors (TFs) can activate, enhance or repress the transcription of genes and thereby regulate the presence of proteins that exert their function within the cell. They are defined as proteins able to bind to DNA in a sequence-specific manner that regulate gene expression. They do so both in a tissue-specific way, as about 1/3 of all transcription factors are expressed tissue-specific, and in response to stimuli¹³³. For example, in inflammation transcription factors will be activated downstream of inflammatory signalling cascades; and these transcription factors will initiate the transcription of genes required for the inflammatory response. This leads to the production of cytokines, chemokines, and other mediators, the priming of immune cells for their effector functions and is required for the onset of resolution as well. Since some transcription factors are only present in certain cell types, different cell types will partially express different genes and hence exert different functions. TFs initiate transcription by binding to their binding site in the promoter region of a gene, which is a stretch of DNA upstream of and very close to the transcription site of the gene. In addition, transcription factors binding to so-called enhancers or silencers regulate transcription by facilitating or disabling, respectively, access of transcription

factors to the promoter and of the RNA polymerase to the DNA, e.g., by moving or disassociating histones from the chromatin and hence opening up the chromatin. While in simpler organisms genes tend to be regulated by singular TFs binding to longer stretches of DNA, in eukaryotes with big genomes genes are usually regulated by several transcription factors in cooperation that each bind to relatively short stretches of DNA¹³⁴. The cooperation of transcription factors can occur as direct protein-protein interaction where dimerization of TFs (or assembly of even more than two) is necessary for binding to the DNA, or as indirect protein-protein interaction where the binding of several transcription factors to DNA is necessary to open up the chromatin by moving or removal from the histone. The latter is called collaborative competition or indirect co-operativity and often requires the TF binding sites (TFBS) to be close. An interesting example of this is the occurrence of repeated TFBS for the same TF close to transcription initiation site, also called homotypic clusters. TFs can however also collaborate when further apart, when binding of one TFs leads to conformational changes of the DNA making a distant stretch of DNA accessible to other transcription factors¹³⁵.

1.4.4 Analysis of transcriptional regulation

There are several ways to analyse transcriptional regulation of genes in the laboratory and *in silico* tools are based on insights and data gained from these experiments. Typically experiments either examine chromatin openness and mechanism relating to it, or the binding and actions of transcription factors themselves. The majority of data online stems from Chip experiments. In Chip experiments, DNA-binding proteins are crosslinked with the DNA and precipitated using antibodies. The genomic locations attached to precipitated TF-antibody complexes will be analysed using either sequencing (ChIPseq), qPCR (ChIP-qPCR), microarray (ChIP-Chip), or sequencing after exonuclease treatment (ChIP-exo)¹³³. Chip experiments allow for identification of genomic locations specific TF bind to and are high-throughput approaches. Transcription factors can also be analysed *in vitro* and *in vivo* using genetic mutation of model organisms or cell lines, for example KO models, knock-in or overexpression of transcription factors followed by transcriptomic analysis of gene expression. This approach will show which genes are regulated by one transcription factor. An example of experimental examination of the openness of chromatin is the mouse atlas of cis-regulatory genes that was established by combining RNAseq for transcriptomics and ATAC-seq for chromatin accessibility on mouse tissues¹³⁶. ATAC-seq relies on the

1 Introduction

preferential binding of a transposase (enzyme) to open chromatin, which fragments the DNA upon binding. The released DNA fragments will be sequenced allowing the identification of open chromatin regions. The analysis of this data and subsequent validation by ChIPseq, allowed the study authors to establish relationships between transcription factors and their expression, open chromatin regions and TFBS therein, and cell types. The data is available to browse via a web-interface that allows for the search of individual genes for their tissue-specific expression, open chromatin regions associated with the gene and predicted transcription factors based on TFBS in these open chromatin regions¹³⁷.

1.4.5 Measurements of metabolites using LC-MS/MS techniques

As part of this work, metabolites will be measured using LC-MS/MS, liquid chromatography – tandem mass spectrometry. In the following I will introduce how chromatography and spectrometry work with a focus on the specific techniques used in this thesis. Chromatography and spectrometry are distinct techniques that were developed separately; and are now often used coupled, as this enhances detection sensitivity. See Figure 9 for the depiction of a LC-MS/MS machine.

1.4.6 Liquid chromatography

In liquid chromatography, components in the sample are separated based on their physical and chemical properties. The sample will be injected onto the column and the column will be flushed with a liquid phase¹³⁸(Figure 9 C, left side). Sample components will adhere to the column and be released into the liquid phase based on their affinity to liquid phase and column. In order to release all components of interest, the composition of the liquid phase will be altered throughout the run. As a result, the components will release from the column at different times. There are different types of physical properties that can be used for separation, e.g., hydrophobicity or charge of the component. For example, if a column is coated with a hydrophobic compound, and the liquid phase is hydrophilic, hydrophobic compounds will have stronger affinity to the column than hydrophilic compounds that might not or barely adhere to the column. The hydrophobic compounds will be slowed down by their adherence to the column, while hydrophilic compounds will travel faster through the column. Hence, hydrophilic compounds will release earlier from the column than hydrophobic compounds. In order to facilitate release of the hydrophobic compounds

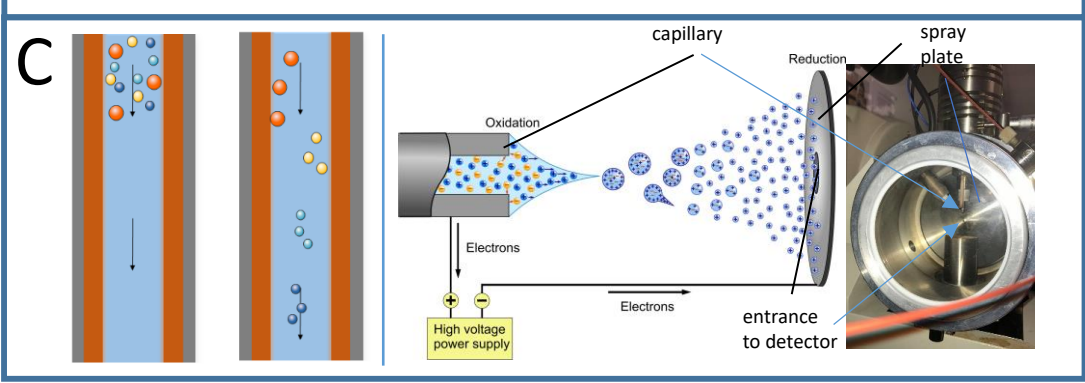
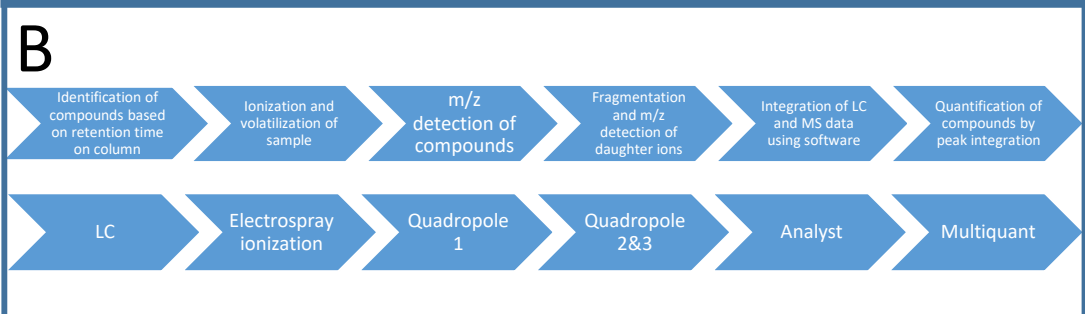
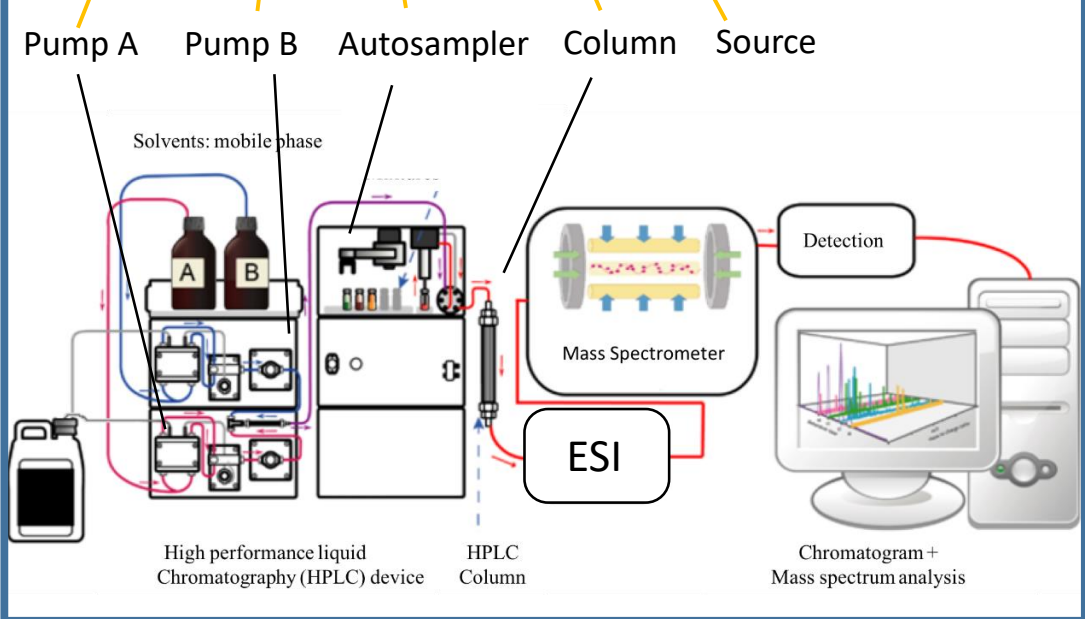
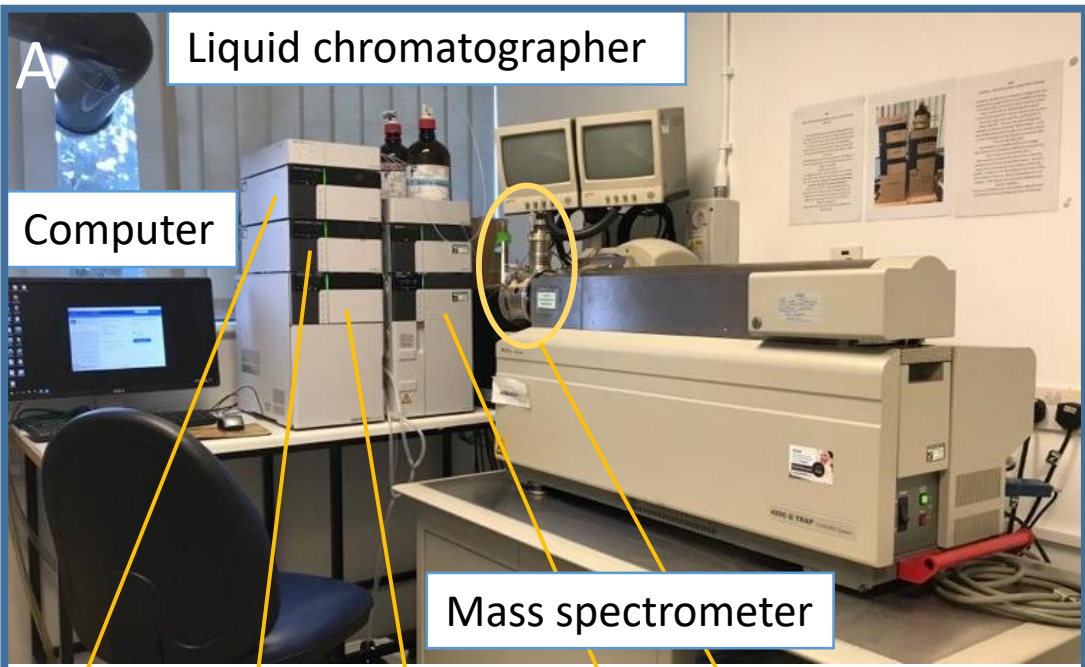
as well, the liquid phase will be adjusted to contain higher amounts of solvents that the hydrophobic compounds are soluble in. The type of columns and liquid phases will be chosen based on the type of compounds to be quantified. The liquid chromatography mechanism is illustrated in Figure 9. To quantify the chosen lipids, a column coated in apolar, hydrophobe C18:0 (stearic acid) was chosen. Hydrophobicity of lipids typically increases with the molecule size, e.g., the chain length of the fatty acids. In other words, smaller fatty acids are rather hydrophilic due to the short chain length and the acid group, and hence have a stronger affinity towards a hydrophilic liquid phase rather than the hydrophobic column coating. Bigger lipids have a stronger affinity to the column than to a hydrophilic liquid phase. By flushing with acidified water first, hydrophilic lipids are released in the beginning of the run. Increasing the percentage of an organic solvent, methanol in this method, enables the release of hydrophobic lipids over time. The time at which the compound is released from the column is called the retention time; since the retention time is specific to the compound within the method and can be used as part for the identification of compounds. Previously, liquid chromatography has been used by itself by connecting a detector behind the column. This approach would result in a chromatogram with peaks of detection signal intensity for every component and the time of release on the x-axis. Coupling liquid chromatography instead with mass spectrometry greatly improves the sensitivity of detection and the ability to distinguish and identify components.

1.4.7 Mass spectrometry

In mass spectrometry, compounds will be sensitively detected based on their mass to charge ratio, i.e., the mass of a compound in relation to its charge will be used to identify it. In addition, the amount of every compound can be sensitively calculated based on the measurements. Consequently, known compounds can be quantified precisely based on their molecular mass. Importantly, the mass spectrometer can only detect ions, in other words charged compounds. The ionization of the sample compounds is hence the first step that prepares the compounds for measurement. In addition, the ions have to be present in gas-phase. When the mass spectrometer is coupled to a liquid chromatographer, the liquid portion of the effluent need to be evaporated to release the ionized compounds of interest into gas-phase before measurement. Commonly in LC-MS, including this approach, electro spray ionization is used. Electrospray ionization turns the liquid sample gaseous and ionizes

1 Introduction

it simultaneously (Figure 9 C, right side). Eluate off the liquid chromatography machine is released in small amounts off the capillary into a chamber, called the source (Figure 9 A, right side of Figure 10 C). The stark voltage difference between the capillary and the walls of the source lead to an electrostatic field that turns the injected liquid into a fine mist of charged droplets. In positive ion mode, electrochemically generated cations migrate away from the positively charged capillary, while in negative ion mode the charges are reversed and anions migrate away from a negatively charged capillary. The ion mode will be chosen based on the compounds to be measured and their preference to lose or accept electrons. In other words, a compound unlikely to accept electrons is hard to detect in negative ion mode since it will not accept the negative charge and hence also not migrate through the chamber. The liquid in the droplets evaporates as the droplets pass through the chamber, in this machine under influence of both high temperature and nitrogen gas, releasing the ions into gas phase. The released ions will be detected by the detection system of the mass spectrometer; in this approach a linear ion trap quadrupole is employed. Every component will be detected with a specific mass to charge value. A component with two charges would for example have a m/z value corresponding to half its mass. Since in this approach all components measured only receive a singular charge, this does not require further consideration. In addition, tandem mass-spectrometry allows for further identification of the components. In tandem MS or MS/MS, the detected component is fragmented after initial measurement and the fragments or “daughter ions” are measured as well. Since the fragments a molecule breaks into are specific to the molecule, this adds certainty to the identification of the molecule. For an overview of the steps in LC-MS/MS, see Figure 9 B. To summarize, the MS-component of the system determines the mass and quantity of every component and its fragments, and the results are computationally combined with the results from the liquid chromatography.



1 Introduction

Figure 9 Overview of a liquid-chromatography-mass spectrometer. A) Photograph of the LC-MS machine used in this study, and a schematic representation of a LC-MS system. (attribution: Cwszot | Dagui1929 | CasJu | YassineMrabet, CC0, via Wikimedia Commons)

B) Flow diagram of the steps involved in quantification of components, and the machine parts and software required to execute these. C) shows the mechanism of liquid chromatography (left side) and the electro spray ionization (middle figure, attribution: Andreas Dahlin, CC BY 2.0 <<https://creativecommons.org/licenses/by/2.0>>, via Wikimedia Commons) and the machine part in which the electro spray ionization is carried out.

The described technical set-up of lipid analysis allows for different analysis approaches with different aims and outcomes. First off, the machine can be set up for targeted and untargeted analysis. In targeted analysis, the user defines which compounds they want to potentially detect and knows their retention times and m/z values. The machine will be programmed to only quantify compounds corresponding to the defined list of compounds, which leads to greater sensitivity of the results. Untargeted analysis, in contrast, will quantify all compounds with all retention times and all m/z ratios the machine is able to detect. This can lead to the identification of new compounds, but especially when analysing complex samples can lead to the presence of many peaks that may overlap. In this analysis, a targeted approach was chosen to focus only on lipids of interest. This means that lipids will be identified due to their retention time in combination with both their derivatized mass and the mass of the fragment.

1.5 Rationale

Sepsis still is and is increasingly a major cause of death all around the globe, with a heavy burden on the neonatal population. Gaps in our knowledge of the underlying complex mechanisms is hampering the progress in diagnosis and treatment urgently needed. To address these gaps requires a scientific approach that connects systematic changes with cellular signalling mechanisms and that bridges the domains of metabolism and immunity. An important player in cellular signalling is G-protein coupled receptors (GPCRs) that transmit signals from the surface to the inside of the cell and can regulate several processes. Two GPCRs have been shown to be significantly upregulated in neonatal sepsis, GPR43 and GPR84, and act as integral members of a gene signature that accurately predicts sepsis. Whereas GPR43 is well-researched and known to bind SCFA originating from the microbiota, GPR84 and its ligands and downstream effects still require further elucidation. GPR43 is an immunomodulatory GPCR upregulated in response to inflammatory stimuli, it dampens the immune response, is chemotactic and is present on neutrophils and T-cells. GPR84 is also upregulated in response to inflammatory stimuli, e.g., LPS, and thought to be pro-inflammatory and chemotactic and has MCFAs as putative, though disputed, ligands. It is expressed by a variety of cells, especially by myeloid cells. Here, I hypothesize that MCFA-GPR84 signalling axis is a key player in the regulation of the sepsis immune response. To shed light on the possible role of this axis in sepsis I will investigate its ligands, where and

1 Introduction

how GPR84 is expressed, and the possible source of MCFAs for downstream GPR84 signalling in an inflammatory context.

While MCFAs have been proposed as ligands for GPR84 and are present in the human body their concentrations in healthy volunteers have been consistently measured lower than concentrations necessary to activate GPR84. To date the measurement of MCFAs in the context of an inflammatory disease has not been reported. This raises a key unanswered question of whether MCFAs are altered systemically in sepsis. To investigate this central question, I will establish a protocol for reliable extraction and quantitative measurement by LC-MS of MCFA from whole blood. This protocol will be used to measure MCFA in healthy controls and septic neonates and adults.

In addition, this protocol will be used to inquire into potential local production of MCFA in inflammation. To date the cellular synthesis of MCFAs in mammals is known to occur in mammary gland epithelial cells and only during lactation. The enzyme responsible for generating MCFAs in mammals is OLAH and is thought to be restricted in its expression to lactating mammary tissue. It is not known whether immune cells have the capacity for regulating OLAH and producing MCFAs. Hence, I will perform further *in silico* investigations into the transcriptional regulation of *OLAH* and in which contexts and on what cells it is expressed, followed by the culturing of a selection of cells to measure MCFA production in the supernatant.

The last and very important question raised in the context of testing my hypothesis aims at better understanding the regulatory network governing expression of *GPR84* in sepsis. Understanding the transcriptional regulation will provide insights into the cell-specific and context specific expression. These investigations will be based on *in silico* analyses of available data sources.

All in all, the outlined approach will give important insights into the GPR84 signalling axis from ligand generation and regulation to key immune cell targets in the context of sepsis.

2 Materials and methods

All centrifugation steps were performed in a microcentrifuge or benchtop centrifuge from Thermo Fisher Scientific (Thermo Fisher Scientific, Waltham, Massachusetts, United States).

2.1 Optimization of a method for reliable lipid detection from blood samples

2.1.1 Materials

All lipids were bought at highest purity grade possible from Sigma-Aldrich (St. Louis, Missouri, United States). All solvents purchased were HPLC grade and purchased from Fisher Scientific (Waltham, Massachusetts, United States). Derivatisation reagents were purchased from Fluorochem (Hadfield, United Kingdom) (3-nitrophenylhydrazine (3-NPH)) and Sigma Aldrich (3-NPH, *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) and pyridine). For the quenching reaction, mercaptoethanol, succinic acid and formic acid were purchased (Sigma Aldrich, St. Louis, Missouri, United States).

2.1.2 LC-MS/MS conditions

All samples were run on a LC/MS system consisting of a Nexera XR LC-20AD XR (Shimadzu, Kyoto, Japan) liquid chromatograph coupled to a SCIEX 4000 QTRAP LC MS/MS system (Shimadzu, Kyoto, Japan) with Turbo Spray ESI.

Declustering potential, collision energy and CX were determined by tuning compounds on the MS. Derivatized 25 μ M stock solutions of every fatty acid were injected onto column using a syringe. The autotune option of the SCIEX analyst software (AB Sciex, Danaher, Washington D.C., U.S.A) automatically tuned and determined the optimal settings.

Software

Peak area calculation by integration of peaks, calculation of amounts on column after calibration curve integration and lipid/IS ratios were performed by MultiQuant software (Sciex, Danaher, Washington D.C., U.S.A). Statistical analysis and visualization of all results was performed in Excel 2016 (Windows Office, Redmond, Washington, U.S.A) and GraphPad Prism version 8.4.3 for Windows (GraphPad Software, San Diego, California USA).

2.1.3 Derivatization optimization

The derivatisation reaction as previously described by Han et al.¹³⁹ was conducted to find optimal reaction conditions. In short, a mix of primary MCFA standards was prepared in

Materials and methods

75% acetonitrile and 75% methanol. The mix was derivatized by addition of 200 mM 3-NPH solution and 120 mM EDC- 6% pyridine solution in either 50% acetonitrile or 50% methanol in a 2:1:1 ratio, and subsequent incubation for either 0.5, 1, 2 or 3h in a waterbath at 40°C. The reaction was stopped by incubation for 1 minute on ice. The reaction mixture was then filled up to 2ml with respectively 75% acetonitrile or 75% methanol and frozen at -20 °C for measurement the next day. On day of measurement samples were defrosted and centrifuged (10 min at 18 000 *g*) and supernatant was transferred to clean HPLC vials. All samples were measured by LC-MS/MS.

Calibration curves were made by generating a mix of primary standards at the highest concentration and dilution with solvent to the lower concentrations required. All dilutions were derivatized as described and measured by LC/MS in ascending order.

Effect and necessity for derivatisation of pyridine was established by derivatisation of a lipid mix with 3-NPH and EDC in the presence and absence of pyridine and subsequent measurement by LC-MS/MS.

The quenching method was chosen by derivatisation of an internal standard mix and subsequently using different quenching methods. The quenching methods constituted in addition of one or a combination of chemicals: 120mM mercaptoethanol solution diluted 1:1 with 75% MeOH, 120mM mercaptoethanol solution and 1M succinic acid solution, 1M succinic acid solution diluted 1:1 with 75% MeOH, 0.5% formic acid solution in 75% MeOH, and subsequent incubation for 30 minutes at 40°C in a waterbath. Levels of primary lipids and internal standards measured by LC/MS were compared among methods to determine reduction in derivatisation of contaminants and the effect of quenching on derivatized lipids.

2.1.4 Detection limits on the machine

To determine detection limits on the LC-MS/MS machine, calibration curves of internal standards were generated spanning a range of concentrations equalling 0.25 pg till 3 to 5000 pg of deuterated lipid injected onto column during measurement by LC-MS. Calibration curve samples were derivatized as described previously and measured by LC-MS/MS in ascending order of concentration.

2.1.5 Extraction from blood in PAXgene buffer

BloodPAX samples were obtained from a healthy volunteer under the ethics SMREC 16/04. Extraction methods were tested by addition of IS mix to a bloodPAX sample or a PBS in PAX sample. For the biphasic layer method, internal standard mix in 75% acetonitrile, 10µl 0.5% formic acid and 200ul ethyl acetate were subsequently added to 50µl sample. The sample was first vortexed for 10 minutes at 4°C, then centrifuged for 10 minutes at 18 000 *g* at 4°C. The top layer was transferred to a glass vial, evaporated in a SpeedyVac and re-suspended in 40µl acetonitrile. This suspension was derivatized and quenched with the above-described method and lipids were measured by LC-MS. For the protein crash method, internal standard mix and 450 µl of ethanol were added subsequently to the sample, followed by the same processing steps as described for the biphasic layer extraction from mixing by vortex till measurement on the machine. For the Bligh Dyer method, the previously published and commonly used protocol was followed¹⁴⁰.

2.1.6 Whole blood extraction

Whole blood samples with EDTA were obtained from a healthy volunteer under the ethics code SMREC 16/04. All extraction methods tested were amended with 10 second of mixing by vortex after addition of solvents and internal standards mix and sonication in a sonicator on iced water for 1 minute. The biphasic layer extraction was performed as described above and in a modified version in which evaporation in the SpeedyVac was replaced by evaporation using a N₂-blower. The protein crash method was modified so that after addition of internal standard mix and 200 µl of methanol, samples were sonicated as described above, mixed by vortex and subjected to centrifugation as previously, and 100 µl of the top layer was transferred for derivatisation without prior evaporation and re-suspension.

An internal standard mix was prepared based on previous recorded measurements as detailed by the human metabolome database¹⁴¹ and used for the following blood to MeOH ratio experiment. The constellation of the internal standards mix is recorded in Table 3. A pooled blood sample of neonatal sepsis cases and healthy controls (n=22) was used. The MeOH to blood sample ratio was optimized by extracting blood samples in differing sample to MeOH ratio's with addition of the same amount of internal standards using the previously established protein crash method. The ratio's used were 10µl sample to 240 µl MeOH, 20 µl sample in 230 µl MeOH, 30 µl sample in 220 µl MeOH, 40 µl sample in 210 µl

Materials and methods

MeOH and 50 μ l sample in 200 μ l MeOH. Extractions were done with 5 replicates and 5 blanks were extracted, in which 250 μ l pure MeOH were extracted. 10 μ l of IS mix was added to all samples. The calculated lipid concentrations resulting from this experiment were used to validate the internal standard concentrations in the previously prepared internal standard mix.

Compound name (all deuterated)	Abbreviated compound	Concentration IS (μM) in internal standard mix
Acetic acid	C2:0-d3	50
propionic acid	C3:0-d2	15
butyric acid	C4:0-d3	15
Hexanoic acid	C6:0-d3	25
Octanoic acid	C8:0-d15	10
Decanoic acid	C10:0-d3	25
Dodecanoic acid	C12:0-d23	10
myristic acid	C14:0-d27	10
palmitic acid	C16:0-d2	25
stearic acid	C18:0-d35	25
arachidonic acid	C20:4-d8	25
Acylcarnitine	C2:0- carnitine-d3	2.5
propanoylcarnitine	C3:0- carnitine-d3	2.5
butyroyl-carnitine	C4:0- carnitine-d3	2.5
hexanoylcarnitine	C6:0- carnitine d3	2.5
octanoylcarnitine	C8:0- carnitine-d3	2.5
decanoylcarnitine	C10:0- carnitine-d3	2.5

Materials and methods

dodecanoylcarnitine	C12:0- carnitine-d3	2.5
myristoylcarnitine	C14:0- carnitine-d3	2.5
palmitoylcarnitine	C16:0- carnitine-d3	2.5
stearoylcarnitine	C18:0- carnitine-d3	2.5

Table 3 The concentration of the internal standards in the internal standards mix.

2.2 Measurement of lipids in plasma of adult sepsis patients and healthy controls

Plasma samples from the ILTIS study were extracted, derivatized and measured according to the method optimized previously using the same materials. The recruitment of sepsis patients for the ILTIS study was approved by the Health and Care Research Wales Research Ethics Committee under reference 17/WA/0253, protocol number SPON1609-17 and IRAS project ID 231993. Recruitment of healthy adult volunteers was approved by Cardiff University's School of Medicine Research Ethics Committee under reference 18/04.

2.2.1 Calibration curve

A calibration curve was prepared with concentrations of lipids in the range expected to be encountered in the samples. The calibration curve included internal standards at the same concentration as when added to the samples.

2.2.2 Sample preparation and measurement

Samples were extracted batchwise in randomized batches. Per batch 5 blanks (10 μ l methanol) were extracted in parallel with the samples. 10 μ l of sample were mixed with 250 μ l of methanol containing internal standards at concentrations as described in

Compound name (all deuterated)	Abbreviated compound	Concentration IS (μM) in internal standard mix
Acetic acid	C2:0-d3	50
propionic acid	C3:0-d2	15
butyric acid	C4:0-d3	15
Hexanoic acid	C6:0-d3	25
Octanoic acid	C8:0-d15	10
Decanoic acid	C10:0-d3	25
Dodecanoic acid	C12:0-d23	10
myristic acid	C14:0-d27	10
palmitic acid	C16:0-d2	25
stearic acid	C18:0-d35	25
arachidonic acid	C20:4-d8	25
Acylcarnitine	C2:0- carnitine-d3	2.5
propanoylcarnitine	C3:0- carnitine-d3	2.5
butyroyl-carnitine	C4:0- carnitine-d3	2.5
hexanoylcarnitine	C6:0- carnitine d3	2.5
octanoylcarnitine	C8:0- carnitine-d3	2.5
decanoylcarnitine	C10:0- carnitine-d3	2.5

Table 3. 3. Sonicated in minute, rpm for 10 then 000 g for 10 supernatant by addition solution and

dodecanoylcarnitine	C12:0-carnitine-d3	2.5
myristoylcarnitine	C14:0-carnitine-d3	2.5
palmitoylcarnitine	C16:0-carnitine-d3	2.5
stearoylcarnitine	C18:0-carnitine-d3	2.5

Samples were iced water for 1 vortexed at 1400 minutes at 4°C, centrifuged at 18 minutes. 100 µl of were derivatized of 50 µl 3-NPH 50 µl EDC and

pyridine solution and subsequent incubation for 30 minutes at 40°C in a waterbath. The reaction was quenched by addition of 100 µl 0.5% formic acid solution in 75% methanol and incubation for 30 minutes at 40°C in a waterbath. Until measurement samples were stored at -20°C. Before quantification on the machine, samples were spun down and supernatant was transferred to be measured. A QC samples was prepared by pooling aliquots from all extracted samples. The QC sample was measured after every 10 other measurements on the machine.

2.2.3 Data analysis

Peaks were integrated using MultiQuant and concentrations in samples were calculated based on internal standards using excel. For lipids with a corresponding deuterated lipid included in analysis, concentrations were calculated based on the lipid to IS ratio and the used internal standard concentration. For other lipids a calibration curve of the lipid to a similar deuterated lipid ratio was used. All further analysis was conducted using R, and GraphPad Prism version 8.4.3 for Windows (GraphPad Software, San Diego, California USA) and MetaboAnalyst 4.0 and higher (metaboanalyst.ca, University of Alberta (Wishart group))¹⁴².

2.3 Lipid measurements in neonatal suspected sepsis cases and neonatal controls

For the nSep cohort, sampled were extracted, derivatized and measured according to the method optimized previously and with no modifications to the method used for the ILTIS cohort. The calibration curves were prepared according to the same protocol as for calibration curves used in the ILTIS cohort. The study was approved by the Wales Research Ethics Committee 2 (Ethics REC reference 19/WA/0008).

2.4 Mapping the transcription factor network associated with *GPR84* expression

To obtain the data to analyse transcriptional regulation in sepsis, two paediatric datasets were downloaded and merged, and the data from a neonatal sepsis study was obtained. The data of two paediatric datasets (GEO accession numbers: GDS4273; GDS4274) were downloaded as raw CEL files from the Gene Expression Omnibus (GEO) from NCBI and RMA-normalized using the limma package in R. Data from the neonatal sepsis dataset was analysed and wrangled using the tidyverse package and the oligo package. Visualizations were made using the ggplot2 and EnhancedVolcano package in R for Volcano plots. Correlations between transcription factors and *GPR84* were computed using the Hmisc package in R.

To obtain insights into transcription factors likely regulating the upregulated genes in sepsis, the ChEA3 tool was used using the browser interface. A list of upregulated genes in paediatric sepsis and neonatal sepsis was obtained respectively by filtering the list of genes as obtained from oligo for a logFC bigger than 1 between controls and sepsis cases in excel. The results from the analysis were downloaded from ChEA3 as an excel file. A figure displaying the association with regulation in different tissues was obtained by using the global network feature, choosing the GTEx TF option, choosing to colour nodes by tissue specific (A) and GO pathways (B) and displaying the top 50TFs.

To find out which transcription factors can bind the regulatory region of *GPR84*, the ENCODE database hits for *GPR84* were obtained from the Harmonizome website (<https://maayanlab.cloud/Harmonizome/>). Hits for *GPR84* in the Signalling Pathways Project database were obtained using the web interface, choosing to browse the cistromics database with the single gene option in human tissues. The results were downloaded as excel file and further filtered based on tissues. The overlap between the databases were computed using the BioVenn web interface.

The tidyverse package in R was used to generate a table of all transcription factors in the analysis with columns denoting if a transcription factor had a hit in the analysis. To generate a Venn diagram, the BioVenn web interface was used.

Functional relationships were identified and visualized using STRING via the web interface. All High Confidence and Medium Confidence were used as input for the “multiple proteins query”. For the first visualization, a minimum confidence of 0.6 was set and only

experiments, neighbourhood, gene fusion and databases were used as possible input. The TFs were clustered using MCL clustering with an inflation parameter of 3 and disconnected TFs were hidden. For the second visualization, the same minimum confidence of 0.6 was used, and only physical interactions were chosen to be visualized. This network was also clustered using MCL clustering with an inflation parameter of 3 and disconnected nodes are hidden.

The promoter region for GPR84 was identified using the European mirror of the UCSC genome browser and the EPD database. Tracks indicating transcriptional regulation were chosen to be displayed and the transcription start site (TSS) was obtained from the EPD database. The chosen promoter region was displayed using the function to add a User-specified track. The nucleotide content of the promoter sequence was obtained using PROMO when searching for TFBS in the promoter region.

PROMO was used to identify TFBS in the promoter region previously identified, setting the factors and species to "Human" and the SearchSites function with the standard setting (Maximum matrix dissimilarity rate = 15). The results were transferred into an excel file. To produce an alignment between mouse and human GPR84 promoter to identify the conserved subpart, Dialign was used in the web interface. The result was downloaded, and the conserved sequence of the human promoter was extracted by choosing all bases that were within the sequence that showed alignment over the majority of the stretch. Genes correlating with GPR84 with a Pearson correlation >0.77 were chosen. Correlations were computed using Pearson correlation in MORPHEUS on a subset of the paediatric sepsis dataset containing only sepsis cases. Before using Pearson correlation, transcriptomics data were normalized using Robust Z-score. The promoters for these genes were obtained using the EPD database and taking the 1300 bp around the transcription start site (TSS) as was done for GPR84. TFBS were identified using PROMO, setting Species and Factors to Human, and the MultiSearchSites option. The standard setting for the dissimilarity matrix was used (15) and TFBS needed to be detected in 12 out of 16 sequences. The overlap between the searches was calculated and visualized using BioVenn.

The human TFBS PWM of all TFs identified as High and Medium Confidence were selected in JASPAR in the online tool. These PWMs were queried against the conserved promoter sequence identified in the previous analysis. The result was downloaded as an excel list and manually examined for pairs and clusters of TFBS corresponding to interacting TFs.

Materials and methods

The tool EnhancerControl was used to establish tissue-specific expression of murine Gpr84 and to find out which TFs bind to the regulatory regions identified in their experiments to regulate Gpr84. Gpr84 was queried using the web interface and the results were downloaded. The location of the open chromatin regions identified by EnhancerControl were compared to the genomic region of the murine TSS as identified by EPD.

Enrichr was used to find out more about potential transcriptional regulation of GPR84 based on co-expression and data from KD and KO experiments.

The timecourse of expression of GPR84 and transcription factors was visualized using timecourse data from an experiment where murine bone marrow derived macrophages were infected *in vitro* with murine cytomegalovirus. The data was supplied in a log-normalized format by a colleague. Fold change to t=0 was calculated in excel and this data was visualized using the R packages ggplot2, viridis, plotly, hrbthemes and kableExtra. Data was wrangled using the tidyverse and the janitor package.

For the *in vitro* timecourse of expression, THP-1 cells were grown in Roswell Park Memorial Institute 1640 (RPMI-1640) medium supplied with 10% foetal calf serum (FCS), L-glutamine, penicillin and streptomycin. THP-1 cells were distributed into 12 well plates 24 hours prior to experimentation. At t=0 cells were treated with either LPS (100ng/ml), dexamethasone (100nM) or both (same concentrations as when separate). Cells were sampled at t=0, 2,4,6,8h by spinning down supernatant in eppendorfs while incubating the bottoms of the wells with RLT buffer with 1% mercaptoethanol. Clear supernatant was taken off and the RLT +mercaptoethanol buffer was transferred from the bottom of the well to the corresponding pellet. The pellet was lysed by pipetting and frozen at -20 degrees till RNA extraction. RNA was extracted using the Qiagen RNeasy Plus extraction kit following the protocol. RNA yield was quantified using Nanopore and cDNA was made using the RT SSIV and random hexamer primers (Qiagen), qPCR was performed using perfeCTa LowRox toughmix (Qiagen) with Taqman primers (Qiagen) on a Quantstudio 3 Real-Time qPCR machine (ThermoFisher Scientific). Expression fold changes were calculated using the delta-delta Cycle Threshold (ddCt) method.

2.5 The biosynthesis of GPR84 ligands

2.5.1 Cell culture - General remarks

For cell culture reagents and media components were either kept at 4°C (cell medium, phosphate buffered saline solution (PBS)) or -20°C (trypsin, fetal calf serum, L-glutamine, hormones and growth factors). All sterile handling steps, i.e., media preparation, experiments on cells and cell strain maintenance, were performed in a Biosafety level 2 (BSL2) hood. Before use the hood was cleaned with both MycoPlasma-Off (Cambio, Cambridge, United Kingdom) and 70% ethanol, and all reagents and instruments used in the hood were cleaned outwardly with 70% ethanol. After use the hood was cleaned using detergent, Mycoplasma-Off and 70% ethanol. After preparation of complete medium, this was kept tightly closed at 4°C. Care was taken to only use complete medium up to 6 months after initial preparation to avoid negative effects from L-glutamine degradation. All cells were grown in Thermo Scientific Nunc EasYFlasks (Thermo Fisher Scientific, Waltham, Massachusetts, United States) with filter cap to ensure air flow. These flasks are treated for cell adhesion and work equally well for suspension cells. Cells were split regularly as required by cell type and experimental planning. When splitting cells, a visual inspection was performed to control cellular health and density. Cells were generally grown in 5-7 ml of medium in T-25 flasks, 20-25 ml medium in T-75 flasks and 30 ml in T175 flasks. The incubator was set to 37°C and 5% CO₂. When a new cell line was introduced into the laboratory, it was incubated in a separate incubator and tested for mycoplasma after reaching sufficient density using the MycoAlert Mycoplasma testing kit (Lonza, Basel, Switzerland).

2.5.2 General techniques

2.5.3 Cell maintenance

Cells were generally maintained according to the specifications set out in ATCC. Cell medium was warmed in a waterbath set at 37°C before use. When obtaining cells frozen, these were quickly thawed in the waterbath at 37°C and transferred into their corresponding cell medium, warmed to 37°C. Cellular suspension was then subjected to centrifugation in a benchtop centrifuge and the cellular pellet was resuspended and transferred into a T-25 flask. When the cellular suspension reached the required density, it was transferred to a T-75 or T-175 flask with the appropriate amount of medium. Cells were

Materials and methods

maintained by regular splitting, i.e., a fraction of the present cells were transferred into fresh medium and incubated until the desired density for another split or for cellular experiments was reached. When splitting suspension cells, i.e., THP-1 cells, a fraction was obtained by taking the appropriate volume of the cellular suspension. For adherent cells, i.e., HUVEC cells, the medium was discarded, and cells were washed with PBS before being incubated with trypsin (both Gibco, Thermo Fisher Scientific, Waltham, Massachusetts, United States). Release of the cells was checked under the microscope until 70% of the cells released into the medium. The disassociation was then stopped using the addition of medium and the cells were subjected to centrifugation in a benchtop centrifuge (180 *g*, 2 minutes, room temperature). The pellet of cells was resuspended in fresh medium and a fraction of this was used to seed cells into a fresh flask.

2.5.4 Freezing

In order to freeze cells, cells were grown to confluency, and using the same techniques as when splitting disassociated and/or collected from the flasks and subjected to centrifugation. Cell pellets were suspended in fetal calf serum containing 25% dimethylsulfoxide (DMSO) and transferred into cryogenic vials (Thermo Fisher Scientific, Waltham, Massachusetts, United States). Vials were transferred into a Mr. Frosty freezing container (Thermo Fisher Scientific, Waltham, Massachusetts, United States).

2.5.5 Charcoal-treatment of serum

For experiments researching lipid synthesis or ligand binding to the receptor, fetal bovine serum (FBS) was subjected to charcoal treatment to decrease any potentially present lipids. The aim of this was to ensure there would be no high background of lipids that could obstruct detection of lipid production, and that there would be no background lipids present that could activate the receptor. Cell-culture grade dextran-coated charcoal was dissolved in 10 mM Tris solution (pH 8.0), at a concentration of 1g/ 16 ml. The resulting slurry was added to FBS, 1ml slurry per 10 ml FBS. Using a magnetic stirrer, the solution was stirred over night at 4°C. Next it was subjected to centrifugation at 1000 *g*, 10 minutes, 4°C. The supernatant was transferred to a new vessel and 1 ml of previously prepared charcoal slurry added per 50 ml starting volume FBS. This solution was incubated at 55°C for 1 hour and then subjected to centrifugation at 1000 *g*, for 10 minutes at 4°C. The

resulting slurry was filtered using a 0.45 um filter and stored in aliquots at -20°C degrees until use.

2.5.6 Sampling for RNA extraction

RNA was extracted using RNeasy extraction kit (Qiagen), cDNA was made using RT SSIV and random hexamer primers (Qiagen), qPCR was performed using perfeCTa LowRox toughmix (Qiagen) with Taqman primers (Qiagen) on a Quantstudio 3 Real-Time qPCR machine (ThermoFisher Scientific).

2.5.7 Sampling for lipid quantification

When sampling for lipid quantification, sampling technique differed based on cells being adherent or not. For lipid quantification from supernatant, in adherent cell types, supernatant was sampled directly from the well in which the cells were growing and transferred into a glass vial. Cells were then scraped using a sterile cell scraper (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and transferred to an eppendorf for centrifugation (2700 *g*, 10 minutes) using a benchtop centrifuge (Eppendorf, Hamburg, Germany). In non-adherent cells the cellular suspension was transferred to an eppendorf for centrifugation (2700 *g*, 10 minutes) using a benchtop centrifuge (Eppendorf, Hamburg, Germany). After centrifugation, the supernatant was sampled (for lipid quantification). All samples were frozen at -80°C after sampling.

2.5.8 RNA sampling and extraction

Depending on the cell type, RNA was either extracted using a modified trizol extraction method as described by Toni et al¹⁴³, or using the RNeasy Plus extraction kit((Qiagen, Hilden, Germany)) according to the supplied protocol. Trizol extraction was used for all primary cells, e.g. peripheral blood mononuclear cells (PBMCs) and neutrophils. The RNeasy Plus extraction kit was used for RNA extractions from cell lines. In both cases all reagents and disposables used were sterile and RNase-free. When sampling into trizol (Thermo Fisher Scientific, Walthamstow, U.S.A), cellular suspensions were subjected to centrifugation (7500 *g*, 5 minutes, room temperature) while 1 ml of trizol was added to the empty corresponding cell well. This was done to ensure collection of all cellular RNA, since PBMCs and neutrophils will be in suspension but can adhere if activated. Cell pellets were then resuspended in the trizol from their corresponding experimental well with repeated pipetting, kept at room temperature for 3 minutes and frozen at -80°C. Before extractions

Materials and methods

all samples were thawed at room temperature. All trizol extractions were performed in a BSL2 hood. For trizol extractions, 200 μ l chloroform (Thermo Fisher Scientific, Walthamstow, U.S.A) was added to the trizol suspension, samples were shaken for 15 seconds and incubated at room temperature for 3 mins, and then subjected to centrifugation (12 000 g , 15 minutes, 4°C). The upper aqueous phase was transferred to a new eppendorf and the chloroform extraction was repeated once on the transferred sample. The RNA was precipitated by mixing the obtained solution with 250 μ l isopropanol, mixing by inverting the tubes and incubating at room temperature for 10 minutes. After centrifugation (12,000 $\times g$, 10min, 4°C) all supernatant was discarded and the remaining pellet was washed with repeated ethanol washes. To this end, 1ml of 75% ethanol solution was added to the pellet, centrifuged (7500 $\times g$, 5 min, 4°C) and the supernatant discarded. This step was repeated twice. The washed pellet was dried and solubilized in RNase free water (Qiagen, Hilden, Germany) and stored at -80°C if not immediately processed.

For RNA extraction using the RNeasy kit (Qiagen, Hilden, Germany), cells were sampled using RLT lysis buffer, either by resuspending cell pellets using repeated pipetting or by adding RLT buffer onto adherent cells after taking off the supernatant. The RNA extraction was performed according to the instructions in the RNeasy kit. DNA was removed using gDNA spin columns, by loading the sample onto the column and centrifuging. The filtered sample was then washed in several consecutive steps in the provided RNeasy spin columns, using the provided buffers. Finally, RNA was eluted off the column using RNase-free water. RNA was quantified using a Qubit™ RNA High Sensitivity (HS) (Invitrogen, Thermo Fisher Scientific, Walthamstow, U.S.A) on a Clariostar plate-reader (BMG Labtech, Germany). To this end, a mastermix containing 0.5 μ l of Quant-iT RiboGreen reagent and 99.5 μ l buffer per sample was prepared and distributed into a black-walled, clear-bottomed 96-well microplate (Thermo Fisher Scientific, Walthamstow, U.S.A). As reference the provided standards were used to generate a calibration curve. RNA concentrations were calculated based on the obtained calibration curve.

cDNA was made using RT-PCR using the SuperScript IV reverse transcription system (Qiagen). RNA was mixed with RNase-free water to obtain the desired amount of RNA per reaction in 11 μ l. 1 μ l random hexamers and 1 μ l dNTPs (10 mM) were added. After incubation for 5 minutes at 65°C a mastermix containing 4 μ l 5x SSIV buffer, 1 μ l DTT, 1 μ l RNaseOUT inhibitor and 1 μ l SSIV Reverse Transcriptase per reaction was added. The qPCR

was performed using Taqman probes and perfeCTa LowRox toughmix (Qiagen). Per well 5 μ l toughmix, 0.5 μ l Taqman probe and 4.5 μ l cDNA at desired dilution were added. qPCR was measured on a quantStudio 3 Real-Time qPCR system (Thermo Fisher Scientific, Waltham, Massachusetts, United States) using the deltaCt program. The Taqman probes used were: beta-2-microglobulin (B2M, Hs00187842_m1), TATA-box binding protein (TBP, Hs00427620_m1), peptidylprolyl isomerase A (Hs03045993_gH), oleoyl-acyl-hydrolase (OLAH, Hs00217864_m1), medium chain acyl-CoA dehydrogenase (ACADM, Hs00936576_m1), G-protein coupled receptor 84 (GPR84, Hs00220561_m1); all with FAM-MGB fluorophore, low ROX.

2.5.9 HUVEC

A human umbilical vein endothelial cell line (HUVEC) that was immortalized using overexpression of the telomerase reverse transcriptase (tert) was obtained from the Cardiff cell biobank. The HUVEC-tert cells were incubated in endothelial cell growth medium MV2 containing Fetal Calf Serum (0.05 ml/ml), Epidermal Growth Factor (recombinant human) (5 ng/ml), Basic Fibroblast Growth Factor (recombinant human) (10 ng/ml), Insulin-like Growth Factor (Long R3 IGF) (20 ng/ml), Vascular Endothelial Growth Factor 165 (recombinant human) (0.5 ng/ml), Ascorbic Acid (1 μ g/ml), Hydrocortisone (0.2 μ g / ml). To establish OLAH expression in full medium and in medium lacking fetal calf serum and hydrocortisone, cells were seeded at a density of 100 000 cells/ cm^2 in either full or reduced medium and incubated for 24 hours at 37°C and 5% CO_2 .

2.5.10 Extraction of immune cells from healthy donor blood using Ficoll-density gradients

Healthy volunteers were recruited and consented under the departmental ethics for blood collection. Per volunteer, 50 ml of blood were collected using butterfly needles. After collection of blood, the blood was immediately processed using the density gradient centrifugation method. The blood was diluted with buffer (phosphate buffered saline, 2 mM EDTA, Sigma Aldrich, St. Louis, Missouri, United States) in a 1:4 blood to buffer ratio. This dilution was slowly layered onto Ficoll-Paque reagent (Histopaque, Sigma-Aldrich, St. Louis, Missouri, United States). A gradient was generated by subjecting these samples to centrifugation for 30 minutes at 400 g at room temperature with deceleration disengaged. The layer containing PBMCs was sampled and purified using several washing steps. Firstly, the layer was transferred to a clean tube and washed after addition of buffer by subjecting

Materials and methods

it to centrifugation at 300 *g* for 10 minutes. The supernatant was removed, and the pellet was washed again by adding 50 ml of buffer and subjecting the cells to centrifugation at 200 *g* for 10 minutes. The latter step was repeated once.

To obtain neutrophils, the pellet containing red blood cells and neutrophils was treated with ACK-lysis buffer. The pellet at the bottom of the tube after centrifugation was treated with twice its volume of ACK lysis buffer. After incubation at room temperature for 5 minutes, the solution was spun at 200*g* for 5 minutes. The supernatant was removed, and the remaining pellet treated again with twice its volume ACK lysis buffer followed by incubation and centrifugation as before. This step was repeated 3-4 times until a relatively clean neutrophil pellet was evident.

The obtained cells were transferred into RPMI cell medium supplemented with charcoal-treated serum, counted and seeded into plates. Cells were immediately treated with either LPS (end concentration 100ng/ml), dexamethasone (end concentration 100nM) or both, and samples were taken after 3 and 6 hours from treated neutrophils and 3 hours, 6 hours and 24 hours for PBMCs.

3 Optimization of a method for reliable lipid detection from blood samples

Fatty acids (FA) are carboxylic acids with a long aliphatic chain and are either saturated (no carbon-carbon double bonds) or unsaturated (one or more carbon-carbon double bonds). They are classified according to their chain length as short chain (2-6 carbons; SCFA), medium chain (8-12 carbons; MCFA) and long chain (>12 carbons, LCFA) fatty acids. In human biology there is a considerable interest in SCFA, as they are thought to be generated by the gut microbiota and directly influence the host's health¹⁴⁴. Research in this domain points to the bacterial fermentation of complex carbohydrates in the gut as pivotal for human health; the resulting SCFA bind to G-protein coupled receptors, most notably GPR43. Some of the effects ascribed to the SCFA-GPR43 signalling are appetite regulation, improvement of the intestinal epithelial barrier and induction of cell death in cancer cells¹⁴⁵⁻¹⁴⁷. Furthermore, SCFA can also bind to receptor GPR109a and GPR41¹⁴⁸. Unlike SCFA, little is known about the role and presence of MCFA in the body, specifically during the inflammatory response. Though MCFA are present in the human diet, e.g. in palm kernel oil, coconut oil or dairy, measurements of MCFA in healthy volunteers have indicated only low, and frequently not detectable, concentrations in the blood¹⁴⁹⁻¹⁵¹. Further research indicates that MCFA are quickly resorbed from the blood stream after a meal high in these fatty acids^{151,152}. Medium-chain fatty acids have been suggested as possible ligands of GPR84⁹⁶. GPR84 expression is strongly upregulated in the blood of sepsis patients³². This leads to the question if MCFA are present in the blood of sepsis patients. In addition, measuring a range of lipids, since technically possible, allows for a better understanding of potential lipid changes in sepsis.

Firstly, the concentrations of different subsets of acylcarnitines, i.e. SCFA, MCFA or LCFA esterified to carnitine, have been found to be significantly altered in sepsis survivors compared to non-survivors⁵⁰. Carnitine is the molecular shuttle that facilitates transport of fatty acids into the mitochondrion for β -oxidation, an important step for energy generation from lipids.⁴⁴ An analysis conducted by Langley et al. (2014) integrating lipidomics with proteomics, lead the authors to conclude that carnitine shuttling in sepsis non-survivors is decreased due to a decrease in fatty acid transport proteins⁵¹. Liu et al. (2016) also found

Results

a significant difference in the LCFA palmitic acid and stearic acid⁵⁰. Accordingly, the method developed here was also optimized to measure SCFA (since GPR43 has been shown to be upregulated in sepsis as well³²); LCFA due to their role in metabolism and inflammation, and acylcarnitines due to previous studies showing altered concentrations in sepsis. In this chapter I developed a method to accurately measure these lipids in patient blood samples. The methods to reliably and sensitively measure metabolites (including lipids) in complex biological mixtures include the use of a gas-chromatography (GC) or liquid-chromatography (LC) coupled to a mass spectrometer^{138,153}. In order to measure lipids, they need to be extracted from the samples, in some cases derivatized and then quantified. In our approach a liquid chromatography-mass spectrometry (LC-MS/MS) system with electrospray ionisation is used. The advantage of using LC-MS/MS over GC-MS/MS is the possibility to measure a wider range of metabolites, shorter run times and better sensitivity^{154,155}. Owing to their small molecule size and consequent hydrophilicity, SCFA and MCFA are usually difficult to detect using LC-MS methods with electrospray ionisation¹⁵⁶. However, derivatisation of the lipids is commonly employed and greatly improves their detection in LC-MS¹⁵⁶. Derivatisation is a chemical reaction which converts a chemical compound into a product of similar chemical structure, called a derivative. As a result, these molecules will have the chargeability and hydrophobicity that enables their efficient ionization during electrospray ionisation, one of the initial and vital processes in mass spectrometry. In addition, derivatisation ensures that no native lipids potentially present in LC-MS solvents (a particular issue with fatty acids) will contribute to background as the non-derivatised analyte is clearly distinguishable. This effectively decreases the background signal when using a targeted LC-MS/MS approach and hence improves sensitivity of detection. When performing LC-MS/MS, derivatized molecules are fragmented in a consistent fashion; the fragments offering an additional mean of identifying the compound. Han et al. (2015) described the detection of SCFA by LC-MS/MS after derivatisation with 3-nitrophenylhydrazine hydrochloride (3-NPH; Figure 10)¹⁵⁷. During the derivatisation reaction 3-NPH is covalently bound to the compound by nucleophilic addition/elimination at the carboxylic acid group. This reaction also requires 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and pyridine which stabilizes the pH of the reaction^{158,159}. This is a suitable method for this approach, since all compounds to be measured have an exposed carboxylic acid group and this method has been used on both

fatty acids and acylcarnitines^{157,160}. The approaches described by Han et al. use either methanol or acetonitrile as solvents to carry out the derivatisation reactions. Due to the lower solubility of MCFA than SCFA, both methanol and acetonitrile were tested in the different reaction solvents to develop this method and the derivatisation time was optimized. Further, the LC-MS/MS method was optimised to obtain good chromatography and to avoid carry-over of analytes between samples. Lastly, the efficient extraction of analytes from the samples is an important first step that required extensive testing and optimization for the sample types used in this study. Optimised extraction of samples eliminates any components present in the sample that could interfere with the measurement of the intended compounds, for example by blockage of the column or thin tubes of the liquid chromatographer¹³⁸. Simultaneously, the compounds to be measured need to be recovered reliably in the sample mixture. Due to short chain and medium chain fatty acids suspected to be present even in HPLC-grade solvents, a quenching reaction after derivatisation has been introduced. The quenching reaction ensures that background levels of fatty acids will not be derivatized after the initial derivatisation has been conducted. For example, Chan et al. (2017) described addition of 2-mercaptoethanol to reduce functionality of EDC and addition of succinic acid to exhaust the derivatisation agent¹⁵⁸. Extraction of lipids has been previously documented from serum, plasma or whole blood. While extracting analytes from whole blood can provide information about the presence of compounds in both cells and the extracellular matrix combined, processing serum or plasma provides information exclusively on their presence in the extracellular matrix. To obtain serum or plasma requires for blood samples to be processed shortly after collection and higher volumes to be collected due to loss of volume during processing. In my studies whole blood was also used in sample cohorts due to organizational restraints on sample processing. Several extraction methods were compared to extract all the lipids earlier specified (SCFA, MCFA, LCFA, acyl-carnitines) from whole blood. The methods tested included biphasic layer extraction with two different solvents and the protein crash method. The biphasic layer extraction relies on the partitioning of compounds between two immiscible layers of liquids of opposing polarity, namely an aqueous and organic layer. The fatty acids are then analysed in the solvent organic layer post collection, evaporation, and reconstitution in solvents suitable for LC-MS injection. The protein crash method relies on extracting analytes and removing proteins by crashing them out using organic solvent

Results

and consecutive centrifugation. This method requires no evaporation and reconstitution of extracts prior to LC-MS/MS analyses¹⁶¹.

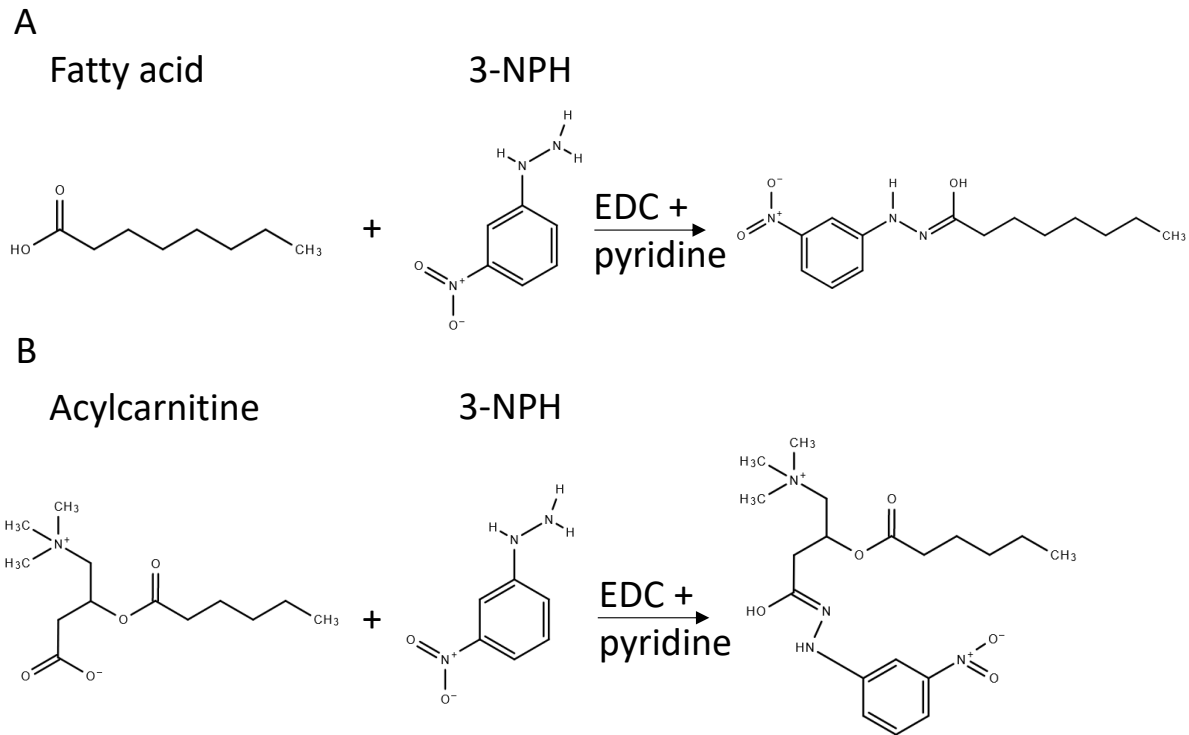


Figure 10 Derivatization of fatty acyls using 3PNH and EDC as described by Han et al (2015)(A) and derivatisation of acylcarnitine as described by Meierhofer (2019) (B). : Catalysed by EDC the carboxylic group of the fatty acid and the acylcarnitine, respectively, binds covalently to the nitrous group of 3-nitrophenylhydrazine. This reaction releases one molecule of water (not depicted). Pyridine is required as it stabilizes the pH of the reaction mixture. Han et al. (2015) found best derivatisation results by heating the samples to 40 °C with 30 minutes of reaction time being sufficient, and both acetonitrile and methanol as solvents were adequate¹⁵⁷. The fatty acid depicted as example is octanoic acid, and the depicted acylcarnitine is hexanoylcarnitine. Chemical structures were generated using chemspider (A. Chemist, ChemSpider SyntheticPages, 2001, <http://cssp.chemspider.com/123>).

3.1 Results

3.1.1 Derivatisation optimization

Firstly, since derivatisation with 3-NPH had been established for SCFA¹⁵⁷, I sought to confirm whether the published conditions (with additional tested variations) were suitable for derivatisation of MCFA, the focus of this study. To this end, 40 µl MCFA standards (C8:0; C10:0; C12:0; C8:0d15; all 50 µM, n=3) were derivatized in 75% aqueous methanol or 75% aqueous acetonitrile for different durations (0.5, 1, 2 and 3h) with 20 µl 3-NPH (200 mM) and 20 µl EDC-pyridine (120 mM, 6% pyridine v/v) at 40°C¹⁵⁷. To account for background levels of MCFA in solvent and, reagents, parallel 'blank' samples (solvent only) were also derivatized. Derivatized MCFA from the various treatments were measured on the LC-MS/MS as described below (see section LC-MS/MS conditions for machine settings).

Regardless of reaction solvent (acetonitrile or methanol), all measured MCFA were successfully derivatised with maximal derivatised lipid achieved after 30 minutes (Figure 11). This is consistent with previous findings¹⁵⁷. However, significantly higher levels of C10:0 and C12:0 MCFA were derivatised when using 75% aqueous methanol compared to acetonitrile as the reaction solvent (t-test, p=0.0001, 30 minutes of incubation), which may be explained by higher solubility of fatty acids with increasing chain length in aqueous methanol compared to aqueous acetonitrile.

We subsequently confirmed these conditions (75% aqueous methanol; 30 min derivatisation) were able to derivatise FA standards of varying length (SCFA and LCFA) over large concentration ranges. Equally, we confirmed they were suitable for the derivatisation of acyl carnitines which has been done previously by Han et al.¹⁶² and explored in detail by Meierhofer¹⁶⁰. The linearity of LC-MS/MS response vs analyte concentration can be observed in Figure 12 for a range of deuterated lipids. More information on the linearity of detection can be found in the paragraph on LC-MS/MS conditions.

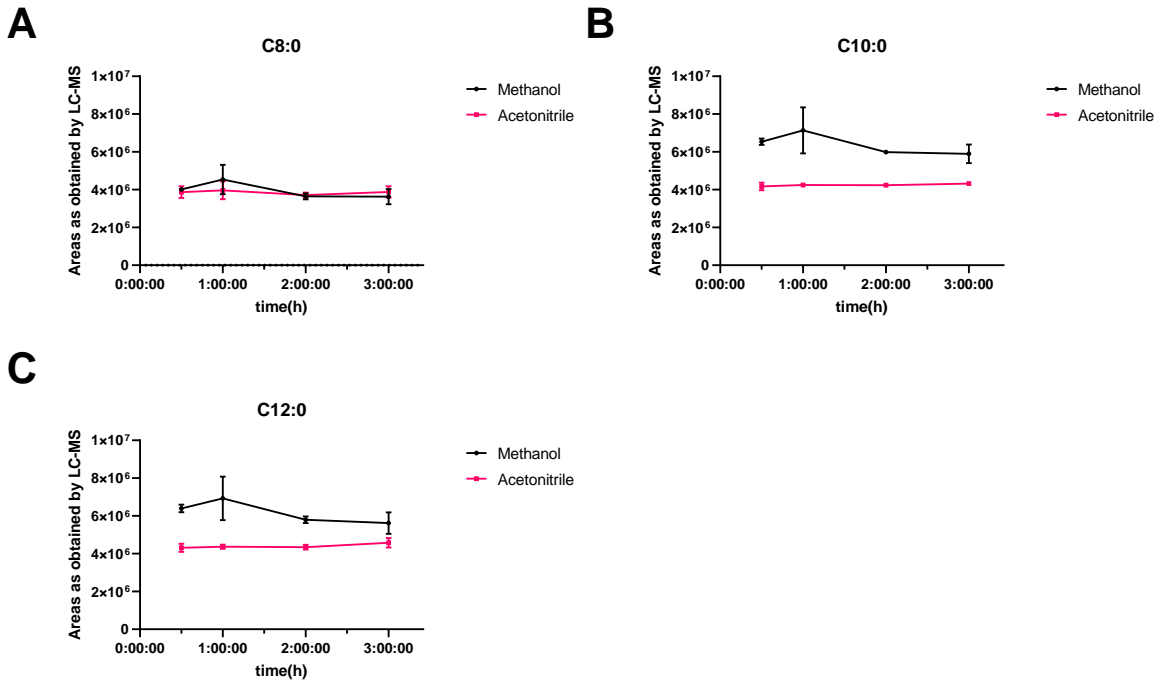


Figure 11 Derivatization of MCFAs under different experimental conditions. LC-MS/MS Peak areas of different compounds derivatized using either 75% methanol or 75% acetonitrile as reaction solvents. The tested time periods ranged from 30 minutes to 3 hours. Derivatization was carried out in a waterbath set to 40°C. Data are averages ($n=3$, error bars indicate standard deviation). Background was measured using both an acetonitrile and a methanol blank that was subtracted from the samples.

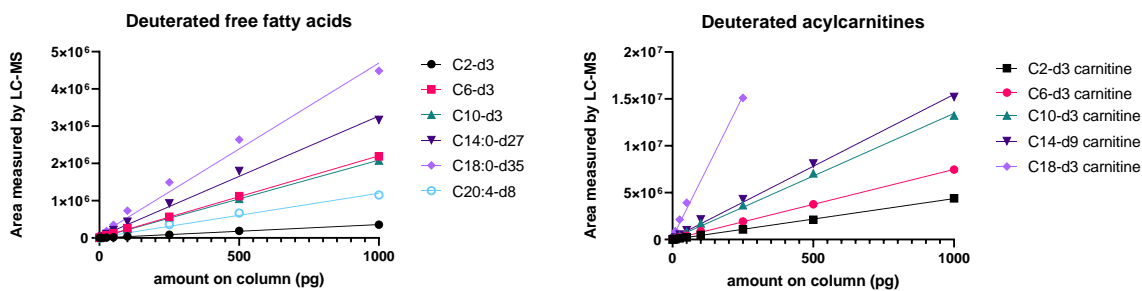


Figure 12 Calibration curves of a selection of deuterated FFA and acylcarnitine standards. The x-axis shows the amount injected on the LC-MS/MS column (picograms; calculated based on the input amount of standards stocks before derivatization). The y-axis shows the areas as measured by LC-MS/MS. The lines indicate the interpolated curve based on the datapoints, the goodness of fit showed R^2 values above 0.99 for all lipids except C18-d35 ($R^2=0.985$).

Results

3.1.2 Quenching optimization

Background levels of several fatty acids including SCFA and MCFA were found to be present in blank solvent samples that underwent the derivatisation process. Potential sources of these background FA can be from reagents and solvents and are often unavoidable when extracting and working up samples prior to derivatisation¹⁵⁸. Hence, they need to be accounted for in analyses of samples for accurate analyte quantification e.g., via the use of blank samples and subsequent subtraction. However, since the derivatisation reagent is used in high concentrations and is present in the final samples, another source of background levels of FA can be further unwanted derivatisation of background FA in LC-MS solvents during the final analysis stage, since derivatisation is possible at a range of temperatures¹⁵⁷. That is why we sought to test and include a post derivatisation reagent 'quenching' step to eliminate this issue.

Here blank samples (HPLC grade methanol) were processed through the derivatisation method and then analysed (LC-MS/MS see section LC-MS/MS conditions) directly or post an additional post derivatisation 'quenching step', and then subsequently compared. To account for possible deleterious effects of quenching reagents on FA recovered from blood samples in future experiments, all blank samples tested also contained a mixture of deuterated internal standards since deuterated FA are not present as background contaminants. Quenching steps tested were addition of mercaptoethanol (20 µl 120 mM mercaptoethanol, 20 µl HPLC grade water (referred to as just "water" from here on)); formic acid (200 µl 0.1% formic acid in water); mercaptoethanol and succinic acid (20 µl 120 mM mercaptoethanol and 20 µl 1M succinic acid); or a high amount of succinic acid alone (20 µl 1M succinic acid, 20 µl water). To the control sample, 40 µl of water were added. Following addition of quenching reagents all samples were incubated at 40°C for 30 minutes^{163,158}. Subsequently, results are presented in Figure 13 as the ratio of the lipid measured in the sample that underwent derivatisation + quenching: to the sample that underwent derivatisation alone. For non-deuterated lipids, a ratio below 1 indicates the quenching reaction reduced background lipid contamination post derivatisation (values >1 indicates that the quenching reaction actually introduced contamination). For supplemented deuterated standards a value below 1 indicates that the quenching reaction has potential deleterious effects on lipid levels following the derivatisation reaction.

Firstly, integratable FA peaks (LC-MS) were recorded for a large number of FA in all blank samples, confirming that background FA contamination would be an issue in sample processing and needs to always be taken into account. Secondly, since additional quenching did not reduce the measured levels of FA in blank samples post derivatisation beyond values of ca 0.5 (50%) for most FA, this suggests that the majority of background FA contamination occurs pre/during derivatisation and can only be taken into account by processing blank samples alongside real samples for deductive purposes.

When comparing the various quenching methods, formic acid quenching reduced the amount of background fatty acids derivatized for all lipids except C10:0 (Figure 13 G,H) but did not negatively affect the deuterated internal standards. For example, background levels of C8:0 and C12:0 were reduced by 7.9% and 21.7%, respectively. C10:0 levels go up by around 20%, but its background concentrations were 3 to 6 times lower than those of C8:0 and C12:0 to start with. In other words, if this quenching method introduces any background contamination of C10:0, then the levels introduced are negligible since the absolute amounts introduced are comparably small. Quenching with mercaptoethanol and succinic acid (Figure 13 C,D) or succinic acid alone (Figure 13 E,F) reduced the amount of derivatized background fatty acids for several fatty acids but reduced the amount of deuterated acetic acid internal standard, suggesting deleterious effects. Formic acid quenching was hence deemed to be the most suitable quenching method.

The final optimised method for derivatisation and quenching is as follows: samples will be incubated with 3-NPH (200 mM in 75% methanol) and EDC-pyridine (120 mM, 6% v/v in 75% methanol) in a ratio of 2:1:1 and then quenched by addition of a 0.1% formic acid solution.

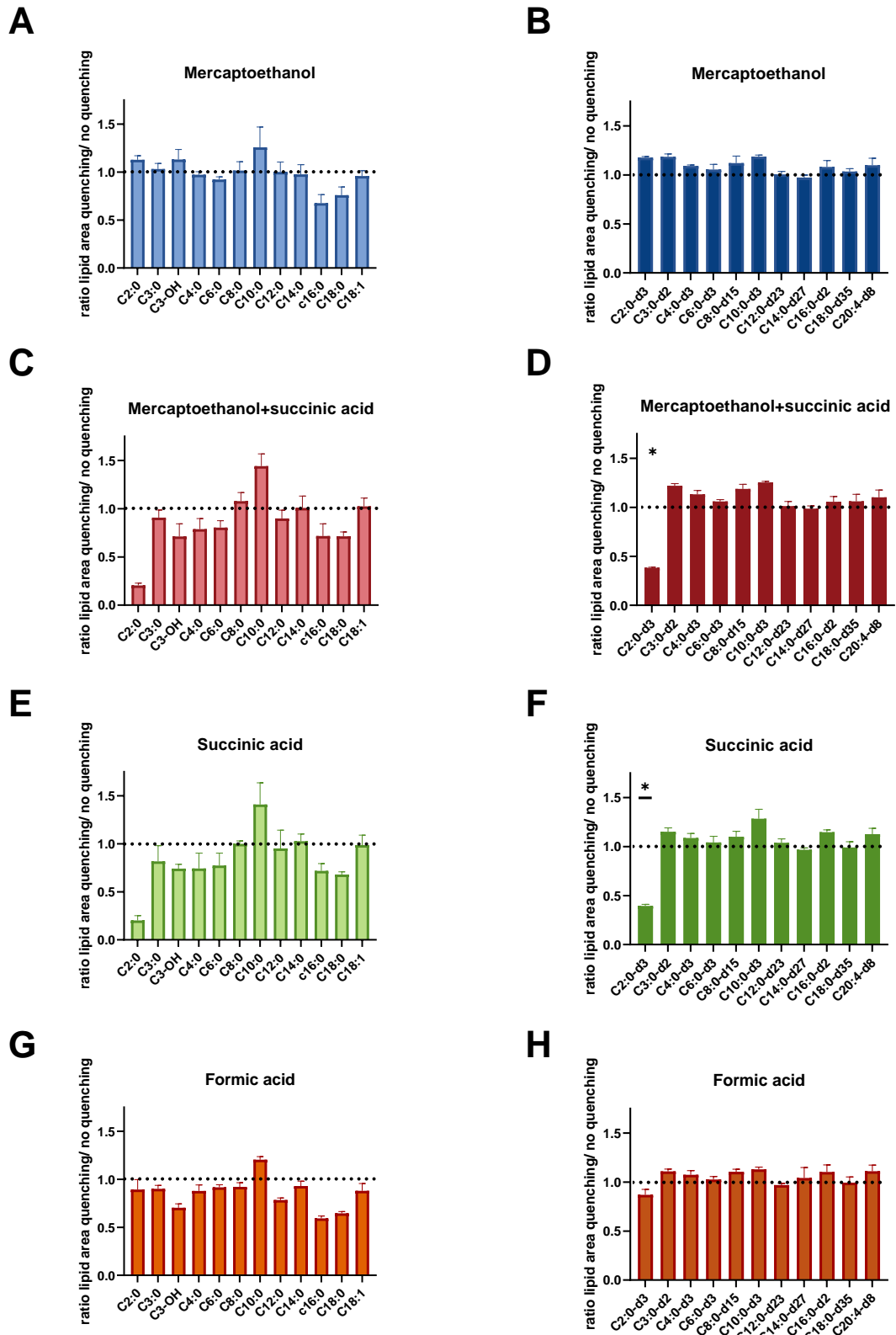


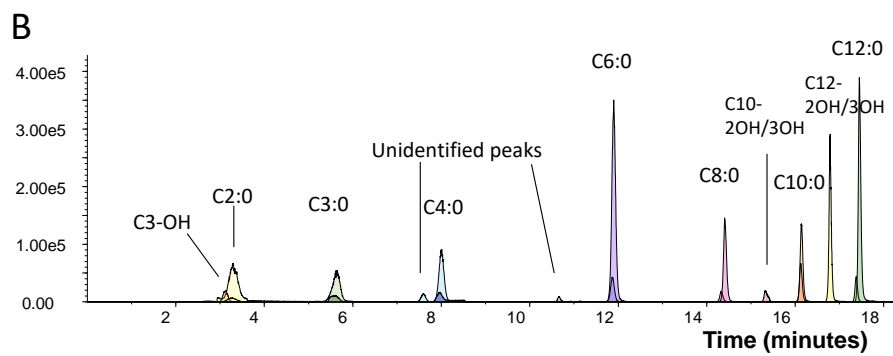
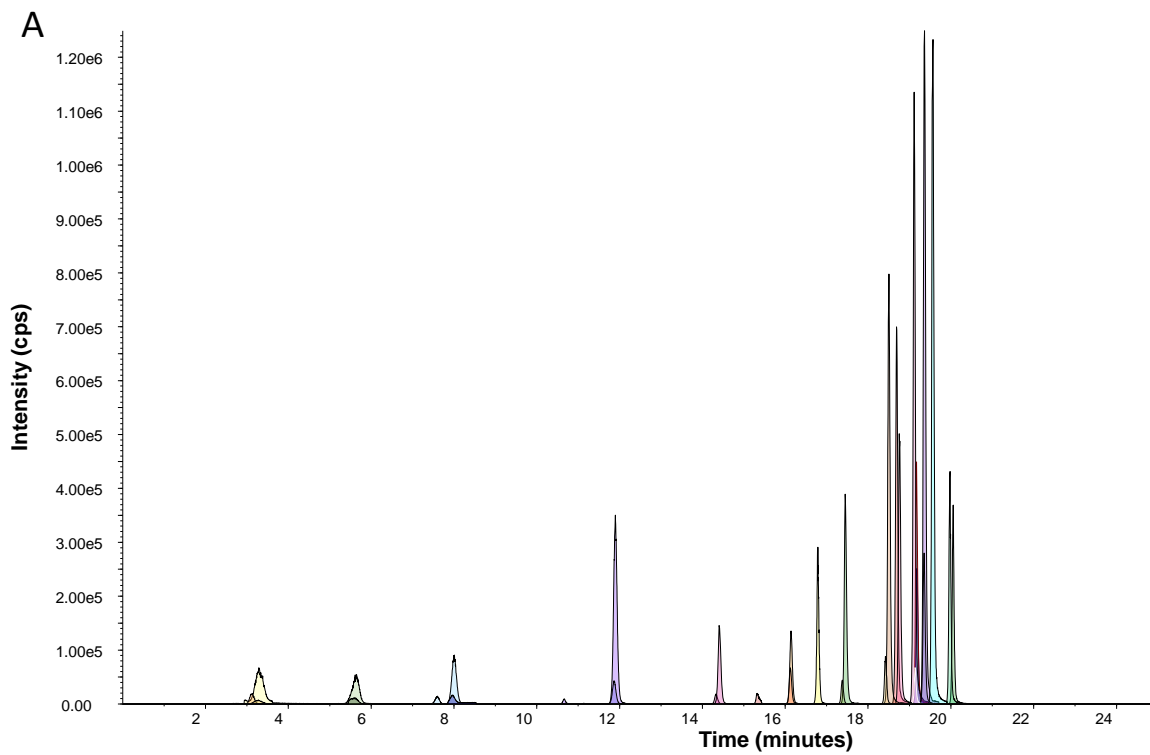
Figure 13 The effect of different quenching methods on contaminating fatty acids. 4 different quenching methods were compared to the control (no quenching) to check which method could reduce background levels of fatty acids measured on the machine. The y-axis depicts the ratio of the lipids measured in the “sample” that underwent derivatisation and

quenching, to the lipids measured in the control – a similar sample that only underwent derivatisation but no quenching. All samples were diluted to the same volume to assure comparability. On the left side all lipids that are measured are contaminants, since no not deuterated lipid standards were added. On the right side, internal deuterated standards are depicted. These were added right before derivatisation and are used here as an indication of deleterious effects of the quenching methods on derivatized lipids

Results

3.1.3 LC-MS/MS conditions

LC-MS/MS conditions were optimized by my colleague Dr. Daniel White based on settings presented by Han et al.¹⁵⁷, for consistent and satisfactory chromatography and to avoid carry-over of sample between samples. LC-MS/MS chromatograms for a range of acylcarnitines and fatty acids and can be seen in Figure 14 and Figure 15, respectively. Analytes were resolved on a C18 column (100mm x 2.1mm x 2.6 μ m, 100 Å; Kinetex by Phenomenex, Torrence, California (U.S.A)), kept at 50°C. Separation was obtained using 100% H₂O + 0.1% formic acid as mobile phase A and 100% methanol + 0.1% formic acid as mobile phase B using a linear gradient: 15% mobile phase B for 2.47 minutes; then mobile phase B increasing to 55% up to 11.12 minutes; then mobile phase B increasing to 100% up to 20 minutes; held at 100% B for 12 minutes, then 10 minutes re-equilibration at 15% mobile phase B. The mobile phase flow rate was 0.2 mL min⁻¹. Samples were kept at 4°C in the autosampler and a 2 μ l injection volume was employed.



Results

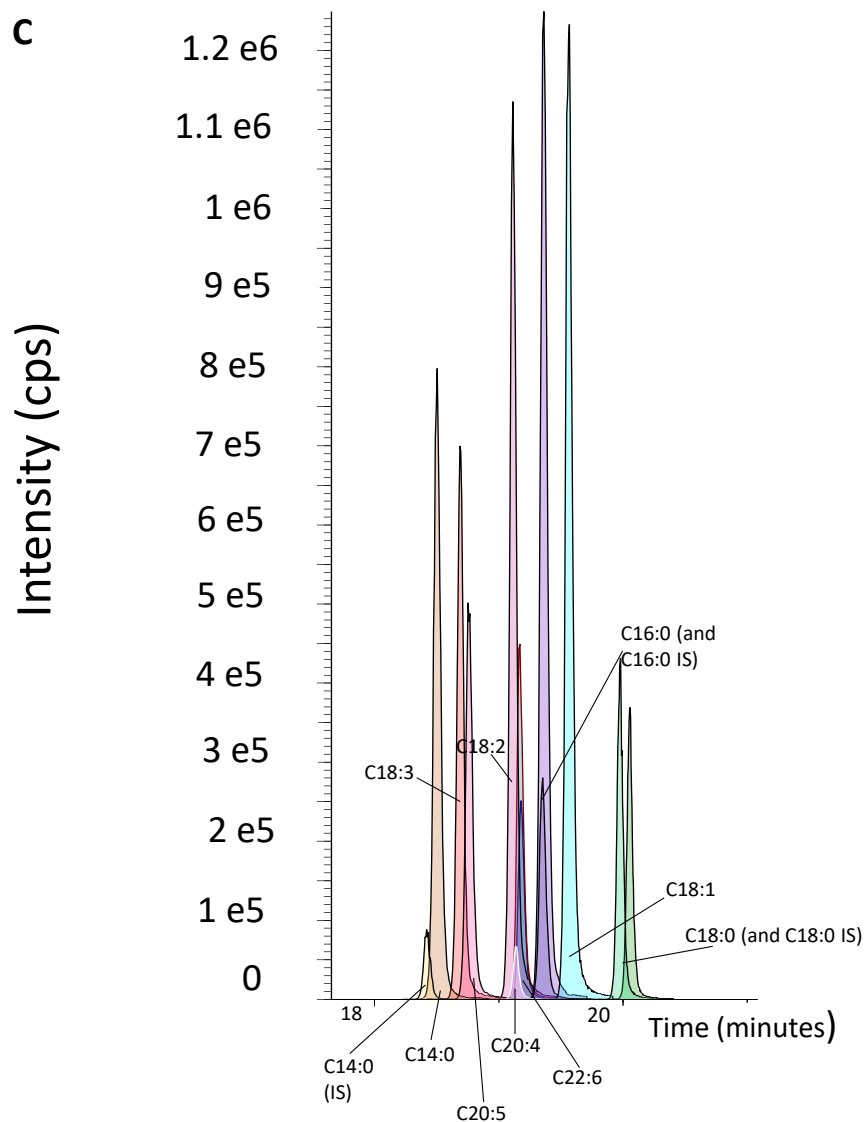


Figure 14 A) Example chromatogram of free fatty acids (standard mix, 50 μ M, the deuterated standards were added as internal standard mix at concentrations as set out in the methods chapter) run on the optimised LC-MS/MS system. For most lipids only the unlabelled standard is indicated, the deuterated standard is typically visible overlapping (often they don't have the exact same RT, there is slight difference) with its corresponding standard. B) shows a closer look at the lipids that elute up to 18 minutes into the run. C) shows a closer look at lipids eluting after 18 minutes. With some exemptions, peaks are broader and flatter for smaller lipids and narrower and higher for longer lipids.

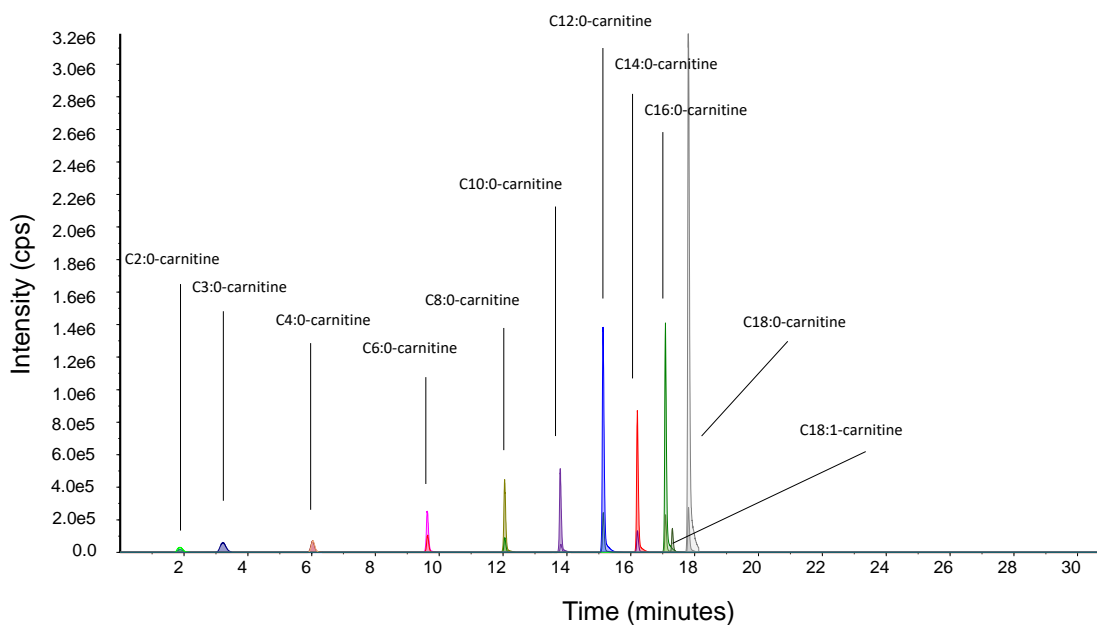


Figure 15 Example chromatogram of acylcarnitines (standard mix) run on the optimised LC-MS system. All carnitines are present at 10 μ M, and the deuterated standards were added as internal standard mix whose concentrations are set out in the methods chapter. For all lipids only the not deuterated standard is indicated, the deuterated standard is typically visible overlapping with its corresponding standard. At this stage of optimization C18:2 and C18:3 standards were not used yet.

Results

Fatty acids were detected in negative ion mode and acyl-carnitines in positive ion mode. Declustering potential and collision energy were determined by direct injection of standards onto the mass spectrometer (Table 4). Retention times of lipids for scheduling were routinely confirmed using a standard mixture of lipids (all at 50 μ M; typical retention times indicated in Table 4). Deuterated internal standards were measured using the same settings as their corresponding standard lipids, with exception of Q1-mass.

To confirm the suitability of both the derivatisation reaction and the settings of the equipment titration curves of lipids were generated. The range of these curves were based on expected amounts in patient samples with a generous range above and below. The range in which we can reliably quantify the compounds is determined by the linear range of detection, i.e., all amounts that when plotted against their measurements fall onto a straight line. For deuterated standards, linearity was achieved for most lipids even in the lower range of the calibration curve (Figure 12). The short chain fatty acids could be detected upwards of 2.5 pg (C3-d2), 10 pg (C4-d3) and 25 pg (C2-d3) on column. FFA with a longer carbon chain could be detected at 0.5 pg on column with exception of C10-d3 that could be detected upwards of 2.5 pg on column. These amounts would correspond to \sim 5 μ M C2-d3; \sim 1.8 μ M C4-d3; \sim 0.5 μ M C3-d2; \sim 0.3 μ M C10-d3; and between \sim 0.08-0.04 μ M for other FFA in samples with the finally established extraction method. Deuterated acylcarnitines could be detected at amounts on column upwards of 0.5 pg (SCFA-carnitines) or 0.25 pg, (MCFA and LCFA-carnitines) corresponding to between \sim 0.027 and \sim 0.046 μ M in samples, a similar range as reported by Han et al.¹⁶². These results confirm great sensitivity achieved by the machine in measuring derivatized lipids with this method. The upper limit for linear detection was generously above the expected concentrations in sample. Calibration curves made of lipid stocks with addition of deuterated lipids indicate the range in which these lipids can be measured reliably in extracts (Figure 16). Minimum and maximum values at which these lipids could be detected in the linear range are shown in Table 5. The table shows the concentration range in which measurements by LC-MS/MS were in the linear range. The calibration curve was prepared by dilution of lipid stocks and subsequent derivatisation, and I calculated to what concentrations of lipids in samples before extraction that would correspond. In addition, the corresponding amount on column (in pg) that was injected per measurement, of each compound, is indicated in the right column.

Comparing the ranges in which free fatty acids can be detected (post blank/background subtraction) as amounts on column to the ranges of corresponding deuterated standards consistently shows higher minimum values for not deuterated fatty acids. This could be caused by background levels of these fatty acids introduced during sample preparation. Background levels of acetate have been described in pyridine, and Zeng and Cao showed that pyridine was not essential for the derivatisation reaction¹⁵⁹. However, derivatisation of MCFA was greatly reduced without pyridine (data not shown). Subsequently, pyridine was deemed necessary for derivatisation. Since acetic and propionic acid are typically present in relatively high concentrations, their detection should fall within the concentration range with the established method. The background problem is not apparent for acylcarnitines, which can be detected in a very low range similar to the ranges in which deuterated acylcarnitines were detected. The range of detection was hence deemed suitable, and I proceeded using this derivatization method and these settings on the machine.

The quantification of lipid concentration was done by obtaining peak areas per lipid using Multiquant software (SCIEX, AB Sciex LLC, Framingham, U.S.A). Only peaks at least 5 times higher than the background and with at least 10 scans per peak were integrated. To quantify lipids, known concentrations of deuterated internal standards were added in the first step of the extraction process. For every lipid that had a corresponding deuterated internal standard available, concentrations in sample were calculated using the area ratio of the measured lipid to the internal standard, taking into account the sample volumes etc. For some lipids, no directly corresponding deuterated internal standards were commercially available. These were C3-OH, C18:1, C18:2, C18:3, C20:5 and C22:6 and C18:1 carnitine. In these cases, amounts of lipid would be calculated using a calibration curve based on a standard of the primary lipid and the closest (carbon number) internal standard available. Regardless of quantification approach, the MS response of fatty acids/acyl carnitines was always confirmed to lie within the linear dynamic range of MS response based on comparisons with standard curves for both deuterated and primary standards (see Figure 12 and Figure 16).

Results

Table 4 MS settings for the detection of fatty acids and acyl carnitines. Declustering potential (DP), collision energy (CE), mass in Q1 and mass in Q3. The retention time (RT) is an indication since retention times vary between runs due to various factors.

lipid	DP	CE	Q1 m/z	Q3 m/z	Indicated RT
Fatty acid					
C2:0	-35	-26	194.1	137.1	3.28
C3:0	-49	-27	208.1	137.1	5.66
C3-OH (LACTIC)	-76	-26	224.1	137.1	3.26
C4:0	-90	-28	222.1	137.1	8.2
C6:0	-100	-30	250.2	137.1	12.05
C8:0	-120	-34	278.2	137.1	14.53
C10:0	-127	-36	306	137.1	16.24
C10-2OH/3OH	-115	-32	322.6	137.1	15.56
C12:0	-134	-41	334.3	137.1	17.55
C12-2OH/3OH	-115	-36	350.3	137.1	17.26
C14:0	-120	-42	362.4	137.1	18.57
C16:0	-137	-44	390.4	137.1	19.37
C18:0	-140	-50	418.5	137.1	20.07
C18:1	-138	-48	416.5	137.1	19.58
C18:2	-140	-44	414.4	137.1	19.12
C18:3	-134	-42	412.4	137.1	18.76
C20:4	-145	-42	438.5	137.1	19.17
C20:5	-135	-42	436.4	137.1	18.8
C22:6	-148	-44	462.3	137.1	19.25
Carnitines					
C2:0 carnitine	76	27	339.3	220.1	1.86
C3:0 carnitine	76	27	353.0	220.1	3.15
C4:0 carnitine	76	35	367.3	220.1	6
C6:0 carnitine	91	29	395.3	220.1	9.62
C8:0 carnitine	101	39	423.4	220.1	12.04

C10:0 carnitine	106	33	451.4	220.1	13.81
C12:0 carnitine	111	35	479.5	220.1	15
C14:0 carnitine	128	36	507.3	220.1	16.21
C16:0 carnitine	129	37	535.3	220.1	17.07
C18:0 carnitine	128	39	563.4	220.1	17.79
C18:1 carnitine	134	44	561.3	220.1	17.3
C18:2 carnitine	136	42	559.3	220.1	16.8
C18:3 carnitine	138	43	557.3	220.1	16.4

Results

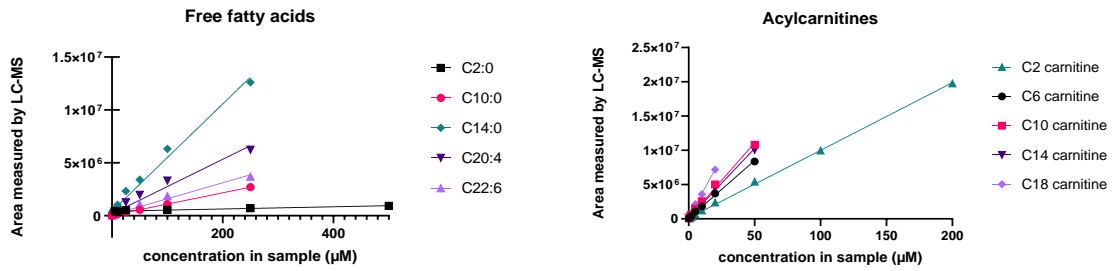


Figure 16 Calibration curves of a selection of free fatty acids (left side) and acylcarnitines (right side). The x-axis depicts the concentration in sample in μM that corresponds to the concentration in a blood sample with the established extraction method calculated from the stock concentration before derivatisation. The y-axis depicts the areas as measured by LC-MS/MS. All depicted standard curves have a R^2 -goodness-of-fit value of above 0.99.

Table 5 The table shows the concentration range in which measurements by LC-MS/MS were in the linear range. The calibration curve was prepared by dilution of lipid stocks and subsequent derivatisation, and I calculated to what concentrations of lipids in samples before extraction that would correspond. In addition, the corresponding amount on column (in pg) that was injected per measurement, of each compound, is indicated in the right column.

	Range of concentration in sample before extraction and derivatisation (μM)	Range of amount on column (pg)
C2:0	5-500	24.9-2487.2
C3:0	2.5-2.500	13.3-2666.7
C4:0	0.5-0.500	2.8-5692.3
C3-OH	25-500	143.6-5743.6
C6:0	2.5-2.500	16.0-6414.4
C8:0	5-500	35.6-3564.1
C10:0	0.5-0.500	3.9-3923.10
C10-2OH	2.5-2.500	20.7-4135.9
C12:0	0.5-0.500	4.3-2141.0
C12-2OH	0.1-0.100	0.9-2245.5
C14:0	1-100	9.3-929.2
C16:0	1-100	10.9-1001
C18:1	0.25-50	2.7-533.9
C18:2	0.025-50	0.3-531.3
C18:3	0.01-25	0.1-264.4
C20:4	0.025-100	0.3-1124.4
C20:5	0.05-1000	0.6-11189.8
C22:6	0.5-1000	5.9-11853.9
Carnitines		
C2-carnitine	0.02-200	0.2-1738.5
C4-carnitine	0.02-100	0.2-905
C6-carnitine	0.01-50	0.09-470.5

Results

C8-carnitine	0.02-50	0.9-506.8
C10-carnitine	0.005-50	0.2-542.8
C12-carnitine	0.02-50	0.5-578.2
C14-carnitine	0.01-50	0.1-614.7
C16-carnitine	0.01-20	0.1-260.1
C18-carnitine	0.01-20	0.1-274.5
C18:1-carnitine	0.01-50	0.1-722.3

3.1.4 Analyte recovery from whole blood

I sought to develop an analyte (FFA/acyl carnitines) extraction procedure from whole blood prior to derivatisation (see above). In addition to measuring endogenous lipids in the samples, extraction efficiency was based on the recovery of deuterated internal standards, C3:0-d2, C8:0-d15, C12:0-d23, C6:0-d3-carnitine, spiked into blood samples prior to extraction. PBS control/blank samples were extracted in parallel to blood samples for the subsequent determination and correction for background levels of lipids (see above).

Several extraction procedures were evaluated. The biphasic layer extraction¹⁶²; a modified biphasic layer extraction; and the protein crash method¹⁶¹, were compared by extracting whole blood (50 μ L; n= 5) and PBS samples (50 μ L n=5) spiked with 11.2 μ L of internal standard mix (50 μ M) containing above mentioned internal standards. As the biphasic layer extraction protocol necessitates the evaporation and reconstitution of FA extracts, the modified biphasic layer extraction protocol aimed to reduce potential volatilization of fatty acids (especially volatile SCFA) by utilizing a gentler method of evaporation (N₂ blower) compared to vacuum evaporation. The protein crash method did not include any evaporation and reconstitution steps, instead a proportion of the resulting supernatant was directly derivatized. Additionally, a brief sonication step was introduced to ensure lysis of all blood cells. An overview of the extraction methods used is provided in Figure 17. Standard recovery results were compared to blank samples (n=3; 100% methanol) spiked with the same levels of deuterated internal standard mix before derivatisation. Internal standard recovery was calculated as the ratio of measured (LC-MS/MS) internal standard concentrations in extracted blood samples: concentration in these spiked blank samples.

While there were no significant differences in recovery of deuterated MCFA (Figure 18 A, C), the protein crash method recovered significantly more deuterated C3:0 and deuterated C6:0-carnitine(Figure 18 B, D). In fact, there was no recovery of deuterated C3:0 by biphasic layer extractions, and it is likely that the volatile SCFA evaporated during the evaporation step. Likewise, there was also no recovery of deuterated C6:0-carnitine using the biphasic layer extraction and very little using the modified method, and this is probably caused by minimal partitioning of this acylcarnitine into the recovered solvent layer. Recovery of C3:0-d2 and C6:0-d3-carnitine by the protein-crash method were on average 38% and 65%, respectively. The protein crash method extracted carnitines above background level for all carnitines, whereas the biphasic layer extraction showed higher extraction at longer chain

Results

lengths of fatty acids indicating that extraction efficiency using this method might depend on the acyl chain length (Figure 20). In addition, the protein crash method was the only method that led to an extraction of significantly higher than background levels for all fatty acids that were detected in quantifiable amounts (Figure 19, two-way ANOVA). C12-2OH/3OH could not be detected with any of the extraction methods. For most fatty acids tested, the background levels were lower when using the protein crash method. It is also noteworthy that there is less variation between replicate samples with this method, as can be seen from the tighter error bars. Finally, C3-OH, i.e., lactic acid, is expected in relatively high amounts in the blood and we saw much higher values for this lipid when using the protein crash extraction. In conclusion the protein crash method was deemed the most suitable extraction method due to its robustness and best recovery.

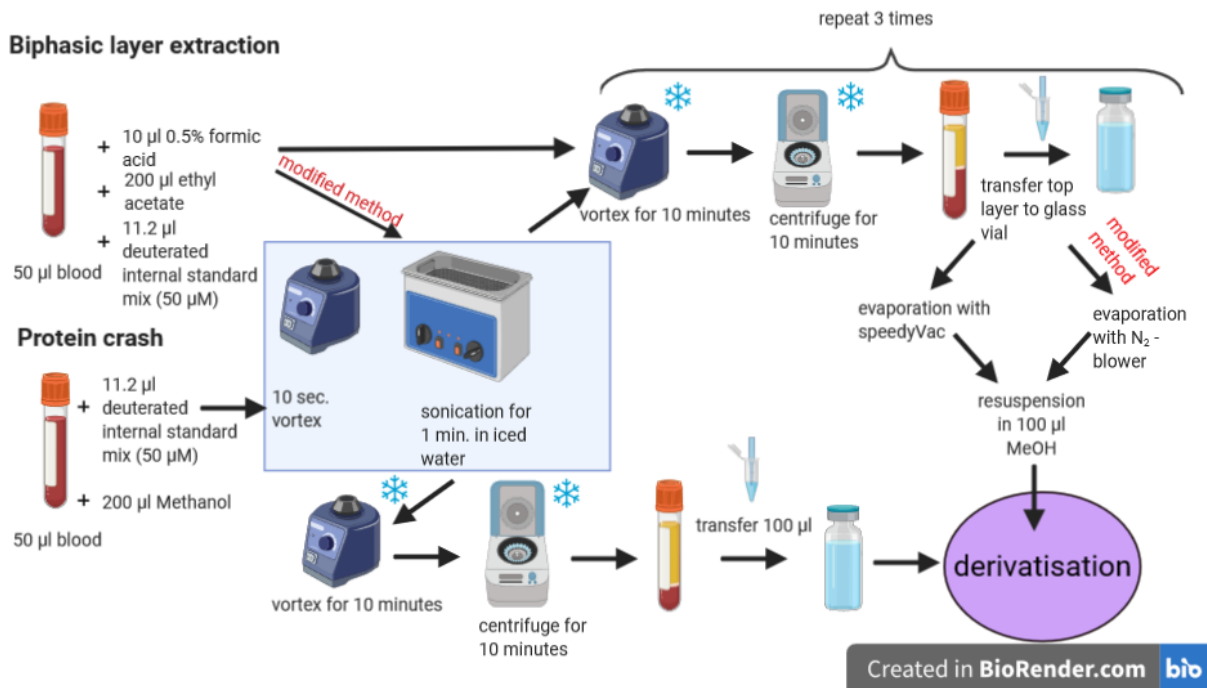


Figure 17 Overview of methods utilized to extract fatty acids and carnitines from whole blood. This overview was created using BioRender.com.

Results

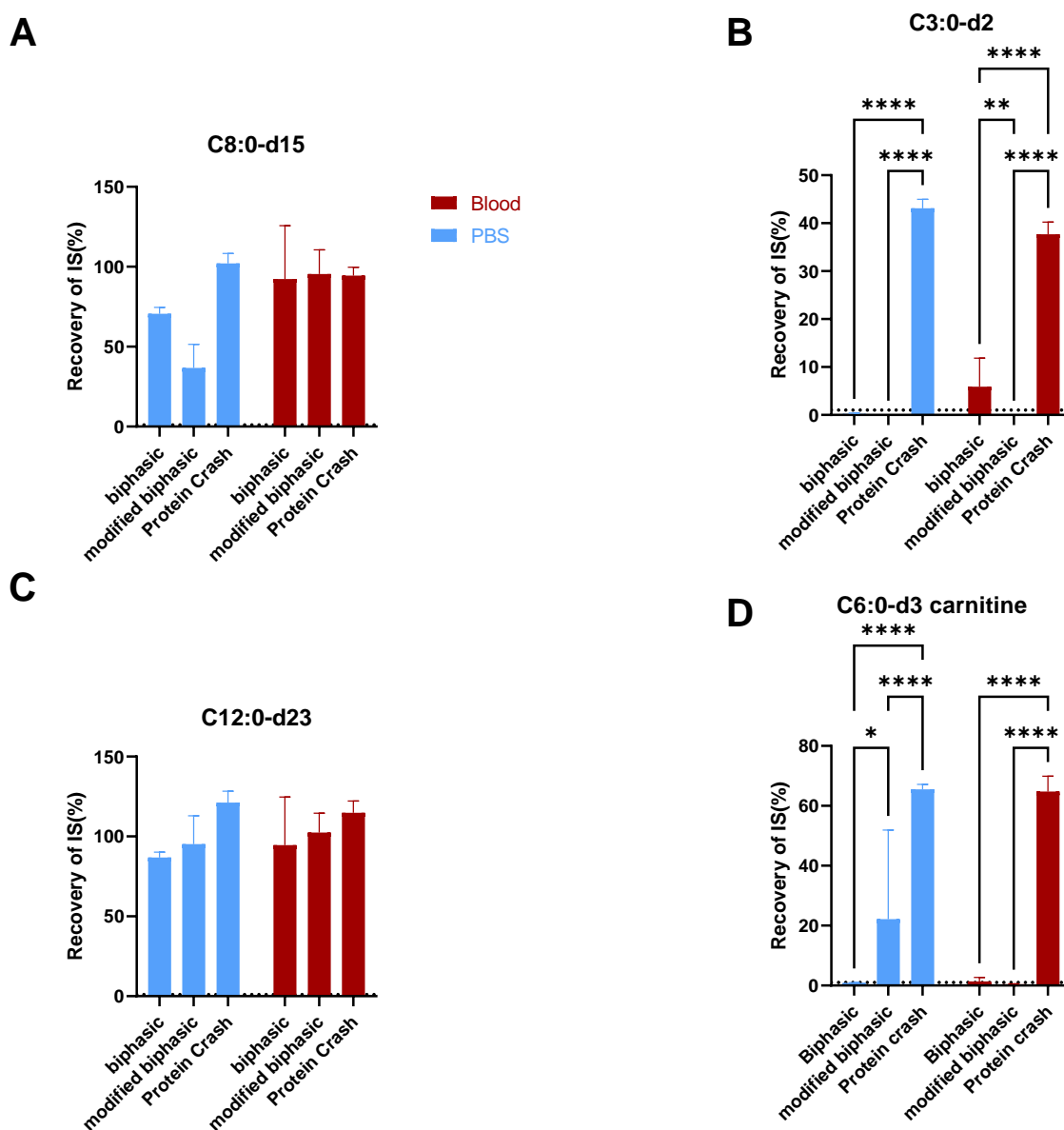


Figure 18 Percentage recovery of deuterated fatty acids and acylcarnitines (spiked internal standards) by different extraction methods from PBS and blood. The extraction methods used were the biphasic layer extraction using ethyl acetate or dichloromethanol (modified biphasic) or the protein crash method. The internal standards were spiked into 50 μ l blood or PBS samples before extraction and compared to peak areas of the internal standards spiked into blank samples (100% methanol) at the same concentration directly before derivatization without extraction. Recovery was calculated as percentage of peak areas from extracted samples by peak area of directly derivatized solvent. All extractions were done with $n=5$. The statistical test performed is a 2-Way ANOVA. Significance levels displayed are: * = $p \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $P \leq 0.0001$.

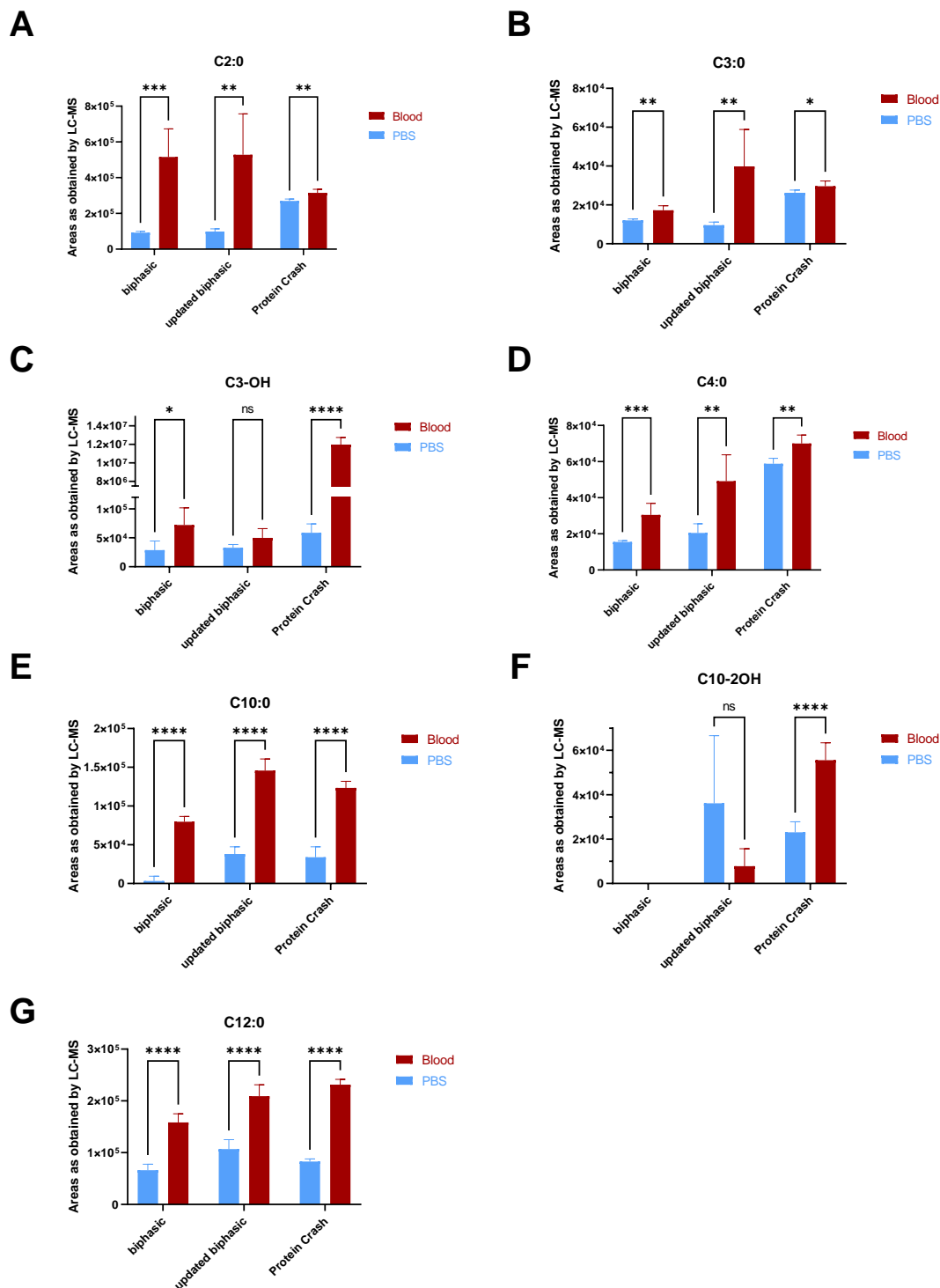


Figure 19 Short- and medium chain fatty acids extracted from 50 μ l PBS or 50 μ l blood using the biphasic method, the modified biphasic method and the protein crash method. The y-axis shows the areas as measured by LC-MS/MS. When comparing areas measured, keep in mind that the methods lead to different fraction of initial sample being derivatised. Significance was calculated using two-way ANOVA. Significance levels displayed are: * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$. All extractions were done with $n=5$.

Results

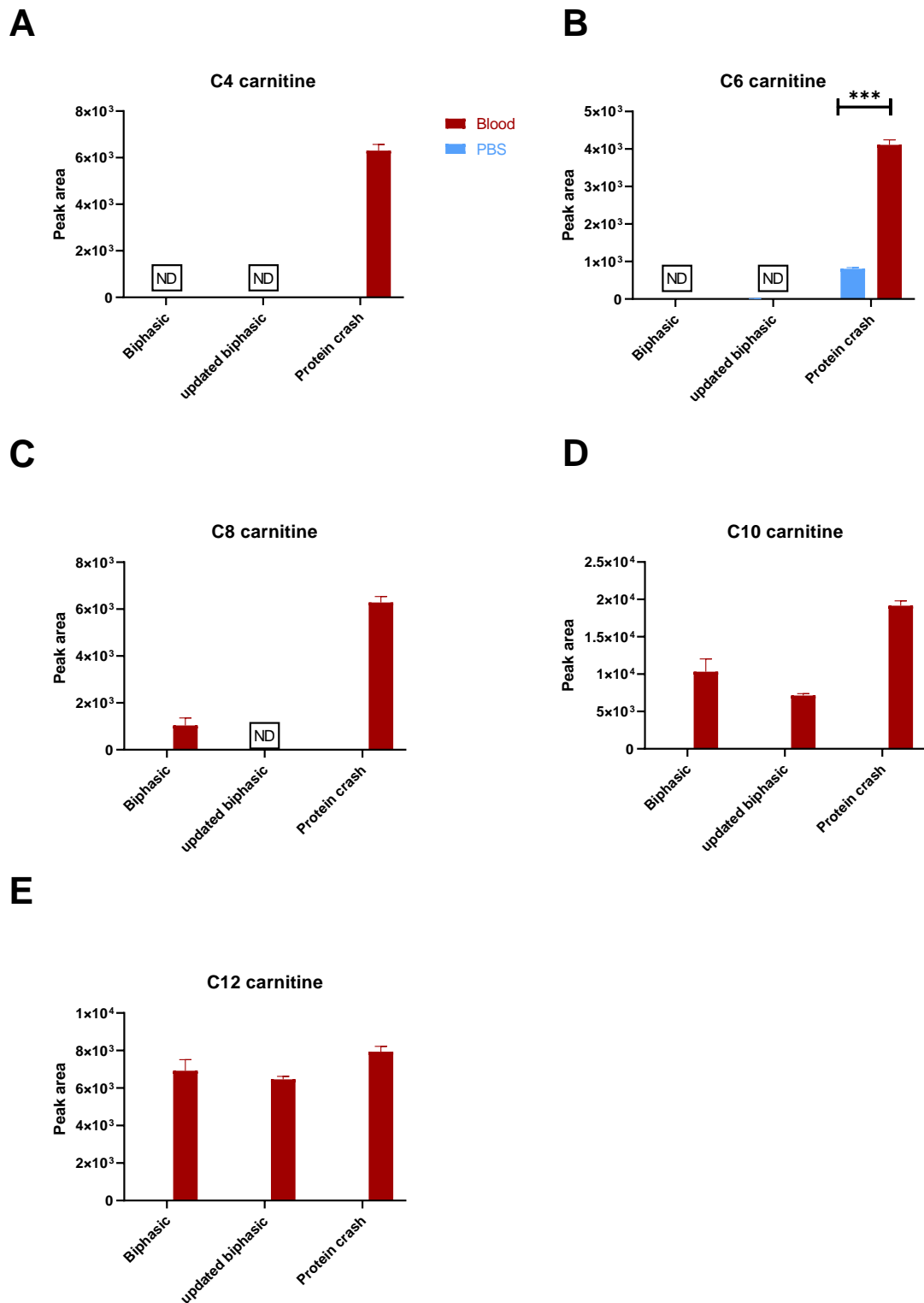


Figure 20 Acylcarnitines extracted from blood and PBS using the biphasic method, modified biphasic method and protein crash method. The y-axis shows the peak area that was measured by LC-MS/MS. Due to the poor recovery of the internal standard, the peak area was deemed a better representation of extraction than the calculated amount in sample. All extractions were done with n=5.

Since the protein crash lipid extraction method is not exhaustive, it is important to investigate potential saturation effects during analyte recovery. By varying the amount of blood added while keeping the total extract volume the same, i.e., by only varying the volume of the methanol extraction solvent, it becomes evident which ratio of solvent to blood is best to reliably extract the maximum lipids from the blood sample without saturating extraction efficiency. Here 10, 20, 30, 40 or 50 μL of blood were spiked with deuterated internal standard mix (10 μL ; concentration and composition set out in materials and methods chapter) and MeOH to reach a final volume of 260 μL at the start of extraction. Blank extractions (50 μL HPLC water instead of blood, $n=5$) were used to account for background levels of FFA. As final step of the extraction, 100 μL of supernatant were transferred for derivatisation.

Post derivatisation samples were analysed on the LC-MS/MS as previously described (see section LC-MS/MS conditions). When comparing the LC-MS/MS analyte areas measured for deuterated internal standards in this experiment, these are very similar for most deuterated lipids and there is no clear reduction/ increase as a function of MeOH to blood ratio (Figure 21). This supports efficient recovery of internal standards regardless of blood volume.

When looking at LC-MS/MS areas for recovered analytes extracting from only 10 μL of blood was sufficient for recovery of all fatty acids and carnitines above background levels (Tukey's multiple comparison's test). As expected, the lipid to IS ratio increases with increasing volumes of blood in the extraction mix. However, it is important to note that the increase in ratio is not proportional to the increase in volume of blood extracted from, i.e., it was not linear beyond circa 20 μL blood in most cases. This would imply saturation of the extraction mechanism and was particularly noted for C3-OH (Figure 22 D), C10:0 (Figure 22 G) or C2:0 carnitine (Figure 23 A).

It is also noteworthy that variation between replicates increases when extracting from a higher blood volume in this experiment, again indicative of potential saturation of extraction effects. To further explore the variability, the % of covariance (%CV) between replicate extractions was calculated for all FFA and carnitines (Figure 24). For both fatty acids and carnitines, the %CV increased with increasing blood volumes. For carnitines, the %CV remained below 10% when extracted from 10 μL whole blood and was above 10% when extracting from higher volumes. The %CV of fatty acid extractions was below 10% for

Results

almost all extractions from 10 μ l, with C10:0, C12:0 and C10-2OH being just above this threshold (11%, 10.2% and 10.1% respectively). Extracting from 10 μ l blood was hence chosen as the most suitable approach.

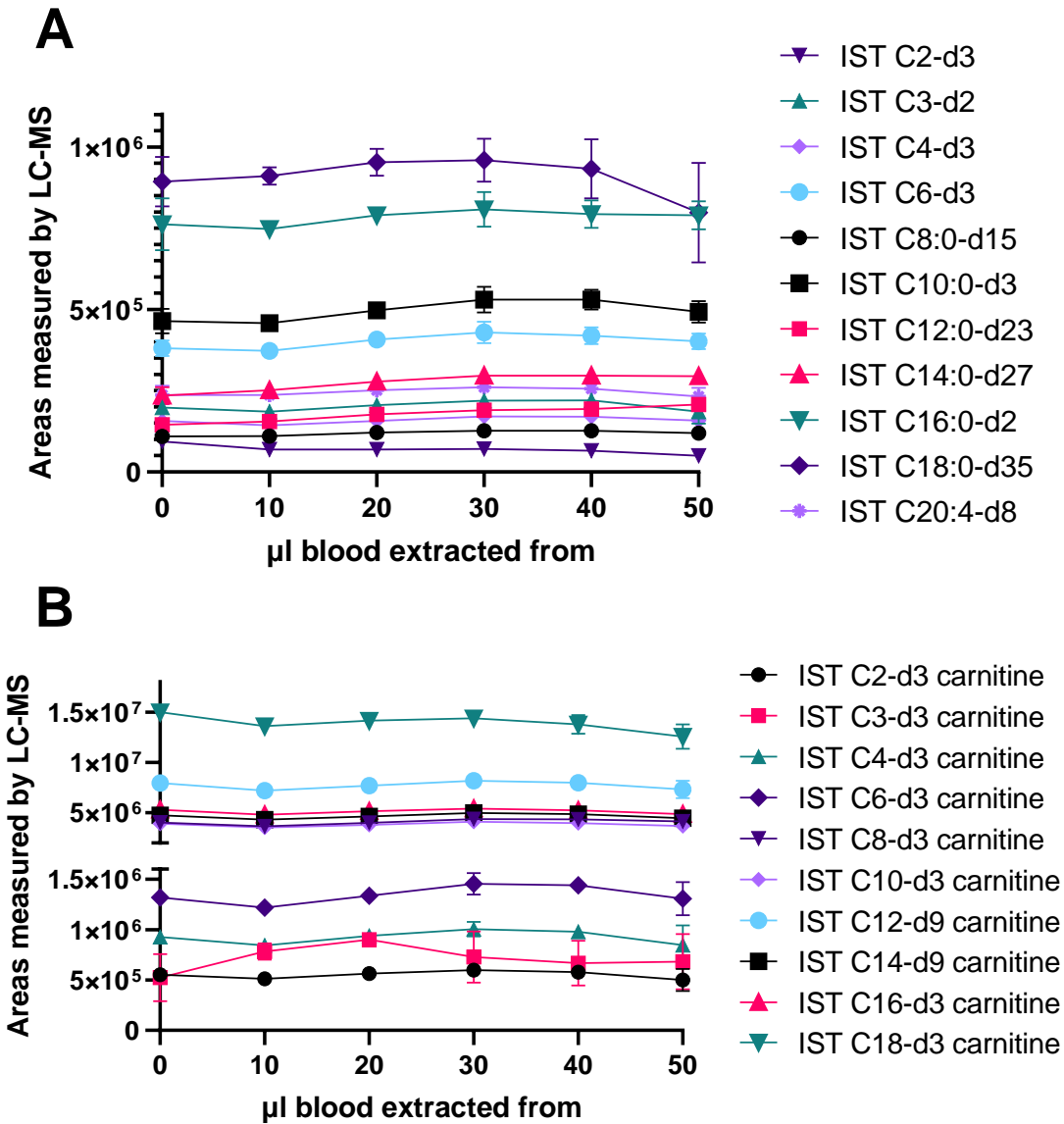
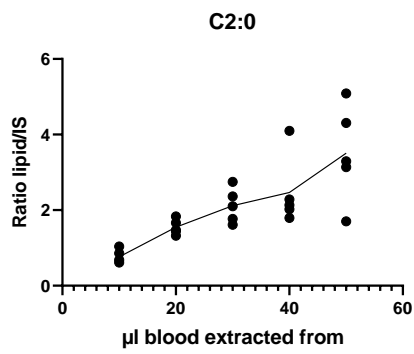
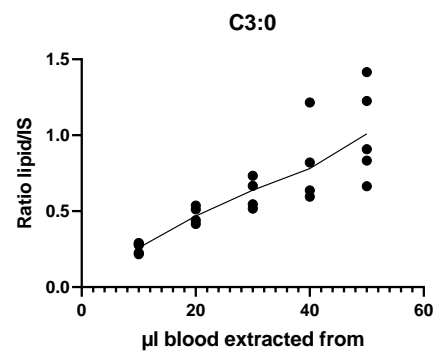
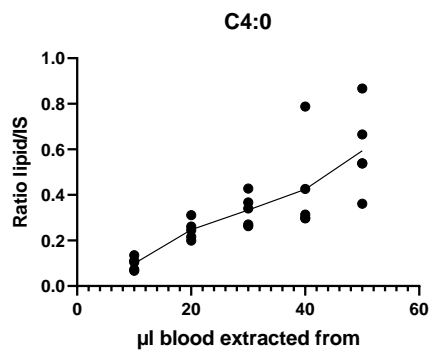
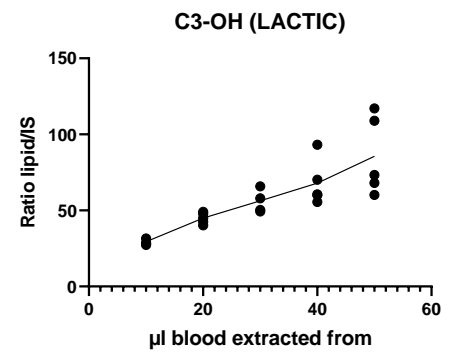
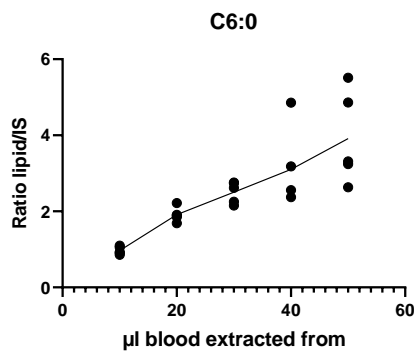
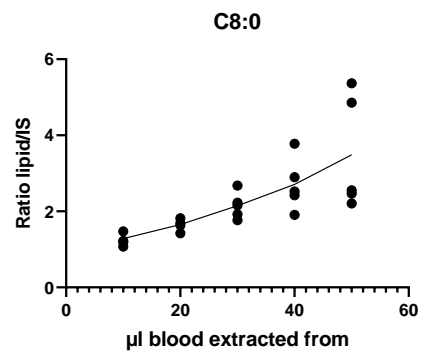
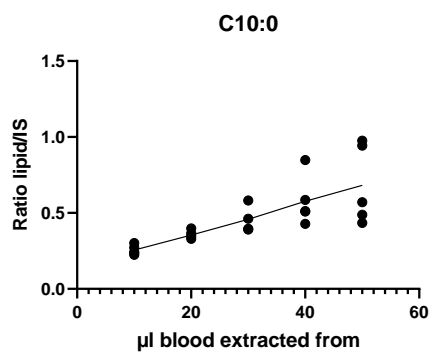
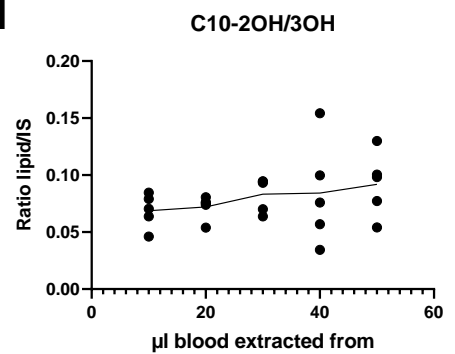
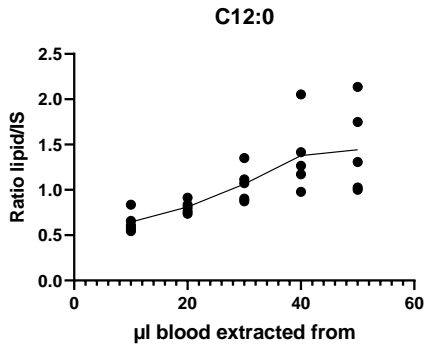
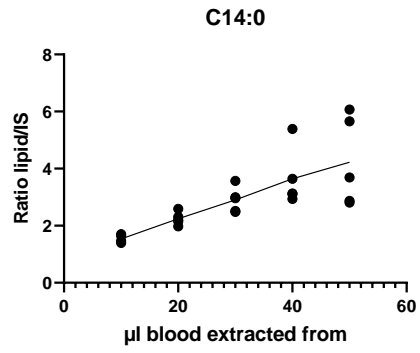
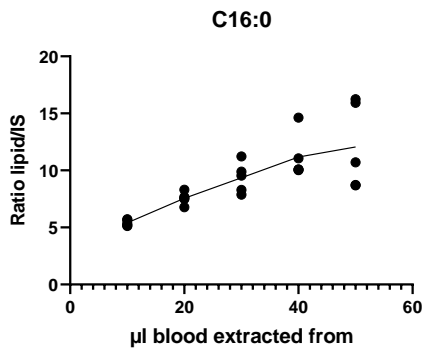
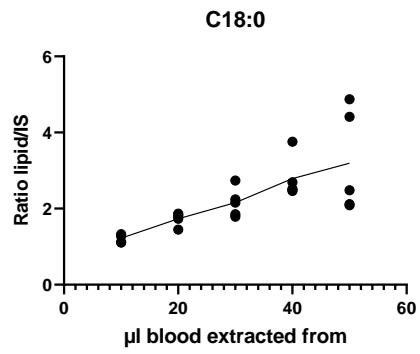
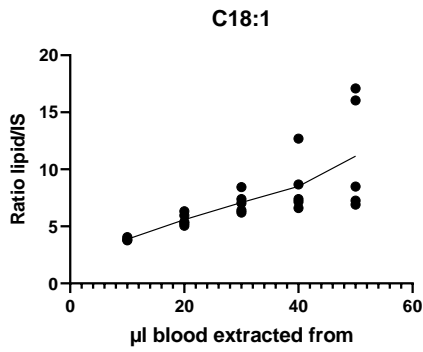
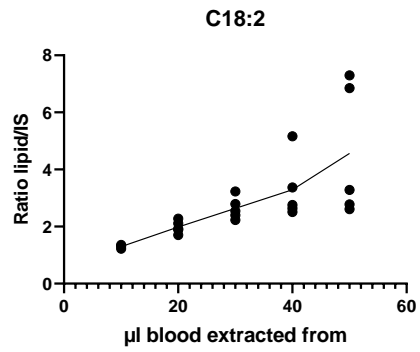
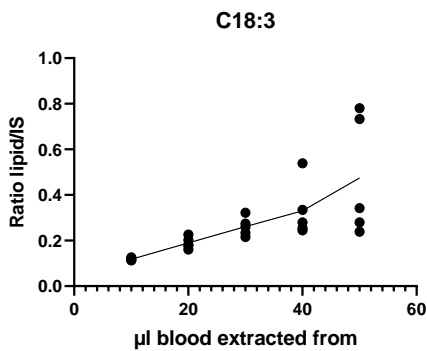
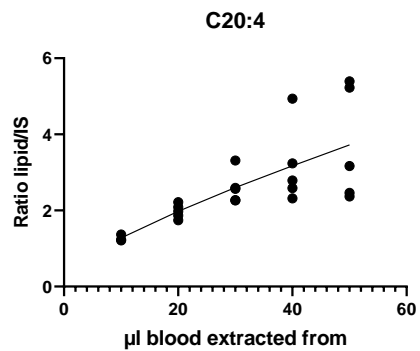


Figure 21. The figure shows the areas as measured by LC-MS/MS for the extracted deuterated standards (FFA, A, Acylcarnitines, B) for all extractions as average of the 5 replicates with standard deviation on the error bars. The x-axis depicts the amount of blood used in the extraction, and the y-axis the area measured by LC-MS/MS.

A**B****C****D****E****F****G****H**

I**J****K****L****M****N****O****P**

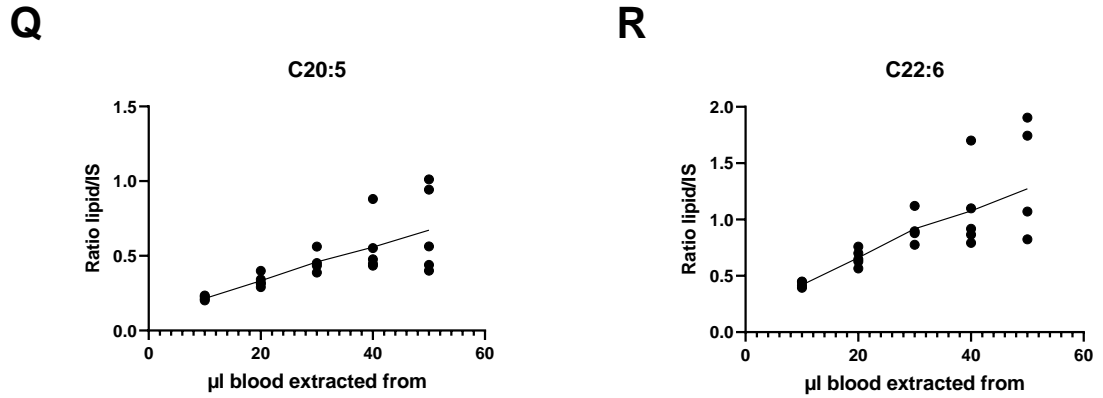
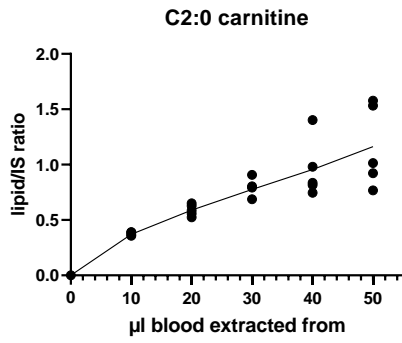
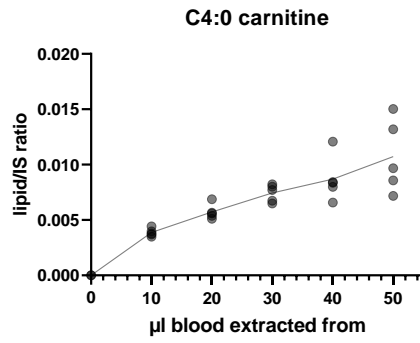
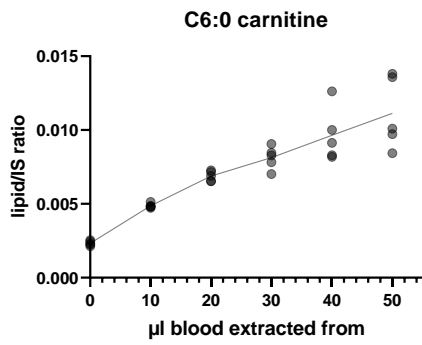
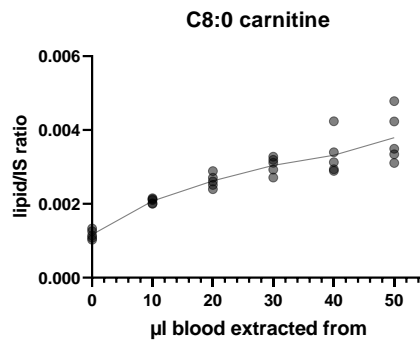
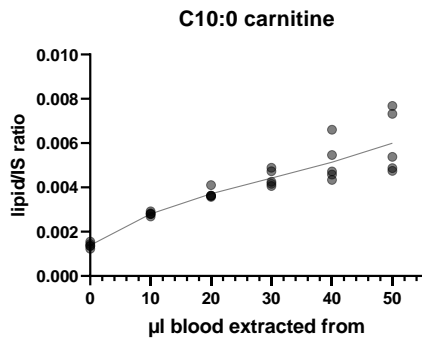
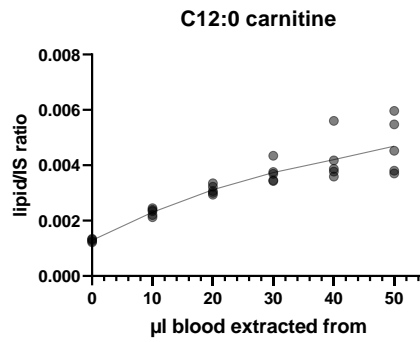
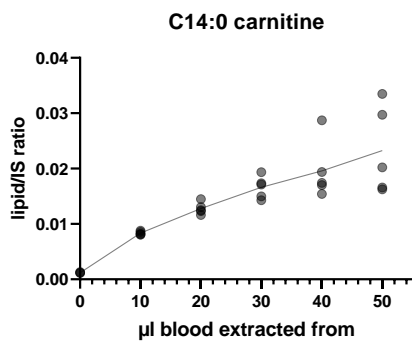
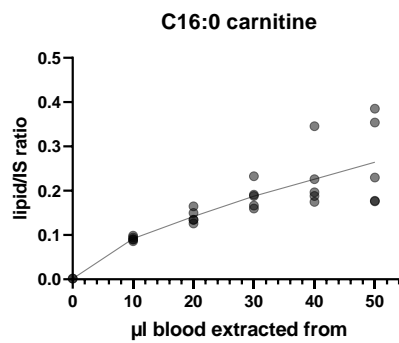


Figure 22 Lipids to internal standards ratios of FFA measured in samples with varying blood to MeOH ratio and the control sample containing no blood. 50, 40, 30, 20, 10 μl or 0 μl blood were mixed with 10 μl internal standard mix and a corresponding MeOH volume to an end volume of 260 μl as part of a protein crash extraction. The ratio of lipid to IS from the 0 μl blood sample was subtracted from the other samples for background correction. The y-axis depicts the ratio of the measured lipid to its corresponding or closest FFA as set out in the explanation above. The line indicates the average, and the dots indicate replicate measurements. This experiment was conducted with $n=5$.

A**B****C****D****E****F****G****H**

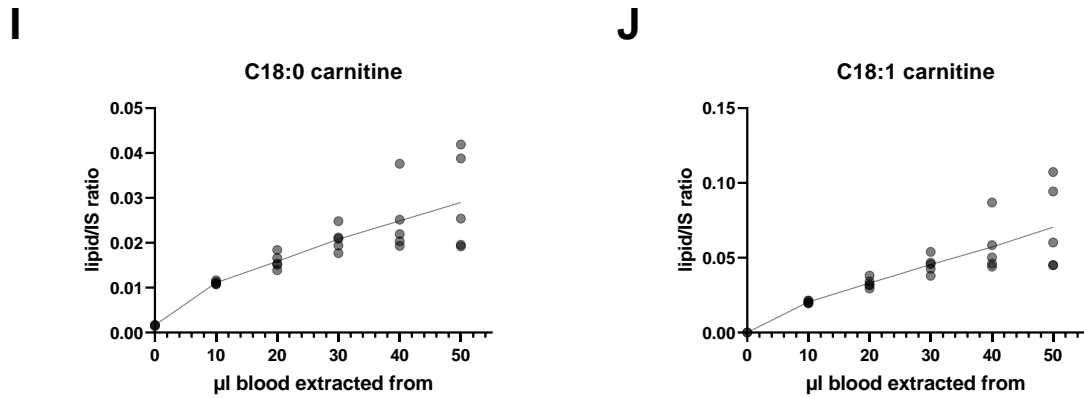


Figure 23 Lipids to internal standards ratios of acylcarnitines measured in samples with varying blood to MeOH ratio and the control sample containing no blood. 50, 40, 30, 20, 10 μl or 0 μl blood were mixed with 10 μl internal standard mix and a corresponding MeOH volume to an end volume of 260 μl as part of a protein crash extraction. The ratio of lipid to IS from the 0 μl blood sample was subtracted from the other samples for background correction. The y-axis depicts the ratio of the measured lipid to its corresponding or closest FFA. C18:1 carnitine to internal standard ratio was calculated based on deuterated C18:0 carnitine. This experiment was conducted with $n=5$.

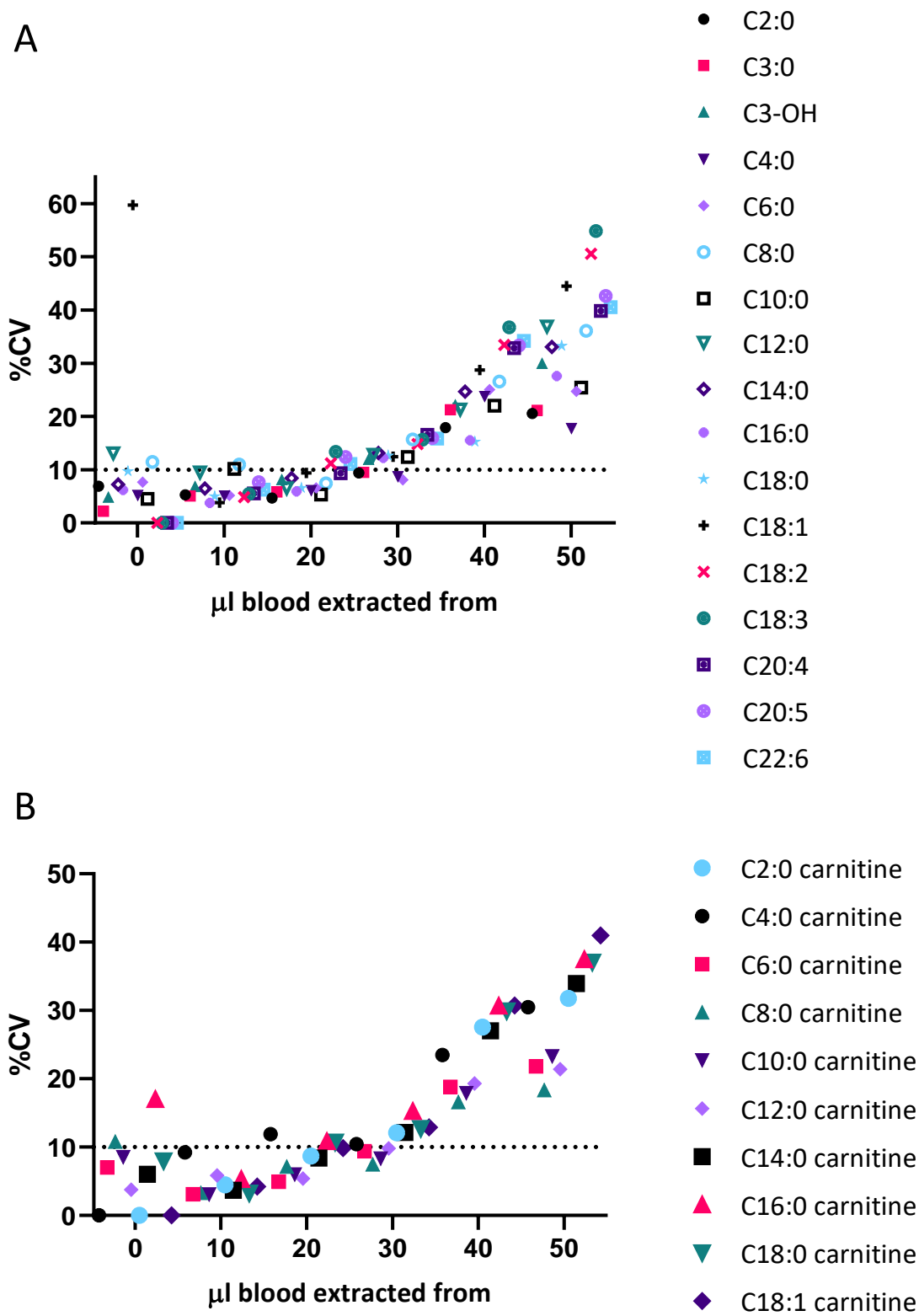


Figure 24 The plot shows the percentage of covariation for every lipid on the y-axis set out against the volume of blood extracted from. The total extraction volume was kept constant. This analysis indicates how much the extraction matrix impacts the reliability of lipid quantification. Fatty acids are shown in A) and acylcarnitines are shown in B).

3.1.5 Efforts to decrease the background in extraction and derivatization

Efforts were made to decrease the background in free fatty acids that could be observed in blanks after extraction. Firstly water and methanol were compared as blank solvents. To this end, aliquots of both HPLC water and methanol were extracted and derivatised according to the established protocols (see above) and levels of lipids (determined by LC-MS/MS) were compared. As is evident from Figure 25, the HPLC water is likely not a source of contamination or at least would contain similar levels of contamination to methanol.

Another potential source of background FFA are the derivatisation reagents themselves¹⁵⁸. Subsequently, since derivatisation reagents are likely used in excess, I tested if I could reduce the levels of derivatisation reagent used. To this end, I serially diluted the derivatisation reagents and used these diluted reagents to derivatize methanol (blanks) aliquots containing internal standards in triplicate. As can be seen in Figure 26, using less derivatisation reagent led to lower observed levels of deuterated internal standards, indicating that derivatisation was reduced. In addition, the backgrounds measured did not reduce proportionally to the reduction in derivatisation reagent, it hence seems likely that the reduction in derivatisation has a bigger impact than a reduction in background introduced by the derivatisation reagents.

The most reliable way of assessing background is to extract from a series of blanks during every extraction batch and then subsequently subtract blank signals (background) from all samples prior to further data processing. This approach was subsequently employed in all studies.

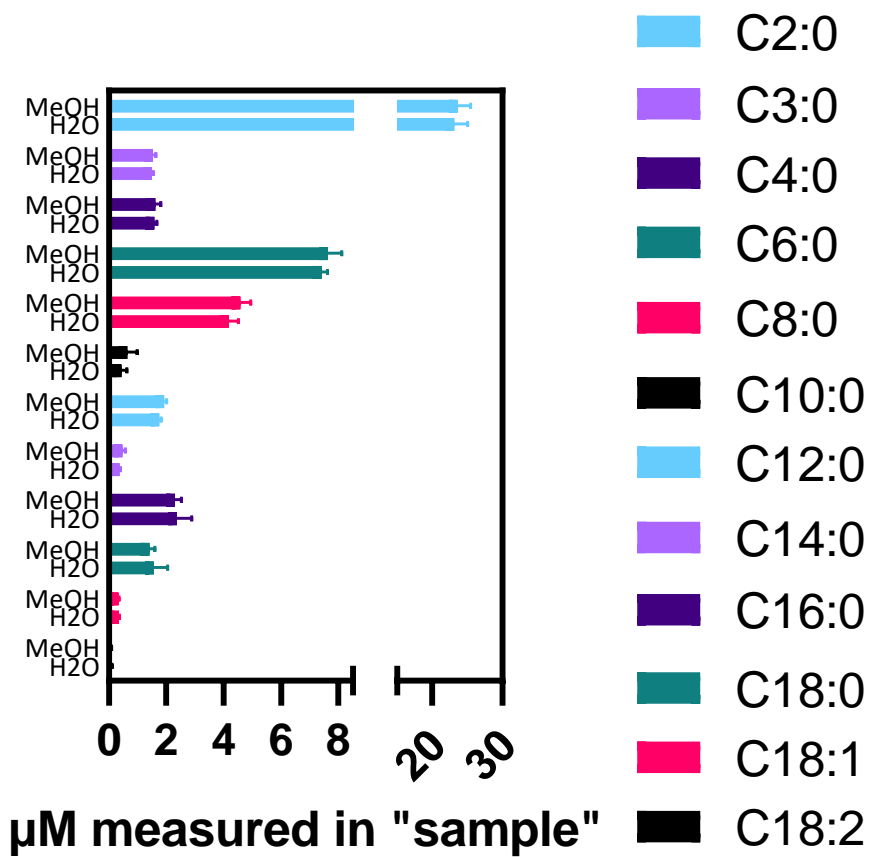


Figure 25 Levels of FFA in blank solvents. Bar plots of background measured when extracting from either a 10 μ l methanol (MeOH) or water (H₂O) "sample" according to the established protocol. There were no significant differences between MeOH and water extracts (2-way ANOVA).

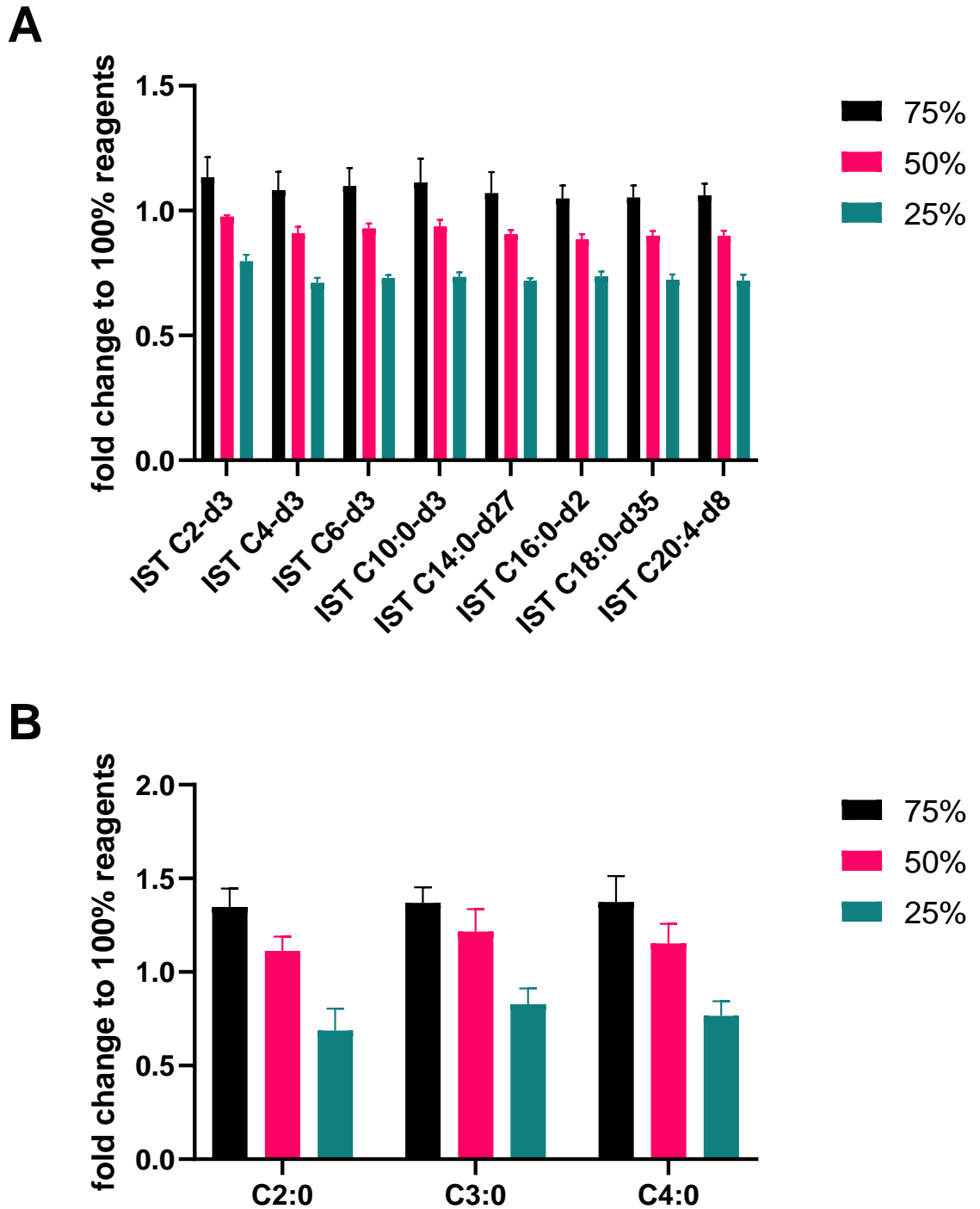


Figure 26 The fold change of measured lipid after derivatisation with different dilutions of derivatisation reagent. The derivatisation reagent was diluted to 75%, 50% or 25% of the concentration in the established protocol. The percentage of derivatisation reagent used is plotted against the areas measured per deuterated lipid. The percentage derivatisation reagent relates to the established protocol (100%).

3.2 Discussion

The aim of this investigation was to optimize a method that would allow for reliable FFA and acylcarnitine quantification from a low volume blood sample using LC-MS/MS. The most important aspect of the method was to achieve reliable measurement of medium-chain fatty acids (MCFA), with the measurement of further fatty acids (SCFA, LCFA) and acylcarnitines being a secondary desirable aim. The final established method consists of three consecutive parts: lipid extraction using the protein crash method, derivatisation using 3-NPH and EDC-pyridine, and the measurement by LC-MS/MS (Figure 27). The establishment of this method will contribute to our understanding of how these lipids change in sepsis but could also be used in future approaches in other conditions and most likely other complex mixtures. This would especially be relevant in relation to the measurement of MCFA, that previously have been neglected in many lipid studies. Both GC-MS and LC-MS machines are prevalent in laboratories performing lipid measurements, and the establishment of a LC-MS/MS method hence allows for a more widespread measurement of MCFA.

To evaluate the quality of the method, its components should fulfil certain requirements. The derivatisation process should lead to sensitive detection of analytes above background levels (present in solvents, reagents etc). The derivatisation with 3-NPH and EDC under the established conditions allows for detection of MCFA and LCFA at a picogram on column range that will be sufficient for detection of biologically relevant concentrations in the blood above background levels. The detection of acylcarnitines (very low/ND background levels) is very sensitive and should be sufficient despite the low expected concentrations in blood samples. The quantification of SCFA showed relatively high background levels and subsequent higher limits of detection as evidenced by the calibration curves. Attempts to address this issue by leaving out pyridine (a reported source of contamination¹⁵⁹) from the derivatisation reaction, led to failure of derivatisation. Also, deuterated SCFA internal standards showed a higher limit of detection than other deuterated lipids despite not detected background levels. This indicates the LC-MS/MS has a lower detection sensitivity for the SCFA compared to other lipids. Acetic acid (C2:0) is reported in relatively high concentrations in the blood but also shows the highest background levels. Propionic (C3:0) and butyric acid (C4:0) are reported in lower concentrations, which come close to the limit

of detection we found in our method evaluation¹⁴¹. The chosen derivatisation method hence appears robust and suitable for the measurement of all targeted lipids from low blood volumes (10 μ L) except for SCFA. However, in all cases blank extractions are required for subtraction of background levels of analytes in sample extracts prior to data processing. The quenching of the derivatisation reaction should reduce the amount of contaminating lipids measured without affecting the detection of already derivatized lipids and internal deuterated standards. Overall, the success of the quenching methods was rather modest and partially showed deleterious effects on measurements of deuterated, hence non-contaminating, lipids. None of the methods reduced the SCFA background without also showing negative effects on C2-D3. As a result, quenching with formic acid was chosen since it does not affect already derivatized lipids and partially reduced the measurement of contaminating LCFA.

The extraction of lipids should lead to robust results with little variation between the lipids extracted per replicate, little sample input volume required, and the successful extraction of all lipids specified. Efficient extraction and derivatisation of analytes from whole blood was most successful using the protein crash method. Though adequate for extraction of MCFA, biphasic layer extraction was poor for recovery of deuterated propionic acid and C6:0-carnitine. This method employs partitioning of compounds into two layers and the evaporation of the lipid-containing layer to obtain the lipids. It is likely that the relatively volatile SCFA evaporate in the evaporation step, whereas acylcarnitines potentially only partially partition into the solvent layer due to their chemical nature e.g., size and polarity. For all extraction methods differences in extraction efficiencies for different lipids were noted; it was hence decided to use corresponding internal standards for all lipids to be measured, contingent on availability. The normalization to internal standards should account for the moderate extraction efficiencies of some lipids, since the extraction efficiency for deuterated and not deuterated lipids from the sample will be the same. Since the input concentration of the deuterated lipids is known, the concentrations in the sample of the measured lipids can be calculated. For lipids where no exactly corresponding internal standard was available, calibration curves were established based on lipid to deuterated standard ratio using the closest deuterated standard available. This was the case for lactic acid (C3-OH), oleic acid (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), eicosapentaenoic acid (C20:5) and docosahexaenoic acid (C22:6). C3-OH was calculated

using deuterated C3:0; C18:1, C18:2 and C18:3 were calculated based on deuterated stearic acid (C18:0); eicosapentanoic acid (C20:5) and docosahexanoic acid (C22:6) were calculated based on deuterated arachidonic acid (C20:4). Another requirement for the method to be successful is the detection of lipids above background. This was evaluated by comparing PBS extractions to blood extractions. Only the protein crash method led to significant extraction above background for most fatty acids evaluated fulfilling this requirement. C10-2OH/3OH and C12-2OH/3OH could not be measured with different extraction approaches and hence might either not be present in the healthy volunteer samples used or not extracted successfully, since measurement of deuterated standards on the machine worked well. Furthermore, it is important to note that in the extractions a relatively high background of SCFA can be observed (Figure 19). This background persisted despite the change from acetonitrile to methanol earlier in the method development process after confirming presence of acetic and propionic acid in acetonitrile (data not shown). Further efforts to decrease the background as evidenced by the experiments outlined were not successful. When reducing the amount of derivatization reagents (a potential source of background levels) derivatization is affected negatively. In addition, background levels were measured at the same level regardless of if a HPLC pure methanol "sample" or a HPLC pure water "sample" was extracted. Apart from these experiments, several precautions during sample processing were established: every pipette tip was washed with methanol before pipetting any solvent and only glass vials were used for every step of the process. Despite these efforts, a background persists in all experiments and this background varies; it is hence not entirely possible to predict if it will be possible to reliably measure SCFA with this method in patient samples. Using a rigorous background correction and quality control however will indicate this. I hence chose to process 5 blank samples of HPLC grade water alongside every batch of samples in order to estimate the background, and the background measured in the blank samples will be subtracted as average from all samples in the batch.

In a last experiment both the optimal blood sample volume and the robustness of the protein crash extraction method was evaluated. When using one step extraction processes (i.e. not repeated solvent addition and recovery) such as the protein crash method, it is important to ascertain whether the extraction solvent is in sufficient volume to blood sample, to fully recover the analytes of interest i.e. it is not saturated. The results show that

a lower whole blood volume (10 μ l) in a total extract volume of 260 μ L led to reliable measurement of lipids above background; least variability between replicates ($\leq 10\%$ CV) and was in a linear extraction range for most FFA/acylcarnitines. The extraction of 10 μ l blood samples with the protein crash method was hence chosen as reliable and suitable method for lipid extraction from whole blood samples.

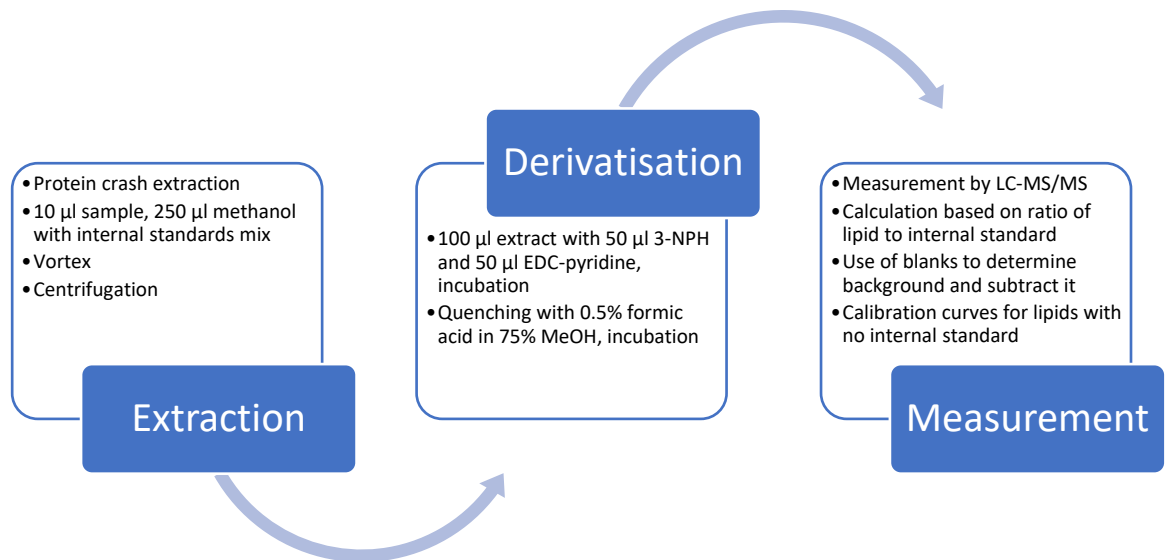


Figure 27 The optimized workflow as established in this chapter for free fatty acid and acylcarnitine quantification in blood samples.

4 Measurement of lipids in plasma of adult sepsis patients and healthy adult volunteers

In this chapter, I present the measurement of lipids in the plasma of adult sepsis patients and healthy adult volunteers. The lipids measured are straight chain fatty acids with a chain length from 2 carbons to 18 carbons, hydroxylated MCFA, the poly-unsaturated fatty acids C18:1 (oleic acid), C18:2 (linoleic acid), C18:3 (α -linolenic acid), C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid), C22:6 (docosahexaenoic acid) and acylcarnitines. Some of these lipids have been shown to change in sepsis in previous research^{50,51}, while some of these lipids to my knowledge have not been measured before in sepsis, e.g. C8:0 (octanoic acid), C10:0 (decanoic acid) and C12:0 (dodecanoic acid). Any changes occurring in the measured lipids would be connected to the origin of these lipids, their function, and cellular and physiological metabolism. The measured lipids can have different origins. Short chain fatty acids are foremost generated by the intestinal microbiota, whereas saturated MCFA and LCFA are more likely obtained from the diet but can also be generated in the body. Generally, most lipids can be synthesized in the body, while only few lipids can be obtained exclusively via the diet, e.g., C18:2 (linoleic acid) and C18:3 (α -linolenic acid). Acylcarnitines present in the plasma would most likely originate from inside the host cells, as dietary acylcarnitines will not directly pass into the blood. Acylcarnitines play an important role in cellular metabolism since they facilitate the transport of lipids into the mitochondria where they can be oxidized for energy generation. Since lipid metabolism has been described as being altered in sepsis it is relevant to measure both the free fatty acids and the acylcarnitines^{48,49,54,63}. Notably, most publications on lipid changes in sepsis cite overall lipid changes, i.e., from measuring non-esterified fatty acids or triglycerides, without specifying separate lipid species. Other studies perform untargeted approaches and only report significantly changed metabolites or the top changed metabolites. As a result, few information on the lipids I measured is published in previous studies, especially regarding the FFA. Previous studies have seen decreases in long-chain PUFAs in sepsis¹⁶⁴ and changes to acetate and lactate^{165,166,167}. I could not find information on the remaining FFA. Rogers et al. measured several of the lipids also measured in my approach, however, they

Results

compared survival in critically ill patients in the ICU and none of the here measured FFA were significantly altered between these two groups¹⁶⁸.

Measuring these lipids in sepsis will address several questions. Firstly, it assesses the viability of the established method for the physiological and pathophysiological measurement of a range of lipids in a well characterised cohort. Since the method was optimized on whole blood, this will also assess how suitable the method is for plasma samples. Secondly, this serves to address the questions about lipid concentrations and lipid metabolism in sepsis patients in contrast to healthy volunteers. The central question is if MCFA are present in adequate concentrations to activate the receptor and if they are elevated compared to healthy controls. Evaluating overall which lipids are decreased and increased in sepsis patients indicates underlying changes in metabolism. In addition, this also provides an initial investigation for exploring any connections between clinical parameters and lipid measurements.

To obtain reliable measurements of lipids by LC-MS/MS it is important to consistently adhere to a set of standard operating procedures for sample analysis. These standard procedures have been set out in the chapter detailing the optimization of the method. When analysing a cohort of clinical samples, it is common to check the performance of the method and prove the reliability of the results in combination of the presentation of the results. Furthermore, only lipids whose measurements are deemed reliable enough will be included in the study. In this study, due to potential backgrounds of certain lipids, the background of each lipid will be measured in every batch of extractions by extracting 5 blank samples of 10 µl HPLC water alongside the plasma samples. This background will be subtracted per lipid from every sample. If a lipid cannot be detected above background levels in the majority of samples, it will be excluded, unless the presence in a subset of samples is an expected possibility from a biological point of view. A final measure is the use of a quality control sample to evaluate machine performance during the run. During longer runs the performance of the machine can be affected by various factors affecting its reliable measurement, e.g., when relatively dirty samples lead to a build-up on the spray plate, hence impacting ionization and subsequently detection sensitivity. Therefore, a quality control (QC) sample is run after every 10 plasma samples and the measurements of the quality control sample over time will be evaluated. Though variability in lipid measurement in the quality control are common, a high variability in measurements of a lipid in the

quality control sample indicates that the machine is not reliably measuring this lipid. The variability is evaluated as coefficient of variation, which is the ratio from the standard deviation to the average measurement of the QC expressed as a percentage. In previous metabolomic studies of human tissues %CV of up to 50% have been accepted^{169,170}, though a %CV of 20% is often cited in LC-MS studies^{171,172}. In addition, a lipid has to be detected in the quantifiable range in all measurements of the quality control sample to qualify as being detected reliably.

The cohort used in this chapter is part of the Innate-like T-cells in Sepsis (ILTIS) study (IRAS ID: 231993). This study enrolled adults on the intensive care unit within 36 hours of admittance who were diagnosed with sepsis based on the sepsis-3 definition. The control group consists of healthy volunteers that are age matched. The data presented is from 52 sepsis patients and 18 healthy volunteers. Although this sample size is small this is a phenotypically well characterised cohort, and any significant alterations would warrant further investigation.

Results

4.1 Results

4.1.1 Demography of study populations

A summary of the demographics of the study population is summarized in Table 6. As can be seen, the median age is well matched, as it is 61 years in the patient population and 55 in the volunteers. The age range of healthy volunteers does not extend as high up as the age range of sepsis patients. The sepsis cohort has a higher percentage of males than the volunteer, 59.6% versus 44.4%. Among the comorbidities were cancer (n=2), chronic heart disease (n=1), chronic renal disease (n=1) and immunosuppression (n=2).

Roughly half of the cases had a confirmed infection (n=27), with several cases of pneumonia and urosepsis. Confirmed infections were overwhelmingly bacterial with only four viral flu cases diagnosed. 29 sepsis patients required mechanical ventilation, for an average of 10 days, and 20 patients received renal replacement therapy in ICU. Mortality in this cohort was at 21% within 60 days, or 11 patients out of 52.

Table 6 Demography of sepsis patients and healthy controls included in this study

	Sepsis patients	Healthy controls
n	52	18
Median age (age range)	61 yrs (19-84)	55 yrs (25-66)
Gender (%male)	59.6%	44.4%
Ethnicity		
- British White	- 49	
- Asian or Asian British	- 2	
Comorbidities	7	0
- Cancer	- 2	
- Chronic heart disease	- 1	
- Chronic renal disease	- 1	
- Immunosuppression	- 2	
Sepsis cause		-
- Post-operation	- 14	
- Bacterial	- 33	
- Flu	- 4	
Type of sepsis		-
- Pneumonia (confirmed/ unconfirmed infection)	- 12 (9/3)	
- Urosepsis (confirmed/ unconfirmed infection)	- 9 (1/8)	
- Sepsis (confirmed/ unconfirmed infection)	- 30 (17/13)	
Average APACHE II score (range)	17.8 (5-33)	-
Mechanical ventilation in ICU (average duration in days)	29 (10, 3-58)	-
Renal RT (replacement therapy) in ICU	20	-
Non-survivors 7 days (average age)	4 (69 yrs)	-

Results

Non-survivors 30 days (average age)	9 (69 yrs)	-
Non-survivor 60 days (average age)	11 (69 yrs)	-
Non-survivor 1 year (average age)	12 (67 yrs)	-

4.1.2 Evaluation of the reliability of lipid measurements in this cohort

Before analysing the results from a biological standpoint, I evaluated if the method to measure the lipids performed reliably, i.e., if the measurement of lipids in this cohort meets scientific standards set out in recent publications^{173,174}. Lipids whose measurements do not fulfil all criteria as set out in the introduction will be excluded. All samples were processed using the method established in Chapter 3 and described in further detail in Materials and Methods section 1. In short, 10 µl of plasma were mixed by vortex with 250 µl of ice-cold methanol containing internal standards, and then subjected to centrifugation at 4000 *g*. A 100 µl aliquot of the supernatant was derivatized with 50 µl 3-nitrophenylhydrazine (3-NPH) and 50 µl 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) before measurement by LC-MS/MS.

To ensure correct measurements, it was evaluated if lipids were quantified above background and if the measurements lay in the linear range. The background was detected by using 5 extracted blanks per extraction batch: 10 µl aliquots of HPLC water were extracted alongside each extraction batch. Most lipids could be detected in the majority of samples in concentrations higher than their measured concentrations in the background samples. C2:0 (acetic acid), C3:0 (propionic acid) and C4:0 (butyric acid) were detected at levels below background in the majority of samples, and the remaining measurements should hence be excluded as well from further consideration. Hydroxylated MCFA, i.e., C10-2OH/3OH and C12-2OH/3OH, were not detected at all and/or not quantified above background, in the majority of samples. The hydroxylated MCFA have been rarely detected before in blood, and if, only at very low concentrations¹⁴¹. In the majority of sepsis patients and healthy controls no hydroxylated C10:0 (decanoic acid) could be detected; only 11 samples showed detectable concentrations. Since hydroxylated C10:0 (C10-2OH/3OH) was initially not detected in the quality control measurements, but appeared in later quality control measurements, it cannot be ruled out that the measurements of hydroxylated C10:0 in these samples are artefacts. For this reason, all C10-2OH/3OH measurements are not considered in further analysis. In contrast, hydroxylated C12:0 (C12-2OH/3OH) could be measured in a larger portion of samples than hydroxylated C10:0. This lipid was detected above background in 30 samples from both patients and healthy controls. However, hydroxylated C12:0 could not be detected in the QC. Hence, C12-2OH/3OH was also excluded from further analysis. The medium-chain fatty acids C8:0 (octanoic acid), C10:0

Results

(decanoic acid) and C12:0 (dodecanoic acid) were measured below background in a smaller subset of samples, however since they were quantifiable in the majority of samples this is acceptable. Three lipids were quantified outside of the higher end of the linear range in most of the samples: C3-OH (lactic acid), C16:0 (palmitic acid) and C18:1 (oleic acid). To measure these lipids 10-fold dilutions of all samples and the QC were prepared and measured again in one consecutive run. The values reported for mean and QC in these lipids are based on the diluted sample measurements (Figure 28) and were calculated taking the dilution factor into account. The calculated concentration hence represents the concentration in the undiluted sample.

An important parameter to evaluate is the machine performance of the LC-MS/MS with the method used. All samples were run on a Nexera XR LC-20AD XR (Shimadzu) liquid chromatograph coupled to a SCIEX 4000 QTRAP LC MS/MS system (Shimadzu) with Turbo Spray ESI in one continuous run. To evaluate machine performance, a quality control sample - a pool of all samples - was measured after every 10 samples measured. Consistent measurement of lipids indicates robust machine performance. Since it is a pool of all samples, the background subtracted is the average of all extracted blanks of all batches. As can be seen in Figure 28, most lipids and carnitines were measured in the quality control in a consistent manner. The lipids C3:0 (propionic acid) and C12-2OH/3OH were not measured above background in most or all samples of the quality control, confirming the previous decision to exclude these lipids. Coefficient of Variability (%CV) was used to assess variability in measurements. Fatty acids present in a relatively low amount in the quality control, e.g., C3:0 (propionic acid), C4:0 (butyric acid) and C6:0 (hexanoic acid), are not quantified reliably as seen by the high %CV values of well above 200% (Figure 28C). These fatty acids are present in a concentration outside the range in which their reliable measurement is possible. Any variability occurring in measurements of these lipids in the tested cohort could hence be due to variability in quantification within this range rather than reflecting actual variability in concentrations. That is why C6:0 (hexanoic acid) will be excluded as well as C3:0 (propionic acid) and C4:0 (butyric acid) which were also already excluded due to the failure to detect them in most samples. The remaining free fatty acids have %CV below 16%, with several showing very low %CV below 5%. When comparing this to the %CV of 20% usually deemed acceptable as set out in the introduction, these values are well below the acceptance level and indicate reliable measurement of these lipids. All

carnitines can be detected with %CV values below 10 and hence the measurements of these analytes can be deemed very reliable as well. In conclusion, for further analysis, I decided not to consider, C2:0 (acetic acid), C3:0 (propionic acid), C4:0 (butyric acid), C6:0 (hexanoic acid), C10-2OH/3OH (hydroxylated decanoic acid) and C12-2OH/3OH (hydroxylated dodecanoic acid), while the quantification of the remaining lipids appears reliable.

Results

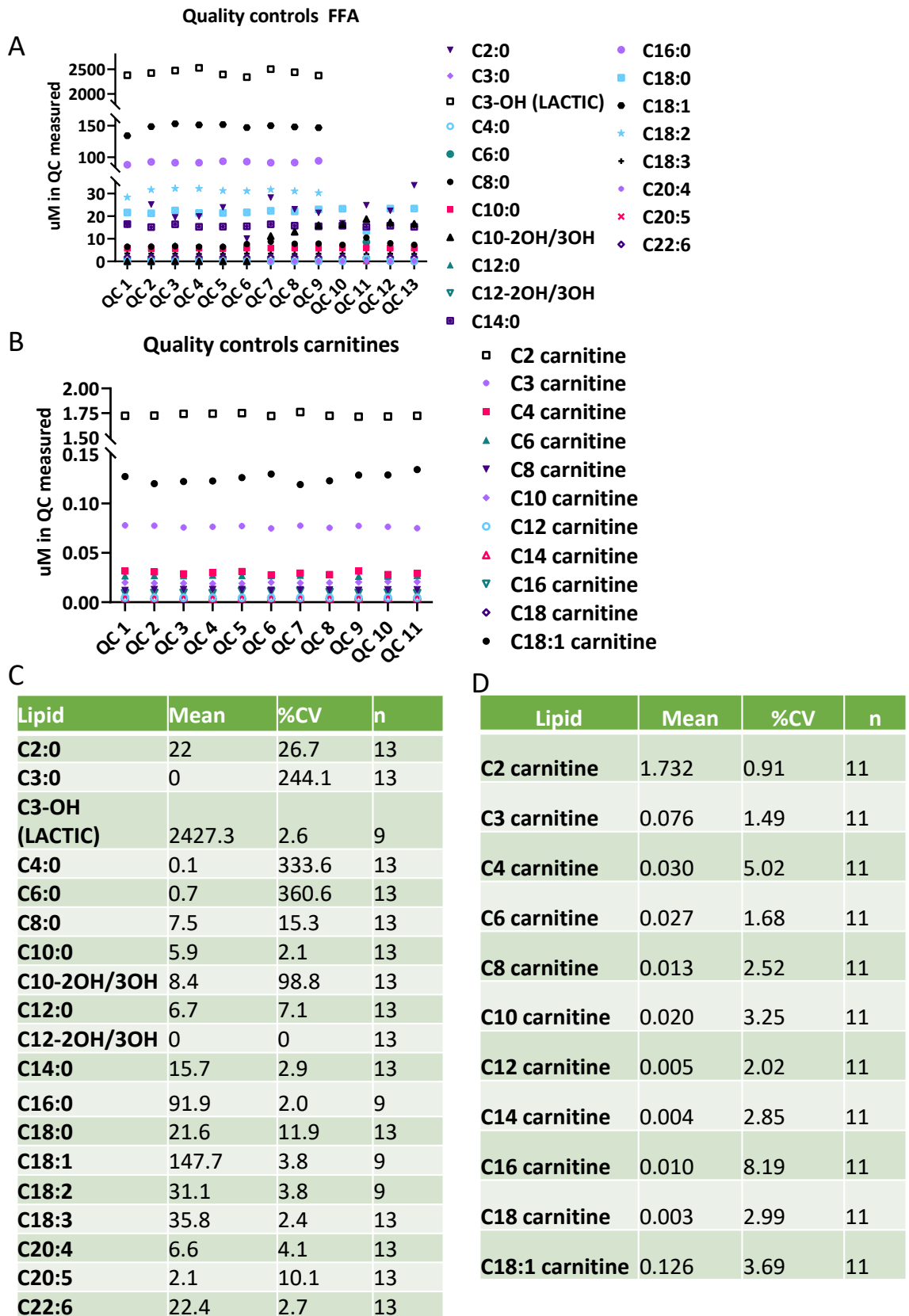


Figure 28 Measurement of lipids in the quality control. A) Measurement of free fatty acids across 13 measurements (undiluted samples) or 9 (diluted samples) of the quality control sample. C) and D) Tables summarizing the mean, %CV and n of quality control

measurements for FFA and acylcarnitines. B) Measurements of acylcarnitines across 11 quality control measurements. QC1 is the first quality control measurement and QC13, 9 and 11 are the last quality control measurement for FFA diluted and undiluted and acylcarnitines, respectively. The difference in how many quality control sample measurements were performed is caused by differences in how many measurements were executed per run, i.e., when sample measurements were repeated due to technical issues such as leaks in the liquid chromatography or moving peaks.

Results

4.1.3 Measurements of potential ligands for GPR84

The first question was whether MCFAs that are able to activate GPR84 are present in plasma and potentially increased in sepsis. Previous studies report that the MCFA C10:0 (decanoic acid), C12:0 (dodecanoic acid) and hydroxylated C10:0 and C12:0 activate the receptor^{96,97,122}. Since the hydroxylated MCFAs could not be detected reliably in plasma, I can only answer this question for C10:0 and C12:0. Firstly, C10:0 (decanoic acid) could be detected in nearly every sample. It is significantly increased in sepsis compared to healthy controls ($p=0.043$, Wilcoxon-test). It is noteworthy that in most patients the C10:0 levels were low and at similar levels to those observed in healthy controls, while a subgroup (13 patients) shows clearly elevated levels at about 4 - 28 times the median (median:1.9 μM , range outliers: 8.5 – 53.6 μM). In Figure 29 dashed lines are shown in the C10:0 and C12:0 plots to indicate the concentration levels required to activate the receptor using as the reference value the EC50 values obtained in cAMP assays by Wang et al. (C10:0 4.5 μM , C12:0 8.8 μM)⁹⁶. As can be seen in Figure 29, a subset of patients had elevated C10:0 levels in their plasma in the range that would activate the receptor in vitro. For C12:0 (dodecanoic acid) on the other hand, concentrations in healthy controls and sepsis patients are very similar and the spread of values in the patient population is less extreme. The subset of patients with sufficiently high plasma levels to activate the receptor is very small. In addition, some healthy controls had sufficiently high plasma levels as well. In roughly 20% of the samples, C12:0 could not be detected. In conclusion, it seems that in a subset of patients the concentrations required to activate GPR84 can be reached systematically.

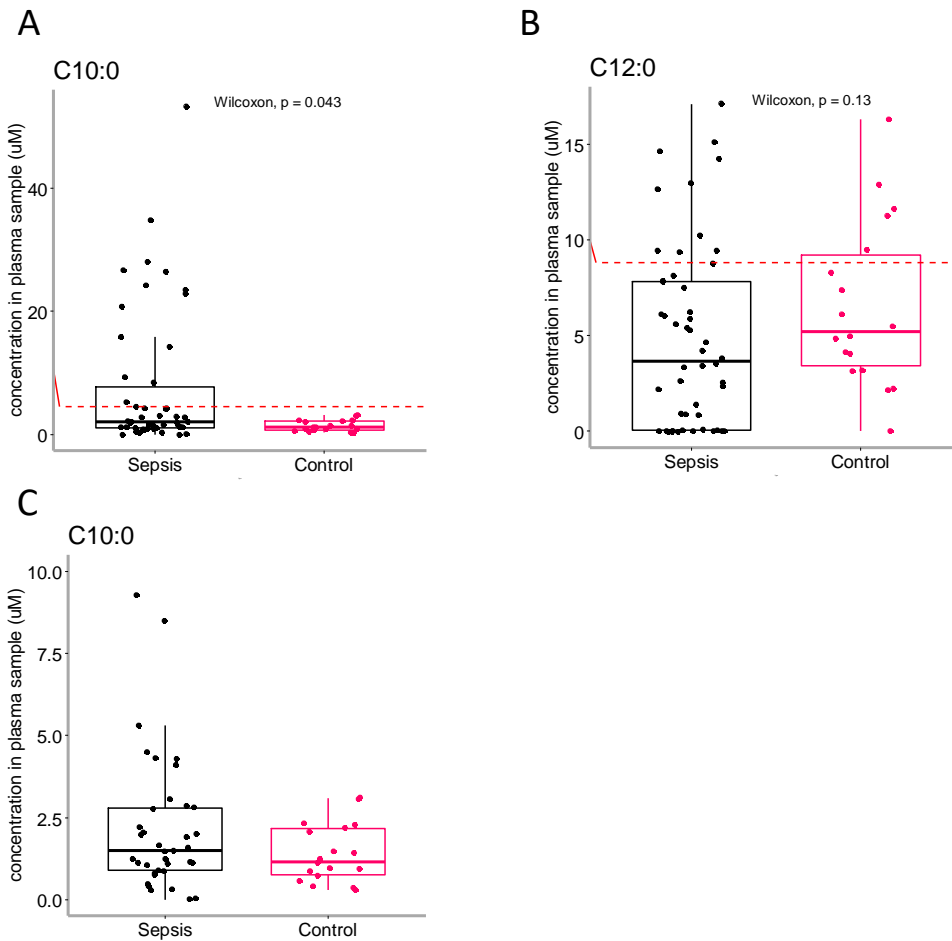


Figure 29 Boxplots of the concentrations in μM of potential GPR84 ligands in the plasma of sepsis patients (n=52) and of healthy controls (n=18); A) C10:0 or decanoic acid, B) C12:0 or dodecanoic acid, C) all samples with C10:0 levels below 10 μM . The dashed line indicates the EC50 concentration required to activate GPR84 as evaluated by Wang et al. using a cAMP assay⁹⁶(C10:0 4.5 μM , C12:0 8.8 μM). The statistical test performed was a Mann-Whitney-Wilcoxon test with a significance level of 0.05. The lower half of each box in the boxplot depicts the 2nd quartile of values, with the upper half depicting the third quartile. The extended lines depict the first and last quartile of values, with any points reaching below or above these lines being statistical outliers.

Results

4.1.4 Measurements of additional straight chain lipids

Further fatty acids were measured in the range from C8:0 (octanoic acid) to C22:6 (docosahexaenoic acid) and C3-OH (lactic acid). These fatty acids were measured in order to understand lipid metabolism better and explore any connections between changes in other fatty acids and the changes observed in MCFA. First, C3-OH or lactic acid, was measured in high concentrations with very similar concentrations measured in controls and in sepsis patients (Figure 30 A). This is surprising since increased lactic acid is proposed as biomarker for sepsis^{3,5}; I hence would have expected to see increased concentrations in at least a subset of sepsis patients compared to healthy controls. The concentrations measured in healthy controls are higher than reported, as lactic acid is typically reported between 500 and 2200 μM according to the HMDB¹⁴¹. Furthermore, a concentration above 2000 μM is an indicator of septic shock as defined in the sepsis-3 definition³, a value that is exceeded by the average for both cohorts. The reason underlying this discrepancy is not readily clear as the quality control measurements show a low %CV of 2.6% and the measurements are in the linear range of the calibration curve. This argues against possible variation due to technical reliability of the measurements and may suggest variation arising from the manner of sample collection for the controls as well as the small sample size of the population. Sample collection and storage has previously been found to significantly affect measured lactate levels from both volunteers and patients¹⁷⁵⁻¹⁷⁷. In this connection, about a third of the controls show levels above the normal range and all below septic shock range of 4mmols/l. In both the sepsis patient group and controls approximately half show sepsis 3 levels (<2 mmols/l) with 7 individuals displaying levels in the septic shock range.

C8:0 or octanoic acid is present in significantly higher concentrations in sepsis patients than in healthy controls ($p=0.0011$, Mann-Whitney-Wilcoxon-test), with concentrations observed in healthy controls being very low or not detectable (Figure 30B and F). The levels observed in the controls is in line with previously published data that medium chain fatty acids are present in low amounts in the blood of healthy fasted and non-fasted volunteers^{150,178}.

C14:0 (myristic acid), C16:0 (palmitic acid), and C18:0 (stearic acid), show no difference in sepsis patients and healthy controls, with a wider spread of values in sepsis patients than

in controls (Figure 30C, D, E). There is a small but non-significant increase apparent in C16:0 (palmitic acid).

C18:1 (oleic acid), is significantly increased in sepsis, while C18:2 (linoleic acid) and C18:3 (α -linolenic acid), are not (Figure 31 A, B, C). The further LCFA C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid) and C22:6 (docosahexaenoic acid) are significantly lower in the plasma of sepsis patients than in healthy controls (Figure 31 D, E, F).

Overall, it seems that straight chain saturated fatty acids with a chain length of longer or equal to 12 are not changed in sepsis, while medium chain fatty acids with a chain length of 8 or 10 are increased in a subset of patients. In terms of unsaturated fatty acids, only C18:1 (oleic acid) is significantly increased while C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid) and C22:6 (docosahexaenoic acid) are decreased, and C18:2 (linoleic acid) and C18:3 (α -linolenic acid) are not significantly affected. It is noteworthy that most lipids show a great heterogeneity with several plasma concentrations quantified outside of the upper quartile of values.

Results

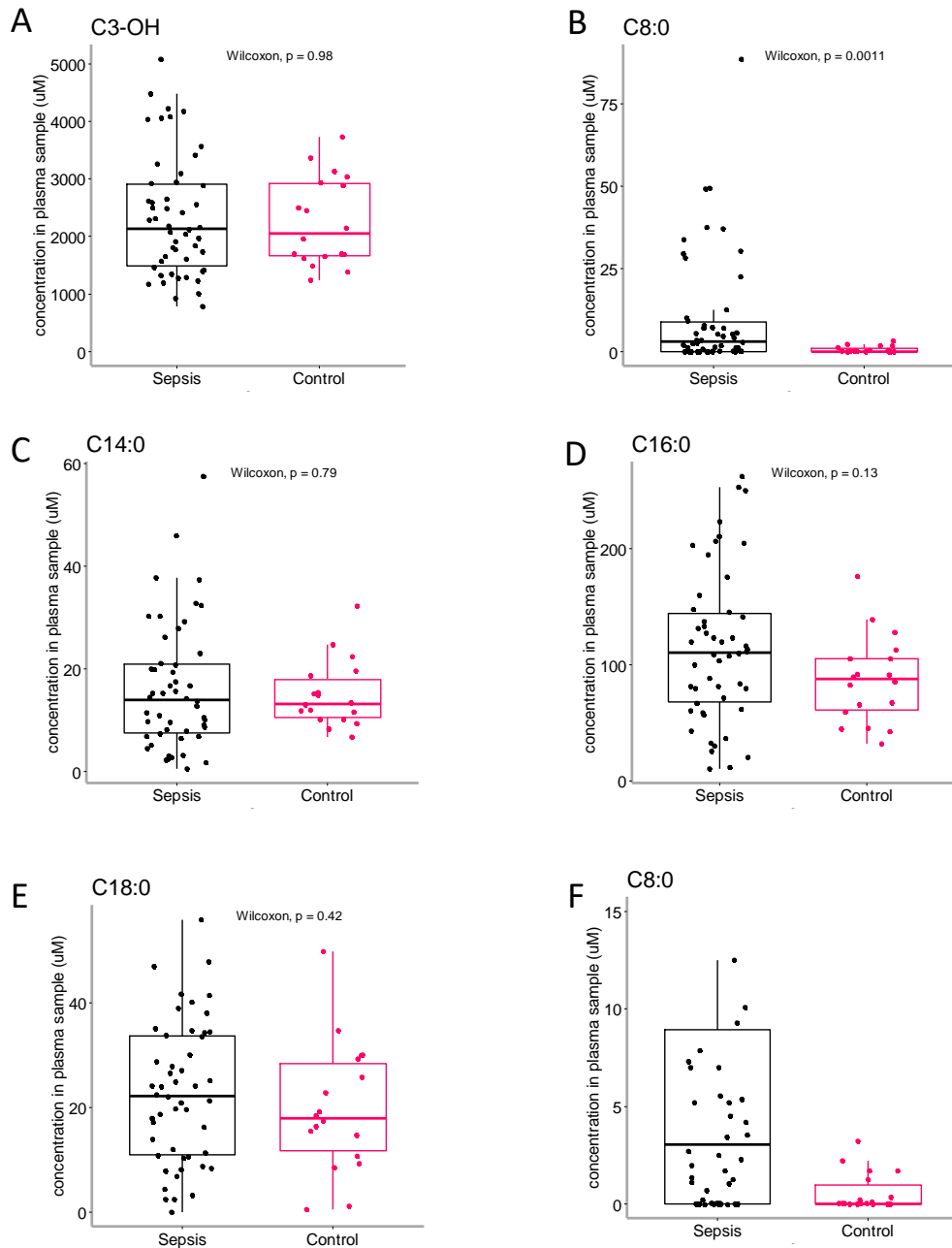


Figure 30 Concentrations in μM in the plasma of sepsis patients and healthy controls; A) C3-OH or lactic acid, B) C8:0 or octanoic acid, C) C14:0 or myristic acid, D) C16:0 or palmitic acid, E) C18:0 or stearic acid, F) C8:0 in all samples with concentrations up to 15 μM . The statistical test performed is a Mann-Whitney-Wilcoxon test with a significance level of 0.05.

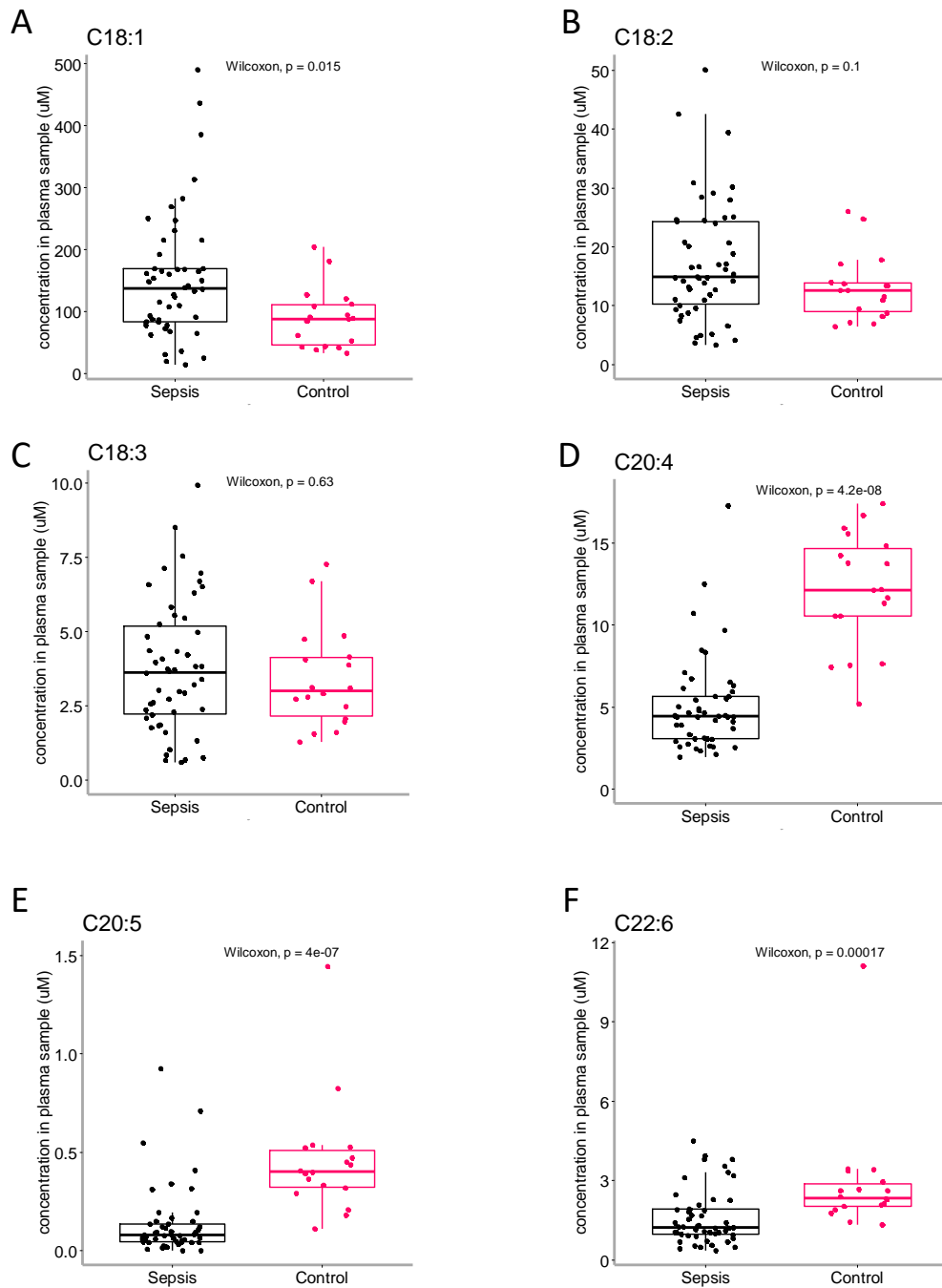
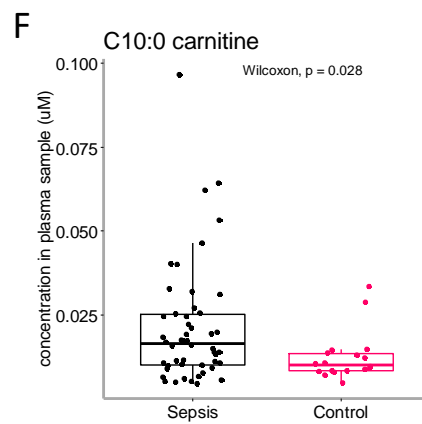
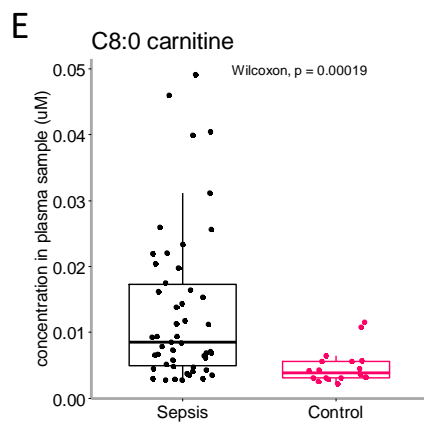
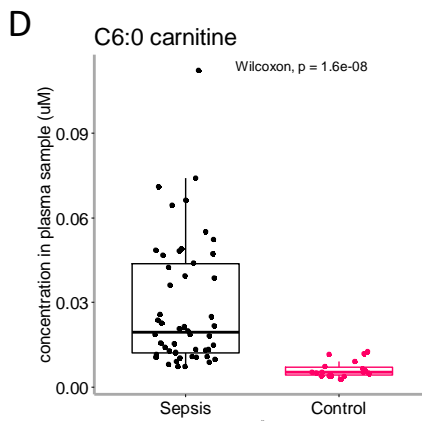
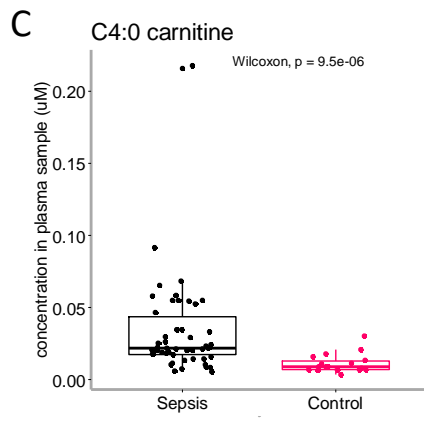
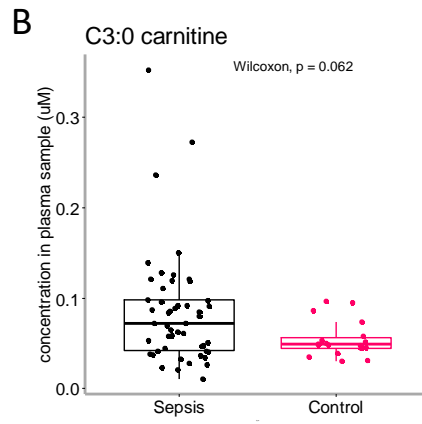
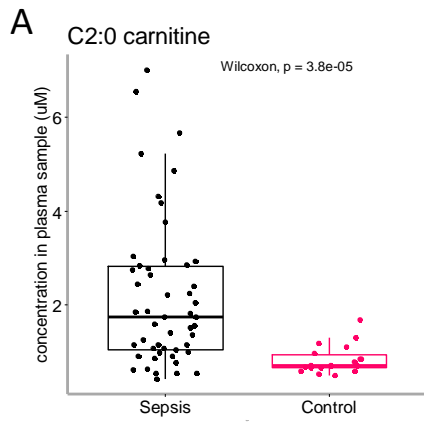


Figure 31 Concentrations in μM in the plasma of sepsis patients and healthy controls; A) C18:1 or oleic acid, B) C18:2 or linoleic acid, C) C18:3 or α -linolenic acid, D) C20:4 or arachidonic acid, E) C20:5 or eicosapentaenoic acid, F) C22:6 or docosahexaenoic acid. The statistical test performed is a Mann-Whitney-Wilcoxon test with a significance level of 0.05.

Results

4.1.5 Measurements of acylcarnitines

Apart from free fatty acids, acylcarnitines with a range from acetylcarnitine (C2:0 carnitine) to oleoyl-carnitine (C18:1-carnitine) were measured. Since acylcarnitines play a pivotal role in facilitating lipid transport into mitochondria, measuring them will give further insights into potentially changed lipid metabolism inside the cell. The transport of lipids into the mitochondria is necessary for both oxidative phosphorylation, the energy-generating breakdown of lipids, and lipid synthesis in mitochondria. Boxplots depicting acylcarnitine measurements in sepsis patients and healthy controls can be seen in Figure 32. Significant differences between sepsis patients and healthy controls can be observed for all acylcarnitines with a chain length between C2:0 and C:10, except C3:0 carnitine (propionylcarnitine), and stearyl carnitine (C18:0-carnitine) (Figure 32 A-F, J). The short- and medium chain fatty acylcarnitines with a chain length between 2 and 10 are increased in sepsis patients, whereas stearyl carnitine (C18:0-carnitine) is decreased in the plasma of sepsis patients. C3:0 carnitine (propionylcarnitine) is near to significance ($p=0.062$). Overall, and as expected, acylcarnitines are measured in low, mostly submicromolar concentrations. The acylcarnitine measured in the highest concentrations is acetylcarnitine with concentrations in the low micromolar range (Figure 32 A).



Results

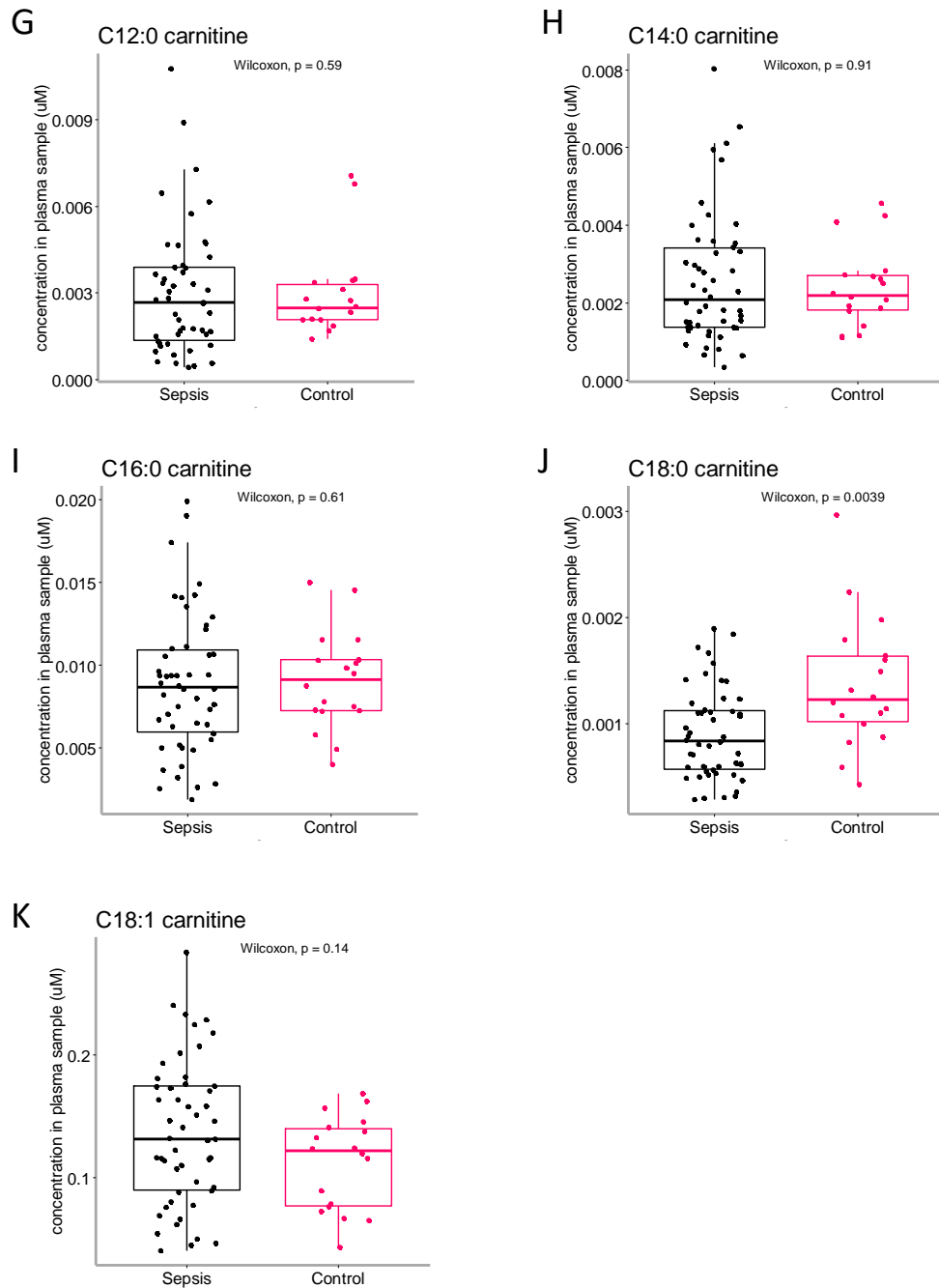


Figure 32 Boxplots depicting the acylcarnitine concentrations measured in plasma samples of sepsis patients and healthy controls. Shown are A) acetylcarnitine (C2:0-carnitine), B) propionylcarnitine (C3:0-carnitine), C) butyrylcarnitine (C4:0-carnitine), D) hexanoylcarnitine (C6:0-carnitine), E) octanoylcarnitine (C8:0-carnitine), F) decanoylcarnitine (C10:0-carnitine) G) dodecanoylcarnitine (C12:0-carnitine), H) myristoylcarnitine (C14:0-carnitine), I) palmitoylcarnitine (C16:0-carnitine), J) stearoylcarnitine (C18:0-carnitine) and K) oleoylcarnitine (C18:1-carnitine). The statistical test performed is a Mann-Whitney-Wilcoxon test with a significance level of 0.05.

4.1.6 Identification of lipids significantly altered in sepsis

An analysis using the metaboanalyst tool¹⁴² was performed to further explore the metabolic changes observed in sepsis. My intention was to generate a better overview and hence understanding of the metabolic changes in sepsis in this study. In addition, I aimed to use any insights as starting point to understand the factors that might be causing the heterogeneity in lipid levels among sepsis patients. Metaboanalyst is a platform for metabolomics analysis enabling various statistical approaches to metabolomics data and visualizations. First off, a Volcano plot provides a concise overview of lipids significantly different and how they are changed in sepsis. It plots the \log_2 -transformed fold change versus the $-\log_{10}$ -transformed p-value and typically only the measurements above the fold change and significance threshold are shown. The statistical test performed was a Kruskal-Wallis-test with multiple testing adjustment on \log_{10} -transformed lipid concentrations. In addition, dashed lines indicate a \log_2 FC of 1, i.e., a fold change of 2, but lipids with a lower fold change are also depicted. As can be seen in Figure 33, C20:4 (arachidonic acid) and C20:5 (eicosapentaenoic acid) were significantly decreased, while several carnitines that bind small-chain to medium-chain fatty acids were significantly increased in the plasma. Also, C8:0 (octanoic acid) and C10:0 (decanoic acid) were significantly increased. All of the lipids that were found to be significantly changed in the previous analysis are also significant in this analysis. It becomes apparent in the volcano plot that some of the significant lipids have a fold change lower than 2. While the boxplots and associated statistical tests provide a more granular view unto lipid changes, the volcano plot identifies the most robustly changed and thus, most likely, the most important lipids.

Results

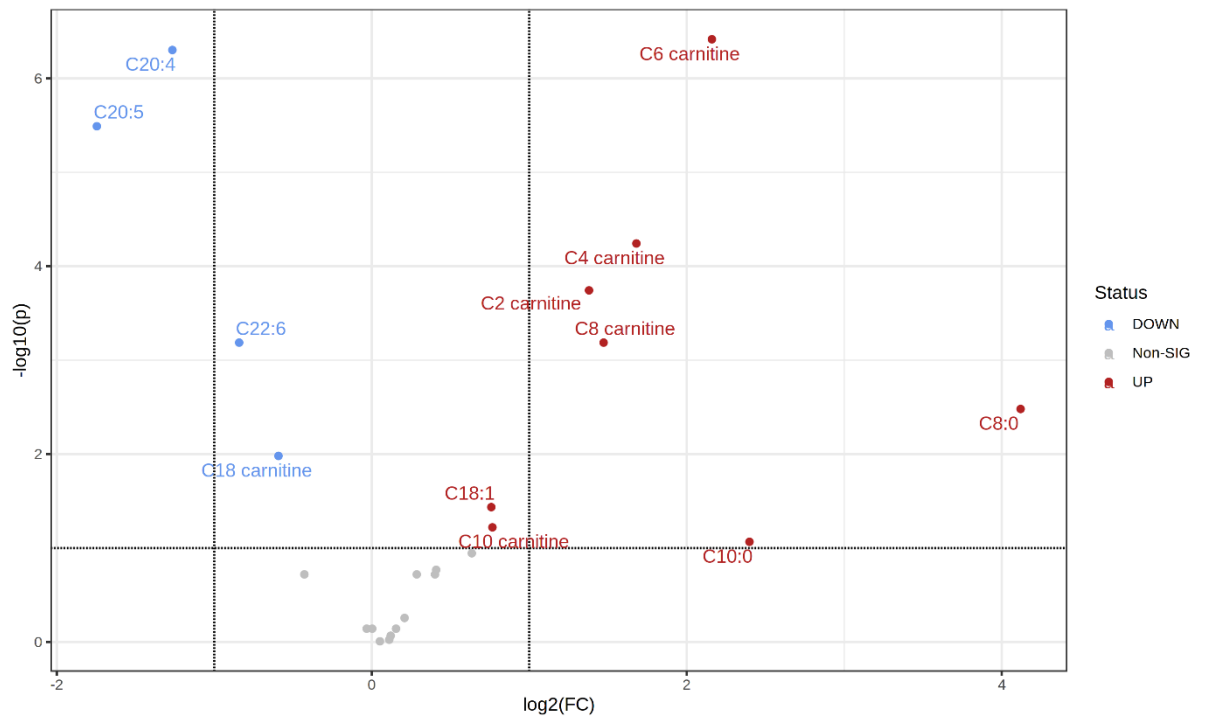


Figure 33 Volcano plot depicting lipid concentrations altered in the plasma of sepsis patients compared to healthy controls. The y-axis depicts the negative log₁₀ of the p-value, hence the higher the value the higher the significance. The x-axis depicts the log₂ of the fold change calculated based on not-normalized lipid concentrations. The dashed lines at log₂(FC) 1 indicate a 2-fold change on raw data. Significance was calculated using the Kruskal-Wallis test on log₁₀-transformed lipid concentrations and is adjusted for multiple testing. Dashed lines indicate a log₂(FC) of 1 and a $-\log_{10} p$ -value that corresponds to $p=0.1$ after multiple testing correction.

4.1.7 Explorative, unsupervised investigation of lipids altered in sepsis

Several of the lipids exhibited a wide range of measurements among the sepsis patients. I hypothesized that the different levels of lipids may be correlated to subsets of patients. To examine if this wide range of measurements indicates separate subgroups of patients, I generated a heatmap of lipid measurements in the patients and controls (Figure 34 A), with the colour bar on top identifying survivors, non-survivors, and healthy controls, and a second bar identifying controls, urosepsis, pneumonia and sepsis cases. The lipid measurements are log-normalized and centre-scaled. Both lipids and samples are clustered using the Ward.D clustering method, and Euclidean distance measures in metaboanalyst. A heatmap could show distinct clusters of patients and controls and using the colour bars on top could help identify the reason for these clusters. However, the heatmap reveals a high level of heterogeneity and the clusters are not very striking. The samples are agglomeratively divided into two clusters, one cluster containing healthy controls and the majority of sepsis patients, and the other cluster containing a smaller subset of sepsis patients. This indicates that the subcluster of sepsis patients on the left has a lipid profile different from both controls and other sepsis patients. Furthermore, a few sepsis survivors cluster with the healthy controls. Sepsis non-survivors do not cluster together but are distributed over several clusters. Looking at the different sepsis causes as indicated by the colours of the second colour bar, it is also apparent that the different types do not form separate clusters. The lipids are first divided into two clusters, with each cluster then containing another two clusters consisting of several subclusters. The bottom cluster contains one cluster with the MCFA C8:0 (octanoic acid), C10:0 (decanoic acid), C12:0 (dodecanoic acid) and all acylcarnitines with a fatty acid chain length between 2 and 10, and another cluster with the remaining acylcarnitines with a fatty acid chain length exceeding 10. The top cluster contains all long-chain fatty acids, with one cluster containing C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid) and C22:6 (docosahexaenoic acid), and the other cluster containing C14:0-C18:0 and C18:1 (oleic acid), C18:2 (linoleic acid) and C18:3 (α -linolenic acid). The relationship between the lipids will be further explored in a separate analysis (Correlation between lipids).

To further explore the clustering of patients based on the lipid measurements, Principal Component Analysis (PCA) was performed in metaboanalyst. The PCA is an unsupervised analysis method that arranges samples that vary in a similar way in an n-dimensional space

Results

based on the provided variables, i.e., measured concentrations of the FFAs, after converting these into abstract principal components. Hence, the principal components are generated based on abstracted, differently weighted values of measurements that are supposed to best explain the variance, not the grouping, in the data. Figure 34 B shows the PCA plot in 3D. The distribution of patients and controls in the 3D plot confirms the impression that differences in lipid profiles are present between different subgroups but due to their heterogeneity do not appear very stark, as no distinct groupings of samples become apparent in the plot. This makes sense as the differences found are heterogeneous, i.e., several lipids have high average fold changes but also show a great spread of values, especially in the sepsis group. We hence see a PCA with samples relatively spread out without establishing distinct groupings.

I decided to further explore the PCA outcome using 2D-plots of the PCA instead, i.e., biplots. In the 2Dplot of PC1 and PC2 (Figure 34 C) one can see that there is a trajectory from healthy controls to sepsis survivors to non-survivors, moving from left to right, i.e., along the x-axis or PC1. Some of the lipids that are part of PC1 are C8:0 and C10:0. When looking at a biplot of PC2 and PC3 (Figure 34 D), a group of 10 patients and 1 control appears to be further apart from the other samples. Evaluating the position of this group in the biplot it is apparent that this group is more dissimilar to the lipid profile of non-survivors than most survivors and controls. This group consists of 1 non-survivor, 9 survivors and 1 control. Among the patients in this group, 2 had pneumonia causing sepsis and 8 had sepsis of undefined aetiology (OUA). Looking at the metadata of this group, I can see that they had fewer mechanical ventilation interventions (3/10) than typical for sepsis and pneumonia (compare Table 7 A), were in the majority female (7/10), slightly younger than the average sepsis patient (average 55.3 years) and stayed on average slightly shorter than the typical pneumonia/ sepsis OUA patient (10.2 days, range 3-26.5). When looking at the loadings plot (Figure 34 E), this group of samples is at a 180° angle to the C12:0 loadings arrow. This indicates that this group is distinct from other samples based on their C12:0 measurements, and that their C12:0 measurements are lower than those of any samples moving in direction of the arrow. These samples are also further to the right than many samples, and with the C8:0 and C10:0 arrows pointing to the left, this would indicate that these samples are also further apart from other samples due to relatively lower concentrations of C8:0 and C10:0. I double-checked the C8:0, C10:0 and C12:0 concentrations for this group of

samples and found that all of these samples had indeed comparably low concentrations of all of these three lipids. I would tentatively suggest that this is a group of patients and 1 control with lipid profiles that might indicate survival, however the data is not very strong. Another possible explanation for the heterogeneity in the sepsis measurements could be the cause or type of sepsis, i.e., pneumonia, urosepsis or sepsis of undefined aetiology (OUA). In the PCA plot of sepsis causes, the urosepsis patients overlap strongly with the healthy controls (Figure 34 F). Pneumonia and sepsis OUA patients overlap partially with the healthy controls but are clearly more distant than urosepsis patients. In addition, pneumonia and sepsis OUA patients show great overlap with each other. It appears likely that part of the heterogeneity of lipid measurements is dictated by the sepsis cause. I will explore this further in one of the following paragraphs after another analysis of survival in relation to the lipid profile.

Results

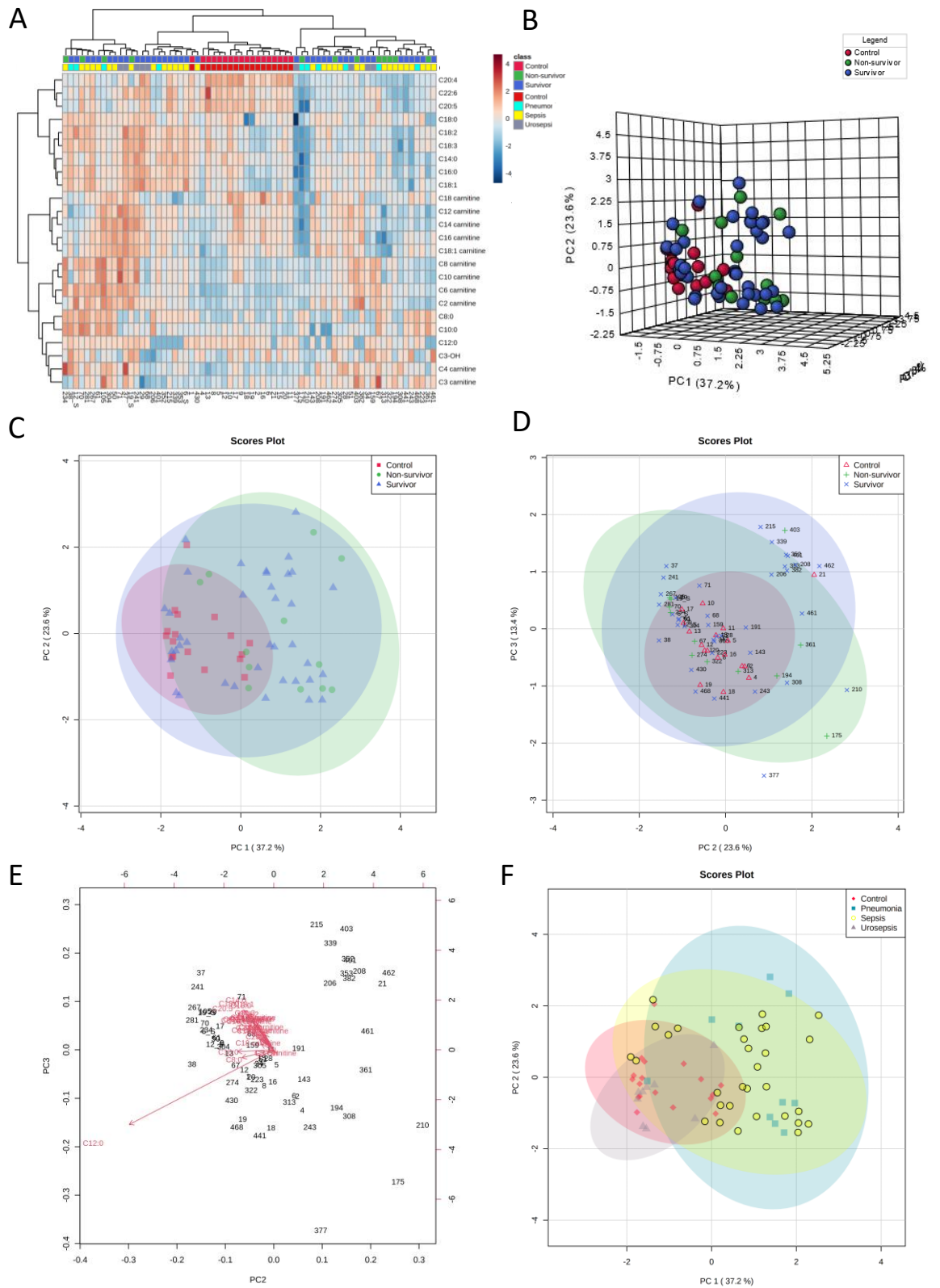


Figure 34 A) shows the heatmap of all samples and all lipids measured (log-normalized and centre-scaled). Both lipids and samples are clustered using D. ward clustering, distances are

Euclidean. The colour scale in the heatmap goes from 4 (increased) to -4 (decreased) and the colour indicates the z-score of the lipid measurement within the row for each sample. B) shows a 3D-plot of a PCA on the data with the colours indicating controls, survivors and non-survivors. C), D) and F) are also PCA plots of the exact same data but shown as biplots. In C) the datapoints are coloured based on their belonging into either the control, survivor or non-survivor group, with PC1 being the x-axis and PC2 the y-axis. D) shows the biplot of controls, survivors and non-survivors with PC2 as the x-axis and PC3 as the y-axis. F) shows the same PCA as C), but the datapoints are coloured based on their belongings into either the control, pneumonia, sepsis or urosepsis group. The coloured "halos" indicate the 95% confidence regions of the groups. E) also shows a biplot with PC2 on the x-axis and PC3 on the y-axis, and it shows both the spread of samples, indicated by their numbers, and the loadings of the principal components as indicated by the red arrows.

Results

4.1.8 Lipid concentrations and survival

Next, I conducted a partial least square – discriminant analysis (PLS-DA) in metaboanalyst. In contrast to PCA, this is a supervised method, i.e., the tool is aware of the group every sample belongs¹⁷⁹. PLS-DA tries to compute which variables best predict group membership using multivariate regression. I chose to do this analysis in order to find out which lipids were affected by survival. Many factors can influence the variance in lipid measurement, which could obscure the potential variance caused by survival in an unsupervised method, i.e., in the PCA I did. Since there is a risk of “overfitting” the data in supervised methods, classification performance will be evaluated using leave-one-out-cross-validation. In the 2D PLS-DA plot (Figure 35A), the survivors and non-survivors were slightly further apart compared to the PCA, and both groups overlapped less with the controls. Looking at which lipids drove the differentiation into groups, we see mostly lipids that were also present in the volcano plot with significant and relatively high fold changes, e.g., C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid), C22:6 (docosahexaenoic acid), carnitines bound to SCFA, MCFA and C18:0 (stearic acid), and the FFA C8:0 (octanoic acid) and C10:0 (decanoic acid). In contrast to previous analysis, C12:0 (dodecanoic acid) and C16:0 (palmitic acid) were also found in this analysis. The coloured bar at the right of the loadings plot (Figure 35B) shows the relative concentrations between the three groups. As indicated by the colours, most lipid levels followed a trajectory from low to high or vice versa from controls to survivors to non-survivors. In contrast to this, C12:0 (dodecanoic acid) and C16:0 (palmitic acid) did not have this clear trajectory. It is important to note that since the colour bar only depicts averages without considering the variance, the colours could be misleading. The method evaluation using cross-validation indicated that the model had a moderate ability to accurately predict the groups, and the results should hence be regarded with caution (Figure 35D). The PLS-DA seemed best at differentiating controls from survivors and non-survivors, and less good at differentiating survivors and non-survivors, since these groups still overlapped a lot. When performing PLS-DA on only survivor and non-survivor samples, the model was overfitted as evident from low and negative Q2 values in the cross-validation (Figure 35 D). In other words, we cannot reliably differentiate survivors and non-survivors in this dataset using PLS-DA. To statistically test if any lipids were altered between survivors and non-survivors, I conducted fold change analysis and non-parametric tests on these groups with and without multiple testing

adjustment in metaboanalyst. As can be seen in Figure 35C, 5 lipids had a fold change bigger than 1.5 between survivors and non-survivors (C8:0 (octanoic acid), C10:0 (decanoic acid), C3:0 carnitine (propionylcarnitine), C8:0 carnitine (octanoylcarnitine), C14:0 (myristic acid)). A further 5 lipids had a fold change above 1.2 (C18:1 carnitine (oleoylcarnitine), C16:0 carnitine (palmitoylcarnitine), C20:4 (arachidonic acid), C16:0 (palmitic acid), C18:1 (oleic acid)). Before multiple testing adjustment, the difference between survivors and non-survivors was significant for C3:0 carnitine (propionylcarnitine) and nearly significant for C10:0 (decanoic acid) in a Mann-Whitney-Wilcoxon-test, but none of these changes were significant after applying a false discovery rate of 0.1. Finally, I decided to see if a logistic regression model based on these changed lipids could potentially predict non-survival within this cohort. To this end, I used the metaboanalyst tool and uploaded only the data of non-survivors and survivors of sepsis. Using logistic regression as model, I generated ROC curves for several lipid combinations from the lipids shown in the table in Figure 36B. This led to models with moderate to good predictive ability within this cohort. The best performing model was based on C3:0-carnitine, C8:0, C16:0-carnitine, C20:4, C16:0 and C14:0. C10:0 was not included since it does not have an added effect when C8:0 is included. As can be seen in

Figure 36, the area under the curve (AUC) in “training” was ~0.9. Since I did not have any data for a validation cohort, and the ILTIS cohort was too small to split it, I used 10-fold cross validation instead. In cross-validation, the AUC was ~0.82. The sensitivity and the specificity had similar values in cross-validation: sensitivity 0.75, specificity 0.84. This means that the model was moderately good at predicting non-survival, with more non-survivors wrongly classed as survivors than survivors classed as non-survivors. I think that this is a good indication that plasma lipids could predict non-survival in sepsis and implicates metabolic changes in poorer outcome in sepsis. However, since I could not validate the data in a separate cohort, this is just a preliminary finding.

An increase in SCFA – and MCFA- acylcarnitines would have been in agreement with a study by Langley et al. (2013), that found an increase in SCFA- and MCFA-acylcarnitines to be associated with non-survival⁵¹. Based on the results I cannot draw any conclusion but cannot exclude that part of the heterogeneity is attributable to survival. It is apparent looking at the heatmap of all measurements (Figure 34A) that factors apart from survival must have contributed to the heterogeneity in measurements. A connection of lipid

Results

concentrations with survival could also indicate a connection with severity of disease or other factors that influence survival and warrants further analysis.

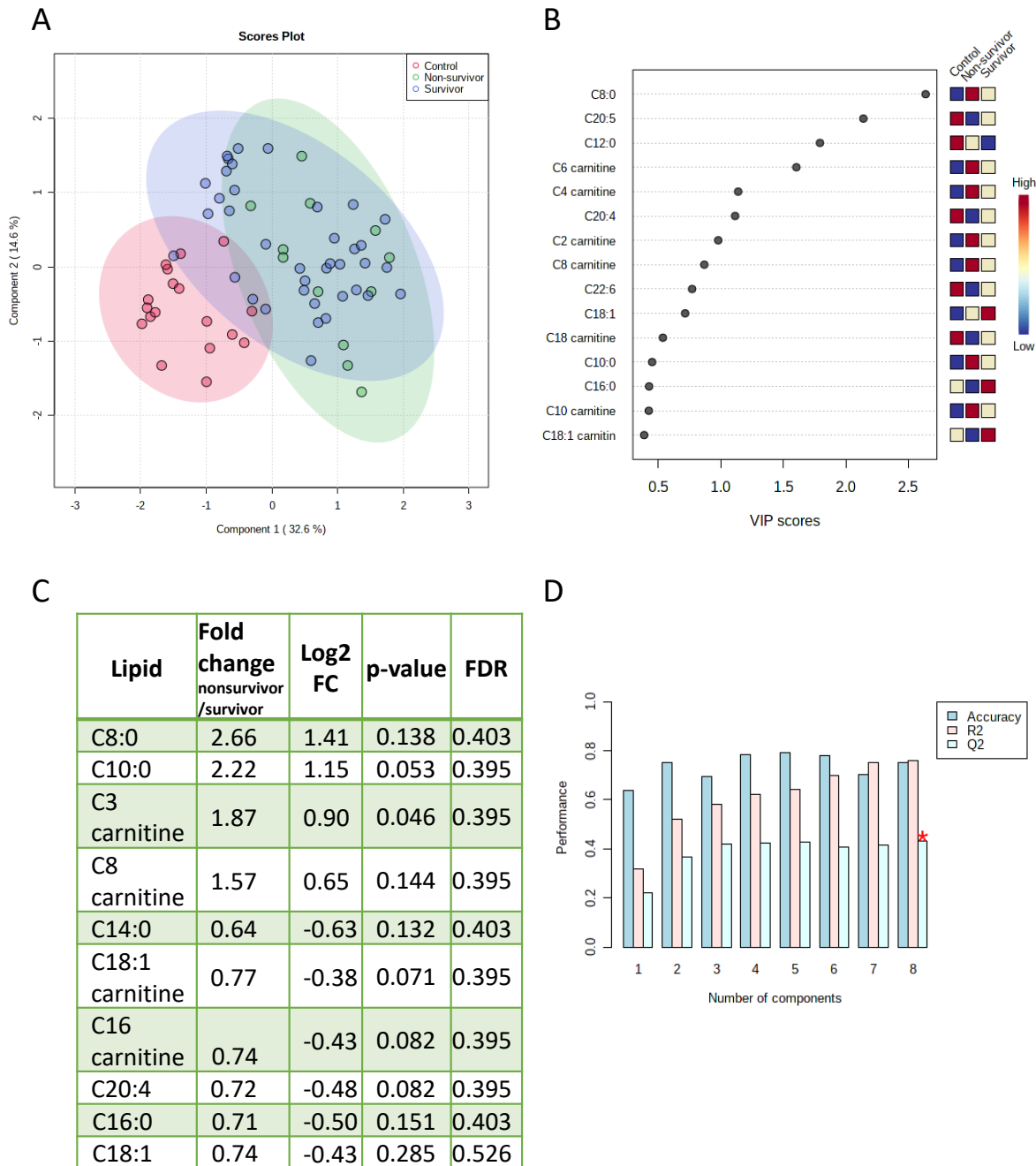


Figure 35 A) shows the 2D-plot of the PLS-DA performed on all data with the groups of controls, survivors and non-survivors. The PLS-DA reduces all data provided into principal components that best separate the groups by combining and weighing the transformed features, i.e., lipid measurements. B) shows the loadings plot of the PLS-DA, indicating the features, i.e., lipids, that showed the strongest effect on the classification of samples into their groups. The x-axis indicates the VIP score, VIP = Variable Importance in Projection, i.e., how much the variable contributes to the projection of the samples in the model. The panel next to it indicates the average concentration in the group in relation to the other groups,

Results

with blue indicating a relatively low average concentration and red indicating a relatively high average concentration. C) shows a table of the lipids with the highest fold-changes between survivors and non-survivors as calculated from average concentration in the non-survivor population divided by the average concentration in the survivor population. The *p*-values are obtained from Wilcoxon tests, and the right column shows the False Discovery Rate (FDR). D) shows the result from the cross-validation of the model which determines which number of components best projects the model. Accuracy, R2 and Q2 examine the model's performance. Higher values indicate better performance with 1 being the maximum.

A

	AUC	Sensitivity	Specificity
Training/Discovery	0.906 (0.874 ~ 0.937)	0.741 (0.658 ~ 0.823)	0.898 (0.866 ~ 0.930)
10-fold Cross-Validation	0.816 (0.672 ~ 0.959)	0.750 (0.750 ~ 0.995)	0.842 (0.726 ~ 0.958)

B

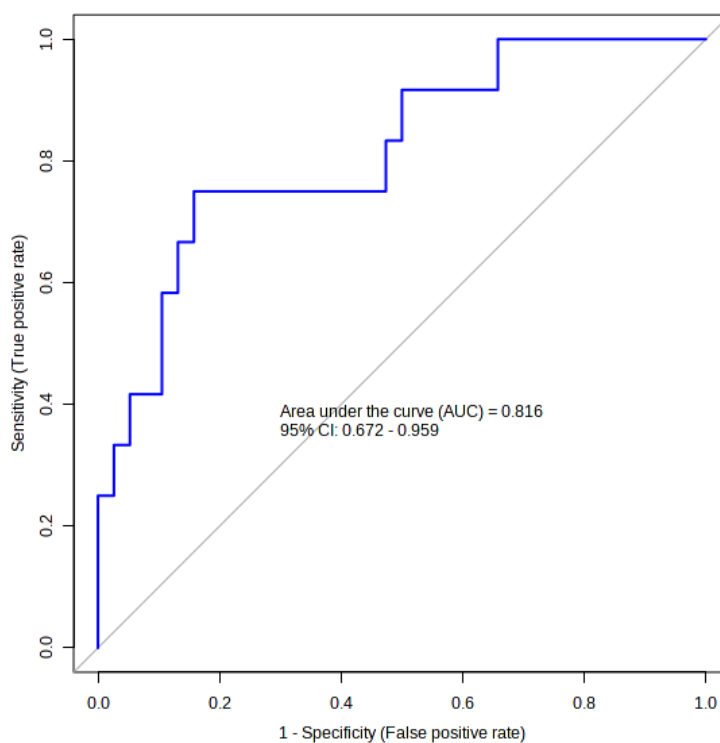


Figure 36 A) shows the table of AUC value, sensitivity and specificity of the model in training and 10-fold cross-validation. B) shows the ROC of the model.

4.1.9 Disease severity and sepsis cause

I previously found that urosepsis patients were more similar to healthy controls than to pneumonia or sepsis patients in their lipid profile. This raised the question how else they differed and specifically if they were maybe less ill than pneumonia or sepsis patients. To better understand this, I created a table containing several variables that indicate, albeit indirectly, the severity of disease. In addition, I created another table with the same variables for the survivors and non-survivors to establish which variables could be connected to non-survival and hence severity. Not all information was available for all patients, that is why I display the data from 49 patients instead of 52 in this table. Looking at Table 7 A, it is evident that all urosepsis patients survived while 30% of pneumonia and sepsis patients died. The difference between survival in urosepsis patients and not urosepsis patients was not statistically significant (Fisher's exact test, $p=0.09$). Furthermore, urosepsis patients stayed for shorter times in the hospital and received less often and for a shorter time mechanical ventilation than sepsis and pneumonia patients (Table 7A). The difference in both frequency and length of mechanical ventilation between urosepsis patients and other patients was statistically significant (Fisher's exact test, $p=0.00029$, Kruskal-Wallis-test with Dunn's correction, groupwise comparison: pneumonia vs urosepsis $p=0.0008$, sepsis OUA vs urosepsis $p=0.0348$). The length of the stay was on average shorter and also shows a smaller range than in pneumonia and sepsis, however the difference is only statistically significant when comparing urosepsis with pneumonia (Kruskal-Wallis-test with Dunn's correction, groupwise comparison: pneumonia vs sepsis OUA $p=0.1209$, pneumonia vs urosepsis $p=0.0024$, sepsis OUA vs urosepsis $p=0.1101$). Taken together this could indicate that sepsis severity in urosepsis patients was less, since this cohort required fewer interventions and on average stayed for shorter times in the hospital. Another difference is the higher age, both on average and looking at the age range, and the lower percentage of males in this cohort (33% compared to 60%). Both these parameters (age, sex) do not differ statistically significant between the cohorts (Kruskal-Wallis-test, Fisher's exact test).

The percentage of urosepsis patients requiring renal replacement therapy is very similar to sepsis and pneumonia patients (44% compared to 40% each) and not statistically significant different (Fisher's exact test). Due to the disease localisation, it seems likely that urosepsis would affect the kidneys, and it is hence remarkable that the percentage of patients

Results

requiring renal replacement therapy is not higher than in other sepsis types. It seems that the urosepsis cases in this cohort were typically less severe than the pneumonia and sepsis cases which may explain the stronger overlap with healthy controls rather than sepsis OUA and pneumonia cases. In order to check how much the used variables would indicate severity, I also made a table showing their values for survivors and non-survivors. Table 7 B shows that the percentage of patients receiving either intervention (mechanical ventilation or renal replacement therapy) is higher in the non-survivors than in the survivors, and that hospital stays and the duration of mechanical ventilation are slightly longer in non-survivors than in survivors. However, the difference in length of stay or ventilation is not statistically significant between survivors and non-survivors (Mann-Whitney-Test $p=0.2439$, $p=0.0937$). The frequency of renal replacement therapy is not statistically significant between survivors and non-survivors, but the frequency of mechanical ventilation is (Fisher's exact test $p=0.0471$). AP2 is short for APACHE II and describes the score calculated in hospitals to predict mortality in patients and should hence also indicate severity. It is calculated in the first 24 hours of a patient's hospital stay based on the worst values for a range of variables that indicate immune system activation and organ dysfunction or damage. In this cohort, the AP2 score seems to perform poorly at predicting mortality, since there is little difference between survivors and non-survivors in AP2 score. Furthermore, AP2 score is highest in the urosepsis cohort that showed the least (none) mortality.

Table 7 A and B show several indirect indicators of severity and the values associated with these for either A) different sepsis causes or B) survivors and non-survivors. AP2 stands for Apache II score, a scoring system for predicting mortality in the hospital.

A

Cause of Sepsis	Total	Age	Gender: Male (%)	Patients did not survive (%)	Patients received renal replacement therapy of respective group	Patients received mechanical ventilation (%)	Average AP2	Average number days mechanical ventilation (range)	Average total stay (range)
Pneumonia	10	59.2 (18-82)	6 (60)	3 (30)	4 (40)	8 (80)	16.2 (0-25)	20.3 (0-58)	27.6 (4-65)
Sepsis	30	59.8 (23-81)	18 (60)	9 (30)	12 (40)	19 (63.3)	17.4 (5-33)	8.5 (0-28)	12 (1-35)
Urosepsis	9	68.2 (42-86)	3 (33)	0 (0)	4 (44.4)	1 (11.1)	19.9 (8-27)	0.3 (0-3)	5.1 (2-11)

B

Survival	Total	Age (range)	Gender: Male (%)	Patients received renal replacement therapy (%)	Patients received mechanical ventilation (%)	Average AP2 (range)	Average number days mechanical ventilation (range)	Average total stay (range)
Non-survivor	12	67 (48-82)	9 (75)	7 (58.3)	10 (83.3)	19.5 (0-33)	14.3 (0-56)	16.3 (2.9-55.1)
Survivor	37	59.4 (18-86)	18 (48.6)	13 (35.1)	18 (48.6)	17 (5-30)	8.6 (0-58)	13.1 (1.1-64.8)

Results

4.1.10 Lipid concentrations in the plasma of patients with different sepsis causes

As it was seen on a global level that lipid concentrations differed between urosepsis and pneumonia and sepsis, I decided to examine this in greater detail by looking at the concentration levels of several lipids in the different sepsis causes. As can be seen in Figure 37, for several free fatty acids the levels were different in urosepsis from pneumonia and sepsis when comparing levels visually. For some, this was confirmed by the statistical tests performed: C8:0 (octanoic acid) concentrations were significantly different in urosepsis and both pneumonia and sepsis, for C12:0 (dodecanoic acid) urosepsis showed significantly decreased concentrations compared to sepsis and in C20:4 (arachidonic acid) and C20:5 (eicosapentaenoic acid) concentrations in pneumonia and sepsis were significantly lower than healthy controls, whereas there was no significant difference between urosepsis and healthy controls, albeit that a non-significant decrease is evident. The acylcarnitines shown on the other hand display no differences between urosepsis and pneumonia sepsis, neither visually nor statistically. The result for C12:0 (dodecanoic acid) caught my attention since the concentrations in urosepsis showed a different trend compared to healthy controls than pneumonia and sepsis, and concentration in urosepsis was significantly higher than in sepsis. These opposing changes could have negated statistical significance when comparing all sepsis cases to the healthy controls. I hence performed a Mann-Whitney-test comparing C12:0 (dodecanoic acid) levels in sepsis to healthy controls after excluding urosepsis and found a significant p-value (0.018).

Sepsis can be caused by different pathogens and hence I next sought to examine the effect of specific pathogens on lipid concentrations. This is shown for the lipid C10:0 (decanoic acid) as an example in Figure 37 I. There was no discernible effect of the pathogen type on the lipid concentrations, however the data lacked power since there were very few data points for some of the pathogen types. Furthermore, the clinical parameters that are routinely measured, e.g., white blood cell counts, bilirubin, etc., did not correlate with the lipid concentrations (data not shown). This would indicate that in this cohort, the lipids did not correlate with different types of organ dysfunction and immune activation markers.

Results

healthy controls (HC). The lines indicate the mean value of the group, and the y-axis depicts the concentration in the plasma sample in μM . The statistical test performed is a repeat Kruskal-Wallis-test comparing all groups using Dunn's multiple testing correction yielding an adjusted p-value with significance level $p=0.05$. Not significant test results are not shown. The significance levels displayed are: $*$ = $p \leq 0.05$, $**$ = $P \leq 0.01$, $***$ = $P \leq 0.001$, $****$ = $P \leq 0.0001$.

4.1.11 Correlation between lipids

In order to gain insight in the pathophysiology of lipid concentrations changes in the plasma of sepsis patients, it could be helpful to understand the relationships between the lipids. Lipids with a shared function and/or shared pathway are likely to be affected in a similar way. To explore this, I computed the Pearson correlation between all lipids and generated a heatmap using the metaboanalyst tool. As can be seen in Figure 38, C8:0 (octanoic acid) and C10:0 (decanoic acid) correlated with each other, but only showed weak correlations with other lipids, including straight chain lipids longer or equal to C12:0 (dodecanoic acid). The correlation between C8:0 (octanoic acid) and C10:0 (decanoic acid) was statistically significant and had a Pearson correlation of 0.71. C12:0 (dodecanoic acid) only showed weak correlation with other lipids. In contrast, all acylcarnitines, except C3:0-carnitine, correlated with acylcarnitines with similar chain length, and correlations were higher the closer in chain length the acylcarnitines were. For example, C10:0-carnitine correlated with C6:0-carnitine ($r=0.73$), C8:0-carnitine ($r=0.92$) and C12:0-carnitine ($r=0.71$). C3:0-carnitine did not show any strong correlations with other acylcarnitines and for example correlates with C2:0-carnitine with an $r = 0.44$, which is its strongest correlation. Acylcarnitines do not correlate with free fatty acids with correlation coefficients above 0.7, though some show weak ($0.4 < r < 0.5$) but statistically significant correlations with free fatty acids with a chain length longer than 14, including lipids with a chain length of 18 and a double bond.

Another group of lipids that correlates with each other is C14:0 (myristic acid), C16:0 (palmitic acid), C18:0 (stearic acid), C18:1 (oleic acid), C18:2 (linoleic acid) and C18:3 (α -linoleic acid). Among these, C18:0 shows significant but relatively weak ($r \sim 0.5$) correlations with C14:0, C18:1, C18:2, C18:3 and a stronger correlation with C16:0 ($r=0.77$). C14:0, C16:0, C18:1, C18:2, C18:3 show significant and strong correlations ($0.76 < r < 0.93$). The last group of lipids correlating with each other is C20:4, C20:5, and C22:6 ($0.77 < r < 0.86$). These lipids also show mostly significant but weaker correlations with the group containing C14:0, C16:0, C18:1, C18:2 and C18:3 ($0.28 < r < 0.65$).

These correlations were computed using Pearson's correlation; and since lipid concentrations are likely dictated by several factors which introduces heterogeneity, it is warranted to additionally use Spearman's correlation. Spearman's correlation is a ranking-based system and can detect non-linear correlations. Spearman's correlation was overall in line with Pearson's correlation findings, though correlation coefficients and significance

Results

found for several lipid pairings were higher in Spearman's correlation. This was especially the case for C20:4 (arachidonic acid), C22:6 (docosahexaenoic acid) and C20:5 (eicosapentaenoic acid) correlations with other LCFA and the correlation between LCFA-acylcarnitines and LCFA.

Overall, the lipids in sepsis can be tentatively divided into groups of similar behaviour: C8:0 and C10:0, short-chain to medium-chain acylcarnitines, long-chain acylcarnitines, and long-chain fatty acids that could be divided into two groups (C14:0 (myristic acid), C16:0 (palmitic acid), C18:0 (stearic acid), C18:1 (oleic acid), C18:2 (linoleic acid), C18:3 (α -linolenic acid); C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid), C22:6 (docosahexaenoic acid) or be seen as one. There might be a shared factor between LCFA and LCFA-acylcarnitines.

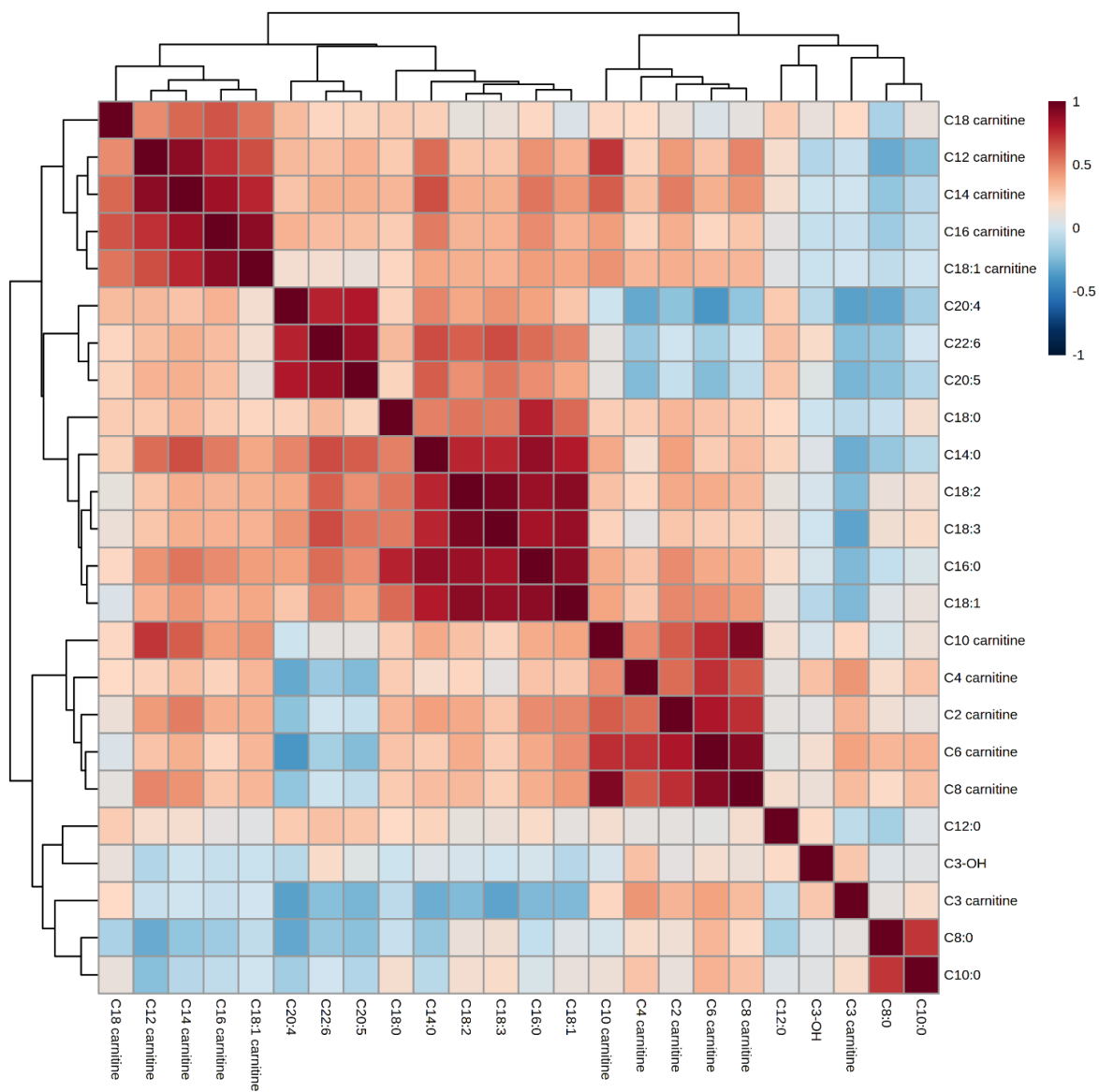


Figure 38 Correlation heatmap of all lipids measured in this analysis. The lipids are log-normalized and centre-scaled, the colour indicates the Pearson's r correlation value ranging from -1 to 1.

Results

4.1.12 Main findings

- C8:0 and C10:0 increase significantly in sepsis
- C12:0 increases in urosepsis compared to sepsis OAU, and is decreased in the sepsis cohort compared to healthy controls if urosepsis cases are excluded from the sepsis cohort
- Several acylcarnitines increase significantly in sepsis: C2:0, C4:0, C6:0, C8:0-carnitine
- C20:4, C20:5, C22:6 decreased
- Less severity and mortality in urosepsis cases in this sepsis cohort

4.2 Discussion

In this chapter I set out to measure a range of lipids in the plasma of a cohort of adults diagnosed with sepsis in the intensive care unit and a group of healthy volunteers as control group. This aimed to answer several questions: 1) if the method performs reliably, 2) if MCFA able to bind GPR84 are present and/or increased in sepsis, 3) how lipid metabolism changes in sepsis and how this relates to any potential differences observed in MCFA concentrations. In this discussion, I would like to address these questions in the order set out in the introduction. This means that I will first discuss how the established technique performed in measuring the lipids and then discuss my findings about lipid metabolism in sepsis. The established method performed at its best for longer fatty acids and carnitines and underperformed for shorter lipids. As a result, all lipids smaller than C8:0 were excluded since those measurements were not deemed reliable. This problem can be explained by a combination of factors. The first factor is the previously mentioned background levels of SCFA for C2:0, C3:0 and C4:0. Background levels of SCFA are commonly a problem and despite my best efforts I could not prevent the contamination of SCFA, as was evident from the detection of SCFA in the blank controls. Background levels were especially high for acetic acid. The second factor are the higher limit of detection for these fatty acids on the mass-spectrometer. The higher limit of detection means that the machine is less sensitive for this type of lipid than for other lipids, e.g., those with a longer carbon chain. Hence, a higher threshold concentration of lipid needs to be present in order to be accurately measured. This problem especially occurred for propionic acid where the limit of detection was just below the anticipated concentrations. The third factor is that the

extraction efficiency is higher for longer chain fatty acids than shorter chain lengths. For SCFA, lower extraction efficiency could contribute to an unfortunate ratio of background lipids to measured sample lipids since much of the background is likely introduced by derivatisation, as is discussed in Chapter 3. The stringent use of extracted blanks for every batch of samples allowed me to detect the high background and hence prevented the reporting of false positives. In this experiment the extracted blanks were aliquots of HPLC water which were extracted alongside samples. The measurement of detected lipids was deemed reliable if they could be detected in the majority of samples and with a coefficient of variability (%CV) within a reasonable range. This was the case for all lipids except hydroxylated MCFA (C10-2OH/3OH and C12-2OH/3OH) that were neither detected in most samples nor showed consistent measurements in the quality control. I would conclude for the evaluation of this method that it performed reliably well for most lipids in this cohort. Further, the most important questions to answer in this chapter were 1) if MCFA are increased in sepsis and 2) if so, if they are increased to a level that could potentially activate GPR84 present on immune cells. The relevant potential ligands are C10:0 and C12:0, which could both be detected reliably. C10:0 is significantly increased in sepsis cases, while C12:0 is not. In the MCFA plots in the results section, the EC50 obtained by Wang et al. is used, as it was the first study to determine the EC50 for MCFA binding to GPR84. However, further studies have performed binding assays for GPR84 and found partially similar and partially quite different values (Table 8). The EC50 for C10:0 in different ligand assays lies between 1.78 and 7.42 μM in 7 out of 10 tested assays, and between 20 and 48 μM in the remaining 3 assays. For C12:0, fewer assays were performed. The range of concentrations found as EC50 values for this lipid ranges from 2.17 and 10.5. The studies and assays in which relatively high EC50 values were found for C10:0 did not test C12:0 or were not able to detect binding by C12:0. It seems likely that the differences in observed EC50 values come from differences in the sensitivity of the tests performed and other experimental factors. Since the EC50 for C10:0 was found in the lower range in the majority of assays, I will assume that range to more likely reflect the actual EC50 of GPR84 and stick to the values measured by Wang et al.(2006)⁹⁶ For C12:0, the range is relatively narrow and to use the value obtained by Wang et al. as guideline is relatively conservative. C10:0 is present in concentrations exceeding the EC50 values described by Wang et al. (2006) in a subset of patients, but not in the healthy controls. C12:0 on the other hand displays

Results

concentrations exceeding said EC50 values in a subset of both patients and controls. When cross-examining the different sepsis causes and C12:0 concentrations, C12:0 is decreased in pneumonia and sepsis OUA and increased in urosepsis, with concentrations in urosepsis being significantly higher than in sepsis OUA. However, due to the spread of concentrations among each group of patients and volunteers, it seems feasible that C12:0 has the potential to activate the receptor in those patients showing levels that exceed the EC50. In healthy controls, the receptor is not expected to be present on the immune cells and the lipids should hence not have any effect mediated via GPR84. An interesting aspect to consider in this regard is also the fact that the concentrations measured are systemic concentrations, while concentrations could differ in different parts of the body. This would depend on the source of these fatty acids, which will be further explored in a later chapter.

Table 8 Overview of the EC50 (and Ki) values obtained for C10 (decanoic acid) and C12:0 (dodecanoic acid) in different ligand binding assays for GPR84. Adapted from Chen et al. (2020)¹³¹

C10:0(μM)	C12:0(μM)	Type of assay	Cell type used in assay	Study
4.6	10.5	[³⁵ S] GTPγS assay EC50	CHO/ human GPR84 cells	Wang et al. (2006) ⁹⁶
25		[³⁵ S] GTPγS assay EC50	Flp-In TREx 293/FLAG-human GPR84-eYFP cells.	Mahmud et al. (2017) ⁷⁷
6.08	3.49	β-arrestin binding	CHO/β-arrestin2/human GPR84 cells	Koese et al. (2020) ¹⁸⁰
30	nd	calcium assay EC50	iCHO/Gqi5/human GPR84 cells	Southern et al. (2013) ¹⁸¹
48		calcium assay EC50	CHO/aequorin reporter/Gα16,Gqs5,Gqo5, and Gqi9/human GPR84 cells	Wang et al. (2006) ⁹⁶
4.5	8.8	cAMP assay EC50	CHO/ human GPR84 cells	Wang et al. (2006) ⁹⁶
	8.87	cAMP assay EC50	CHO/ human GPR84 cells	Koese et al. (2020) ¹⁸⁰
7.42		cAMP assay EC50	CHO/β-arrestin2/human GPR84 cells	Pillaiyar et al. (2017) ¹⁰³
20		cAMP assay EC50	Flp-In TREx 293/FLAG-human GPR84-eYFP cells.	Mahmud et al. (2017) ⁷⁷
48		cAMP assay EC50	Sf9 insect/Gai/human GPR84 cells. hCHO/aequorin	Nikaido et al. (2015) ¹⁸²
1.78	2.17	radioligand binding assay Ki	CHO/ human GPR84 cells	Koese et al. (2020) ¹⁸⁰

Results

In this cohort, several lipids are significantly different in sepsis cases than in controls, but most lipids are not changed significantly. The most significantly decreased lipids in sepsis are C20:4 (arachidonic acid) and C20:5 (eicosapentaenoic acid). These are the precursors for immune mediators⁴². The decrease of these precursors and an assumed increase in immune-lipid mediators is in line with the dynamics expected in pro-inflammatory immune activation and is consistent with the sepsis patients exhibiting an inflammatory reaction. Previous studies have found increases in the lipid mediators originating from these precursors, e.g., eicosanoids and resolvins¹⁸³, while the decrease in C20:4, C20:5 and C22:6 in the plasma of sepsis patients has also been documented¹⁸⁴. A study focusing on arachidonic acid (C20:4), on the other hand, finds a slight increase in arachidonic acid in whole blood samples of sepsis patients compared to healthy volunteers. Their measurements of gene expression of relevant genes, arachidonic acid metabolites and arachidonic acid analogues however revealed several significant changes, indicating the altered arachidonic acid metabolism in sepsis¹⁶⁴. It is interesting to note that previous studies indicate that the supplementation of sepsis patients with oil preparations containing C20:5 and C22:6 would have a beneficial effect on survival, as indicated by a recent meta-study¹⁸⁵ and studies in murine sepsis models¹⁸⁶. Comparing the C20:4 and C20:5 levels for different types of sepsis, it seems that both lipids are less decreased in most urosepsis patients, as their concentrations are not significantly lower than in healthy controls, even though they are also not significantly higher than in pneumonia and sepsis OUA patients. This could mean that most urosepsis patients have either a more localized inflammation, hence affecting the systemic lipid precursor levels to a lesser extent, or a lower inflammatory response compared to other sepsis patients.

Acylcarnitines as a lipid class show the most changes in sepsis. All acylcarnitines binding a C2:0 to C8:0 fatty acid, with exemption of C3:0-carnitine, are increased significantly and with a fold change above 2. C3:0-carnitine is increased with near-significance ($p=0.063$). Acylcarnitines play an important role in cellular metabolism since the binding of fatty acids to carnitine is required for the transfer into the mitochondria where fatty acids can be oxidized for energy generation^{44,45}. Inside the mitochondria acylcarnitines are converted to acetyl-coA-esters and then available for lipid catabolism or anabolism⁴⁴. An increase in acylcarnitines outside of cells is thought to originate from mitochondrial shutdown or overloading^{46,48,51}. C3:0-carnitine is thought to originate from branched chain amino acid

catabolism, which also occurs in mitochondria¹⁸⁷. The different origins might explain why C3:0 carnitine is the only SCFA-acylcarnitine not significantly increased in sepsis and not correlating with any other acylcarnitines measured. Concentrations of acylcarnitines in the plasma of healthy controls are very low, in line with previous publications. Furthermore, C18:0-carnitine is decreased in the plasma of sepsis patients, though with a fold change less than 2-fold. The occurrence of increases in SCFA- to MCFA-acylcarnitines in sepsis have been cited previously and were found to be associated with mortality in sepsis^{51,50,188}. Previous publications finding increases in SCFA- and MCFA-acylcarnitines explain these by perturbation of cellular lipid metabolism which is then assumed to be detrimental for survival.

Finally, the MCFAs C8:0 (octanoic acid) and C10:0 (decanoic acid) are significantly increased in sepsis with a FC above 2. This finding has not been previously reported in the literature and as far as I know has not been previously researched. The literature typically cites food as source for MCFA, however this seems unfeasible in this setting since it requires a radical dietary change in a subset of patients. Albeit possible, i.e., by enteral or parenteral feeding; it seems unlikely since samples were taken within 36hrs after diagnosis of sepsis. In addition, it is unlikely that C8:0/C10:0 levels but not C12:0 levels would be affected by enteral or parental feed. Furthermore, enteral feed typically contains unsaturated LCFA C18:1 (oleic acid), C18:2 (linoleic acid), C18:3 (α -linolenic acid) as their main lipids, and no correlation between MCFA and these lipids was found. This potentially points to a biological mechanism occurring in sepsis, rather than a dietary effect, leading to the increase in these lipids in a subset of patients.

Many of the lipids display great heterogeneity in measurements within the patient group. Generally, sepsis is a disorder with great heterogeneity, e.g., disease severity, disease cause, affected organs and site of infection can differ greatly³. I examined which of these factors might contribute to the heterogeneity in lipid measurements, especially with a focus on C8:0 (octanoic acid) and C10:0 (decanoic acid). Sepsis cause was found to have a significant effect on lipid concentrations for some of the free fatty acids (e.g., C8:0, C12:0), but not for acylcarnitines. For several of the lipids where differences between sepsis causes were not significant, a trend seems to be visible. In contrast, no significant correlations with clinical measurements indicating organ dysfunctions were found; also none with AP2 scores, and there were no significant differences in lipids for causative microorganism and

Results

mortality. In case of mortality a trend can be observed for some lipids including propionylcarnitine (C3:0-carnitine) and a significant difference has been reported by other groups in case of acylcarnitines⁵¹ including C3:0-carnitine. This study likely lacked the power to find significant differences in lipids between survivors and non-survivors. Based on the results obtained I would expect that in a bigger study the lipids C3:0-carnitine (propionylcarnitine), C8:0 (octanoic acid), C10:0 (decanoic acid), C8:0-carnitine (octanoylcarnitine) and C14:0 (myristic acid) would be significantly different between survivors and non-survivors. This is also implicated by the logistic regression model that yielded acceptable sensitivity and specificity in differentiation survivors from non-survivors within the ILTIS cohort. The data on mortality in the ILTIS cohort needs to be regarded as preliminary, since the logistic regression model was only preliminary and there was a lack of power. It does however indicate that this could be an interesting subject for further studies.

The differences and similarities in lipids between the different sepsis causes indicates that there are shared and divergent changes occurring in these patients. A potential hypothesis I would like to propose is that the localization of infection plays a role. A meta-analysis into site of infection and sepsis outcome indicated that the site of infection had a strong effect on survival. In this meta-study, urogenital infections lead to better survival rates than for example respiratory tract and intra-abdominal infections¹⁸⁹. Additionally, it is described in the sepsis-3 paper as a sepsis key concept that “specific infections can cause a local organ dysfunction without eliciting a dysregulated systemic host response”³. Simultaneously, a SOFA score sufficient to identify a patient as having sepsis according to the Sepsis-3 definition can be reached with the dysfunction of one organ alone. If this was the case for these urosepsis patients, it would explain why we observe fewer changes to systemic lipid levels in these patients than in other sepsis patients, with the kidney/urethra being a focus of infection. In other words, it is possible for a patient to be classed as having sepsis according to the sepsis-3 criteria, without the elicitation of a dysregulated systemic host response, if the infection is localized and causes local organ dysfunction. This could hypothetically be the case in urosepsis patients in this cohort. However, without additional information on the difference between urosepsis and pneumonia and sepsis on a patient-level, this cannot be further explored in this study. Since the percentage of urosepsis patients in this study is comparable to other sepsis studies¹⁸⁹, understanding the

differences between urosepsis and sepsis from other causes might be possible from existing data and relevant to other studies as well.

Looking at the correlation between the lipids measured, several groups of lipids that correlate with each other become evident. It is likely that correlating lipids share a common pathway or origin. In fact, several of these tentative groupings are known to share a common pathway, e.g. the eicosanoid pathway in case of C20:4, C20:5, C22:6. It is interesting that C20:4 and C20:5 correlate positively despite being competitive substrates for the enzyme COX⁴². C14:0 (myristic acid), C16:0 (palmitic acid), C18:1 (oleic acid), C18:2 (linoleic acid), C18:3 (α -linolenic acid) also correlate with each other. C14:0, C16:0, C18:0 and C18:1 could be synthesized by the cells itself, while C18:2 and C18:3 are known to have a dietary origin⁴³. These lipids show no significant differences between sepsis and controls but exhibit heterogeneity within each group. One exception is C18:1, which is significantly increased, while C16:0 and C18:2 show a non-significant increase in sepsis. The pathway or origin of these lipids is likely only partially affected in sepsis or at least at the early timepoint at which samples were taken. Changes in levels of C8:0 (octanoic acid) and C10:0 (decanoic acid) are highly correlated, which could indicate that their origin is distinct from the pathways affecting LCFA and PUFAs. Since C8:0 and C10:0 concentrations are only significantly altered in a subset of sepsis patients, it is feasible that their pathway is only affected in a subset of sepsis patients.

In conclusion, this small cohort study yielded interesting insights into lipid metabolism in sepsis and hence requires to be confirmed in a separate cohort of sepsis patients and controls. Preferably, a bigger cohort size should be obtained to increase the power of the study. Furthermore, using a different cohort, i.e., neonatal or paediatric patients, would provide insights if differences are age-dependant. Apart from this, some interesting findings and questions have been generated in this study. Notably, the increase in MCFAs in sepsis and the differences between urosepsis and pneumonia and sepsis. It would be beneficial to look at other types of data from sepsis patients, e.g., transcriptomic or proteomic data, to generate hypotheses for underlying mechanisms, or perform *in vitro* or *in vivo* experiments to explore any potential mechanisms.

5 Lipid measurements in neonatal suspected sepsis cases and neonatal controls

In this chapter I present the measurement of lipids with the previously established method in a cohort of suspected sepsis cases in neonates and neonatal controls. Neonates are among the most vulnerable populations to sepsis as described by the WHO and have a relatively high mortality rate compared to other populations¹⁰. The developing neonatal immune system is thought to react differently to infection compared to immune responses in older patients^{31,53}. A theory proposed by Harbeson et al. stated that the different and higher energy demand in early life affects the immune response⁵³. That is why it is especially relevant to examine lipid changes in sepsis in this cohort. Furthermore, the diagnosis of sepsis in neonates is especially challenging, as they may show fewer or different clinical signs to other age groups when developing sepsis. In this cohort, blood samples were taken as timely as possible after suspicion of sepsis and 24 hours after the initial sample. All samples presented in this chapter were sampled 24 hours after the initial sample. The study design aimed to classify all suspected cases retrospectively into confirmed sepsis, clinical sepsis and not sepsis; however, this classification had not been completed by the time I analysed my data¹⁹⁰. That is why I will present the data labelled as suspected sepsis and controls. The reason to present the unclassified data is that it allows for an insight into the ranges of lipids present in the blood and, using clustering and PCA, can show us if there are groups within this population distinguished by their lipid profiles. In contrast to the ILTIS study presented in chapter 4, all samples in this cohort were whole blood samples. This means that any lipids measured may originate from both the intra- and the extracellular milieu in the blood.

5.1 Results

5.1.1 Cohort

In total, 88 suspected sepsis cases and 22 controls were sampled. Controls were sampled from neonates who had their blood taken for routine testing, i.e., to measure blood sugar levels or test for jaundice. All neonates were recruited in the neonatal intensive care unit (NICU). Neonates were included as suspected sepsis cases when their blood was taken for sepsis screening (e.g., blood cultures, CPR testing) and before they were started on

antibiotics¹⁹⁰. Clinicians on the study estimated that about 10-20% of suspected sepsis cases would develop sepsis. This would hence mean that between 8 and 18 neonates in this cohort had sepsis.

5.1.2 Quality control

The samples were processed exactly the same as the ILTIS samples were processed. Extraction and derivatisation were performed in two separate batches, the background of each batch was measured by extraction and derivatisation of 5 blank samples. The background was subtracted from all obtained measurements. In addition, per batch 1 quality control sample was prepared. As previously, this sample was measured by LC-MS/MS after every 10 samples submitted. In Table 9, the quality control measurements and the associated %CV is visible. C4:0 (butyric acid) and C4:0-carnitine (butyrylcarnitine) were excluded from further consideration since the %CV exceeded 15% and the lipids could not be detected in every quality control sample. C8:0 (octanoic acid) and C3:0-OH (lactic acid) were excluded due to technical difficulties leading to insufficient reliable measurements for these lipids. All other lipids were detected reliably.

Results

Table 9 The coefficient of variation in quality control 1 and 2. QC1 was measured 7 times and QC2 was measured 6 times. C8:0 and C3:0-OH were not measured due to technical difficulties.

	QC1, %CV	QC2, %CV
C2 carnitine	1.90	1.83
C3 carnitine	2.25	6.88
C4 carnitine	4.41	141.81
C6 carnitine	5.77	2.10
C8 carnitine	3.73	3.82
C10 carnitine	3.91	5.53
C12 carnitine	2.55	1.87
C14 carnitine	0.59	2.09
C16 carnitine	1.27	1.31
C18 carnitine	0.91	2.14
C18:1 carnitine	2.23	4.60
C2:0	11.90	4.32
C3:0	9.67	3.95
C4:0	70.08	36.61
C6:0	6.26	3.80
C10:0	3.89	9.49
C12:0	6.45	10.80
C14:0	1.92	2.25
C16:0	0.98	5.78
C18:0	1.11	2.47
C18:1	3.00	2.82
C18:2	3.68	5.53
C18:3	6.11	7.18
C20:4	1.60	2.66
C20:5	1.69	8.00
C22:6	1.49	1.47

5.1.3 Measurements of free fatty acids

In order to see the spread of values in suspected cases and controls, boxplots of all lipids were prepared. Firstly, the observed concentrations were within the expected micromolar range. The measurements show great heterogeneity in both groups and bigger heterogeneity in the suspected cases cohort. Significant differences between suspected sepsis cases and controls were detected for C6:0 (hexanoic acid, Mann-Whitney-Wilcoxon-test $p=0.0083$), C10:0 (decanoic acid, Wilcoxon-test $p=6.5^{-5}$), C12:0 (dodecanoic acid, Wilcoxon-test $p=2^{-9}$), C18:2 (linoleic acid, Mann-Whitney-Wilcoxon-test $p=1.7^{-7}$), C18:3 (linolenic acid, Mann-Whitney-Wilcoxon-test $p=5^{-7}$). C14:0 (myristic acid) and C16:0 (palmitic acid) had near-significant changes (Mann-Whitney-Wilcoxon-test $p=0.082$, $p=0.1$ respectively). C6:0 and C16:0 concentrations were increased, whereas C10:0, C12:0, C14:0, C18:2 and C18:3 concentrations were decreased.

Results

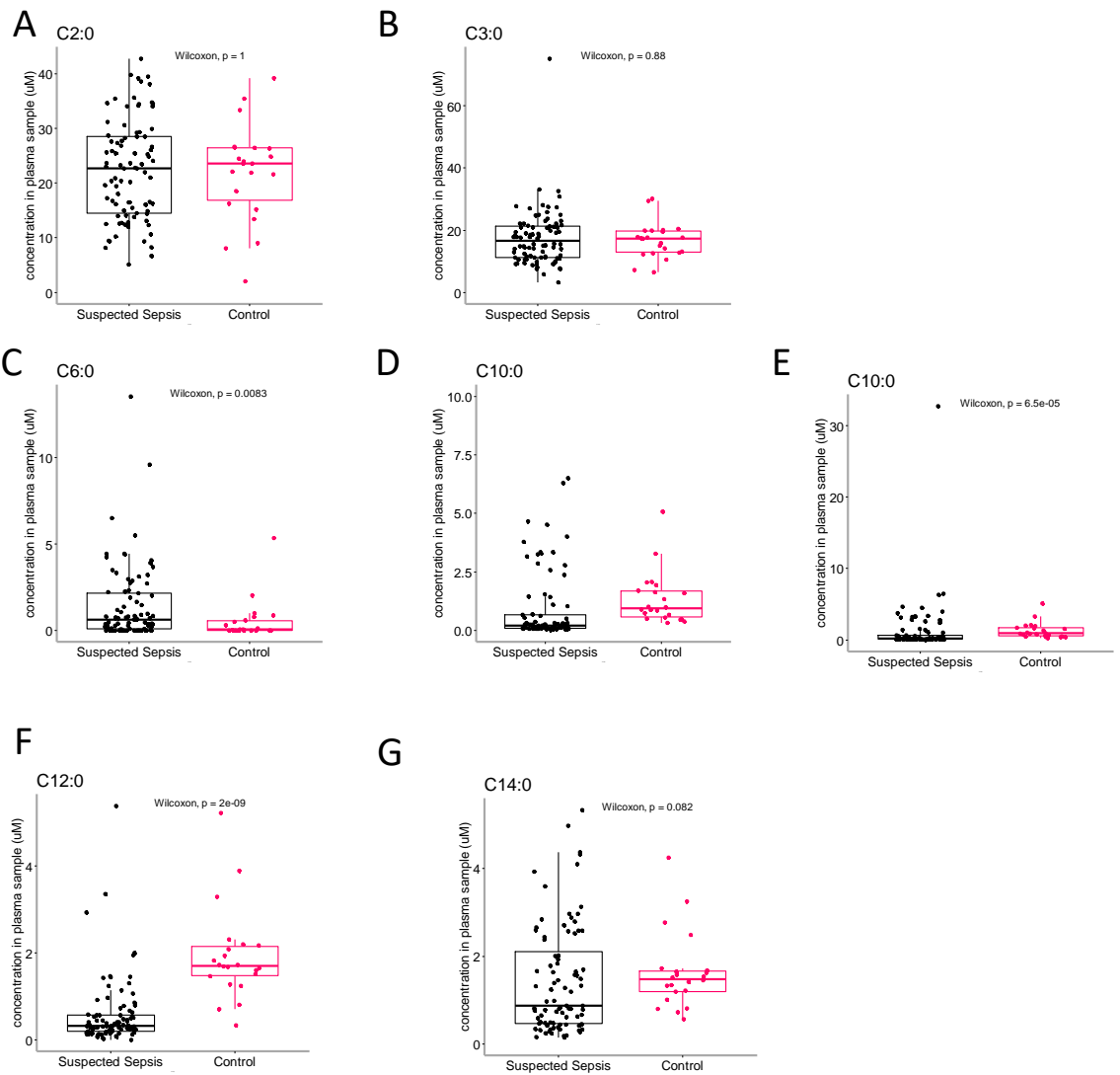


Figure 39 Boxplots of short chain- and medium chain fatty acids measured in whole blood samples of a cohort of suspected sepsis cases and a control cohort. For C10:0 two boxplots are shown (D and E) in order to be able to show the full range (E) and the lower range (D) containing the majority of measurements. The statistical test performed was a Wilcoxon test with a significance level of 0.05. The lower half of each box in the boxplot depicts the 2nd quartile of values, with the upper half depicting the third quartile. The extended lines depict the first and last quartile of values, with any points reaching below or above these lines being statistical outliers.

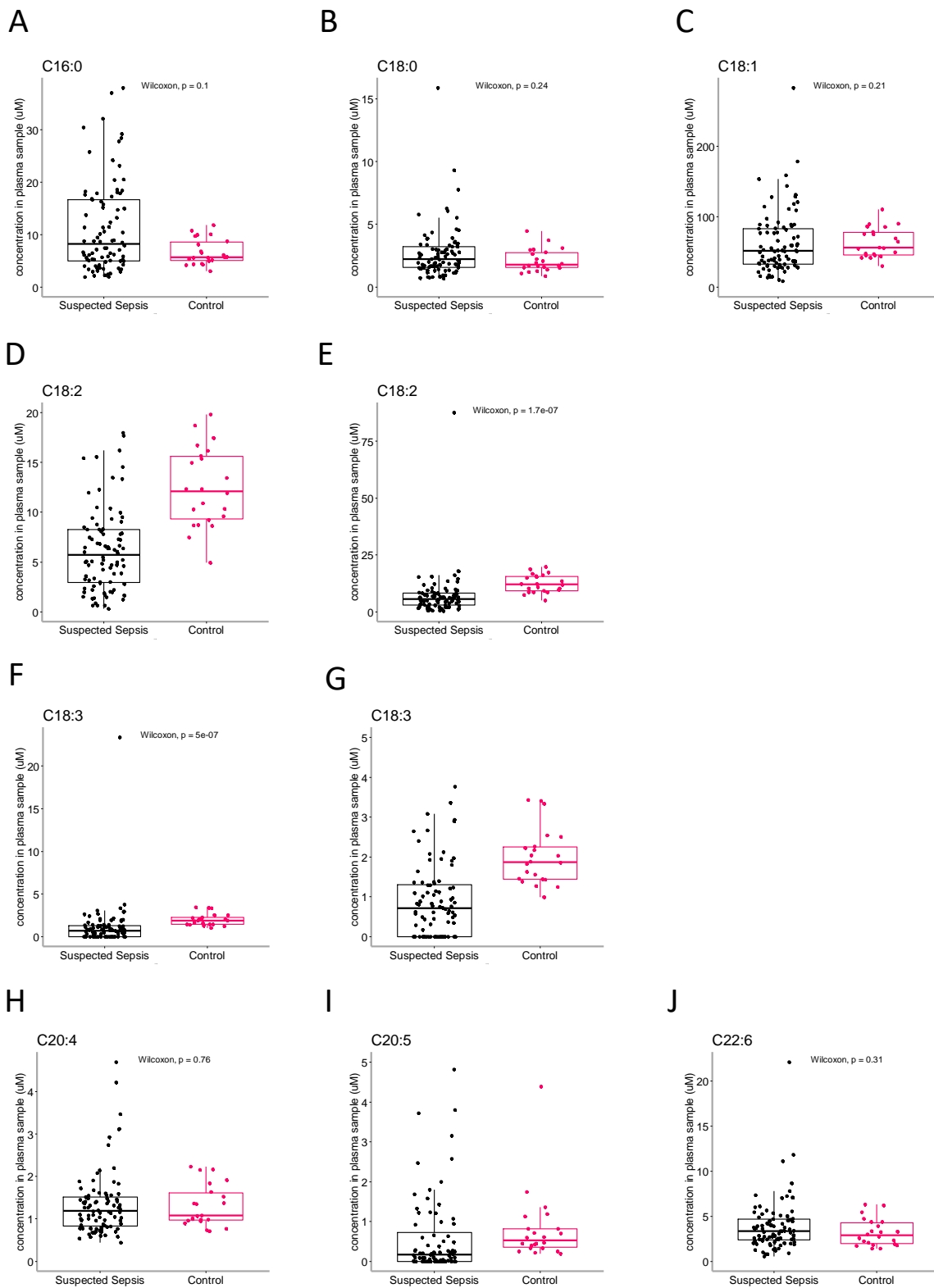


Figure 40 Boxplots of long chain fatty acids measured in whole blood samples of a cohort of suspected sepsis cases and a control cohort. For C18:2 (D, E) and C18:3 (F,G) two boxplots are shown in order to show the full range and the lower range containing the majority of

Results

values. The statistical test performed was a Wilcoxon test with a significance level of 0.05. The lower half of each box in the boxplot depicts the 2nd quartile of values, with the upper half depicting the third quartile. The extended lines depict the first and last quartile of values, with any points reaching below or above these lines being statistical outliers.

5.1.4 Measurements of acylcarnitines

Boxplots of acylcarnitines were prepared for the same reason as for the previously shown lipids: to evaluate the spread of values. Firstly, the concentrations are in the expected range of between submolar to the low micromolar range, with only C2:0-carnitine showing concentrations in a higher micromolar range (~2-25 μM). C8:0-carnitine, C10:0-carnitine, C16:0-carnitine, C18:0-carnitine were significantly changed in suspected sepsis compared to controls (Wilcoxon-test, $p=0.0084$, $p=0.00092$, $p=1,3^{-6}$, $p=2.8^{-7}$, respectively). C8:0- and C10:0-carnitine were decreased and C16:0- and C18:0-carnitine were increased.

Results

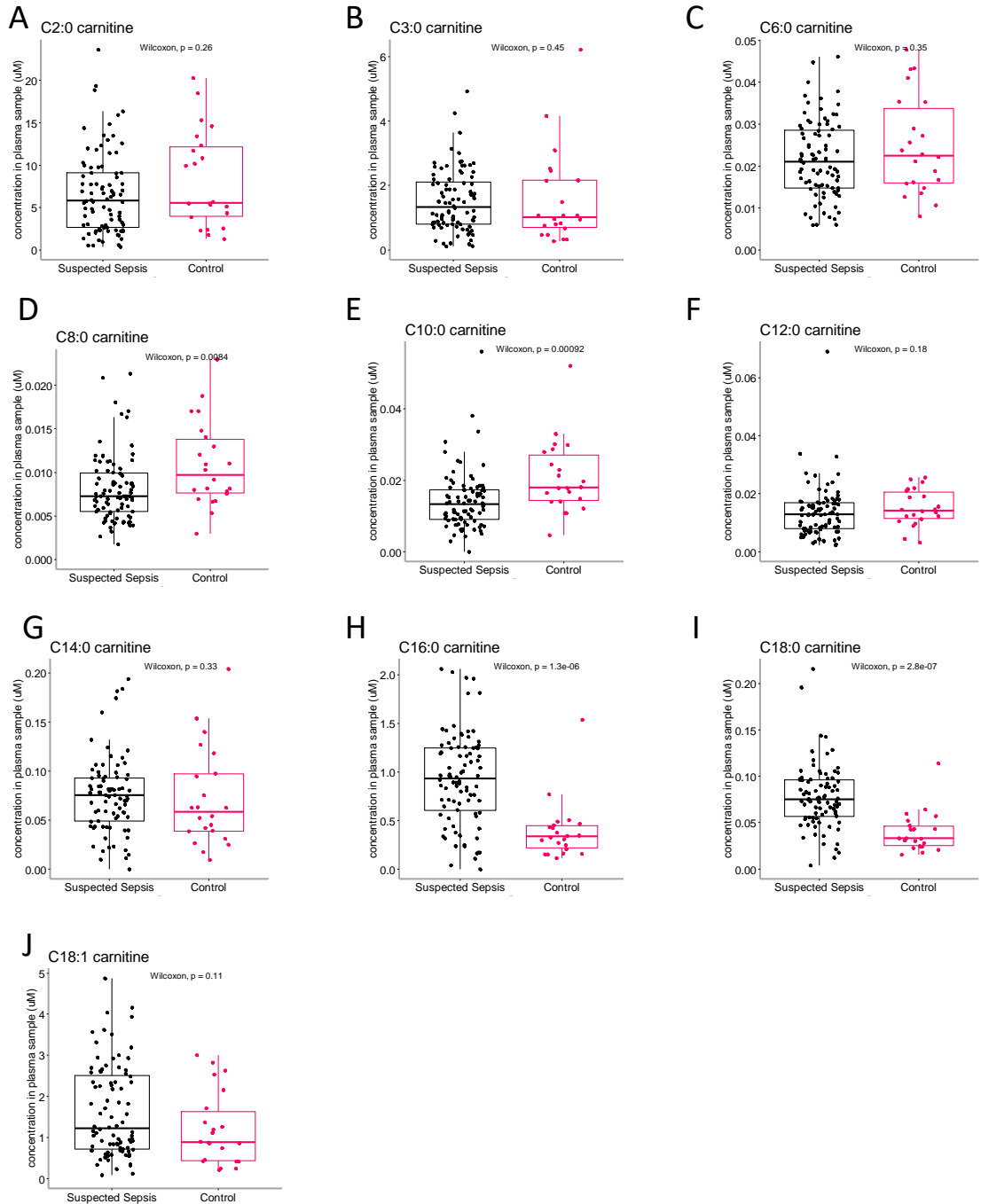


Figure 41 Boxplots of acylcarnitines measured in whole blood samples of a cohort of suspected sepsis cases and a control cohort. The statistical test performed was a Wilcoxon test with a significance level of 0.05. The lower half of each box in the boxplot depicts the 2nd quartile of values, with the upper half depicting the third quartile. The extended lines depict the first and last quartile of values, with any points reaching below or above these lines being statistical outliers.

5.1.5 Lipid profile changes in suspected sepsis cases

Next, I looked at global changes in lipids. Firstly, I generated a Volcano plot using metaboanalyst to depict the lipid changes within one plot (Figure 42). In contrast to the usual volcano plot, also the lipids that do not have a fold change above the desired threshold are visible. This shows that most lipids found to be significant in the previous analysis are significant according to this analysis as well, but that the fold changes for most lipids are relatively low and for several quite close to no change at all, i.e., a $\log_2FC=0$. Next, I generated a heatmap of the lipids in all samples. This heatmap is clustered. The objective of this heatmap was to show any groupings of samples in connection with lipids. Ideally, a distinguishable cluster of suspected sepsis cases would become apparent, with the possibility that these are sepsis cases. The presence of such a cluster would mean that there is the possibility to distinguish sepsis from not sepsis based on the lipid profile in whole blood. As can be seen in Figure 43 A, the samples are mainly clustered into three to four clusters with several subclusters. One of those clusters contains most of the control samples (most right cluster). The cluster most left in the heatmap is clustered differently from the other three clusters. Comparing the colours of the heatmap for the samples in this cluster to those of all other clusters, the cooler colours for the MCFA and LCFA clusters seem the most consistent difference. So possibly there is a subgroup of samples defined by lower MCFA and LCFA concentrations in this cohort. Exploring this further using principal component analysis, one subgroup further away from most samples is visible in the 3D PCA plot (Figure 43B). To understand better which principal components are driving this separation, I also generated 2D PCA plots (Figure 43 D,E,F). As can be seen from the different PCA plots, there is little separation of the samples into subgroups along PC3. Taken together PC2 and PC1 lead to separation of subgroups, and the subgroup visible in the 3D plot was also visible in the 2D plot of PC2 and PC1. In this plot it seems to be more apparent that there were two subgroups rather than one further apart from a main group of samples. Looking at the 2D PCA plot that contains the loadings, it is visible that C6:0, C20:5, C10:0 and C18:3 had the biggest effect on the spread of samples within the plot, as these are the longest arrows (Figure 43 C). From the table of loadings, it is apparent that the main loading of PC2 was C6:0 (-0.88), followed by C20:5 (-0.38), C2:0-carnitine (0.14) and C10:0 (-0.14). For PC1, the main loadings were C20:5(-0.68), C18:3 (-0.41), C6:0(0.41), C10:0 (-0.35) and C18:2 (-0.2). As the subgroups were separated along these two principal

Results

components, it seems indicated that this is a subgroup of patients (and controls) with a different lipid profile regarding these lipids specifically.

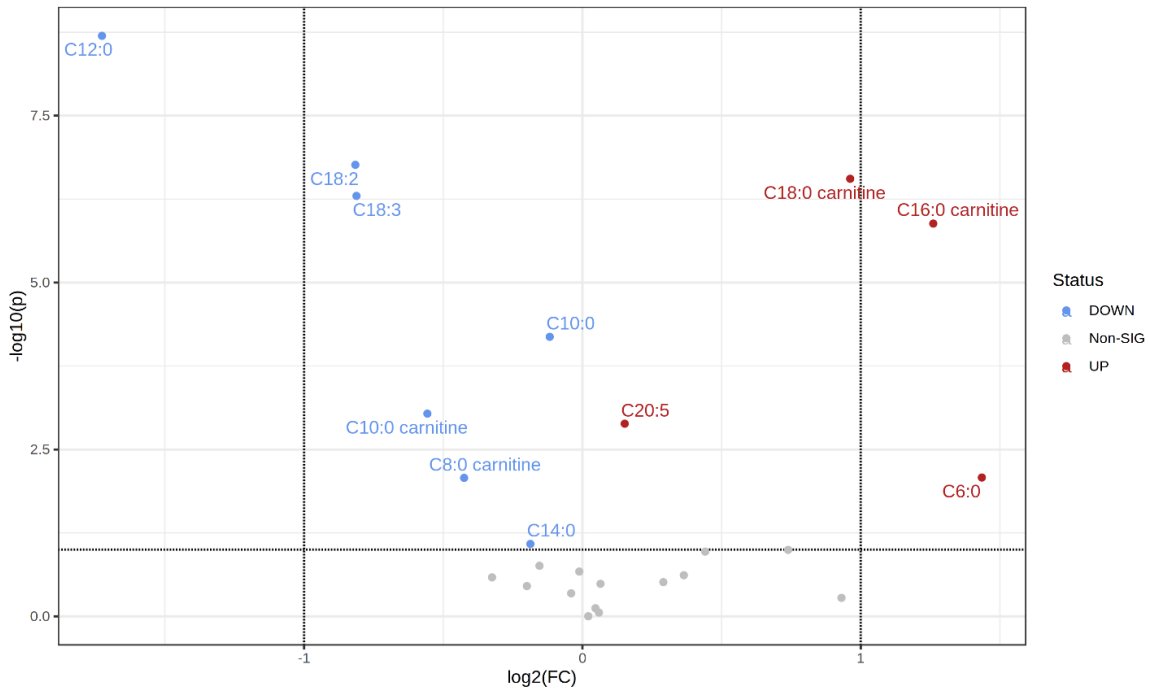


Figure 42 Volcano plot of lipids altered in sepsis. All labelled lipids are statistically significantly altered. The y-axis depicts the negative log10 of the p-value, hence the higher the value the higher the significance. The x-axis depicts the log2 of the fold change calculated based on not-normalized lipid concentrations. The dashed lines at log2(FC) 1 indicate a 2-fold change on raw data. Significance was calculated using the Kruskal-Wallis test on log10-transformed lipid concentrations and is adjusted for multiple testing. Dashed lines indicate a log2(FC) of 1 and a $-\log_{10} p$ -value that corresponds to $p=0.1$ after multiple testing correction.

clustering, distances are Euclidean. The colour scale in the heatmap goes from 4 (increased) to -4 (decreased) and the colour indicates the z-score of the lipid measurement within the row for each sample. B) 3D PCA plot of all samples. C) 2D PCA plot with red arrows indicating the loadings. Instead of dots sample names are shown. D, E, F show the 2D PCA plots with respectively PC1 and PC2, PC2 and PC3 and PC1 and PC3 as axes. The coloured "halos" indicate the 95% confidence regions of the groups.

Results

5.1.6 Correlations between lipids

The final question I wanted to explore with this data, was about the relationship between the lipids. To explore this, I generated a heatmap of correlations between the lipids using Pearson's correlation in metaboanalyst. This analysis can give insights into which lipids behave similarly and hence might be affected by the same pathways or come from the same origins. As can be seen in Figure 44, there are three main clusters of correlating lipids. The first cluster (going from left to right) contains C10:0, C20:5, C18:2 and C18:3. The next cluster contains C8:0-, C10-, and C12:0-carnitine, C12:0 – C18:0, C20:4 and C22:6. The furthest to the right cluster contains C2:0, C3:0, C6:0, and the remaining carnitines, i.e., SCFA-acylcarnitines and LCFA-acylcarnitines. In the middle cluster, the MCFA-acylcarnitines show strong correlations between each other and weaker correlations to the LCFA in the cluster. The MCFA-acylcarnitines in the middle cluster also show correlations with C6:0-carnitine ($r = 0.55-0.77$, $p=0-3.96^{-10}$) and a weak correlation with C2:0-carnitine ($r=0.22-0.34$, $p=0.000189-0.02$). C12:0-carnitine correlates with C14:0-carnitine ($r= 0.61$, $p=5.66^{-13}$). The cluster furthest to the right contains several lipids that do not strongly correlate with most other lipids in the cluster. A good example for this observation is C6:0, which was a dominating factor in the PCA. C6:0 only shows weak correlations with all other lipids. The strongest correlation is with C2:0-carnitine ($r=-0.41$, $p=6.4^{-6}$). Potentially interesting also is the moderate correlation between C2:0 and C3:0-carnitine ($r = 0.5$, $p=1.47^{-14}$) and the weak negative correlation between C2:0 and C18:1 carnitine ($r = -0.41$, $p=6.47^{-14}$).

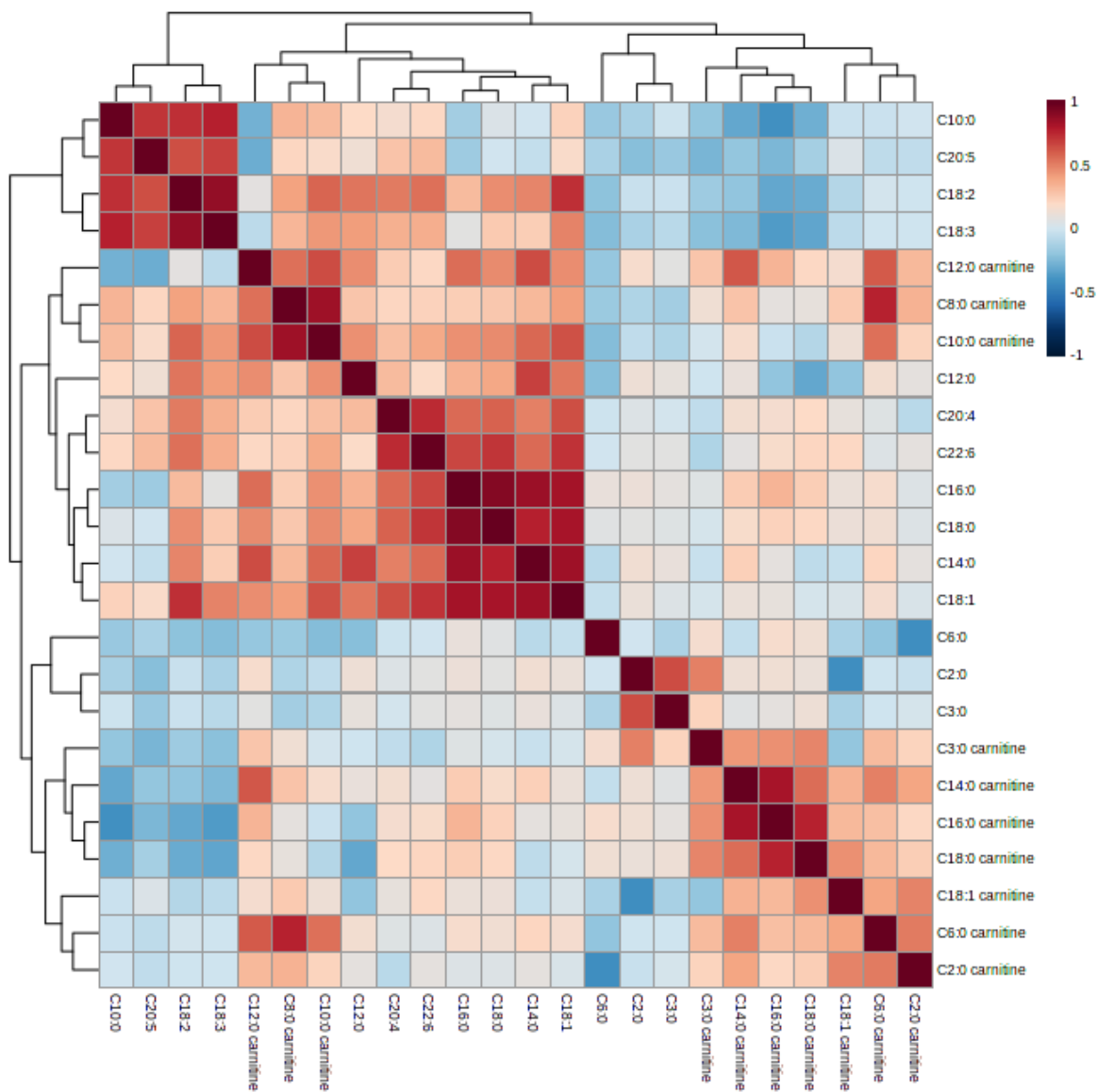


Figure 44 Heatmap of correlations between lipids measured in this cohort. Lipids were log-normalized and centre-scaled before correlations were computed. The correlations were calculated using Pearson's correlation, The colour indicates the r-value and ranges from -1 to +1.

5.2 Discussion

Due to the lack of a defined clinical diagnosis of sepsis for this data it is not possible to draw unequivocal conclusions about sepsis from the results in this chapter. I think however that it is possible to achieve insights into slightly different related questions. Firstly, the reliable measurement of almost all lipids in this cohort shows that the established method works for whole blood samples in addition to the plasma samples in chapter 4. In this chapter the measurement of C2:0, C3:0 and C6:0 was successful, but not of C4:0. A determining factor in this is likely the higher abundance of C2:0, C3:0 and C6:0 in the samples. In future endeavours, measurement of C4:0 might be possible with a machine that shows higher detection sensitivity for the SCFA. Another important finding in this chapter is that several lipids were significantly different between control neonates and neonates suspected of sepsis. To understand this finding better I would like to explain the cohort better. Neonates in the NICU are closely monitored and hence potential signs of a development towards sepsis are found early, yet the clinical deterioration towards sepsis is not a uniform process. This also means that the suspected sepsis cases in this study are a heterogeneous population of infants with varying levels of unwellness or sickness. Since the signs of sepsis can be hard to diagnose, many false positives are likely. Treatment, i.e., antibiotics, are usually started very quickly. Depending on several factors, including the potential antibiotics resistance of any infecting pathogen, this treatment might be effective quickly or not effective at all. As the samples I analysed were post antibiotic treatment I cannot rule out that this was a confounding factor when compared to controls. Although it is worth noting that all suspected cases were treated, hence the sub-stratification of lipid levels in this group would unlikely be due to the antimicrobial treatment. In addition, in terms of suspected cases that do not develop sepsis, it is not possible to distinguish between suspected sepsis cases that were not on a sepsis trajectory and those cases that did not develop sepsis due to their response to the early treatment. Since the samples I used in this study were taken 24 hours after the initial suspicion of sepsis and hence after the commencement of treatment, some of the neonatal patients could potentially already see an improvement of their health. This adds to the heterogeneity of the cohort. It is interesting that despite this assumed heterogeneity within the cohort, significant differences can be observed compared to the controls. One possible interpretation is that

even (relatively mild) unwellness leads to changes to the lipid profile in the blood. This seems feasible, since in boxplots and PCAs no subgroup of the cohort consisting of potential sepsis cases is distinctly visible. In the future when the samples are classified and clinical data can be connected to the samples, it would be interesting to not just analyse if lipids are significantly different between sepsis and non-sepsis, but also if there is a trajectory in the lipid profiles that matches a trajectory in severity from unwellness towards sepsis.

C10:0 and C12:0, of special interest due to being GPR84 ligands, are present in lower concentrations in this cohort compared to the other study. C12:0 is present in the suspected sepsis cases in decreased amounts that are all below the previously handled cut-off value of 8.8 μM . C10:0 is present in sufficient amounts to activate the receptor in only six samples ($<4.5 \mu\text{M}$).

When looking at the PCA of all samples based on the lipid measurements, there were one or two subgroups that appeared differently from other samples. Among the lipids influencing the distribution of the samples were lipids found to be altered in the plasma of sepsis patients in the previous study in chapter 4, i.e., C10:0, C2:0-carnitine and C20:5. However, comparing the ILTIS study with the nSep study is problematic due to both the difference in age group and, mostly, the difference in sample type. Indeed, in this study, some of the significantly changed lipids show changes contrary to those in ILTIS, while some lipids significant in this study were not significant in ILTIS and vice versa. Of course, it is possible that, due to heterogeneity of measurements, lipids would be significantly decreased or increased in sepsis in the nSeP cohort, and that this is masked by the suspected sepsis cases that are not sepsis. In other published work that have conducted an analysis of metabolites that change in whole blood samples of bacterial sepsis from adult patients admitted to emergency room; the only lipid measured to be significantly altered that was also measured in the nSep cohort was C14:0 (myristic acid), which was increased¹⁹¹. In the nSep cohort, C14:0 on average appears to be decreased in the suspected sepsis cases. However, this lipid shows great heterogeneity and there are several top outliers.

The acylcarnitines measured in this study show trends different to what is reported in the literature and in the ILTIS study. In sepsis, SCFA- and MCFA-acylcarnitines have been found to be increased, especially when comparing non-survivors to survivors of sepsis^{51,188,50,192}. However, all of these studies employed either plasma or serum. In this cohort, there are

Results

no significant increases in this group of acylcarnitines. Instead, C8:0- and C10:0-carnitine are even significantly decreased, albeit with a low fold change. Additionally, the LCFA-acylcarnitines C16:0- and C18:0-carnitine are increased. In the ILTIS cohort, C16:0-carnitine showed no significant changes and C18:0-carnitine was significantly decreased. These differences could be due to the sample type. It would be interesting to explore differences in acylcarnitine changes in plasma or serum and whole blood or cell pellets in future studies. This could lead to a better understanding of the changes in acylcarnitines relating to cellular metabolism and secretion. In addition, it could aid in understanding if changes are caused by cellular metabolism or physiological metabolism. Stringer et al. researched differences in the serum and whole blood metabolome in healthy volunteers and found that not all metabolic pathways could be analysed from serum¹⁹³.

Finally, the relationship between the lipids was analysed using correlation analysis. One interesting cluster is C10:0, C20:5, C18:2 and C18:3. C20:5, C18:2 and C18:3 are commonly derived from the diet as only C20:5 can be produced endogenously and only in low amounts⁴³. C10:0, in the neonatal cohort, is quite likely present in the diet as it is contained in breast milk^{152,194}. A correlation between these lipids could hence reflect a shared origin from the diet. However, further fatty acids, including C12:0, C14:0 and C16:0, should also be present in breast milk, and these do not correlate with this cluster of lipids. Another cluster observed that contains the LCFA C14:0, C16:0, C18:0 and C18:1 could potentially reflect the synthesis of these in the body, as these fatty acids can be produced in the same pathway. In this cohort, many correlations were found that were not present in the ILTIS cohort. This could be both due to the different age, and hence different diet and metabolism, or due to the fact that this is whole blood instead of plasma. Future studies comparing changes inside and outside of cells in sepsis would hence also be interesting relating to these lipids.

6 Mapping the transcription factor network associated with *GPR84* expression in sepsis

In this chapter I aim to elucidate the transcriptional regulation of *GPR84* to gain insight into the regulatory pathways driving expression of *GPR84* in sepsis. The expression of *GPR84* is for the most part undetectable or at very low basal levels of gene expression in cells suggesting that it is not required for normal constitutive or tissue specific cellular functions^{100,110,69,111,113,114}. However, *GPR84* gene expression consistently shows upregulation in both human and murine cells in response to inflammatory stimuli, especially in myeloid cell types^{100,110}. Apart from myeloid cells, expression has also been shown in fat tissue, kidney, liver, intestines, T-cells, B-cells and brain after inflammatory stimuli¹⁰⁰. Most studies examining *GPR84 in vitro* use primary cells or cell lines that are monocyte- or macrophage-like cells, neutrophils, or adipocytes^{69,109,123}. One previously published paper mentions analysis of transcriptional regulation of *GPR84*, disclosing only the role of NFκB as finding¹⁰⁹. They further confirmed a role of NFκB in transcriptional regulation of *GPR84* in an experimental approach in a murine adipocyte cell line. Other transcription factors that regulate *GPR84* expression have not been identified yet. In sepsis, *GPR84* expression is significantly increased in cases versus healthy controls³². Elucidating the transcriptional factor network interacting with *GPR84* and its promoter structure is informative in identifying the cellular and molecular pathways governing its expression. Certain transcription factors regulate cell-specific expression; while others act in response to stimuli, integrating hormonal and cellular signalling pathways. Their activity can be temporally defined, leading to distinct waves of gene expression in response to a stimulus. In the first hours after a stimulus, these are defined as the primary response genes, delayed primary response genes and secondary response genes¹⁹⁵. Genes of the primary response genes often code for transcription factors that regulate genes of the secondary response. This signalling hierarchy leads to temporal patterns in gene expression in response to a stimulus. Which TFs regulate a gene allows for a prediction of its temporal expression which can be confirmed using laboratory techniques. Firstly, timecourse experiments into the expression of *GPR84* in cells treated *in vitro* will elucidate the temporal order of its

Results

expression. In combination with the *in silico* analysis I seek to test whether GPR84 is part of the early immune response.

There are many tools and databases available online to query the transcriptional regulation of a gene. I evaluated different approaches, leading to a combinatorial method resulting in a High Confidence network of transcription factors regulating *GPR84* in sepsis.

Most databases and analysis tools for transcriptional regulation use a compendium of data from different experimental approaches. The Signalling Pathways Project and the ENCODE consortium curated online databases containing experimental data on transcriptional regulation that can be queried in the browser^{196,197}. The Signalling Pathways Project initially focused on nuclear receptor regulated processes in transcriptomic data, but since expanded to the several layers of signal transduction in cells and tissues, and included a cistromic database, that is data on cis-regulatory mechanisms of transcriptional regulation¹⁹⁶. The cistromic data was mined mostly from Chip-Atlas processed SRA datasets from NCBI, but also includes data from smaller scale experiments. Their “ominer” tool allows for the search of transcription factors in combination with cofactors for singular genes, among other options, and displays results as ChIP-atlas generated MACS2-peak values per transcription factor within 10kb of the presumed transcription start site of the queried gene. The ENCODE Consortium is a big collaborative effort between several international research groups with the aim to annotate the human genome further using several high-throughput approaches¹⁹⁷. These approaches seek to understand the spatial organization of the DNA and the regulation of transcription. Functional units that are annotated include enhancers, promoters, silencers, DNA and histone methylation sites and transcription factor binding sites. ENCODE-originating annotation of the human genome can be accessed in the UCSC genome browser, using the ENCODE website or looked up on the Harmonizome website which provides a comprehensive compendium of data-sources with information on genes and proteins¹⁹⁸. Data from ENCODE is used in the online tools remap and Chea3. Remap is an online tool that allows for the search of enrichment of transcription factor binding peaks from curated ChIP-seq, ChIP-exo, DAP-seq experiments from the NCBI, ENCODE and other databases. Rather than providing transcription factors regulating a gene as output, it uses the Chip-derived data to annotate regulatory regions in query sequences¹⁹⁹. Chea3 uses data from several sources, including GTEx co-expression,

Enrichr queries, remap ChIPseq data, ENCODE ChIPseq data, to predict transcription factors regulating the response to a stimulus²⁰⁰. As an input a set of upregulated genes in response to an experimental stimulus or a clinical condition is required; integrating various data sources on transcriptional regulation, the transcription factors most likely regulating the response will be computed, using a scoring system. Another *in silico* approach analyses datasets for co-expression of genes and transcription factors, since the expression of a transcription factor can be upregulated in situations where the transcription factor is active. The power in this approach is gained by using a vast amount of data but limited by the fact that not all transcription factors need to be expressed to exert their function or might be expressed timewise differently from the genes they regulate. ARCHS4 utilizes this approach in its compendium of data analysis tools²⁰¹, data from ARCHS4 is integrated into ChEA3.

6.1.1 Transcription factor binding site analysis

The promoter sequence can be annotated with predicted transcription factor binding sites (TFBS) based on primary sequence using the following *in silico* approach. Transcription factors bind to relatively specific nucleotide sequences, e.g. NFκB1 in humans is thought to bind to nucleotide sequences starting with 3 to 4 guanines and ending in 3 to 4 cytosines with 3 to 5 other bases inbetween²⁰². The nucleotide sequence a TF typically binds to is usually depicted as a motif or sequence logo that reflects the position weight matrix of preferred binding by the TF for each base in each position¹³³. These motifs can be queried against a promoter sequence generating hits of TFBS using a scoring system, using online tools like TRANSFAC or JASPER. These tools do not take into account other factors influencing transcriptional factor binding, e.g., chromatin accessibility and transcription factor interactions, or validate the binding sites with experimental data. All hits are hence predictions of potential binding sites, and the false positive rate is typically high^{203,204}. An overview of the analysis performed in this chapter can be seen in Figure 45.

Results

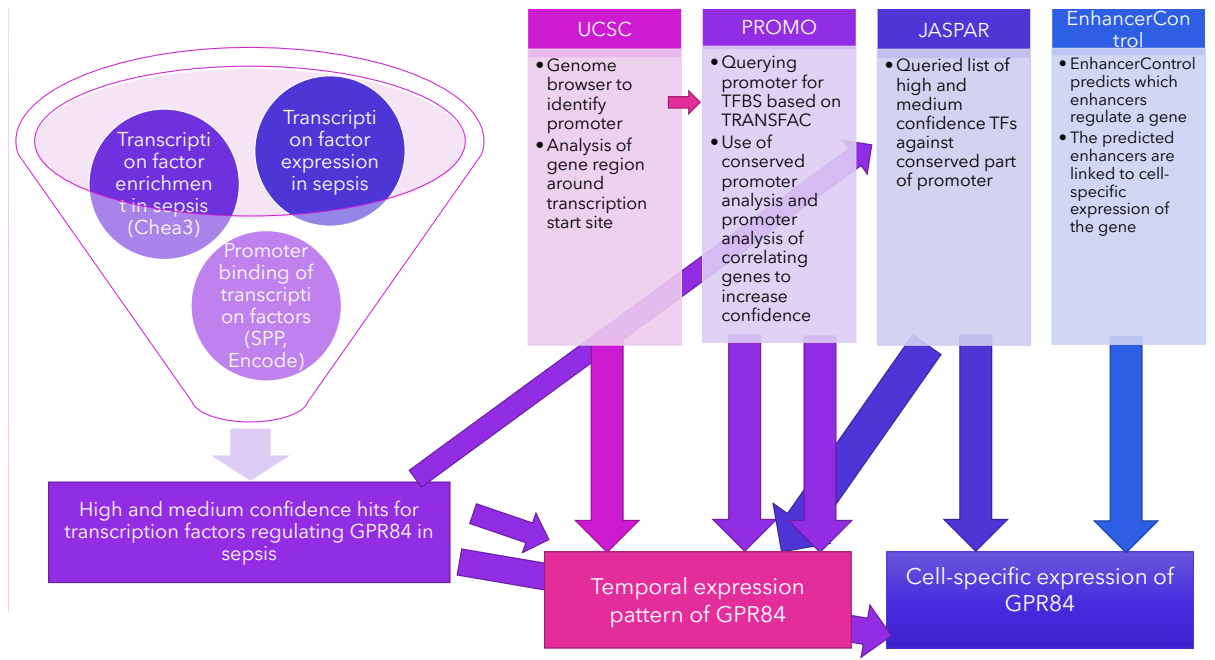


Figure 45 A flow diagram of the analysis steps performed in this chapter and the resulting interpretation of GPR84's transcriptional regulation they contributed to.

6.2 Results

6.2.1 Transcriptional regulation in sepsis

Transcription factor gene expression in sepsis

I first sought to characterise transcription factor expression changes in sepsis. For this purpose, I have chosen to use a transcriptomics dataset of paediatric sepsis cases and healthy controls and to independently confirm the results in a neonatal sepsis dataset. I downloaded the raw data of the paediatric sepsis dataset from GEO and normalized it. The normalized data of the neonatal sepsis dataset was provided to me by a colleague. First, transcription factor gene expression altered in sepsis was established. Lambert et al. have curated a list of over 1600 human transcription factors that was used to select genes in the dataset¹³³.

Several transcription factors are significantly up- or downregulated in both datasets as is apparent in these volcano plots (Figure 46, Figure 47). Among the transcription factors most strongly altered in paediatric sepsis are *LTF*, *DACH1* and *BCL6*, while in neonatal sepsis it is *SPI1*, *JUNB*, *GATA1* and *NFIX*. In paediatric sepsis, 558 TFs are significantly altered, with 15 TFs upregulated with a $\log_{2}FC > 1$, and 30 TFs downregulated with a $\log_{2}FC < -1$. In neonatal sepsis, 571 TFs are significantly altered, with 61 TFs upregulated with a $\log_{2}FC > 1$ and 55 downregulated with a $\log_{2}FC < -1$ (Table 10). Most upregulated TFs in paediatric sepsis cases are also upregulated in neonatal sepsis (11/15), but there is a vast difference in overall numbers of upregulated and downregulated TFs, with more TFs up- and downregulated in neonatal sepsis. A possible explanation is that the transcriptional response in neonates and older children might differ, an observation that has also been made by Wynn et al. based on a paediatric sepsis dataset that contains a neonatal cohort³¹. The paediatric sepsis dataset used here is a combination of two datasets including the dataset used by Wynn et al. and hence also contains a few neonatal sepsis cases. In addition, neonates in Wynn's cohort were full-term babies, while the neonatal sepsis cohort consists exclusively of pre-term babies. The transcription factors upregulated contain several TFs involved in the immune response and associated with immune cells, including *CEBPB*, *SPI1* and *NFE2*.

Next, I sought to identify co-regulated transcription factors with *GPR84* in sepsis. The correlation of expression of TF and *GPR84* in sepsis cases was computed in R using Pearson

Results

correlation for the paediatric sepsis dataset. Pearson's correlation was chosen as appropriate method due to the sufficient size of the cohort. Statistically significant Pearson correlations were found but none above or below the reasonable correlation threshold of 0.7 or -0.7 respectively. The strongest positive correlation between GPR84 and a TFs was found for *DACH1* (0.68), *ZNF438* (0.65), *BATF* (0.63), *RUNX1* (0.63) and *KLF7* (0.63). The strongest negative correlation can be found with *MYCL* (-0.54), *ZNF225* (-0.53) and *ZXDA* (-0.53). Since these are rather weak correlations, no conclusions can be drawn.

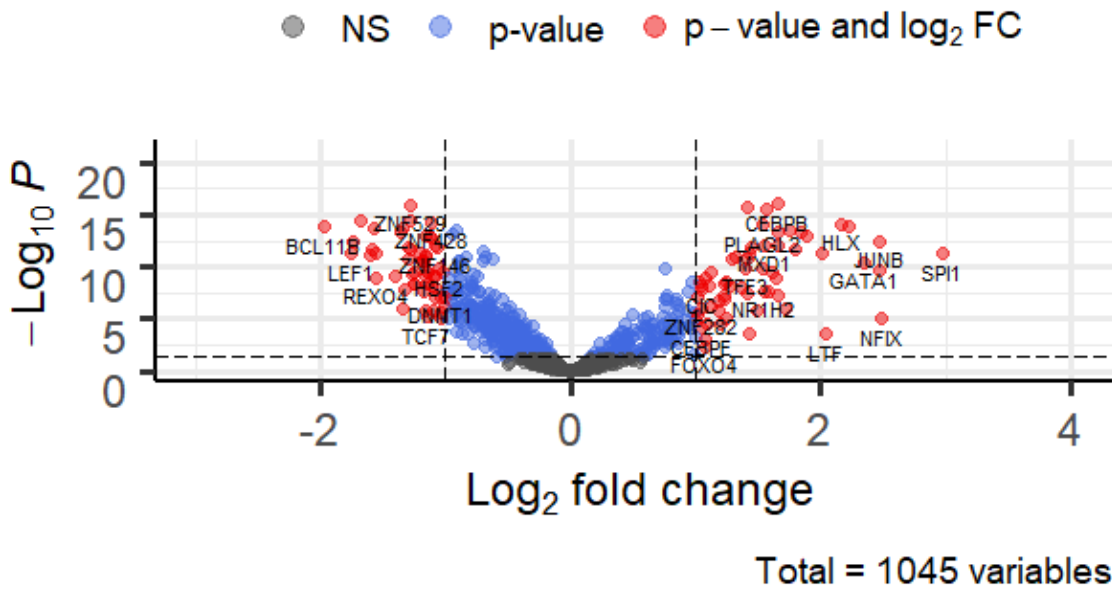


Figure 46 Volcano plot of transcription factor expression in a neonatal sepsis dataset showing the $-\text{Log}_{10}$ of the significance on the y-axis and the Log_2 Fold change comparing sepsis cases to healthy controls. The dashed lines indicate a Log_2 fold change of 1 and a p-value of 0.05 as $-\text{Log}_{10}$ value, respectively.

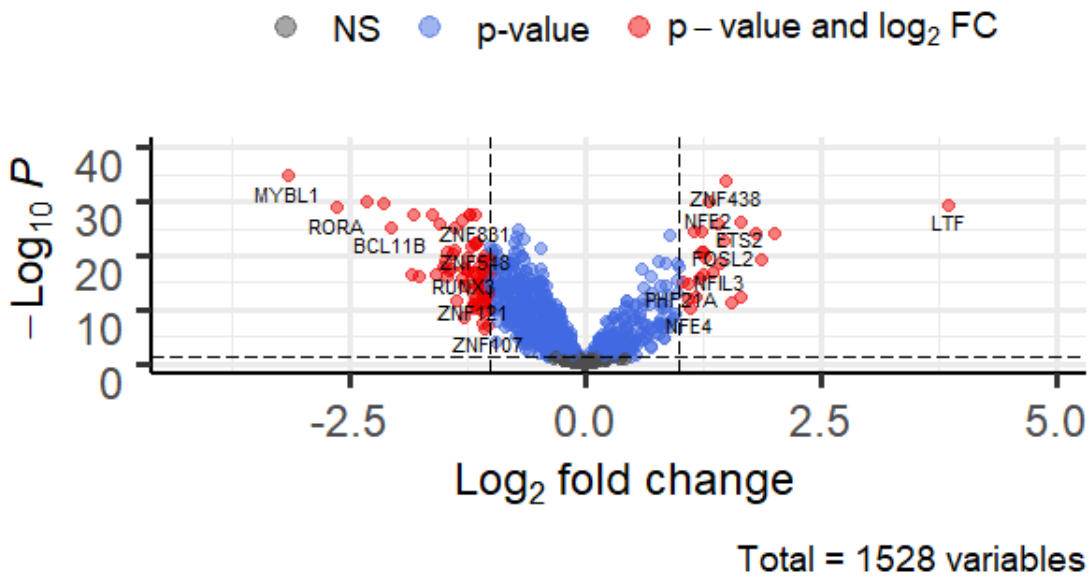


Figure 47 Volcano plot of transcription factor expression in a dataset of paediatric sepsis showing the $-\text{Log}_{10}$ of the significance on the y-axis and the Log_2 Fold change comparing sepsis cases to healthy controls. The dashed lines indicate a Log_2 fold change of 1 and a p-value of 0.05 as $-\text{Log}_{10}$ value, respectively.

Results

Table 10 This table shows the significantly up- and downregulated transcription factor genes identified in the paediatric sepsis transcriptomics data and the neonatal sepsis transcriptomics data.

Downregulated in paediatric sepsis	Log2 Fold change	Upregulated in paediatric sepsis	Log2 Fold change	Downregulated in neonatal sepsis	Log2 Fold change	upregulated in neonatal sepsis	Log2 Fold change	Overlapping TFs
<i>EOMES</i>	-1.83	<i>NFE4</i>	1.11	<i>BCL11B</i>	-1.97	<i>SPI1</i>	2.98	<i>CEBPB</i>
<i>ZNF600</i>	-1.50	<i>LTF</i>	3.86	<i>LEF1</i>	-1.76	<i>NFIX</i>	2.49	<i>MTF1</i>
<i>NR1D2</i>	-1.39	<i>DACH1</i>	1.96	<i>ZNF831</i>	-1.74	<i>IRF7</i>	2.47	<i>SPI1</i>
<i>STAT4</i>	-1.37	<i>BCL6</i>	1.68	<i>MXD4</i>	-1.68	<i>JUNB</i>	2.47	<i>CEBPD</i>
<i>HMGN3</i>	-1.31	<i>PLSCR1</i>	1.52	<i>ETS1</i>	-1.60	<i>GATA1</i>	2.35	<i>ZNF438</i>
<i>BACH2</i>	-1.31	<i>NFIL3</i>	1.44	<i>RORA</i>	-1.58	<i>CEBPD</i>	2.23	<i>ETS2</i>
<i>SATB1</i>	-1.22	<i>NFE2</i>	1.32	<i>KDM2B</i>	-1.56	<i>HLX</i>	2.16	<i>NFE2</i>
<i>RORA</i>	-1.20	<i>ETS2</i>	1.28	<i>REXO4</i>	-1.55	<i>LTF</i>	2.04	<i>NFIL3</i>
<i>GATA3</i>	-1.20	<i>ZNF438</i>	1.25	<i>ZNF816</i>	-1.55	<i>NFKB2</i>	2.01	<i>PLSCR1</i>
<i>ZNF83</i>	-1.19	<i>CEBPD</i>	1.24	<i>HOXB2</i>	-1.40	<i>MTF1</i>	1.89	<i>BCL6</i>
<i>ZNF121</i>	-1.17	<i>CEBPA</i>	1.23	<i>ZNF439</i>	-1.36	<i>ETS2</i>	1.84	<i>LTF</i>
<i>DNMT1</i>	-1.12	<i>SPI1</i>	1.21	<i>CENPB</i>	-1.34	<i>RELB</i>	1.80	<i>EOMES</i>
<i>ZNF266</i>	-1.12	<i>MTF1</i>	1.16	<i>ELF2</i>	-1.34	<i>NFIL3</i>	1.75	<i>STAT4</i>
<i>CXXC5</i>	-1.11	<i>CEBPB</i>	1.16	<i>EOMES</i>	-1.33	<i>KLF1</i>	1.71	<i>BACH2</i>
<i>BCL11A</i>	-1.11	<i>PRDM5</i>	1.15	<i>BCL11A</i>	-1.30	<i>ZFPM1</i>	1.66	<i>SATB1</i>

<i>E2F5</i>	-1.09			<i>ZNF721</i>	-1.30	<i>RFX2</i>	1.65	<i>ROR A</i>
<i>ZNF573</i>	-1.09			<i>ZNF529</i>	-1.28	<i>BCL6</i>	1.65	<i>DNMT1</i>
<i>ZNF420</i>	-1.08			<i>ZNF22</i>	-1.28	<i>CEBPB</i>	1.65	<i>E2F5</i>
<i>ZNF439</i>	-1.08			<i>ZNF101</i>	-1.28	<i>CREB5</i>	1.65	<i>LEF1</i>
<i>KDM2B</i>	-1.08			<i>ZNF827</i>	-1.27	<i>MXD3</i>	1.61	<i>TCF7</i>
<i>ZFP62</i>	-1.07			<i>SATB1</i>	-1.25	<i>PLSCR1</i>	1.58	
<i>LEF1</i>	-1.05			<i>ZNF559</i>	-1.25	<i>ARID3A</i>	1.57	
<i>ZBTB41</i>	-1.04			<i>KLF12</i>	-1.24	<i>NR1H2</i>	1.56	
<i>MEF2C</i>	-1.04			<i>ZNF789</i>	-1.20	<i>MXD1</i>	1.55	
<i>ZFP3</i>	-1.03			<i>STAT4</i>	-1.19	<i>PLAGL2</i>	1.54	
<i>ZNF260</i>	-1.03			<i>ZNF683</i>	-1.18	<i>TRAFD1</i>	1.53	
<i>TCF7</i>	-1.03			<i>ZFP90</i>	-1.17	<i>BATF2</i>	1.49	
<i>TSHZ1</i>	-1.03			<i>ZNF30</i>	-1.17	<i>MBD6</i>	1.45	
<i>ETS1</i>	-1.00			<i>BACH2</i>	-1.16	<i>FOXO3</i>	1.43	
<i>ZNF329</i>	-1.00			<i>NFX1</i>	-1.16	<i>ZNF438</i>	1.42	
				<i>ZNF33B</i>	-1.16	<i>STAT3</i>	1.42	
				<i>ZBTB25</i>	-1.16	<i>CUX1</i>	1.41	
				<i>ZNF285</i>	-1.15	<i>TCF7L2</i>	1.41	
				<i>TCF7</i>	-1.15	<i>TFE3</i>	1.39	
				<i>FOXO1</i>	-1.14	<i>RBCK1</i>	1.36	
				<i>RUNX3</i>	-1.14	<i>TEF</i>	1.32	
				<i>ZNF234</i>	-1.14	<i>ELF4</i>	1.30	
				<i>ZNF544</i>	-1.12	<i>KMT2B</i>	1.27	
				<i>ZNF428</i>	-1.12	<i>STAT1</i>	1.25	
				<i>ZBTB14</i>	-1.11	<i>GFI1B</i>	1.25	
				<i>REPIN1</i>	-1.10	<i>RELA</i>	1.23	
				<i>ZSCAN26</i>	-1.10	<i>SP2</i>	1.23	
				<i>RXRβ</i>	-1.09	<i>ZNF366</i>	1.23	
				<i>THAP1</i>	-1.09	<i>NFE2</i>	1.20	
				<i>ZNF146</i>	-1.08	<i>FOS</i>	1.19	
				<i>ZNF32</i>	-1.08	<i>ARID5A</i>	1.18	

Results

				<i>ZNF512B</i>	-1.07	<i>MAFG</i>	1.12	
				<i>ZBTB24</i>	-1.07	<i>EPAS1</i>	1.12	
				<i>HSF2</i>	-1.05	<i>BATF</i>	1.11	
				<i>CENPX</i>	-1.04	<i>PA2G4</i>	1.09	
				<i>ATMIN</i>	-1.03	<i>TFDP2</i>	1.08	
				<i>DNMT1</i>	-1.03	<i>NCOA1</i>	1.08	
				<i>ZSCAN2</i>	-1.02	<i>FOXO4</i>	1.07	
				<i>E2F5</i>	-1.01	<i>PBX1</i>	1.05	
				<i>CXXC1</i>	-1.00	<i>ZNF282</i>	1.05	
						<i>CIC</i>	1.04	
						<i>CEBPE</i>	1.04	
						<i>ATF6</i>	1.04	
						<i>PPARG</i>	1.02	
						<i>TAL1</i>	1.01	
						<i>IRF3</i>	1.01	

6.2.2 Functional enrichment of transcription factors in sepsis

Since a transcription factor does not need to be upregulated in order to be active, transcription factors active in sepsis in both datasets were predicted using ChEA3. As input ChEA3 requires a list of upregulated genes in response to a stimulus²⁰⁰. I hence obtained the top upregulated genes in both neonatal and paediatric sepsis and provided these as query lists into ChEA3 in two separate analyses. ChEA3 then generates a list of statistically enriched TFs, among other types of output, computed using a background list of control genes using Fisher's exact test to compare the hits for the control list and the provided query list. The hits are found when ChEA3 utilizes a comprehensive approach querying the query and the control gene lists against several databases with different types of experimental data on transcriptional regulation, including data from ENRICH and ENCODE. Every hit per upregulated gene in these databases is taken into account. The output list of transcription factors is ranked depending on the type and number of libraries the hits for a transcription factor were found in and the number of genes the hit was found for.

The lists generated by ChEA3 of statistically enriched TFs in paediatric and neonatal sepsis had much overlap, e.g., ~74% among the 250 highest ranking TFs. The transcription factors found were strongly associated with immune cells of the blood and the immune response (Figure 48A, B). This association is automatically computed by ChEA3 based on the top ranked TFs and GTEx tissue data, a database for gene expression and regulation in different tissues²⁰⁵. This prediction is hence based on the expression of the enriched TFs in different tissues. Apart from association with blood and immune response, association with fat tissue was also found (Figure 48 A) and lymphocyte differentiation, epidermis development and central nervous system neuron tissue development (Figure 48B). Filtering the 250 top ranked TFs for TFs with at least one hit in one of the libraries for GPR84 lead to a list of 60 TFs based on the paediatric sepsis gene list and 62 TFs based on the neonatal sepsis gene list (Supplementary table 1). The overlap between these hits was ~82% or 51 TFs in common. Among these transcription factors were SPI1, CEBPB and NFKB1. When comparing TFs enriched according to this analysis with TFs that are significantly up- or downregulated in sepsis, 113 out of the top 250 enriched TFs are also significantly different expressed, of which 81 are upregulated and 32 are downregulated. This aligns with the previous statement that some transcription factors do not need to be upregulated to be

Results

active. As expected, the transcription factors enriched in the sepsis response are important in regulating the immune response and several are immune cell specific.

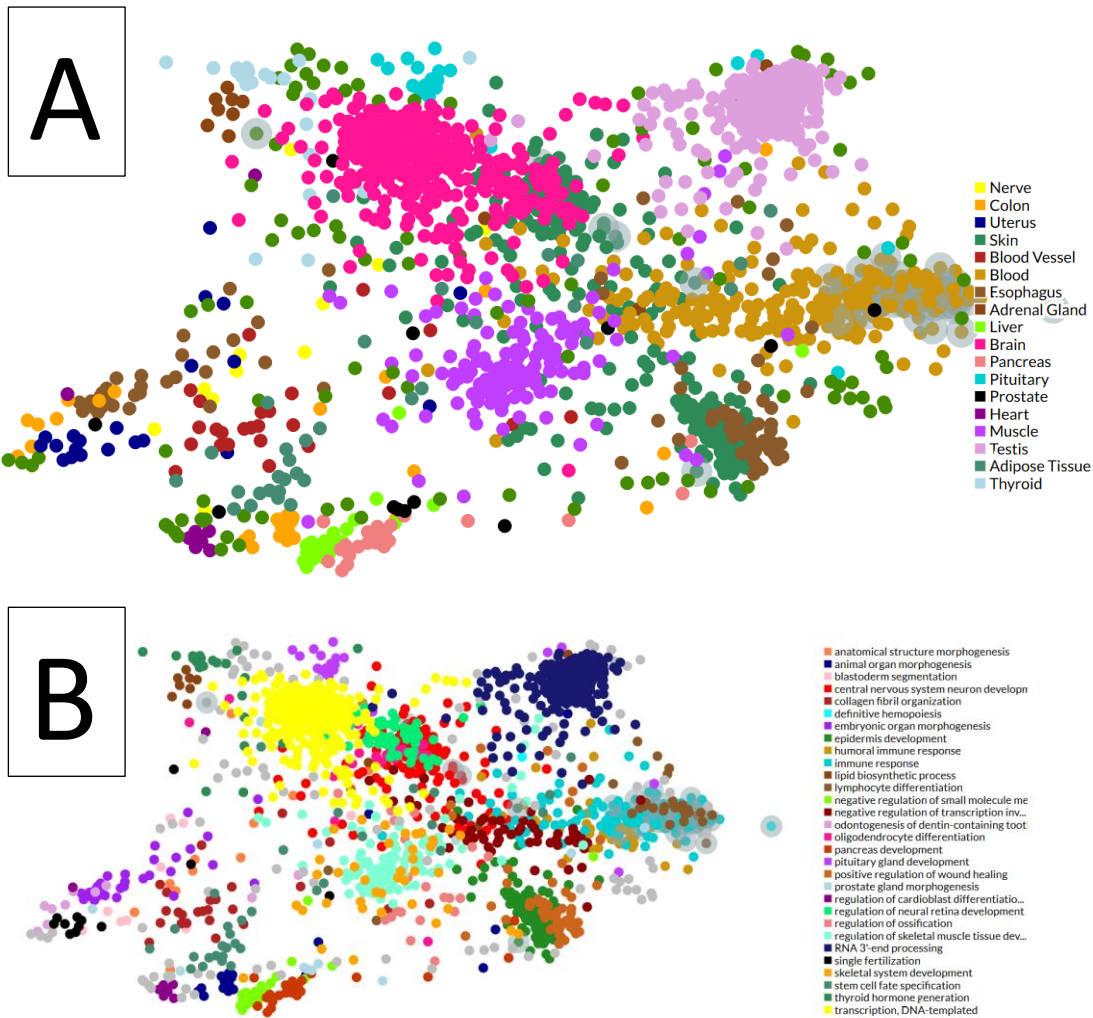


Figure 48 Depiction of transcription factors and their association with tissues (A) and GO pathways (B) based on GTE:X - TF data. Figure generated by Chea3 based on TFs predicted to regulate upregulated genes of a transcriptomics dataset of paediatric sepsis cases. Each dot represents a TF, and the colour of the dot indicates the tissue (A) or GO pathway (B). The top 50 ranking TFs enriched in sepsis are surrounded by a grey circle.

Results

6.2.3 Physical transcription factor interactions with the *GPR84* promoter according to existing Chip data

Since the binding of transcription factors to gene regulatory regions is a pivotal component of transcriptional regulation, I then sought to find out what transcription factors have been shown to physically bind the *GPR84* regulatory region. To do this I queried publicly available data from high throughput Chip experiments. As previously mentioned TFs likely binding to a putative gene's promoter according to ENCODE can be looked up using Harmonizome¹⁹⁸. ENCODE defines the putative promoter region of a gene as the 1000 bp around the TSS. This yields a list of 32 TF. The Signalling Pathways Project yields an excel file with MACs-peak values and yields 153 distinct TF for all cell types and 77 TF for immune cells specifically (Supplementary table 2)¹⁹⁶. 17 TF overlap between these two databases (Figure 49). Filtering TFs identified by SPP by cell type of the ChipSeq experiment, shows that the majority of overlapping TFs (16/17) have been identified in experiments on immune cells. Among the TFs found to overlap between the datasets are SPI1, CEBPB, MEF2A, STAT3, ZNF384, CTCF and MYC. Among the transcription factors found are both common transcription factors and transcription factors specific for immune cells and the immune response.

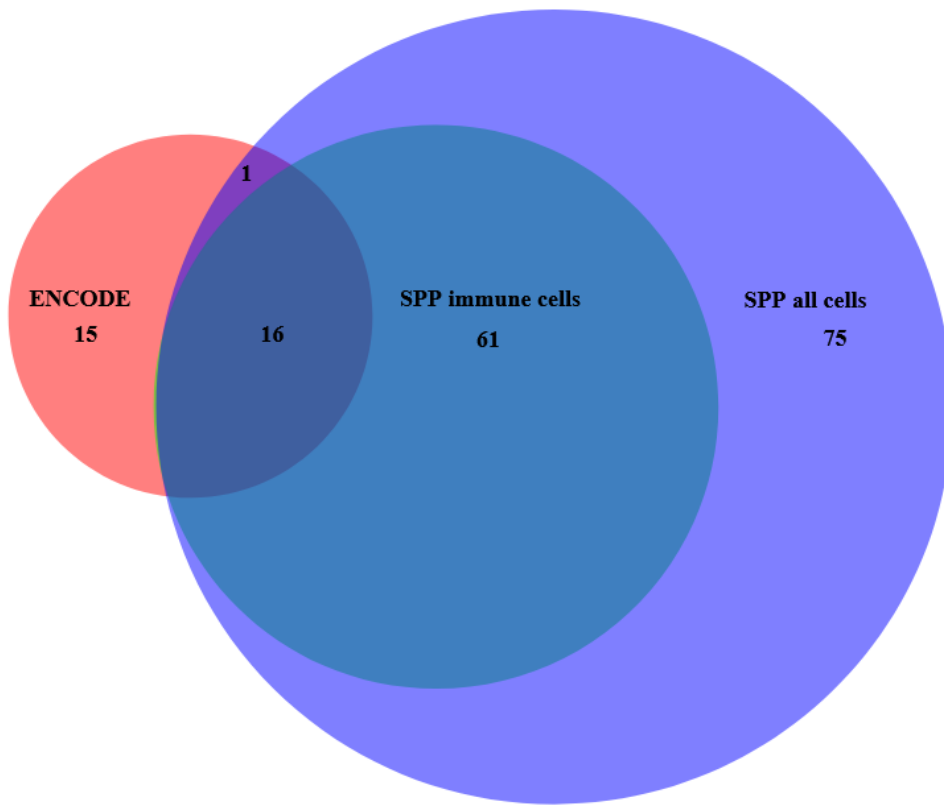


Figure 49 Venn diagram of Chip hits found in ENCODE and the Signalling Pathways Project (SPP) databases using online tools Harmonizome and Ominer. Venn created using BioVenn²⁰⁶

Results

6.2.4 High confidence networks of transcription factors regulating *GPR84* in sepsis

To identify which TFs interact with *GPR84* and to rank them as High and Medium confidence hits, the information from previous analyses was aggregated with exclusion of the TFBS analysis. Results from the TFBS analysis were excluded because a focus on experimentally derived data to yield functionally relevant results with high confidence was sought. Supplementary table 3 showcases a list of all TF identified by one or several approaches, with exemption of the TFBS analysis, with a separate column indicating if a TF had a hit in the corresponding analysis. The overlap between the data is also visualized in a Venn diagram (Figure 50). Taken into account this compounded information, High Confidence and Medium Confidence TFs that likely regulate *GPR84* expression in sepsis were identified. If a TF is enriched according to ChEA3 and had a hit in one of the libraries for *GPR84*, this means that the TF is probably regulating gene expression of several upregulated genes in sepsis including *GPR84*. Among the libraries used by ChEA3 are Chip based libraries, but also co-expression libraries. A hit in the SPP database in combination with a hit in ChEA3 hence adds additional confidence that a TF is active in sepsis and regulates *GPR84* by binding to its regulatory sequence. TFs that with High Confidence regulate *GPR84* in sepsis were hence identified by both ChEA3 and SPP. Medium Confidence hits were required to be predicted as enriched by ChEA3 without a hit for *GPR84* but with a hit in the SPP Chip data. Medium Confidence are quite likely to regulate *GPR84* since they are likely active in sepsis and able to bind the promoter. Some of the TFs with a hit for *GPR84* in the ChEA3 libraries do not have a hit in the SPP database, however ChEA3 uses ENCODE Chip data. As seen in previous analysis, ENCODE identified some Chip hits that were not present in SPP data. The TFs enriched in sepsis with a hit in the ENCODE database were sought to be identified as Medium Confidence hits, but no additional TFs were identified this way. High Confidence and Medium Confidence hits are shown in Table 11. TFs that were identified in the SPP database but are not enriched in sepsis are likely TFs that can regulate *GPR84* expression in conditions that differ from sepsis. TFs identified by ChEA3 with a hit for *GPR84* are relatively likely to regulate *GPR84* expression in sepsis, however they were not included as Medium Confidence hits in an effort to be conservative. To identify functional relationships between the transcription factors found, networks were visualized using STRING²⁰⁷(Figure 51, Figure 52). Looking at the first network a large immune TF cluster is evident containing FOS and JUN, JUN-like transcription factors JUNB

and JUN, FOSL1 and FOSL2. In this cluster also contained are ATF3, TCFL2, IRF1, GTF2B, IRF4, EGR1, CREB1, HIF1A and CEBPA. Several of these transcription factors are known to be part of the immune response. Further clusters containing STAT1, STAT2, STAT3, STAT5A and ZEB1; NFKB1, RELA and POU2F2; GATA1, GATA2, TAL1, MYB, RUNX1, TCF12, ELF1 and MEF2A are apparent. When looking at clusters in transcription factors with physical interaction, a smaller cluster containing FOS, JUN, JUNB, JUN, FOSL1, FOSL2 and ATF3 presents itself rather than the bigger cluster based on functional interactions. Most clusters from functional analysis undergo only small changes compared to clusters based on physical interaction, e.g. the cluster containing NFKB1 and RELA loses POU2F2 and clusters with CEBPB instead. This indicates that many of the functional interactions detected in the first network rely on physical interaction. Among the transcription factors identified are several genes associated with gene expression in immune cells specifically, e.g. SPI1 and CEBPB. Several of the transcription factors are involved in immune responses, e.g. NFKB1, STAT transcription factors, GATA1 and GATA2. NFκB was previously found to regulate *GPR84*¹⁰⁹, and was also found in this analysis as high confidence TF hit.

Results

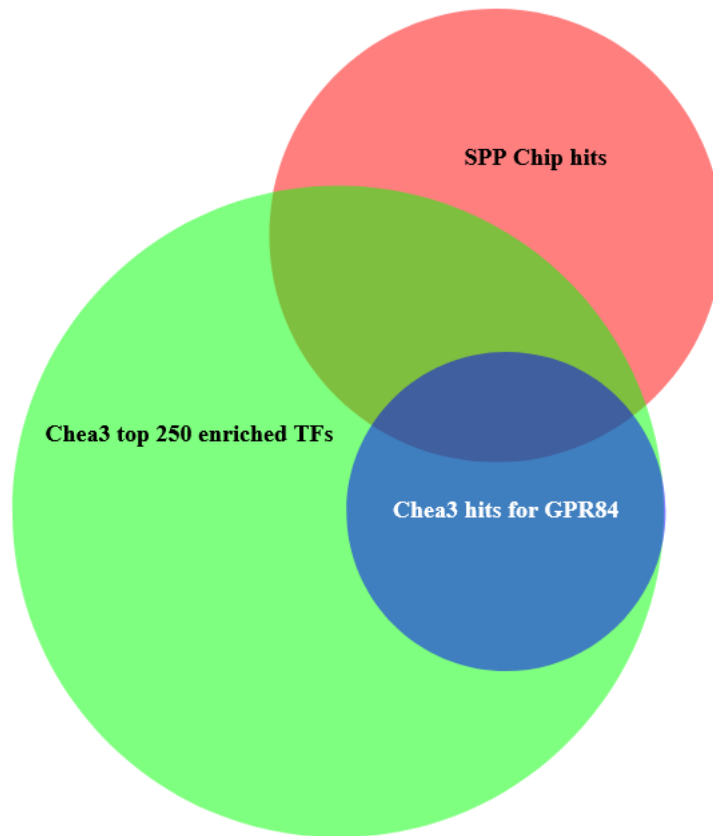


Figure 50 Venn diagram showing the overlap between the results from two analysis approaches.

Table 11 High Confidence and Medium Confidence TFs driving GPR84 expression in sepsis

High Confidence	Medium Confidence
RUNX1	FOS
CEBPB	JUN
STAT1	IRF1
CREB1	MAX
SPI1	MYB
SP1	TCF12
GATA1	ETS1
FOSL1	CEBPA
ELF1	ELF3

ATF3	USF1
SOX9	JUND
IKZF1	FOSL2
NFKB1	RELA
FOXP1	MXI1
MEIS1	MEF2A
	STAT3
	SREBF1
	STAT5A
	HIF1A
	BHLHE40
	SRF
	KLF11
	TCF7L2
	EGR1
	MAFK
	HMGN3
	GATA2
	TAL1
	POU2F2
	SP2
	LYL1
	ERG
	JUNB
	GTF2B
	ARID3A
	IRF4

Results

	STAT2
	ZEB1
	NFYC

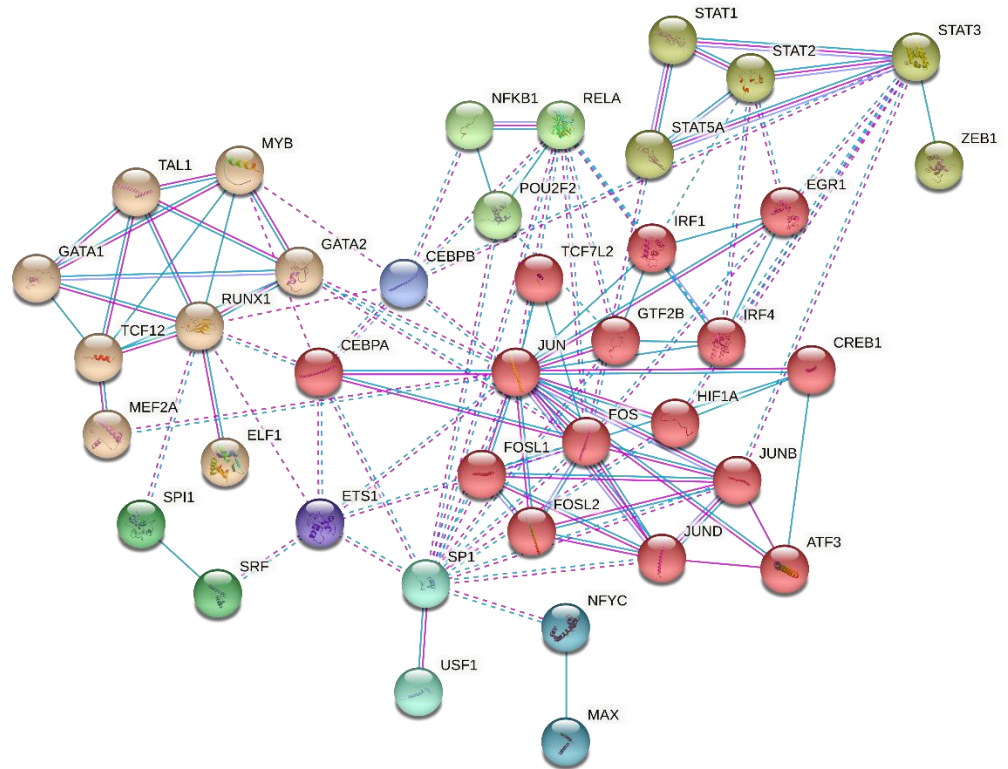


Figure 51 STRING network of transcription factors predicted to regulate GPR84 in sepsis with high and medium confidence. The edges indicate the evidence for interaction, confidence of interaction is minimum 0.6. Turquoise edges indicate interactions from curated databases, purple indicates interactions that were experimentally determined. The nodes were clustered using MCL clustering with an inflation parameter of 3. The edges indicate that proteins have a specific joint shared function, independent of if they physically interact.

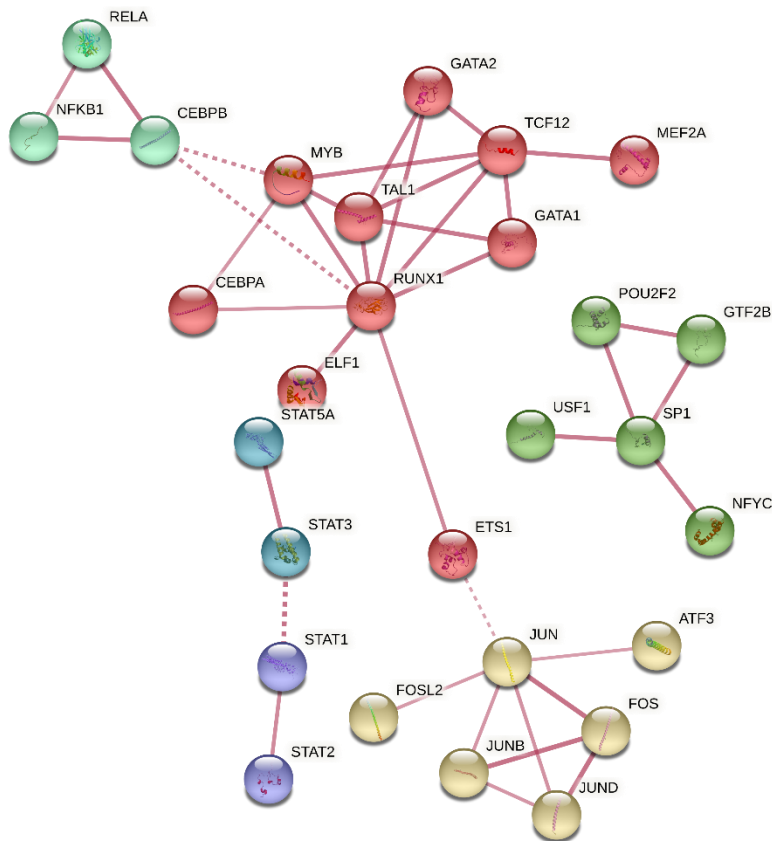


Figure 52 STRING network of physical associations between High Confidence and Medium Confidence TFs. Edges indicate confidence, minimum confidence being 0.6. Edges stand for physical interaction between the proteins. Nodes were clustered using MCL clustering with inflation parameter of 3.

6.2.5 The role of regulatory regions in the transcriptional regulation of GPR84

Identification of the promoter and regulatory hallmarks adjacent to the transcription start site

The regulatory regions and their architecture can provide important insights into the transcriptional regulation of a gene. To define the genomic region around the transcription start site of *GPR84* including regulatory regions and chromatin landmarks, the UCSC genome browser was used since it provides a visual summary for these features²⁰⁸. I also used this to identify the putative promoter sequence. The *GPR84* gene sequence is situated on the negative strand on the 12th chromosome. There are two GenCode Transcripts for the gene, one being identical to the curated RefSeq gene. Both transcripts span the same genomic region. Visualizing it in the UCSC genome browser shows promoter sites predicted upstream on the negative strand by ENCODE, Genehancer and ORegAnno (Figure 53). Most *in silico* tools use a putative region of 1kb (=1000 base pairs) or 10kb around the transcription start site (TSS) as putative promoter. To obtain a putative promoter sequence the EPD website in conjunction with the UCSC browser was used. EPD annotates a TSS very close to *GPR84*'s first exon. The TSS predicted by EPD is in a different position than the TSS predicted by Genecards (visible in Figure 53). When looking at the area around the TSS predicted by EPD, several other features of a promoter are apparent (Figure 54). Firstly, a promoter sequence needs to be open so that TFs can bind. Adjacent to the TSS is a stretch of DNA with DNase I hypersensitivity. DNase I is an enzyme that cuts DNA and can only do so if the DNA is accessible, but not when it is condensed. In addition, a stretch with ChIPseq TF clusters are found around the TSS. Lastly, one can see that the sequence around this TSS is more highly conserved than the sequence further away from the TSS. Evolutionary conservation of a sequence indicates that the sequence is functional. The majority of promoters contain CpG islands, which means that stretches of the promoter harbour cytosine and guanine nucleotides at percentages that are relatively high compared to the rest of the genome²⁰⁹. Depending on the stringency used, a sequence will be a CpG island at >55%, >60% or >65% GC-content²¹⁰. Notably no CpG islands were annotated for this promoter. The percentages of A, C, G and T in the sequence were ~27.6%, ~19.6%, ~25.9% and ~26.9%, hence the GC-content is markedly lower than 55%. Based on these findings taken together, I chose the DNA stretch of 1000 base pairs upstream and 300bp downstream of the TSS predicted by EPD, annotated in Figure 54 as black bar titled "User

supplied track”, for further analysis. It appears that the promoter is likely a relatively closed structure, which could indicate the requirement of opening the promoter before transcription can be induced by transcriptional factors.

Results

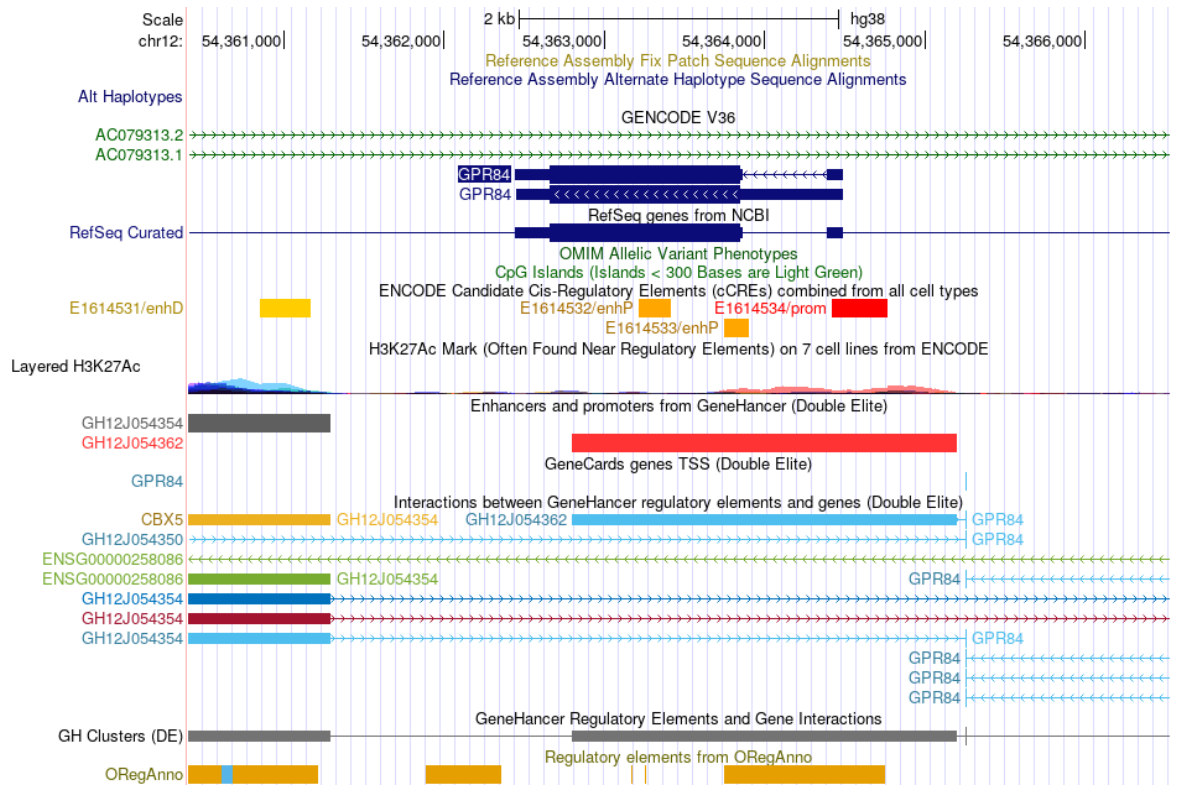


Figure 53 Annotation of the genomic region around *GPR84* from UCSC genome browser²¹¹. Tracks relating to gene expression regulation were chosen including predicted promoter sites and enhancers from GeneHancer, Encode and ORegAnno.

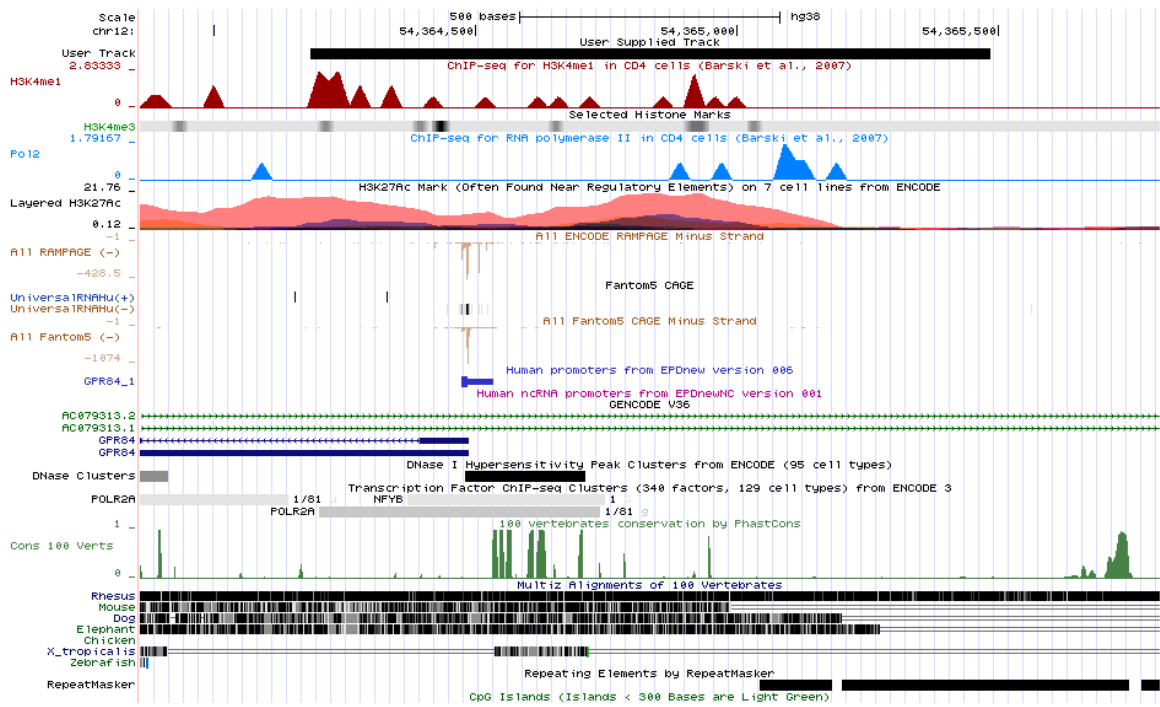


Figure 54 The genomic region upstream of *GPR84* visualized in the UCSC genome browser with tracks relating to transcriptional regulation including DNase I Hypersensitivity, sequence conservation, CpG Islands and TF ChipSeq clusters.

Transcription factor binding sites in the putative promoter sequence

Querying a primary sequence for TFBS based on weighted position matrices from the TRANSFAC database allows for the identification of putative transcription factors regulating a gene for which no experimental data is available. This analysis was focused on the previously identified putative 1300 bp promoter region for human *GPR84*. Information about TFBS can be used to establish the promoter architecture which allows for further insights into the transcriptional regulation of a gene. Using PROMO²¹², 552 putative TFBS from 65 different TF were found in the human *GPR84* promoter. To reduce the number of false positives in this approach, evolutionary conserved TFBS and TFBS in co-expressing genes were analysed. The presence of TFBS in promoters of genes correlating with *GPR84* indicates the TFBS to be actually functional, since genes with correlating expression are likely to be regulated by the same transcription factors. In addition, TFBS that are functional are more likely to be evolutionary conserved than randomly occurring motifs. To find evolutionary conserved TFBS, the promoters of human and murine *GPR84* were aligned using Dialign²¹³. The 500bp long subpart of the human promoter that aligned with the murine promoter was extracted and queried in PROMO for TFBS. This yielded a list of 359 TFBS of 44 different TFs. To analyse TFBS in promoter sequences of correlating genes, genes correlating with a Pearson correlation higher than 0.77 in the paediatric sepsis dataset were obtained, yielding a list of 16 genes including *GPR84*. To obtain their promoter sequences from EDP, the 1300 bp around the TSS were used as was done for *GPR84*. PROMO was used to find TFBS present in at least 75% or 12/16 sequences. This search resulted in 57 different TFs found. There is a strong overlap between the three searches performed (Figure 55), with 44/65 TFBS found in the *GPR84* promoter also found only in the conserved subpart, and 40 of the TFBS in the conserved subpart also found consistently in correlating gene promoters. shows the TFBS that were found. Among these are hormonal receptors (progesterone receptor, PR and glucocorticoid receptor, GR) and transcription factors involved in immune cells (CEBPB, PAX5).

A second approach to identify and confirm TFBS was performed using JASPAR. JASPAR allows the user to query TFs against a promoter sequence and predicts potential TFBS based on TFBS motifs. The previously established lists of High Confidence and Medium Confidence TFs was queried against the conserved subpart of the promoter identified in the previous analysis. This search yielded a list of 213 TFBS for 44 TFs, which was manually

examined for potential homotypic and heterotypic clusters. Among potential TFBS clusters identified are GATA2 TFBS pairs, a cluster of TFBS of Ets family TF, and STAT family TF pairs. Two NFkB1 binding sites were identified, in contrast to previous PROMO analysis of the same sequence.

Results

Table 12 shows the TFBS found in the human GPR84 promoter, the TFBS found in the majority of genes correlating with GPR84 in sepsis, and the TFBS found in the promoter sequence that is conserved in humans and mice.

Human GPR84 promoter TFBS	TFBS in promoters of correlating genes	TFBS in conserved subpart of human GPR84 promoter
GR-alpha	GR-alpha	GR-beta
AP-2alphaA	AP-2alphaA	C/EBPbeta
RXR-alpha	RXR-alpha	NF-1
YY1	YY1	ENKTF-1
AhR	ENKTF-1	XBP-1
AhR:Arnt	GR-beta	YY1
ENKTF-1	C/EBPbeta	SRY
GR-beta	NF-1	TCF-4E
C/EBPbeta	TFII-I	GR
NF-1	c-Myb	GR-alpha
TFII-I	ER-alpha	FOXP3
c-Myb	ATF3	PR B
ER-alpha	Pax-5	PR A
ATF3	p53	c-Ets-2
IRF-2	TFIID	HNF-3alpha
Pax-5	NFI/CTF	C/EBPalpha
p53	HNF-3alpha	NF-Y
TFIID	GR	SRF
NFI/CTF	USF2	TFII-I
HNF-3alpha	MEF-2A	c-Myb
GR	HNF-1C	AR
USF2	FOXP3	LEF-1
MEF-2A	HOXD9	VDR
HNF-1C	HOXD10	PXR-1:RXR-alpha
FOXP3	IRF-1	STAT4
HOXD9	NF-AT1	c-Ets-1

HOXD10	PR B	Elk-1
IRF-1	PR A	AP-2alphaA
NF-AT1	XBP-1	TFIID
PR B	C/EBPalpha	USF2
PR A	c-Jun	NFI/CTF
XBP-1	T3R-beta1	RXR-alpha
C/EBPalpha	POU2F1	T3R-beta1
NF-Y	GATA-1	PEA3
c-Jun	VDR	GATA-1
T3R-beta1	PXR-1:RXR-alpha	HOXD9
POU2F1	STAT4	HOXD10
GATA-1	c-Ets-1	NF-kappaB1
VDR	RAR-beta	Ik-1
PXR-1:RXR-alpha	LEF-1	ELF-1
STAT4	SRY	PU.1
c-Ets-1	TCF-4E	E2F-1
RAR-beta	c-Ets-2	RAR-beta
LEF-1	AR	POU2F1
TCF-4	Elk-1	
SRY	MAZ	
TCF-4E	PEA3	
c-Ets-2	NF-AT1	
SRF	STAT1beta	
AR	AP-1	
Elk-1	NF-kappaB1	
MAZ	Ik-1	
PEA3	E2F-1	
NF-AT1	NF-AT2	
STAT1beta	EBF	

Results

AP-1	PPAR-alpha:RXR-alpha
NF-kappaB1	ETF
Ik-1	
ELF-1	
PU.1	
E2F-1	
c-Fos	
NF-AT2	
GATA-2	
COUP-TF1	

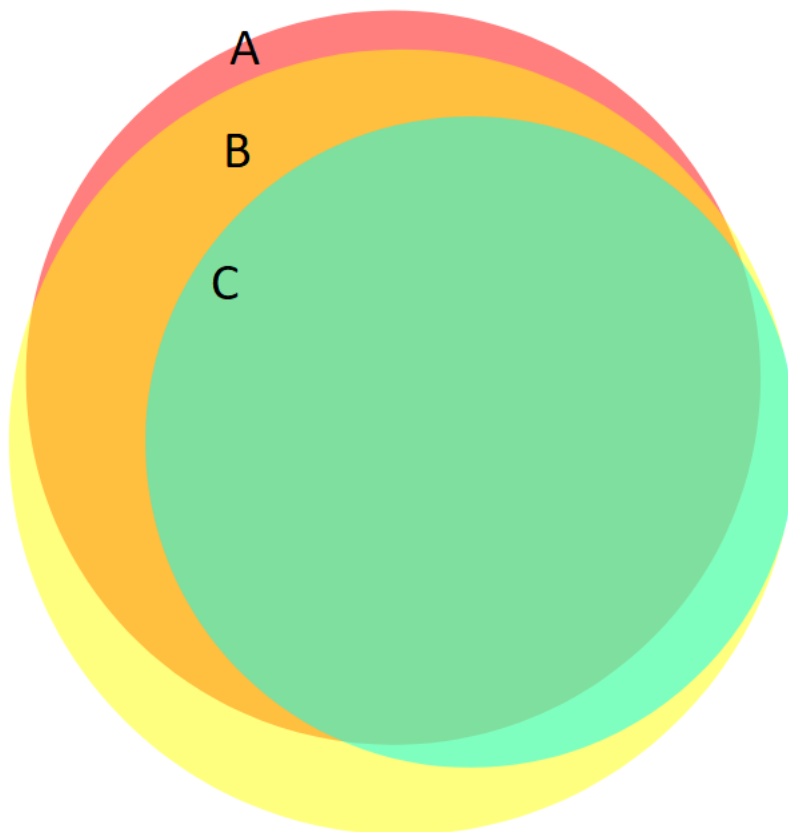


Figure 55 Overlap in TFBS between the human GPR84 promoter sequence (B), a conserved subpart of the human GPR84 promoter (C), and TFBS found in at least 75% of genes correlating with GPR84 (A).

The role of enhancers in regulating GPR84

Genes are regulated by promoters and enhancers, with some genes depending more on enhancers or promoters than others. It is hence important to not solely focus on promoters in analysing transcriptional regulation. That is why I sought to understand better the role of enhancers in the regulation of *GPR84*. Yoshida et al. used extensive mouse experiments, in which they performed ATAC-seq, ChIPseq and microarray analysis on 86 cell types obtained of mice. Using the obtained data, they were able to draw conclusions about the relation between open chromatin regions, transcription factors and gene expression in these cell types, i.e., the expression of a gene in different cell types can be linked with the transcription factors bound in open chromatin regions that are typically open when this gene is expressed. Yoshida et al. computed that immune processes, GPCR signalling and lipid transport are foremost distal enhancer regulated processes, and that *Gpr84* is regulated to ~32% by its promoter and to ~64% by distal enhancers¹³⁶. Yoshida et al. provide an online tool that allows for the querying of the data they established from their mouse experiments, EnhancerControl (<http://rstats.immgen.org/EnhancerControl/>). For a gene queried against their data, the tool provides a prediction in which cell types and tissues the gene is expressed based on the open chromatin regions (OCRs) its expression is associated with. The tool also lists transcription factors binding to open chromatin regions associated with the gene based on ChIPseq. As can be seen in Figure 56, *Gpr84* in mice is mostly expressed in granulocytes, monocytes, macrophages and dendritic cells. Based on the relative proximity (~100-500bp) of two of the associated open chromatin regions to the TSS, it is likely they are within the promoter sequence, whereas the third OCR is further from the TSS (~6000 bp) and hence likely a distal enhancer. TFs found to bind in the distal open chromatin region include Klf transcription factors, which are typical for distal enhancers, Sfp1 which is the mouse homolog to SPI1, Rela, Jund, Nfe2 and Fos. In the other two open chromatin regions, Relb, Rela, Nfkb1, Gata1, Gata2, Elf1 and TCF12 have been found. These transcription factors are in line with expression in immune cells in response to inflammation.

Enrichr is an online tool that allows for finding of “knowledge” on a gene including co-expression among vast databases (ARCHS4²⁰¹) and data from cell line and mouse experiments, e.g. KD and KO of transcription factors and their effect on gene expression and vice versa²¹⁴. The knowledge obtained from this website is shown in Supplementary

Results

table 4. *GPR84* is co-expressed with several TF including *CEBPB*, *RELB*, *NFκB1* and *NFκB2*, several leucine zipper TF (*BATF*, *ATF3*, *BATF2*, *ATF5*) and zinc finger proteins (*ZNF385A*, *ZNF438*, *ZNF267*).

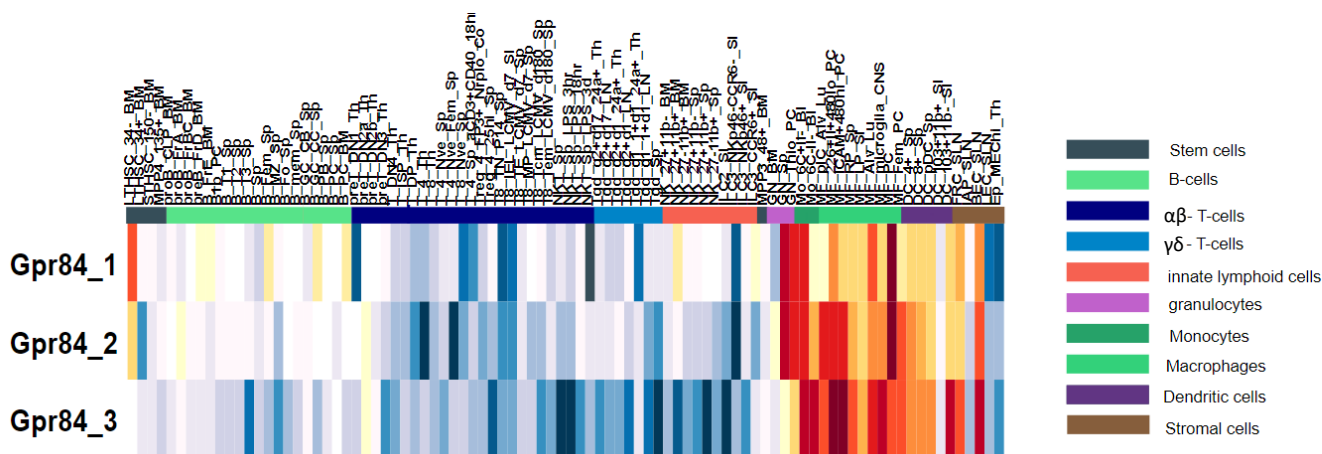


Figure 56 Cell types predicted to express *Gpr84* based on three open chromatin regions (*Gpr84_1*, *Gpr84_2*, *Gpr84_3*) associated with *Gpr84* expression in mice, as established by the EnhancerControl tool based on data from Yoshida et al.¹³⁶, legend added in Inkscape.

Results

6.2.6 Experimental data on cell-specific and temporal expression patterns

Time course of expression in several *in vitro* set-ups

Transcription factors follow a temporal hierarchy in their expression that dictates a temporal hierarchy of the expression of genes in the response to a stimulus, e.g., the inflammatory response to an infection. This is most evident in the initial hours of the response. In order to understand where in this temporal hierarchy *GPR84* is expressed, time course data was analysed. A dataset of normalized gene expression in bone marrow-derived macrophages from mice infected with murine cytomegalovirus (MCMV) or treated with IFN γ was previously generated in our research group. In these two datasets, expression was measured every 20 minutes using microarrays. To examine this in the human system, THP-1 cells were treated with LPS and with dexamethasone in a 24-hour time course measuring after 2,4,6,8 and 24 hours (Figure 57 A). Dexamethasone is thought to inhibit NF- κ B regulated genes¹⁰⁷ and is often used as sepsis treatment²¹⁵, in addition, it is a glucocorticoid, and glucocorticoids are commonly secreted in sepsis²¹⁶. Expression of *GPR84* was measured using qPCR. Another experiment was performed on neutrophils and PBMCs extracted from healthy volunteer blood using Ficoll gradients and ACK lysis of red blood cells. After extracting the cells, they were treated with LPS and dexamethasone for time course experiments with measurements at 3 hours, 6 hours in case of neutrophils and 3 hours, 6 hours and 24 hours in case of PBMCs. These experiments mostly aimed at examining the cell-specific expression of *GPR84* and give limited insight into temporal expression patterns (Figure 57 B, C).

In MCMV infection of murine bone macrophages, expression of *Gpr84* is not induced until after the first 2 hrs have passed in this model (Supplemental figure 1). After 2 hours of infection, there is a sudden increase in expression that peaks after 6 hours with a Log₂ FC of around 2.5. The gene expression then slowly decreases over time but stays on a relatively high level. In the IFN γ treated cells, there is a small increase within the first 2 hours that goes down again (Supplemental figure 1), however it is comparably small and likely insignificant. After 2 hours of IFN γ treatment, there is a strong increase in expression very similar to the increase in expression in MCMV infected cells. It is important when looking at this data to keep in mind that this experiment only had an n=1. Looking at the THP-1 cells (Figure 57 A), it is apparent that expression starts to increase after 2 hours and peaks at 4 and 8 hours in LPS-treated cells, Overall, a 2-Way ANOVA revealed an effect of both time

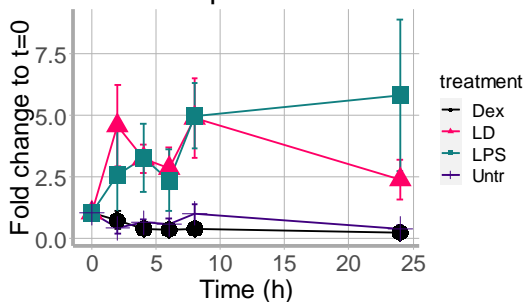
and treatment on the expression of *GPR84* in THP-1 cells. A follow-up Tukey's multiple comparisons test reveals that the difference in expression is significant at $t=4$ between dexamethasone and LPS and dexamethasone treated cells, and between all treated cells at $t=6$. In $\text{IFN}\gamma$ treated murine cells, a peak is present after around 8 hours. Adding dexamethasone in combination with LPS does not inhibit the upregulation of *GPR84* in response to LPS in THP-1 cells, as can be seen by the comparable expression levels in these cells. Dexamethasone treatment alone does not induce *GPR84* expression. This is in line with expectations based on previous studies, since *GPR84* is usually upregulated in response to inflammatory stimuli. The expression after the peak does not seem to decrease in LPS-treated cells within 24 hours but decreases in LPS and dexamethasone treated cells. In neutrophils and PBMCs, *GPR84* seems to behave similarly as in THP-1 cells. In neutrophils, a mixed effects analysis showed that the time elapsed had significant effect on *GPR84* expression ($p=0.0074$), while the treatment had a near-significant effect (mixed-effects-analysis, $p=0.06$). Visually, the biggest increase can be observed in response to LPS after 3 hours. In neutrophils that are simultaneously treated with LPS and dexamethasone, there is a trend that expression by dexamethasone treated cells is lower than in only LPS treated neutrophils. In PBMCs, a mixed-effects-model showed a significant effect of both time and treatment. A follow-up multiple-comparisons-test comparing treatments at every timepoint found significant differences between LPS and dexamethasone at $t=3$ ($p=0.0009$) and $t=6$ ($p=0.011$). After 24 hours, there was a significant difference between LPS and dexamethasone compared to LPS alone, dexamethasone alone and untreated cells ($p=0.047$, $p=0.021$, $p=0.025$). This treatment shows the highest peak in PBMCs of *GPR84* expression after 24 hours. This result is inconsistent with the expected decrease in expression over time after a stimulation.

The expression of *GPR84* in untreated cells is relatively high in cells obtained from the blood; it is at similar levels to LPS and dexamethasone treated cells. Dexamethasone treated cells show lower expression of *GPR84* which is only slightly elevated above baseline. This shows that the processing of the cells in the laboratory induced gene expression and that this gene expression is downregulated by dexamethasone. The stress that cells experience when handled, e.g. during centrifugation, can lead to the upregulation of inflammatory genes.

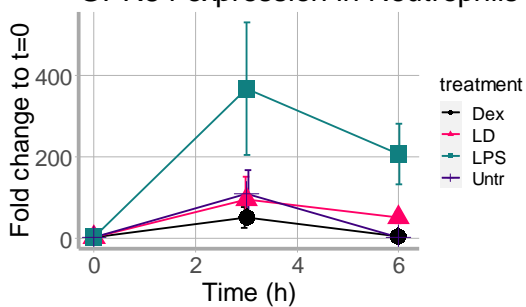
Results

The expression pattern of MCMV infected BMDMs show no upregulation of *Gpr84* before 2 hours of treatment. A classical primary response gene is upregulated within 2 hours of the stimulus, hence *GPR84* does not seem to be a classical primary response gene. The upregulation of *GPR84* in response to inflammation relatively early but not immediately after a stimulus can also be seen in THP-1 cells, as the expression peaks around 8 hours. Overall, the expression pattern seems similar to the MCMV infected BMDMs and IFN γ treated BMDMs. The time course of *GPR84* expression in PBMCs and neutrophils does not allow for any firm conclusions about the time course of expression in these cells due to the lack of granularity. The peak at the 3-hour timepoint points to an early expression of *GPR84*, though the actual peak expression time may be earlier or later, though likely before 6 hours. Also, expression in these cells starts to decrease earlier than in THP-1 cells or infected murine BMDMs. While the general expression pattern shows many similarities, the magnitude of response is bigger in both cell types obtained from volunteer blood, as the fold changes were in the tenfold (PBMCs) to hundredfold range (neutrophils), compared to a 5 times fold change in THP-1 cells and a Log₂ FC of ~ 2.7 in MCMV infected BMDMs. Since THP-1 is a cell line it is not surprising that these cells behave slightly different to blood derived cells. There are many potential explanations for this, and the only direct comparison one can fairly make is between PBMCs and neutrophils. Comparing these, it seems apparent that neutrophils express higher levels of *GPR84* than PBMCs. One of the transcription factors found to potentially regulate *GPR84* expression was NF κ B, which was also found in an experimental investigation by Nagasaki et al.¹⁰⁹. NF κ B is typically active downstream of TLRs, that sense viral and bacterial infections, e.g by binding LPS. It can also be activated as a result of IFN γ signalling. The finding in this study that LPS, IFN γ and viral infection upregulate *GPR84* would hence align with the idea of it being a NF κ B regulated gene. In addition, glucocorticoid signalling is thought to interfere with NF κ B transcriptional regulation, and in this study the glucocorticoid dexamethasone seems to downregulate *GPR84* expression.

A GPR84 expression in THP-1 cells



B GPR84 expression in Neutrophils



C GPR84 expression in PBMCs

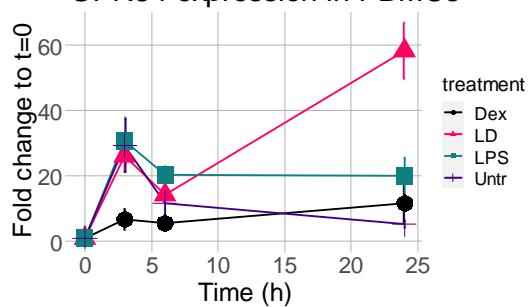


Figure 57 GPR84 expression in A) THP-1 cells, B) Neutrophils and C) PBMCs treated with either dexamethasone (Dex), LPS and dexamethasone (LD), LPS or untreated (Untr) as fold change to time-point 0 with the time in hours on the x-axis. Data from THP-1 cells was obtained in 3 separate experimental repeats (n=3), whereas Neutrophil and PBMC data originated from the cells from 4 healthy volunteers(n=4).

6.3 Discussion

Here I identified a network of TFs that likely regulate the expression of *GPR84* in sepsis. *GPR84* is a highly inducible gene upon immune stimulation, e.g., infections, LPS, or inflammatory cytokines, from a low or undetectable level of basal gene expression. In line with this and previously published experimental evidence, several transcription factors were found that act downstream of TLR activation. The temporal expression pattern of *GPR84* can be expected to be early in the infection based on the transcription factors regulating it, since several are known to be active early in the inflammatory response, e.g., NFκB, CREB, FOS and JUN. This is confirmed by previously published observations and the time course analysis of *GPR84* expression in MCMV infected murine cells. Genes expressed early in the infectious response can be divided into primary and secondary response genes. These two types differ in the transcription factors they are regulated by, their temporal expression patterns and their promoter architecture^{195,210,217}. Several features of the *GPR84* promoter architecture were elucidated. In the following I will link the transcriptional network and the architectural features of *GPR84* as presented in this chapter to propose that it has the characteristics of a secondary response gene. Further down, the implications of the transcriptional network and the regulatory regions for the expression of *GPR84* in different tissues will be discussed. The identification of immune cell specific transcription factors, e.g. SPI1 and CEBPB, indicates the expression of *GPR84* especially in cells of the innate immune system. Firstly, the method used to identify transcription factors regulating *GPR84* in sepsis will be reviewed.

Discussion of the method

Reliable identification of transcription factors regulating a gene using *in silico* methods contains several pitfalls, which requires careful consideration of methodologies and a combined approach to gain reliable results. In order to answer the question what transcription factors regulate *GPR84* in sepsis, it is not just necessary to identify transcription factors likely present and able to bind the promoter sequence, but actually likely to do so causing an effect. This requires the binding site to be accessible and often the binding of other factors to form a complex or cluster of transcription factors. Thus, in this approach emphasis was put on the identification of clusters of TFs, and the

combination of several data sources to establish binding, presence, and possible functional effect of TFs.

Analysis of TF binding to the promoter region

Transcription factors potentially able to bind the *GPR84* promoter were analysed using either Chip databases or by analysing TFBS in the promoter. Differences in TF found between Chip databases can be partially explained by the fact that ENCODE looks at 1000bp upstream of the TSS, whereas SPP looks at a region within 10kb of the TSS, hence a 10x longer stretch of DNA. With regards to the SPP method, it is very unlikely that the promoter region of *GPR84* is actually 10kb long, as the maximum length of a promoter has been calculated to be 6 kb²¹⁷. This means that the SPP database likely leads to the identification of some false positives. Another contributing factor to the differences in TF found is that most TF in ENCODE were only tested in a limited number of cell lines, primary cells and tissue, most of which were epithelial, whereas the Signalling Pathways Project contains many datasets on immune cells¹⁹⁷. *GPR84* is typically not expressed in epithelial cells, and its promoter region is hence likely not accessible for TF in epithelial cells. The use of Chip databases for identification of TFs regulating a gene limits the potential hits hence to TFs that were experimented on in the right cell type.

This constraint is not a problem when using TFBS analysis of a putative promoter sequence. For the analysis of TFBS, motifs or position-weight-matrices are queried against the promoter sequence. These motifs are saved in dedicated databases, notably JASPAR²⁰³ and TRANSFAC²¹⁸. The JASPAR database can be accessed and queried on the JASPAR website. TRANSFAC is a database that is partially commercial, the public database can be queried using the PROMO tool²¹². Since transcription factors bind only to specified nucleotide patterns; presence of such a pattern in a promoter can hence indicate the TF that regulates the gene associated with it. The pitfall to this approach is that many binding sites are short sequences that can occur by chance or these putative binding sites may not be functional in the physiological situation due to further regulatory mechanisms¹³³. It is estimated that out of 1 million TFBS identified per analysis, only 1000 are actually functional²¹⁹. There are different approaches to mitigate this high number. First of all, the analysis of promoter sequences of the same gene in evolutionary distant species to find evolutionary conserved TFBS. If a TFBS is functional and used, then it is also more likely to be evolutionary

Results

conserved. To do this, the promoter sequences can be aligned and TFBS can be queried in parts of the sequence that overlap. That is why I decided to align the mouse and human promoter and use the part of the sequence conserved in the human promoter to find TFBS. Despite the conservation of transcriptional regulation of a gene, the actual position and spacing of TFBS has been shown to be much less conserved²¹⁹. That is why it is likely that TFBS annotated on the conserved human sequence might not be found or located differently in the murine sequence. Using this approach might lower the number of false positives, but has been shown previously to still generate false positives²¹⁹.

Another potential approach to reduce the number of false positives is to analyse a set of genes that are correlating in a transcriptomic dataset. The logic behind this approach is that correlating genes are likely to be regulated by the same transcription factor; a transcription factor binding site found consistently in this gene set is hence more likely to be functional. Even though these approaches will reduce total number of false positives, it is estimated that this reduction is not sufficient to reach trustworthy results²¹⁹. In a TFBS analysis approach it is more likely to find TFBS that are short rather than longer ones, and this bias will not be overcome by analysing several promoters. A cut-off value needs to be chosen in how many promoters a TFBS has to be present to count as a hit, and this cut-off value directly affects the number of false positives versus false negatives. In addition, when TFs bind as part of a cluster, they may bind to sequences different from their typical motifs. In this analysis, several TFs were found that were not identified in other analyses of transcriptional regulation of *GPR84*. NFkB, which was mentioned in the literature, was not found when using PROMO, potentially due to its relatively long TFBS. It is likely that this approach yielded false positives as well as false negatives, based on the constraints mentioned above. Notably, when querying the same sequence used in PROMO using JASPAR, two binding sites for NFkB were annotated. This might indicate a difference in stored PWMs or a flaw in PROMO. Noticeably all three analyses yielded a fair amount of overlap. The TFBS that overlapped could be both true positives, and false positives that the analysis is skewed to find, e.g., short TFBS. In conclusion, further analysis with a different approach is warranted.

Analysis of gene regulatory regions

Transcription factors for a gene regulate its transcription by binding to its respective promoter and/or distal regulatory regions, generally called cis-regulatory modules (CRM).

Common cis-regulatory modules are enhancers and silencers²¹⁹. Looking at Figure 53 one can see that three different research consortia identified the promoter of *GPR84* with a small to moderate amount of overlap but diverge considerably on the overall length of the promoter. This is because the reliable identification of gene regulatory regions is tricky and the laboratory methods to do so have only been developed in recent years²¹⁹. The identification of promoters is easier than the identification of CRM due to the knowledge of typical motifs in promoters and their typical position just upstream of the TSS. Since promoters are always very close to the genes they regulate, it is easy to identify which promoter belongs to which gene. Several online tools use the sequence surrounding the transcription start site as putative promoter, using a size of between 1000bp up to 10 000 bp. Most TFBS will occur close to the TSS, and Hurst et al. calculated that 3000 bp away from the TSS, the frequency of TFBS equals the background frequency caused by random hits²¹⁷. They hence propose 6000 bp as maximum promoter size; it is important to note though that this is an extreme estimate. In their further computational analysis Hurst et al. use sequences up to 1000 and 2000 bp distance from the promoter. As mentioned, the promoter is characterized by several features, including an open chromatin structure and TFBS. When evaluating the sequence around the TSS predicted by EPD, a DNA stretch with DNase I hypersensitivity is apparent very close to the TSS. DNase I hypersensitivity is an indicator that the chromatin is opened, since the DNase enzyme is only able to cut open chromatin.

CRM, in contrast to promoters, can be far from the genes they regulate, up- or downstream and can even be present in exons and introns. Among the possible ways to identify CRM is the computation of enriched TFBS clusters, the experimental identification of regions with increased TF binding using Chip methods, including the identification of so-called “Highly Occupied Target” (HOT) regions, or Chip methods aimed at co-factors typically binding to CRM, and post-translational modifications of histones. To identify distal enhancer regions has been attempted among others by the Ma’Ayan group leading to the Genehancer tool, and by the ENCODE consortium^{197,220}. The enhancers they identified can be seen partially in Figure 53 1 as they are part of the UCSC genome browser. Despite the potential merit in analysing enhancer sequences due to their likely heavy involvement in regulation of *GPR84*, potential human enhancer sequences were not taken into account for further analysis. This was prompted by the general difficulties in reliably identifying enhancer regions, and

Results

linking them to a specific gene²¹⁹. Instead, EnhancerControl was chosen as a reliable source of information since it relies on extensive mouse experiments combining several approaches. The pitfall of the data is that it is murine instead of human. A compendium of data on transcriptional regulation is accessible via Enrichr. The transcription factors *GPR84* is co-expressed with among a multitude of datasets include CEBPB, which is a myeloid transcription factor, and NFkB1 and NFkB2. This is in line with previous findings as *GPR84* is upregulated by LPS, likely regulated by NFkB and is present in myeloid cells. The downside to the co-expression data is that the type of relationship and the direction of relationship is unclear. Potential explanations for a correlation between *GPR84* and a TF could be two genes being subject to the same transcriptional regulation in response to a similar stimulus, the expression in the same cell type, or an increase in TF expression due to effects downstream of *GPR84* activation. The co-expression of *GPR84* with NFkB indicates the expression in an inflammatory environment without further information on the relationship between *GPR84* and NFkB. *GPR84* is co-expressed with ZNF385A whose genomic location is in fact upstream of the *GPR84* gene; the co-expression in this case is hence very likely a co-regulation. Other data that is available in Enrichr is the difference in expression of *GPR84* in KD and KO models of genes encoding transcription factors. This is a very limited dataset which provides robust insights; it is important to keep in mind that regulatory dynamics can be complex and contain compensatory mechanisms.

Analysis of transcriptional regulation in sepsis

Focusing on the transcriptional regulation of *GPR84* in sepsis allows for the combination of several approaches in order to improve predictive strength of the results. In addition, working on a sepsis dataset means that results obtained will be directly relevant to understanding the role of *GPR84* in sepsis. As part of the response to sepsis the expression of transcription factors is strongly changed, as can be seen in Figure 46 and Figure 47. Correlations of *GPR84* with transcription factors in sepsis were analysed. In addition to the constraints on correlation analysis mentioned above, many transcription factors can be active without the need for de novo synthesis and hence transcription of the gene. As a result, an absence of correlation in gene expression does not indicate an absence of relation between the transcription factor and the gene. In fact, no correlation between TFs and *GPR84* was found strong enough to warrant any implication of a relationship and the results of this analysis were hence not further taken into account. The reason for not

finding any strong correlations could be differences in the temporal expression patterns of the genes or other confounding factors.

Chea3 was used to predict transcription factors regulating the transcriptional response to sepsis, using a list of upregulated genes in sepsis as input. Since the data that Chea3 uses is not limited to Chip data, transcriptional regulation independent of a promoter can be found by this approach. In addition, since transcription factors found are enriched based on a large set of upregulated genes, it is likely that they are functionally active. Only some of the transcription factors found in this approach are also upregulated in sepsis, indicating as well that TFs either don't require to be upregulated or might not be upregulated at time of sampling. A second list of TFs was generated based on the Chea3 output with enriched TFs by filtering for TFs with a hit for *GPR84* in at least one of the libraries Chea3 uses.

In order to identify High Confidence and Medium Confidence hits the results from previous analysis specific for sepsis were combined with the Chip results from SPP. Since Chea3 predicts TF based on, among others, ENCODE and ARCHS4 data, this data was not included again into this analysis. To summarize the combined approach, if a TF has a hit in Chea3 it is likely present and functionally active in sepsis; if it has a hit for *GPR84* as well in Chea3 it is relatively likely to regulate *GPR84*. Since Chea3 hits are not necessarily Chip hits, a hit with SPP makes it likely that the *GPR84* promoter in fact has a binding site for the transcription factor found, adding further certainty to the functional regulatory relationship of the TF with regards to *GPR84*. Despite the compounded confidence by this approach, false positives are still possible, e.g. due to the long DNA stretch taken into account in the analysis performed by SPP. If a hit is found in Chea3 without a hit for *GPR84*, it is likely a transcription factor associated with the body's response to sepsis without directly regulating *GPR84*.

To understand the putative TF network that regulates *GPR84* better, TFs were visualized and clustered using Stringdb. Since human transcription factors typically bind as hetero- or homotypic clusters, the identification of clusters adds a layer of certainty to hits found.

Annotating these clusters as promoter architecture using JASPAR to find TFBS in the putative promoter sequence prove to be a less promising endeavour than hoped. As previously mentioned, sequence-based analysis is riddled by false positives. In this approach, more than 200 TFBS were found in a 500bp sequence, many of which overlapped strongly, e.g., because many TFBS are highly similar to each other. To overcome annotating

Results

false positives, the identification of clusters was attempted. Clusters found are described in the results section. It is noteworthy that several of these clusters consist of TFs with shorter TFBS, e.g., a hit for a GATA2 binding site can consist of only 5 bp. Some cooperating TF pairs might not have been found because one or several of the TFs might typically bind in the enhancer region rather than the promoter region. An example of this is the cooperative binding of SPI1 and CEBPB that has been observed in the regulation of IL1 β expression. While SPI1 binds in this case in the promoter, CEBPB binds in a distal enhancer; their interaction is facilitated by long distance DNA looping. Finding a cluster does not necessarily indicate a cluster, due to the high amount of hits the likelihood of finding an “accidental” cluster is increased as well. A more refined approach is warranted, preferable overcoming the pitfalls of sequence-based analysis. Notably the scientists providing the JASPAR tool warn themselves against the pitfalls in using TFBS analysis with reference to a review²⁰⁴ and the recommendation to analyse only conserved sequences.

To conclude, a network of TFs consisting of several TF clusters was found that likely regulates *GPR84*. Using a combination of data sources allows for high confidence in these results, however further experimental validation is still warranted.

***GPR84*: Primary or secondary immune response gene?**

Inducible genes can be divided into primary and secondary response genes. Primary response genes are the genes that are expressed directly after a stimulus, typically within ~30 minutes of the stimulus occurring¹⁹⁵. These genes are also often called immediate early genes. These genes are defined as being expressed without *de novo* protein synthesis occurring, since they are regulated by transcription factors that are present in the cell and activated, e.g., by phosphorylation, downstream of the initial stimulus, e.g., NF κ B or CREB. Delayed primary response genes are expressed 2-4 hours after the initial stimulus without requiring protein synthesis. Primary response genes are usually mediators of the cellular response, many are transcription factors regulating the response to the stimulus, or mediators regulating the cellular response as e.g. receptors or cytokines that work in an autocrine fashion¹⁹⁵. Secondary response genes are expressed typically later than primary response genes and require *de novo* protein synthesis to be expressed. Delayed primary response genes and secondary response genes are typically effectors of the response to the stimulus. Many secondary response genes are regulated by transcription factors that are primary response genes. In summary, in the inflammatory response, the activation of

TLR receptors leads to the activation of transcription factors inducing expression of primary response genes, that in turn further regulate the inflammatory response by activating transcription of secondary response genes, which encode the effectors of the inflammatory response^{210,221}.

Primary and secondary response genes differ in their promoter architecture. The architecture of the promoter determines in part its accessibility. Features that play a role in this are CpG islands and TATA boxes. About 70% of promoters have CpG islands. Even though CpG islands theoretically allow for methylation and hence inaccessibility of a sequence, CpG islands in the genome of healthy cells are rarely methylated²¹⁰. On the contrary, CpG islands are associated with higher accessibility of the sequence since they are less likely to be stably bound by nucleosomes. CpG islands are commonly found in constitutively expressed genes but are also present in induced genes. Among induced genes, CpG island promoters are more common in primary response genes than in secondary response genes²²¹. To better understand the promoter architecture, I compared the *GPR84* promoter to the promoter of a known immediate early gene, *EGR1*. Firstly, as shown in Figure 54, part of the *GPR84* promoter sequence seems to have high accessibility since it is DNase I hypersensitive. Looking at the accessibility of the DNA around the TSS of *EGR1*, it is apparent that much longer stretches are DNase I hypersensitive. These stretches surround the gene. The DNA of the gene appears to be relatively DNase I hypersensitive too, though lesser so than the probable promoter region. Comparing CpG islands, *EGR1* has a long CpG island around its TSS, whereas *GPR84* has none. Taken together this indicates lesser accessibility to the *GPR84* promoter than to the *EGR1* promoter. Another feature of promoters of immediate early genes is the pre-loading of polymerase II, allowing for rapid start of transcription^{217,221}. Both *GPR84* and *EGR1* have Chip peaks for polymerase II in their promoters, however the peak for *EGR1* is far higher (~7 compared to ~0.2). The pre-loading of Polymerase II to promoter sequences appears to require the presence of high affinity TATA boxes.^{221,217} Concomitantly, presence of TATA boxes is more prevalent in primary than secondary response genes. TATA boxes are DNA sequences with a consensus sequence starting with TATA, followed by TATA/ AAAA, AATA, TAAA, followed by A or G²²². They typically occur ~30 bp away from the TSS and bind the transcription factor TATA-binding protein (TBP). At a p-value cut-off of 0.001, EPD does not predict a TATA box in the area around the *GPR84* TSS, but it finds two potential TATA boxes for *EGR1*. This is in line

Results

with the higher Chip peaks for polymerase II in the TSS region of *EGR1*. To conclude, *GPR84*'s promoter has the architectural features of the promoter of a secondary response gene, with its chromatin probably more inaccessible and little to none polymerase II pre-loading compared to the promoter of a primary response gene.

As mentioned before, primary response genes are regulated by TFs that do not need to be transcribed first. Among these TFs are NFkB, SRF and CREB, which were also among the hits found for *GPR84* regulation. Several of the TFs found as hits for *GPR84* however are primary response genes, including *EGR1*, *FOS* and *JUN*, and variations of *FOS* and *JUN*, and several members of the ETS-family (*ETS1*, *ELF1*, *ELF3*). This would indicate that *GPR84* is a secondary response gene. In addition, genes with a bound nucleosome at the promoter need to move the nucleosome before being transcribed. This necessity is more common in secondary than in primary genes and in low CpG promoters than CpG island promoters, as was researched by Ramirez-Carrozzi et al²¹⁰. They found that primary and secondary response genes could be further divided by their dependence on Switch/ Sucrose non Fermentable (*SWI/SNF*), a complex that moves nucleosomes, for gene expression. In an approach using KD of genes coding for proteins involved in transcriptional regulation in a cell line, coupled with LPS induction, they distinguished for a set of 67 primary and secondary response genes which ones depended on IRF3 or *SWI/SNF* presence to be expressed. Most primary response genes did not rely on *SWI/SNF*, whereas most of the secondary genes required *SWI/SNF*. Transcription factors that might be able to induce nucleosome remodelling by *SWI/SNF* include TFs of the *STAT* family and of the *IFN* family. Both of these families had several hits for *GPR84* (*IRF1*, *IRF4*, *STAT1*, *STAT2*, *STAT3*). The regulation by primary response genes is another indication that *GPR84* is a secondary response gene. It is likely that expression of *GPR84* requires nucleosome re-modelling, however this warrants further investigation.

To further examine the temporal pattern of *GPR84* expression, a transcriptomics time course dataset from murine bone marrow derived macrophages infected with murine CMV was analysed. As apparent in Figure 57, *GPR84* expression sharply increases from 2 hours on. For several primary response genes, peaks are visible within the first two hours (data not shown). I also conducted time course experiments on THP-1 cells, and PBMCs and neutrophils extracted from volunteer blood. Since these time courses had less measurement timepoints within the first few hours within stimulation, I cannot conclude

from these experiments how *GPR84* is expressed during the initial response. However, there were similarities in peak times and the temporal dynamics to the expression pattern observed in CMV infected murine macrophages.

To conclude, based on its architecture, transcriptional regulation and temporal expression pattern, *GPR84* is most likely a secondary response gene. However, this could only be proven using an experimental approach detecting if its expression is sensitive to protein synthesis inhibition.

Tissue specific expression of *GPR84*

Some TFs are active in a tissue-specific way and can lead to the tissue-specific expression of genes. These TFs can often bind to promoters and enhancers. By binding to enhancers, they change the chromatin structure to make promoters of tissue-specific genes accessible. A well-known example of this is the TF SPI1 whose expression is required for macrophage, B-cell and neutrophil differentiation. During the developmental stage of macrophages, SPI1 has been shown to mark enhancers with H3 lysine 4 monomethylation (H3K4me1), facilitating binding of p300 when cells were stimulated with LPS²²³. Binding of SPI1 has also been shown to lead to chromatin remodelling, and the finding that it leads to H3K4me1 marks in enhancers has been confirmed by Heinz et al²⁰⁹. SPI1 was consistently found to regulate *GPR84*, including according to ChIPseq data, the ChEA3 analysis, the analysis of TFBS and the analysis using EnhancerControl of TFs binding to regulatory regions in murine cells. According to EnhancerControl Sfp1, the murine homolog of SPI1, binds in the distal enhancer, which is in line with the tissue-specific expression explained above. The cell types predicted by EnhancerControl are in agreement, since they were neutrophils, macrophages and dendritic cells. EnhancerControl did not predict B-cells, which can be regulated by SPI1, or other cells of the adaptive immune system. In their paper Yoshida et al. analysed which TFs might regulate tissue specific expression. Many of the TFs found to regulate alveolar macrophages, PC macrophages, neutrophils have been found to regulate *GPR84* in sepsis. Furthermore, some of the transcription factors regulating cDC, pDC and microglia were found. When looking at the TFs binding *GPR84* regulatory regions according to EnhancerControl, in agreement with its own prediction of cell types, the TFs predicted to regulate aforementioned cell types expressing *GPR84* were found. Reason for this is probably the tendency of tissue-specific TFs to bind enhancers rather than promoters.

Results

Looking at the expression of *GPR84*, it is induced strongest in neutrophils from volunteer blood, but it is also induced in PBMCs. This is in line with findings that *GPR84* is strongly expressed in granulocytes, but also in other cells of the immune system, especially activated monocytes and macrophages^{110,100}.

Expression response to LPS and dexamethasone

Upregulation of genes by NFκB can be glucocorticoid-sensitive, however different modes of interaction between the glucocorticoid receptor and NFκB have been observed²²⁴. This is a relevant factor in sepsis, as glucocorticoid levels can be increased in sepsis patients. In addition, glucocorticoids may be used as treatment. To establish the effect of co-occurring LPS and dexamethasone on *GPR84* expression, gene expression in THP-1 cells, volunteer blood-derived neutrophils and PBMCs treated with LPS and dexamethasone was measured. In THP-1 cells and PBMCs, dexamethasone does not affect the upregulation of *GPR84* by LPS. In neutrophils upregulation of *GPR84* is affected by dexamethasone as gene expression was lower in cells treated with the combination compared to cells treated with only LPS. This is in line with the anti-inflammatory regulatory effects of glucocorticoids, potentially by inhibiting NFκB mediated activation²²⁵.

7 Biosynthesis of ligands for GPR84

In the following chapter I aim to elucidate the potential origin of the putative ligands for GPR84. As was seen in chapter 4, MCFAs are present in the blood of sepsis patients and healthy controls, with a significant increase of decanoic acid in sepsis patients. This increase raises the question where the increased amounts of decanoic acid (and octanoic acid) stem from. Testing whether there could be a systemic immune cell origin of MCFAs would enable us to better understand the MCFA-GPR84 axis in sepsis.

Potential sources for MCFA could be the diet, the intestinal microbiota, pathogens present in the blood of the patient or host cells' metabolism. A steady source like the diet or the microbiota would provide the baseline levels of MCFAs observed in both sepsis patients and healthy controls but might not explain the difference in MCFAs concentrations in healthy controls and sepsis patients. Nevertheless, in scientific articles concerning GPR84, MCFAs are often referred to as derived from the diet¹⁰¹. Potential dietary sources are palm kernel oil, coconut oil and milk and dairy products, with MCFAs most likely present as triglycerides. Measurements of decanoic acid in the plasma of fasting and non-fasting healthy volunteers showed no detection of decanoic acid in fasting individuals and only low micromolar levels (0.1-1.6 μM) in a subset of non-fasting volunteers¹⁵⁰. A study into triacylglycerol ingestion and digestion in rats as part of weaning has shown that released MCFA (C8:0-C12:0) are absorbed rapidly in the upper intestinal tract and transported to the liver²²⁶. In humans, MCFA ingestion has been deemed to be likely relatively low^{101,149}. One important exception are epilepsy patients who ingest a fat-rich diet as part of their treatment¹⁵¹. MCFA concentrations have been measured in the blood of paediatric epilepsy patients treated with a medium chain triglyceride emulsion in a 24-hour time course with regular measurements. The emulsion constituted around 50-60% of their dietary caloric intake, mainly C8:0 and C10:0 (81% and 15% of the emulsion, respectively). The increase of MCFA in the blood after a meal was marked, with peaks at up to $\sim 880 \mu\text{M}$ for C8:0 and $21 \mu\text{M}$ for C10:0 and declined overnight when no feeding occurred. These experiments show that diet derived MCFA can contribute a low baseline of MCFA concentrations in the blood with spikes related to mealtimes. It seems unlikely but not impossible that the observed increases in MCFA C8:0 and C10:0 in the sepsis patients in the ILTIS cohort would stem from the diet.

Results

Another study measuring C6:0-C12:0 in plasma of fasting individuals that were either healthy volunteers, suffered from ulcerative colitis, or had cancer or adenoma, found low baseline MCFA levels in healthy sera that were elevated in colorectal and/ or high-grade dysplasia adenoma, e.g. to around 3 μM for C10:0¹⁷⁸. The MCFA decanoic acid allowed for discrimination of colorectal cancer from ulcerative colitis, breast cancer and healthy controls with a sensitivity of >80%. Furthermore, dodecanoic acid was significantly elevated to ~6 μM in the sera of high-grade dysplasia adenoma cases compared to healthy controls, breast cancer patients and colorectal cancer patients. This illustrates that either the malignant process in colorectal cancer or cancer associated inflammation or, alterations to the microbiome might affect MCFA levels. Regardless, it sets a precedent of increased MCFA concentrations in relation to the health status. In addition, early studies in mouse and rat models of adenocarcinoma found MCFA production by grafted tumours in connection with expression of thioesterase II or *OLAH*²²⁷. This would support the notion that MCFA production in cancer may also be derived from pathophysiological human tissue.

How do human cells produce fatty acids? Human cells synthesize LCFA in the cytosol by repeated elongation of a shorter fatty acid; in this process an elongation cycle that adds 2 carbons to the initial fatty acid is repeated until the fatty acid has the required length. It is released from the cycle by an enzyme that cleaves the synthesized fatty acid off the carrier molecule. The enzymes responsible for this operate in a chain-length specific manner, i.e., different enzymes are responsible for hydrolysis of fatty acids of different chain length ranges. The hydrolysis enzyme specific for MCFA is called oleoyl-ACP-hydrolase (OLAH) or thioesterase 2 (TE2). Previously, its expression has been documented to be specific to the lactation process with a localised production in mammary tissue that is only switched on after delivery of the baby²²⁸. The *de novo* synthesis of medium chain fatty acids by mammary tissue has been proven by the measurement of deuterated MCFAs in breast milk from women ingesting deuterated compounds^{228,229}. The high ratio of MCFA in breast milk to sera concentrations illustrates the strict localisation of the synthesis process and secretion during lactation^{230,228}. The concentration of MCFA C8:0, C10:0 and C12:0 in breast milk is relatively high, ~14%²³¹ and 11.2%²³² of total lipids per weight in studies of fatty acids in breast milk. Another study evaluated medium-chain triglycerides to make up 12 to 16% of total fat in breast milk, depending on the diet of the lactating mothers²³³.

Epithelial cells extracted from breast milk were grown *in vitro* supplemented with radioactively labelled acetate leading to the production of labelled lipids with a chain length shorter than 16 carbons²³². Of total moles of lipids produced by cells from breast milk (i.e., labelled lipids present in cells after incubation with labelled acetate), 3.6% were C8:0, 11.9% were C10:0 and 35% were C12:0 proving that MCFA in breast milk are *de novo* synthesized by epithelial cells. This opens the question if the enzyme OLAH could be present in other cell types that might contribute toward MCFA levels in sepsis, and if the observed concentrations are hence resulting from immune cell MCFA synthesis.

Other potential MCFA sources have been sparsely investigated. I was not able to find any studies that would confirm or even test the production of medium-chain fatty acids by the gut microbiota. A study utilising untargeted metabolomics analysing the colonic luminal metabolome of matched germ free and conventional mice sought to identify and distinguish bacterial and host metabolites. Among the metabolites identified were several lipids, including hexanoic acid and PGE₂, but no medium-chain fatty acids with a chain length between C8 and C12 were reported²³⁴. There is evidence however that bacteria should be able to produce medium-chain fatty acids, and since medium-chain fatty acids can be a part of metabolism or bacterial cells, it is feasible that both commensal and pathogenic bacteria could produce medium chain fatty acids^{98,235,236}. First off, hydroxylated MCFA have been shown to be contained in LPS preparations and in the LPS of specific, albeit not all, gram-negative bacteria^{235,237}. A paper by Kutschera et al. investigating a plant receptor binding hydroxylated MCFA found activation of the receptor by some, but not all LPS preparations and confirmed the presence of hydroxylated MCFA in those LPS preparations that caused a response⁹⁸. This is explained by the fact that the lipid component of LPS contains three fatty acids whose chain length and hydroxylation are species-specific, and hence some species but not all harbour MCFAs in their LPS. The production of MCFAs by bacteria is also of industrial interest as a cheaper method of manufacturing these lipids²³⁶. All in all, both commensal and pathogenic bacteria appear to be potential sources of MCFA; however, the extent of bacterial MCFA production both in quantity and commonality, remains unknown. As a result, current knowledge does not allow for speculation if bacteria derived MCFAs might play a role in activating the receptor or not in a physiological setting. But this does not exclude that possibility.

Results

Another way how host cells could contribute to the MCFA levels in the blood is the release of lipids by adipocytes. However, the human body typically stores fat as esterified LCFA²³⁸. This possibility remains open until tested.

In this chapter I performed both experimental work and *in silico* analysis focusing on the potential production of MCFA by cellular metabolism. I sought to examine the expression of OLAH in different conditions and cells and its transcriptional regulation *in silico* with follow-up experiments *in vitro* into expression of OLAH and MCFA production.

7.1 Results

7.1.1 Systemic expression of OLAH in paediatric and neonatal sepsis patients.

Since the enzyme OLAH is pivotal in MCFA synthesis, the expression of its gene in sepsis was analysed in two transcriptomic datasets (neonatal sepsis: GSE25504³², paediatric sepsis: GSE26378³¹). These are the same datasets previously used in chapter 3. The neonatal sepsis dataset consists of gene expression data of 35 control neonates and 26 infected neonates. The paediatric sepsis dataset studied gene expression in a cohort of neonates, infants, toddlers and children up to the age of 10; 180 were patients diagnosed with sepsis and 52 were controls. The expression of *OLAH* is significantly upregulated in both paediatric sepsis and neonatal sepsis (Figure 58, Kruskal-Wallis-test, $p=6.7^{-11}$, $p=9.6^{-6}$ respectively).

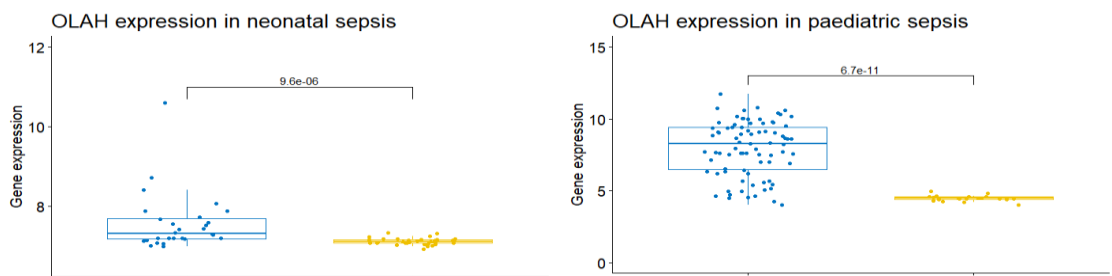


Figure 58 OLAH expression in neonatal sepsis (A) and paediatric sepsis (B). The gene expression was measured by microarrays, data is RMA-normalized. Significance was calculated using the Kruskal-Wallis test.

Results

7.1.2 The promoter region of *OLAH*

The promoter region of a gene can give insights into the expression pattern of the gene. Exploring the promoter region of *OLAH* in the UCSC genome browser, one can see that the region upstream of the transcription start site has no CpG island and only very short stretches that are DNase I hypersensitive. Furthermore, none of the transcription factor ChIPSeq clusters queried: neither Pol2 in CD4+ cells nor for the set of transcription factors and cell types queried by ENCODE3 (Figure 59, also accessible under the link: <https://genome-euro.ucsc.edu/s/LindaMoet/OLAH>). Exploring in which cell types the DNase I hypersensitive stretches were found for *OLAH*, it is evident that they were foremost detected in epithelial, fibroblasts, endothelial cells and epithelial cells of different tissues. When a promoter region is closed (transcriptionally repressed state), the nucleosome may be methylated. Depending on which amino acids in histones are methylated different accessibility of the promoter region can arise. An H3K4me3 modification (trimethylation of histone 3 at lysine position 4) would increase accessibility of the promoter, yet this mark is also not detected near the *OLAH* gene. This indicates that the promoter is typically closed, and that gene transcription requires the opening of the promoter region.

7.1.3 Enhancers in the transcriptional regulation of *OLAH*

Since the promoter of *OLAH* appears to be quite closed, it is interesting to analyse the role of any enhancers on *OLAH* expression. To this end the FANTOM5 database was used that predicts enhancers regulating a gene based on CAGE data. The expression of the enhancers in different cell types is presented and gives an indication in which cells the gene regulated by the enhancer is expressed. For *OLAH*, two enhancers are predicted; the associated tables of cell-specific expression are shown in Table 13 and Table 14. This analysis indicates the expression of *OLAH* in different epithelial cells, immune cells including monocytes, neutrophils, dendritic cells, T-cells and B-lymphocytes.

7.1.4 Cell specific expression of *OLAH* in sepsis

Since the expression of *OLAH* is seen in datasets coming from transcriptomic data from blood samples, it is a feasible assumption that cells present in the blood express *OLAH*. Which cells express *OLAH* is both relevant for further experimental design and to understand the metabolic response in sepsis. Data available from the EBI gene expression

atlas²³⁹ indicates a differential expression of *OLAH* in monocytes, neutrophils, and to a lesser extent lymphocytes. This data comes from 2 separate experiments in which cells of sepsis patients were sorted and gene expression was measured. Gene expression was compared to sorted cells of healthy volunteer blood in both studies. One study reported a whole blood FC in *OLAH* expression in meningococcal sepsis compared to healthy controls of 5.1, while another adult sepsis study reported a FC of 6.9. The former study reported a FC of 5 in monocytes at time point zero and a FC of 4.4 after 8 hours. For lymphocytes a fold change of 2.2 after 8 hrs was reported as earliest available time point. The other sepsis study reported a FC of 4.5 in neutrophils and no fold change in monocytes, CD8+ cells and CD4+ cells. All FC values were Log₂-normalized. Due to different gating strategies, not all cell types were included in both studies. These results indicate that the fold change in *OLAH* expression in sepsis is higher in monocytes and neutrophils than in lymphocytes. For further studies into lipid production and *OLAH* expression, the monocyte-like cell line THP-1 was chosen, and immune cells obtained from healthy volunteers. Also, further research into cell-specific expression of *OLAH* was required, whose results are shown further below.

Results

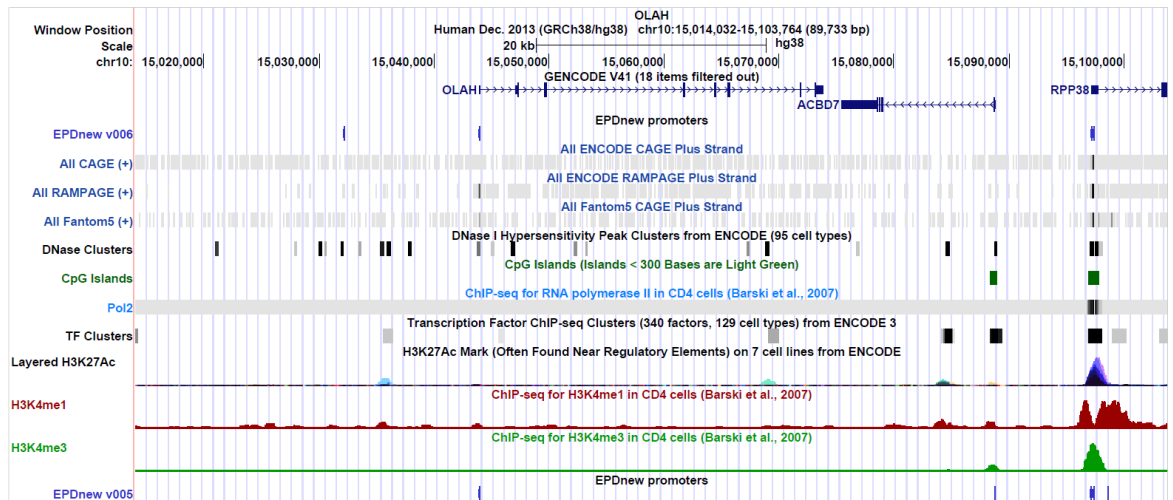


Figure 59 The genomic region around OLAH as shown in the UCSC genome browser. On the left the names of the displayed tracks are shown. Every track indicates a different genomic feature. The region coding for the OLAH gene is indicated in the top track called GENCODE V41. The arrows indicate the reading direction of the gene. EPDnew v006 is a database for experimentally validated promoters based on transcription start site mapping from experimental high-throughput data. RAMPAGE indicates the transcription start site based on cDNA sequencing data. Fantom5 also predicts, in this case, the transcription start site.

Table 13 Expression of the first enhancers found to be likely regulating OLAH according to FANTOM5. The table shows the expression of the enhancers in different cell types, both as percentage of total expression of the enhancers and as tags per million.










































Name	Percentage Bar	Percentage of Expression	Tags Per Million
hepatocyte		15.50%	0.423
lymphocyte of B lineage		13.92%	0.379
neutrophil		9.16%	0.250
fibroblast of tunica adventitia of artery		8.83%	0.241
acinar cell		6.32%	0.172
mesothelial cell		6.13%	0.167
tendon cell		5.65%	0.154
dendritic cell		4.90%	0.133
neuronal stem cell		4.66%	0.127
fibroblast of pulmonary artery		4.16%	0.113
mast cell		3.37%	0.092
retinal pigment epithelial cell		3.14%	0.085
T cell		2.70%	0.074
epithelial cell of prostate		1.48%	0.040
skeletal muscle cell		1.16%	0.032
monocyte		1.10%	0.030
mammary epithelial cell		1.09%	0.030
ciliated epithelial cell		1.07%	0.029
chondrocyte		1.06%	0.029
fibroblast of periodontium		0.92%	0.025
sensory epithelial cell		0.86%	0.023
astrocyte		0.77%	0.021
skin fibroblast		0.62%	0.017
mesenchymal cell		0.54%	0.015
preadipocyte		0.54%	0.015
respiratory epithelial cell		0.37%	0.010

Table 14 Expression of the second enhancer found to be likely regulating OLAH according to FANTOM5. The table shows the expression of the enhancers in different cell types, both as percentage of total expression of the enhancers and as tags per million.

Name	Percentage Bar	Significantly Overrepresented	
		Percentage of Expression	Tags Per Million
neuronal stem cell		40.08%	0.651
osteoblast		12.75%	0.207
lymphocyte of B lineage		12.06%	0.196
T cell		7.46%	0.121
reticulocyte		6.75%	0.110
gingival epithelial cell		5.60%	0.091
mast cell		4.33%	0.070
enteric smooth muscle cell		2.42%	0.039
lens epithelial cell		2.38%	0.039
ciliated epithelial cell		1.80%	0.029
vascular associated smooth muscle cell		1.62%	0.026
dendritic cell		0.91%	0.015
preadipocyte		0.86%	0.014
respiratory epithelial cell		0.59%	0.010
blood vessel endothelial cell		0.39%	0.006

Results

7.1.5 Transcriptional regulation of *OLAH* in sepsis

The transcriptional factors regulating *OLAH* expression were analysed in order to understand its function and expression better. Since *OLAH* is upregulated significantly in sepsis, it is logical that transcription factors regulating *OLAH* are activated in sepsis. The approach chosen largely followed the approach described for the analysis of *GPR84* expression. Transcription factors enriched in sepsis were computed using the online tool ChEA3. Additionally, transcription factors likely to bind the *OLAH* promoter were extracted from the signalling pathways web tool¹⁹⁶(<http://www.signalingpathways.org/ominer/query.jsf>). The overlap of these two tools contains the transcription factors likely to regulate *OLAH* in sepsis as they can likely bind the promoter and are also likely to be present. The result is visualized as a network generated by STRING in Figure 60. In contrast to the previous analysis of *GPR84* expression, TFs enriched in sepsis were not filtered for the gene of interest. This is due to a very low number of transcription factors having a hit in ChEA3 for *OLAH*. Among possible explanations for this are firstly a transcriptional regulation of *OLAH* distinct from most upregulated genes in sepsis or secondly a transcriptional regulation pattern hard to detect in the studies that provide data to the ChEA3 database. Using this approach, several hormone receptors are found including the androgen receptor (AR), progesterone receptor (PGR) and the glucocorticoid receptor (GR/NR3C1). These hormone receptors function as transcription factors when bound by their respective hormones. Furthermore, transcription factors that play a role in the immune response were found, including CREB1, IRF1 and IRF4, NFKB and RELA¹³⁶. These transcription factors would be active downstream of, for example, TLR signalling or, TNF- α (NFKB and RELA) or, interferon signalling (STAT transcription factors, IRF1, IRF4). The transcription factors PPARG, RXRA and RARA are also found to interact with the *OLAH* promoter and are expected regulators of lipid metabolism²⁴⁰. The results found seem to align with the differential gene expression of *OLAH* as reported by the differential gene expression from the EBI gene expression atlas. The tool reports differences in expression in testis in different conditions, upregulated expression in response to the glucocorticoid dexamethasone in various experiments, and differential, increased gene expression in infectious diseases including previously mentioned sepsis studies, COVID-19, and bacteraemia. Table 15 shows the top 25 differentially expression experiments or datasets in which *OLAH* expression was the most

changed. Furthermore, a paper analysing the effects of glucocorticoid medication in cancer using two datasets found *OLAH* to be significantly upregulated in the glucocorticoid treated group in a prostate cancer cohort²⁴¹. Some of the transcription factors found are typical for innate immune myeloid cells, i.e. SPI1, CEBPB, CEBPA¹³⁶, which would be in line with the increase in systemic expression observed in sepsis patients and the increased expression found in sorted immune cells of sepsis patients. Hence, despite being less rigorous in the analysis approach than previously, the results of the transcriptional regulation analysis seem biologically indicative of the signalling of immune cell transcriptional regulation. I used these results as basis for testing hypotheses about *OLAH* expression. Specifically, I wanted to examine whether *OLAH* is upregulated in response to inflammatory stimuli and glucocorticoids on immune cells.

I first sought to identify if glucocorticoid treatment in sepsis leads to higher *OLAH* expression since glucocorticoids are used as a treatment in sepsis²⁰. The paediatric dataset generated by the Hector Wong group used previously in chapter 4 contains the information on the use of glucocorticoids within the cohort²⁴². In this cohort, *OLAH* expression is only slightly and insignificantly increased in glucocorticoid treated patients compared to patients not treated with glucocorticoids (Figure 61). This indicates that *OLAH* expression is not the result of glucocorticoid medication. However, the presence of endogenous glucocorticoids in patients' blood was not measured, and glucocorticoids can be released by the body itself during the immune response¹. Hence, the induction of *OLAH* by endogenous glucocorticoids is still feasible as potential mechanism regulating its expression²¹⁵.

Next, I wanted to look at the role of inflammation and infection in regulating *OLAH* expression in immune cells in other independent studies. I identified two datasets containing different immune cells stimulated with different stimuli. The first transcriptomics dataset, GSE46903, is focused on macrophages but also includes monocytes, B-cells, dendritic cells, T-cells and NK-cells²⁴³. In this dataset, macrophages treated with various inflammatory stimuli (IFN γ , TNF α -PGE₂-P3C, IL-4, LPS) did not upregulate *OLAH* (Figure 64). T-cells showed a moderate increase in *OLAH* expression in response to CD3-CD28 stimulation, whereas dendritic cells did not show *OLAH* upregulation in response to LPS. NK-cells, B-cells and monocytes were not stimulated in this dataset and show a similar baseline *OLAH* expression as macrophages and T-cells (Figure 64). Another

Results

dataset, GSE17762, explores the transcriptomic response to IFN γ by B-cells, monocytes, NK cells, CD8 positive cells and CD4 positive cells (Figure 63), and the effect of different interferons (IFN γ , IFN α , IFN β , IFN δ , IL-12, TNF α) on PBMCs isolated from human blood (Figure 62)²⁴⁴. In this dataset, no increased expression of *OLAH* was found in response to any of the interferons in any of the cell types. This shows that *OLAH* very likely is not an interferon or NF κ B regulated gene.

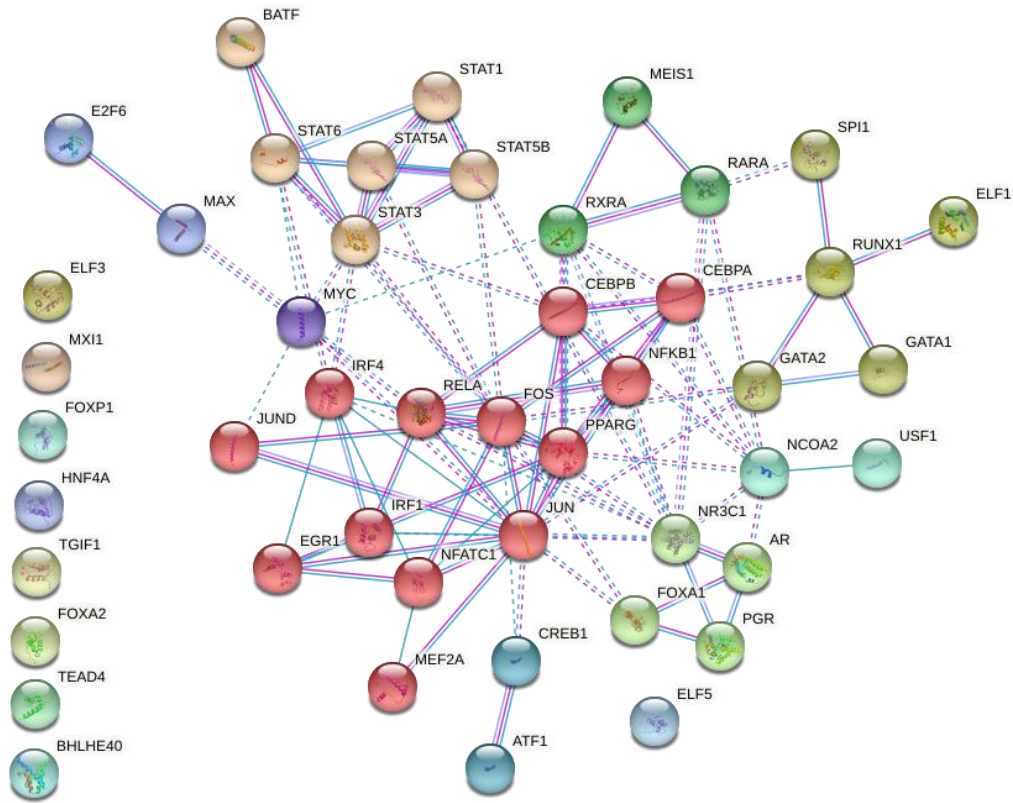


Figure 60 TF network likely regulating OLAH expression in sepsis, generated using STRING. The confidence of interaction between proteins was set to a minimum of 0.6, the colours of the connecting lines indicate the type of evidence for interaction. Turquoise indicates the interaction stems from curated databases; purple connections indicate experimentally determined interactions. The interaction indicates a shared function rather than a physical interaction. Transcription factors were clustered using a MCL inflation parameter of 3.

Results

Table 15 shows the 25 (experimental) conditions in which OLAH expression is the most affected. The data is obtained from the ebi gene expression atlas, accessed 24/10/2022.

Comparison	Experiment name	log ₂ fold change
'dexamethasone; 1 micromolar' vs 'vehicle'	RNA-seq in GR18 cell line to identify genes regu	8.6
'sepsis' vs 'normal' in 'whole blood'	Next generation sequencing of human immune	6.9
'ejaculatory azoospermia; Johnsen score count 5' vs 'normal; Johnsen score count 10'	Transcription profiling of human testis samples	-5.1
'meningococcal sepsis; blood' at '24 hour' vs 'normal; blood' at '0 hour'	Transcription profiling by array of human peripl	5.1
'valproic acid; 15 millimolar' vs 'control' at '2 day'	RNA-seq of Primary Human Hepatocytes (PHH)	5
'monocyte; meningococcal sepsis' at '0 hour' vs 'monocyte; normal' at '0 hour'	Transcription profiling by array of human peripl	5
'untreated; air-liquid interface (day 14)' vs 'untreated; submerged (day 8)'	Genome-wide expression profiling of an in vitro	4.7
'TGFbeta3' vs 'control' at '7 day'	Transcription profiling by array of human bone	4.7
'ejaculatory azoospermia; Johnsen score count 3.2' vs 'normal; Johnsen score count 10'	Transcription profiling of human testis samples	-4.6
'dexamethasone; 100 nanomolar' vs 'ethanol' at '24 hour'	Transcription profiling of human acute lymphob	4.5
'sepsis' vs 'normal' in 'neutrophils'	Next generation sequencing of human immune	4.5
'monocyte; meningococcal sepsis' at '8 hour' vs 'monocyte; normal' at '0 hour'	Transcription profiling by array of human peripl	4.4
'valproic acid; 15 millimolar' vs 'control' at '3 day'	RNA-seq of Primary Human Hepatocytes (PHH)	4.3
'wild type; 4-hydroxytamoxifen; senescent cell' vs 'wild type non-senescent cell'	Analysis of the senescent transcriptome upon e	4.3
'interleukin 13; 100 nanogram per milliliter; air-liquid interface (day 14)' vs 'untreated; submerged (day 8)'	Genome-wide expression profiling of an in vitro	4.3
'dexamethasone; 1 micromolar' vs 'ethanol'	RNA-seq in GR18 cell line to identify genes regu	4.2
'breast carcinoma' vs 'normal'	RNA-seq of blood platelets from six tumor type	-4.2
'ejaculatory azoospermia; Johnsen score count 2' vs 'normal; Johnsen score count 10'	Transcription profiling of human testis samples	-4.2
'CD3/CD28 beads' vs 'none' in 'none'	RNA-Seq of T cells sorted from human PBMCs,	4.1
'squamous cell carcinoma' vs 'normal'	RNA-seq of a panel of non-melanoma skin cance	4
'wild type ZFP36L1; 4-hydroxytamoxifen; senescent cell' vs 'wild type non-senescent cell'	Analysis of the senescent transcriptome upon e	3.9
'12 hour' vs '0 hour' in 'MCF-10A-H1047R'	Knock-in of PIK3CA-H1047R into MCF-10A	-3.9
'24 hour' vs '0 hour' in 'MCF-10A-H1047R'	Knock-in of PIK3CA-H1047R into MCF-10A	-3.8
'died; Pneumococcal Meningitis and HIV negative' vs 'alive; normal and HIV negative'	Transcription profiling by array of children bloo	3.8

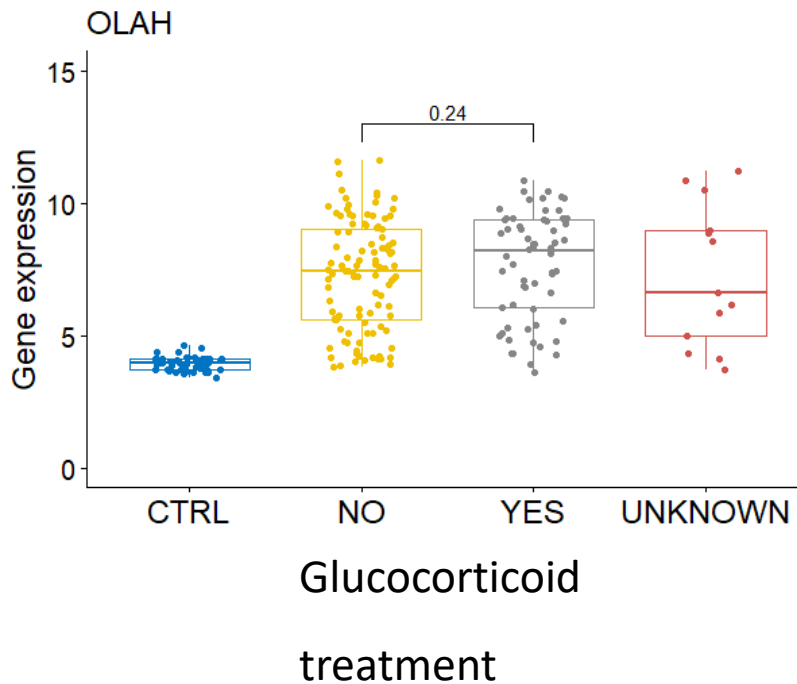


Figure 61 The OLAH expression in healthy controls (CTRL), corticosteroid treated sepsis patients (YES) and patients not treated with corticosteroid treatments (NO). For a subset of patients, it was unknown if they were treated with corticosteroids (UNKNOWN). The gene expression is RMA-normalized and log-transformed. The lower half of each box in the boxplot depicts the 2nd quartile of values, with the upper half depicting the third quartile. The extended lines depict the first and last quartile of values, with any points reaching below or above these lines being statistical outliers. The statistical test performed was a Kruskal-Wallis-test.

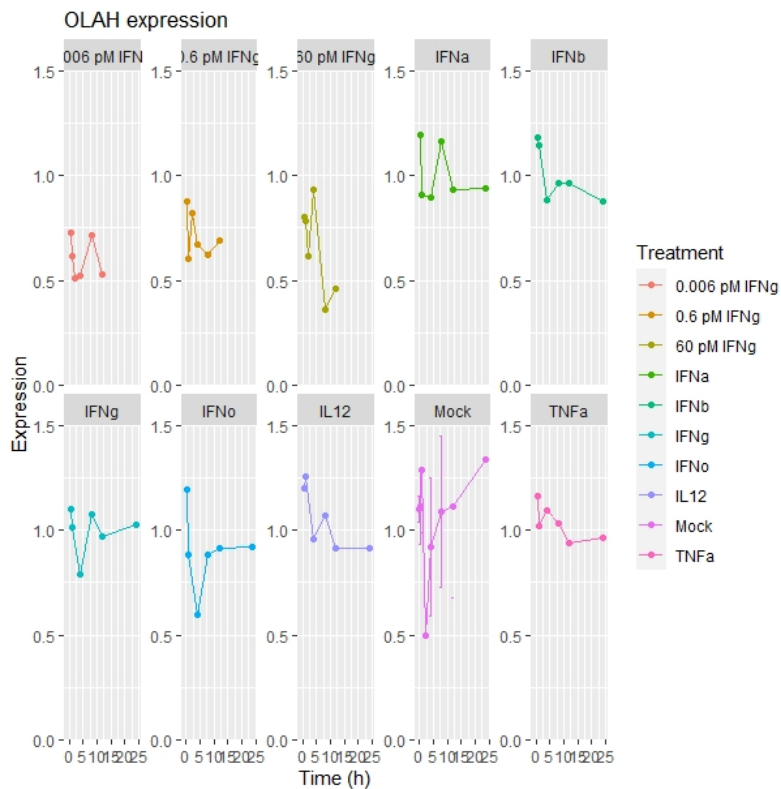


Figure 62 The expression of OLAH in PBMCs treated with different cytokines. The data was obtained from a public repository and stems from an experiment on blood derived PBMCs from healthy volunteers²⁴⁴. The error bars represent the standard deviation. Apart from time-point 0, there was only one replicate per timepoint and treatment. Gene expression is RMA-normalized and log-transformed.

Results

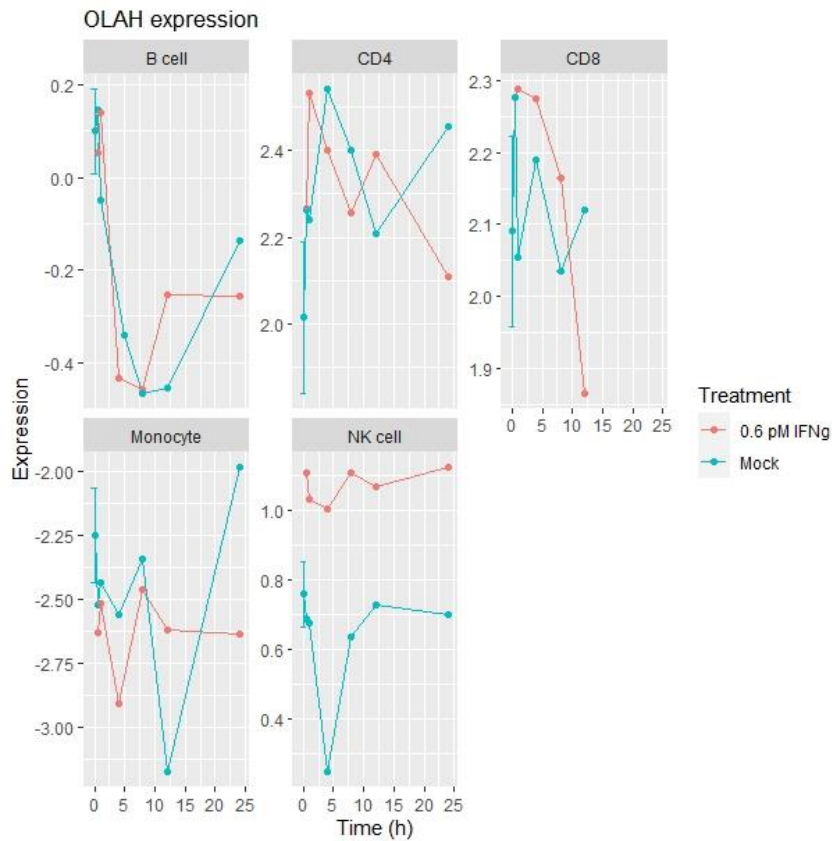


Figure 63 The expression of OLAH in different cell types (B-cell, CD4+ T-cell, CD8+ T-cell, Monocyte, NK cell) in a time course with either mock treatment or 0.6 pM IFN γ . The cells were obtained from healthy volunteer blood²⁴³. The error bars represent the standard deviation. Apart from time-point 0, there was only one replicate per timepoint and treatment. Gene expression is RMA-normalized and log-transformed.

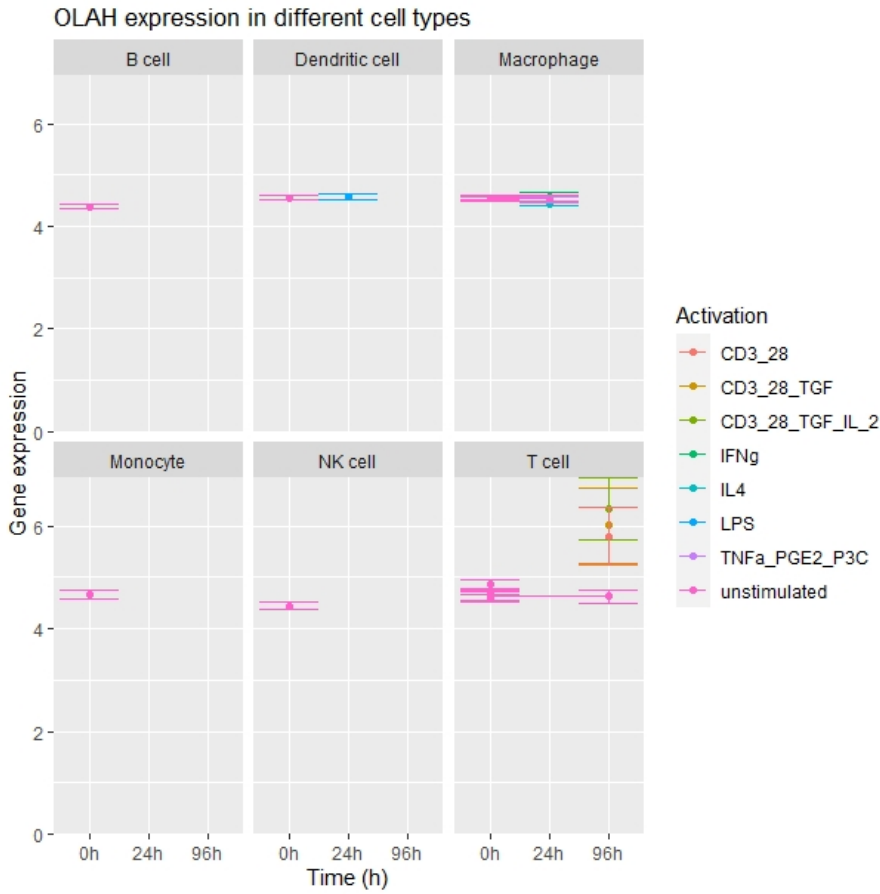


Figure 64 The expression of OLAH in different cell types at 0h, 24h and 96h after treatment with different activating compounds²⁴⁴. Cells were obtained from healthy volunteer blood. The error bars represent the standard deviation. Gene expression is RMA-normalized and log-transformed.

Results

7.1.6 Induction of *OLAH* expression in the THP-1 cell line

To examine if *OLAH* expression is induced by glucocorticoids and how this is affected by inflammatory stimuli, THP-1 cells were treated with LPS (100ng/ml), with dexamethasone (10 μ M) and with a combination of the two (100ng/ml LPS, 10 μ M dexamethasone) for a time course of up to 24 hours. The RPMI medium contained charcoal-treated foetal calf serum, a treatment intended to remove all lipids in the serum. Cells were harvested after each timepoint for RNA extraction prior to gene expression analysis, and the data was collected from 3 repeat experiments. THP-1 cells are a monocytic-like cell line derived from the blood from a patient suffering from acute monocytic leukaemia²⁴⁵. Dexamethasone is a glucocorticoid used to treat inflammatory conditions that binds the glucocorticoid receptor. Gene expression was measured after t=2, 4, 6, 8, 24 and compared to the baseline at t=0. The results are depicted in Figure 65. Untreated cells were used as control for every timepoint to confirm that gene expression was not affected by the handling of the cells and any vehicles used. One can see for this experiment that there were no gene expression changes in the untreated cells. The strongest increase in expression was observed in cells treated with both dexamethasone and LPS (up to 7.8 fold change), followed by cells treated with dexamethasone alone (5 fold change). Treatment with LPS alone did not lead to any changes in gene expression. It is evident that gene expression changed quickly as expression is increased after 2 hours of stimulation in both dexamethasone and LPS and dexamethasone treated cells. Expression peaks 8 hours after stimulation and remains above baseline after 24 hours, especially for cells treated with both LPS and dexamethasone. Since *OLAH* expression was only induced in treatments containing dexamethasone, it seems that this is the decisive factor. In other words, *OLAH* is likely upregulated in response to dexamethasone.

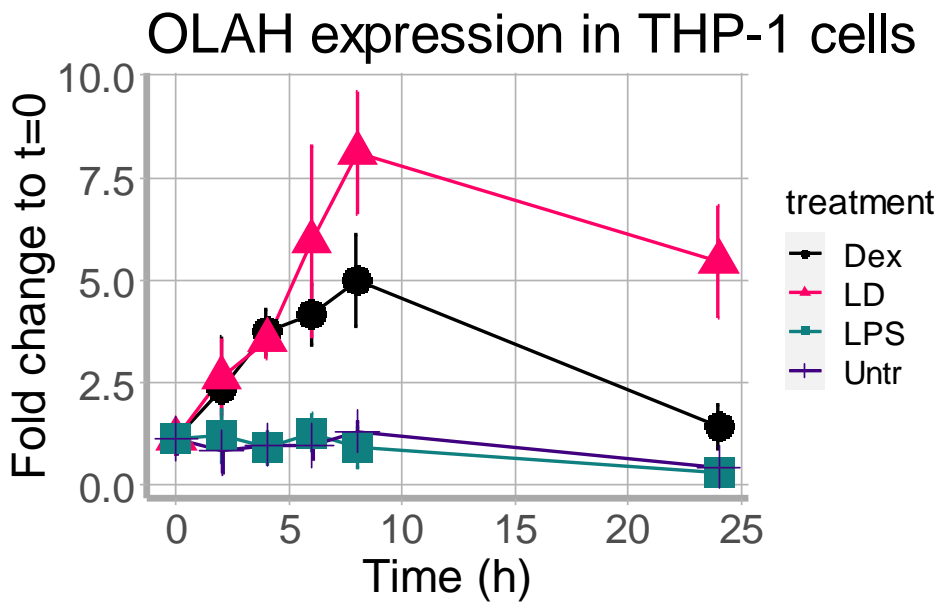


Figure 65 OLAHA expression in THP-1 cells after treatment with different stimuli: Dexamethasone (Dex), LPS and dexamethasone (LD), LPS and no stimuli/untreated (Untr). Depicted is the time course of expression with time in hours on the x-axis and the fold change compared to the baseline expression at t=0 on the y-axis. The expression data was collected in three separate experimental repeats. The error bars depict the SEM.

Results

7.1.7 MCFA production by THP-1 cells

Since *OLAH* was upregulated in THP-1 cells I hypothesized this would lead to MCFA production by these cells. To examine lipid production in THP-1 cells supernatant was harvested during the previously described time course experiment and prepared for lipid measurement using the previous established method for lipid measurement in whole blood.

To this end, THP-1 cells were induced with LPS, dexamethasone and both, as previously. In initial experiment lipid concentrations were measured in the supernatant after 24 and 48 hours. The concentrations of C8:0 and C10:0 were detectable in the supernatant, however they were not in the quantifiable range. Lipids were also sampled in the repeated time course that was conducted for gene expression. Again, C8:0 and C10:0 measurements were not in the quantifiable range. C12:0 concentrations were at the lower end of the quantifiable range in 2 of 3 experimental repeats. The background of C12:0 as measured in extracted blanks was relatively high, albeit still lower for several samples, than the concentrations measured in samples. There was a high variability of measurements between the experimental repeats and hence also no statistical significance (data not shown). As a consequence, THP-1 cells were not further used as model cells for lipid metabolism, and further experiments into lipid production by immune cells were conducted with primary blood cells extracted from healthy volunteers.

7.1.8 *OLAH* expression in endothelial cells and MCFA production

Endothelial cells line the vasculature and hence can release molecules into the bloodstream. These cells are epithelial cells, and as discussed previously mammary tissue which produces MCFAs during lactation utilise *OLAH* present in epithelial cells. I hence decided to examine *OLAH* expression and MCFA production by endothelial cells. To this end, human umbilical vein endothelial cells (HUVEC) were obtained from the Cardiff Cell Bank and treated in a time course experiment similar to the previous experiment. Cells were seeded into plates and treated with LPS, dexamethasone or both in combination for up to 48 hours. The RNA was sampled, and gene expression was analysed using qPCR. The strongest upregulation was found in dexamethasone treated cells, with a modest upregulation of a fold change of 2 compared to baseline levels. Simultaneously to sampling RNA, cell pellets and supernatant were sampled for lipid analysis. To this end supernatant was sampled first, then the remaining supernatant was removed, and the cells were scraped in 50 µl HPLC water. Lipids were extracted and quantified using the previously established method. No MCFAs could be detected in the linear range and/or above background in the supernatant of these incubations. Hence, no further experiments into *OLAH* expression and MCFA production by endothelial cells were conducted. Since this experiment was only conducted in two experimental repeats, this warrants caution.

Results

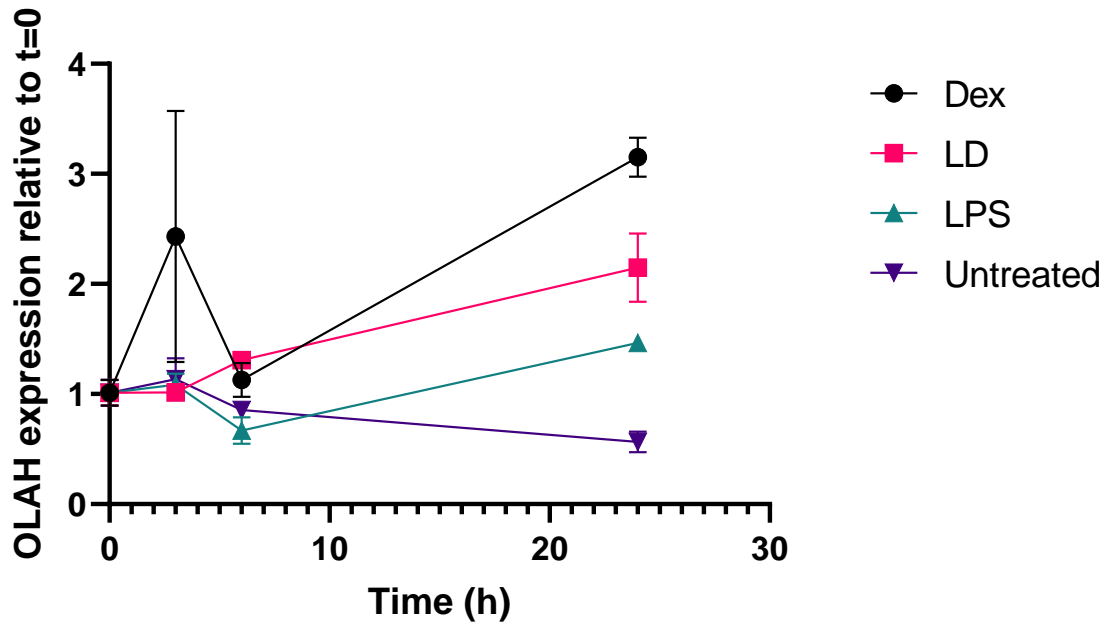


Figure 66 The OLAH expression in endothelial cells shows as the fold change compared to the expression at timepoint 0. The x-axis depicts the time in hours. This data comes from 2 experimental repeats.

7.1.9 *OLAH* expression blood cells and MCFA production

Since results from my previous analysis indicated that both neutrophils, monocytes and T-cells (lesser extend) express *OLAH* in sepsis, I decided to perform experiments on neutrophils and PBMCs extracted from healthy volunteer blood. The obtained cells were transferred into cell medium and were immediately treated with either LPS, dexamethasone or a combination of both, and samples were taken after 3 and 6 hours from treated neutrophils and 3 hours, 6 hours and 24 hours for PBMCs.

These timepoints were chosen due to the limited lifespan of neutrophils and PBMCs after extraction from the blood. At every sampling time point, samples were taken for RNA extraction and for lipid extraction from cell pellets and supernatant. As can be seen in Figure 67 A,B, *OLAH* expression increased strongly in several of the treatments. In neutrophils, the strongest increase was in the dexamethasone treated cells, followed by cells treated with both LPS and dexamethasone. The increase in *OLAH* in LPS treated neutrophils was comparably negligible. The *OLAH* expression after 6 hours is significantly higher in dexamethasone treated cells compared to LPS and untreated cells (Kruskal-Wallis-Test with follow-up Dunn's multiple comparisons test, all p-values adjusted: Dex vs LPS $p=0.0159$, Dex vs LD $p=0.325$, Dex vs Untr $p=0.0389$). In PBMCs, the strongest increase was also seen in dexamethasone treated cells followed by cells treated with both LPS and dexamethasone. The difference between untreated cells and stimulated cells is not significant after 24 hours (Kruskal-Wallis-Test with follow-up Dunn's multiple comparisons test, all p-values adjusted: Untr vs Dex $p=0.0843$, Untr vs LD $p=0.843$, Untr vs LPS $p=0.99$). The expression was highest at 6 hours in neutrophils and 24 hours in PBMCs.

When looking at MCFA concentration changes in the supernatant of treated neutrophils (Figure 67 D), large variation is evident. Despite moderate increases in C10:0 and C12:0, the high variability of measurements means that there is no significant increase in these fatty acids. Hence, neutrophils do not produce any meaningful amounts of MCFA in this experiment. Due to the variability further experiments are required to assess any significant changes.

When looking at MCFA concentration changes in the supernatant of treated PBMCs (Figure 67 C), it is evident that C8:0 does not increase. C10:0 is significantly increased after 24 hours compared to timepoint 0 when aggregating the datapoints from all treatments and experimental repeats at this timepoint to compare with timepoint 0 (Mann-Whitney test,

Results

$p=0.0003$). When comparing the lipid concentrations in different treatments at 24 hours to timepoint 0, then only LPS +dexamethasone treated cells and Untreated cells reach significance, while the other two treatment types are more or less close to significance (Kruskal-Wallis-test with follow-up Dunn's multiple comparisons test, all p -values adjusted: LPS $p=0.0813$, LD $p=0.0238$, Dex $p=0.0573$, Untr $p=0.0397$). There is no significant difference between the different treatments after 24 hours (Kruskal-Wallis-test, $p=0.9679$). This means that PBMCs release a significant amount of C10:0 within 24 hours, but that this is independent of the treatment of the cells. Analysing the increase in C12:0 in PBMCs, this is not significant at either 6 hours or 24 hours, neither in any treatment compared to untreated cells nor compared to timepoint 0. Further experiments into MCFA production could corroborate any tentative findings made in this experiment.

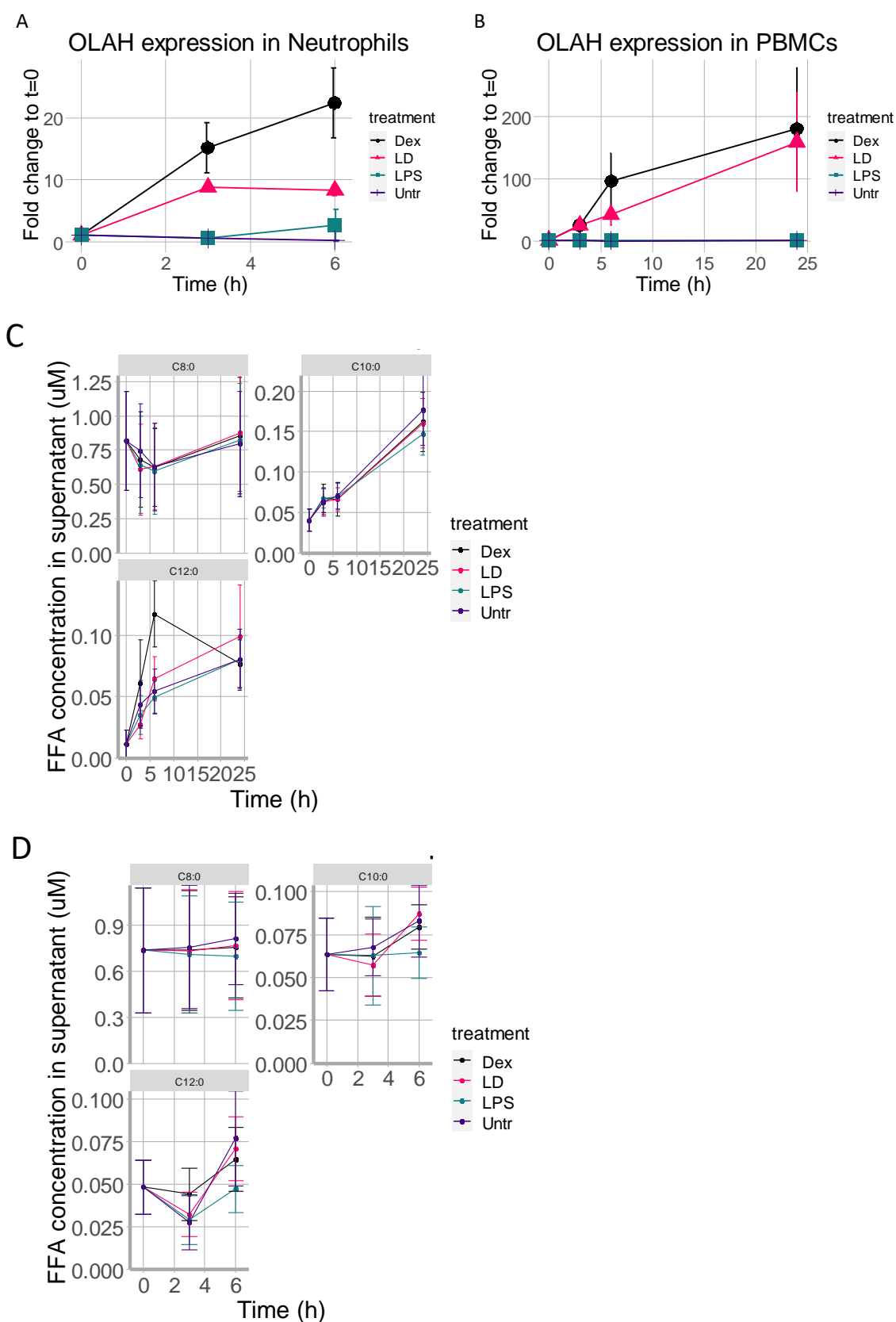


Figure 67 Expression of OLAH as fold change compared to $t=0$ in A) neutrophils and B) PBMCs extracted from healthy volunteer blood ($n=5$) in a time course of A) 6 hours and B) 24 hours. MCFA were extracted from supernatant samples taken simultaneously and

Results

quantified by previous established method yielding concentration in μM as result. Results shown show C) MCFA in PBMCs and D) MCFA in neutrophil time course experiments. Error bars depict SEM in all plots. Dex = Dexamethasone, LD = LPS + Dexamethasone, Untr = Untreated.

7.1.10 ACADM expression

MCFA can be oxidized in the mitochondria of the cells. This requires the activity of the enzyme medium-chain acyl-CoA dehydrogenase (MCAD), encoded by the gene *ACADM*. The oxidation of longer chain fatty acids requires a different enzyme: long-chain acyl-CoA dehydrogenase (LCAD), encoded by the gene *ACADL*. I hypothesized that any changes in production of MCFA might be accompanied by changes in lipid oxidation as well, either in sepsis or other conditions affecting lipid metabolism. An increased consumption of MCFA could offset an increased production, while a decrease in consumption should magnify MCFA levels. To this end I explored the *ACADM* expression in the same datasets I examined for *OLAH* expression. The transcriptional regulation of *ACADM* has been examined before, and I hence chose to not analyse this myself. In neonatal sepsis, expression of *ACADM* is slightly but significantly decreased (t-test, $\log_2FC = -0.995$, adj. $p = 1.187e-08$). In paediatric sepsis, *ACADM* did not change significantly. In transcriptomic datasets of different cell types, no differences in *ACADM* expression are evident. When looking at the information on human *ACADM* on the European bioinformatics (ebi) website, it appears evident that *ACADM* is expressed in many tissues. Differential expression is found for cell differentiation (e.g., activated B-cells vs. memory B-cell, $\log_2FC = 3.2$), different infections models (e.g., *Francisella tularensis novicida* infection of human PBMCs vs. uninfected, $\log_2FC = -2.7$), LPS treatment in THP-1 cells (100ng/ml LPS treatment vs. none on WT THP-1, $\log_2FC = -1.3$) and in neutrophils in sepsis ($\log_2FC = 1.9$). To explore this further, I analysed the *ACADM* expression in the previous time course experiments I conducted. I found that my results for *ACADM* expression are not in line with the results in the ebi datasets. In my experiment, the expression increased in THP-1 cells in all treatments to similar levels, in blood cells it increased in untreated cells and dexamethasone treated cells but not in LPS or LPS + dexamethasone treated cells (Figure 68). The fact that expression increases in all treatments in THP-1 cells indicates that a different factor is responsible for the increase in expression that is common to all treatments, potentially the handling of the cells. It is also worth noting that all tested cell types had detectable *ACADM* expression before treatment, indicating that this is likely a constantly expressed gene.

Results

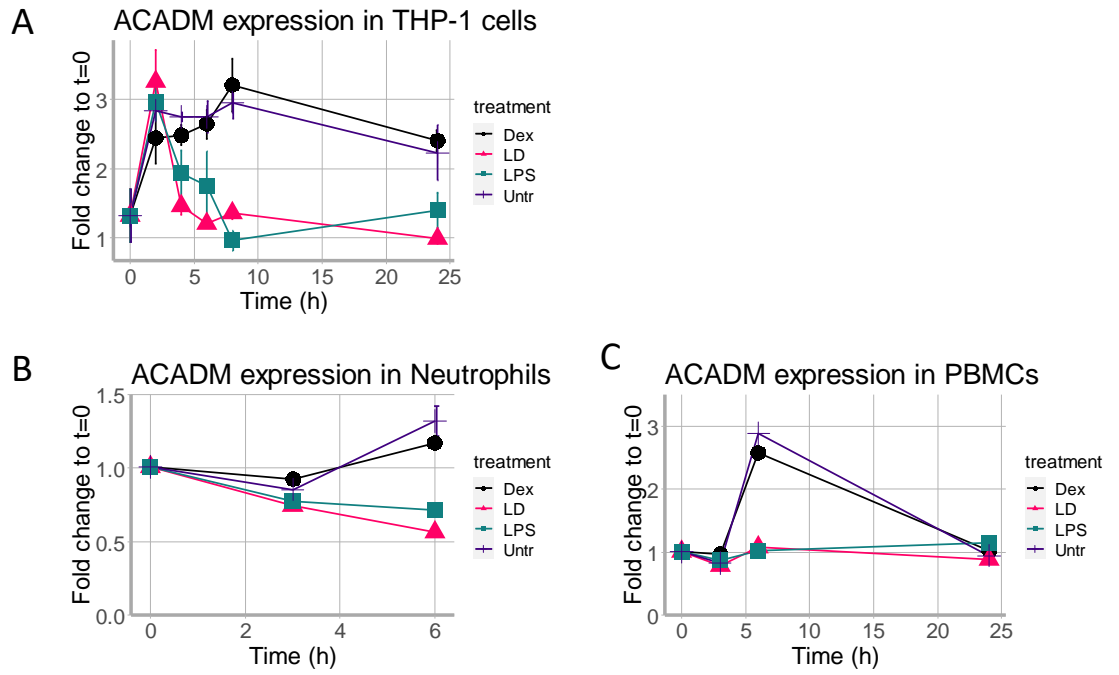


Figure 68 ACADM expression in A) THP-1 cells (n=3), B) Neutrophils (n=2) and C) PBMCs (n=3). The time is depicted in hours on the x-axis, and the y-axis shows the fold change of ACADM expression to t=0. The error bars depict the SEM.

7.1.11 Expression of genes involved in lipid metabolism pathways in sepsis

Finally, I decided to evaluate overall gene changes relating to lipid metabolism pathways in sepsis to see if there is a general shift in gene expression skewing metabolism towards net production of MCFA by immune cells. To this end I used the previously used dataset of neonatal gene expression and filtered it for genes involved in lipid metabolism. I generated this list of genes using REACTOME. Figure 69 shows gene expression changes in genes involved in lipid metabolism in neonatal sepsis. Firstly, many genes are changed. 75 out of 209 genes show significant changes (~35%). I examined the roles of genes significantly different with fold changes above 1.5 using gene descriptions on genecards (accessible via <https://www.genecards.org/>)²⁴⁶. In the mitochondrial metabolism pathway, the majority of significantly changed genes is downregulated. Having a closer look at mitochondrial metabolism, it is noticeable that *ACADM* is decreased while *ACADVL* is increased. *ACADL* is not significantly affected. Two genes encoding enzymes required for LCFA and VLCFA β -oxidation are increased, *ACADVL* and *HADHB*. Paradoxically, *HADHA* is decreased. *HADHA* and *HADHB* are genes encoding the α and the β subunit of a protein involved in the β -oxidation of LCFA. Furthermore, genes involved in β -oxidation of SCFA and MCFA are decreased, i.e., *ACADM*, *HADH*, *ECHS1*. In addition to mitochondria, peroxisomes can β -oxidize fatty acids, a mechanism that is thought to occur increasingly when mitochondria are overloaded. Looking at gene changes within peroxisomes, several genes involved in β -oxidation are increased: *ABCD1*, *ACOX1* and *CRAT*. The *CRAT* encoded protein is responsible for the synthesis of short, medium and branched chain acylcarnitines. However, *CROT*, which is decreased, plays a role in β -oxidation of C6-C10 acylcarnitines. This indicates that while β -oxidation of LCFA is likely increased, β -oxidation of SCFA and MCFA is likely hampered. This could lead to an accumulation of acylcarnitines of this kind since the synthesis of acylcarnitines binding these fatty acids is increased. Another important factor in β -oxidation is the carnitine shuttle. The carnitine shuttle plays a role in both mitochondrial and peroxisomal import of fatty acids. *CPT2*, one of the genes involved in transport of fatty acids into mitochondria and peroxisomes is also increased. Another gene involved in transport of lipids is *SLC27A2*, which can transport LCFA with <22 carbons into cells, peroxisomes and the endoplasmatic reticulum. The increased expression could lead

Results

to sequestering of C22:6 and other LCFA of applicable length into the cells, but also their import into the mentioned cell organelles.

Next, I looked at production of lipids and uptake/ storage of lipids as triglycerides. *DGAT2* and *LPIN2* are upregulated. Their encoded proteins play a dual role in triglyceride synthesis and in regulating lipid metabolism. On the contrary, also *MGLL* is upregulated that separates acylglycerols into glycerol and fat, and *LPIN1*, that has a similar role to *LPIN2*, is downregulated. It is hard to say if triglyceride metabolism is skewed towards catabolism or synthesis. In terms of biosynthesis of lipids, many genes are significantly changed. Many of these genes encode proteins with the same function but slightly different substrate specificity. Several long-chain-acyl-CoA-synthetases are upregulated. These catalyse the reaction of free fatty acids into their acyl-CoA, an important step prior to further processing in the cells including β -oxidation and synthesis. The upregulated *ACSL* can mediate the conversion of FFA with a minimum chain length of C12:0 up to C20:0, including arachidonic acid (C20:4) and arachidonic acid derived mediators (epoxy-eicosatrienoic acid). A further type of genes differentially regulated are *HACD1-3*, these encode 3-Hydroxyacyl-CoA Dehydratases and play an important role in the LCFA elongation cycle during lipid synthesis²⁴⁷. *HACD1* is upregulated, whereas *HACD2* and *HACD3* are downregulated. *HACD2* and *HACD3* are thought to be redundant²⁴⁷. *ELOVL* are genes encoding further protein in this LCFA elongation cycle, with the different proteins being lipid-specific in their function. *ELOVL4-7* are differentially regulated, with *ELOVL4* being downregulated and *ELOVL5-7* upregulated. *ELOVL4* leads to the synthesis of ultra-long chain fatty acids, whereas *ELOVL5* and *ELOVL6* play a role in elongating C20:4, C20:5 and C12:0-C16:0, respectively²⁴⁸. Notably the *ACSL*, *HACD* and *ELOVL* genes are all endoplasmatic reticulum associated. Lipid synthesis in the endoplasmatic reticulum leads to the production of lipid mediators and membrane lipids²⁴⁸. There is a separate lipid biosynthesis reaction that takes place in mitochondria and cytoplasm, involving *OLAH* (upregulated) and *FASN* (expressed but not upregulated) that can lead to the production of MCFA. Finally, several genes in the arachidonic acid pathway are upregulated. This may lead to an increased production of lipid mediators from arachidonic acid.

Overall, the analysis of gene expression of lipid metabolism genes indicates points toward LCFAs being increasingly used in sepsis, both as precursors for lipid mediators and to generate energy from β -oxidation. SCFA and MCFA seems to be less used in these

pathways; however, MCFA production is potentially increased. However, caution should be used in any inference as gene expression does not necessarily predict protein or lipid levels

–

7.2 Discussion

In this chapter I aimed to elucidate a potential source of MCFA in the context of sepsis. I focused on the potential production of MCFA by the body, more specifically cells present in the blood and endothelial cells. In this regard the expression and transcriptional regulation of *OLAH* was thought to be pivotal and hence also investigated. This investigation was used as starting point for experiments into lipid production. Finally, since my interest in MCFAs stems from their potential role as ligands for GPR84, I tried to establish an assay that could confirm or dismiss their role.

Previous studies have firmly established that cellular MCFA production requires the enzyme *OLAH*^{227,249,250}. We find the gene for this enzyme expressed and statistically significantly upregulated in both whole blood studies of sepsis patients and in flow sorted immune cells of sepsis patients. I analysed the transcriptional regulation of *OLAH* to establish the potential mechanism that leads to increased expression of *OLAH* in sepsis. In this chapter I used the same approach for this analysis as in chapter 6 but I used a less stringent rule to choose the transcription factors in the final step of analysis. The analysis hence has a higher likelihood of false positives which warrants the additional confirmation using further datasets. I found both inflammatory transcription factors and the glucocorticoid receptor as potential candidates for transcriptional control of *OLAH*. These transcription factors are likely to play an important role in sepsis. In addition, different hormones were found and transcription factors from the PPAR-family. I did not further pursue the latter two findings, since I did not find suitable datasets (PPAR transcription factors) for this or had no further indications that it might be relevant in sepsis (hormones). Further analyses both *in silico* and *in vitro* into the expression of *OLAH* established that glucocorticoids, but not inflammatory signals alone lead to upregulation of *OLAH*. This confirms that most likely those transcription factors found related to the inflammatory response, e.g., IRF1, IRF4, NFKB, RELA, were either false positives or contribute a modulatory role to its regulation. The transcription factors (SPI1, CEBPB, CEBPA) found indicated that *OLAH* was expressed in myeloid immune cells. Looking at the data of immune cells treated with inflammatory signals, it is apparent that *OLAH* is expressed at detectable levels in various immune cells before stimulation. In my own *in vitro* experiment, *OLAH* is expressed at detectable levels in the cells obtained from healthy volunteers before

Results

stimulation as well. Looking at the data on differential gene expression from the ebi gene expression atlas, *OLAH* differential expression in different epithelial cell lines is presented. This further supports the prediction that *OLAH* is expressed in epithelial cells and immune cells is correct. This is also in line with the data from Fantom5 that predicts the enhancers of a gene and its gene expression based on the enhancers. Since endothelial cells could also contribute to systemic lipid levels, I tested the *OLAH* expression in endothelial cells. However, I found it to be lower than in blood or THP-1 cells with a lesser upregulation in response to dexamethasone. Any contribution to lipid levels by endothelial cells hence seems unlikely. I think it is likely that the transcriptional regulation of *OLAH* is stringently and dominantly regulated via enhancer control rather than a promoter-led transcriptional regulation. Firstly, the predicted promoter region shows the phenotypic features of a closed region, which would hence require to be opened. This phenotypic characterisation stems from the UCSC genome browser that accesses different data sources to visualize the genome and its features. Opening of a genomic region for transcription is commonly initiated by enhancers; the activity of cell-specific enhancers would explain the cell-specific expression of *OLAH*. My *in silico* analysis found the glucocorticoid receptor as transcription factor likely regulating *OLAH*. My *in vitro* experiments support this, since in both THP-1 cells and immune cells obtained from healthy volunteer blood, *OLAH* expression is significantly upregulated in response to dexamethasone, a glucocorticoid receptor ligand. This provides experimental support for the *in silico* analysis of transcriptional regulation.

Next, I sought to examine the lipid levels in my time course experiments of different cells in response to LPS and dexamethasone. Due to the constraints in working with cells obtained from healthy volunteer blood, only relatively short time courses could be performed. Taking into account the time difference from increased expression of a gene until increased presence of the protein, it might be possible for the later times of the time course of PBMCs that the protein is present in increased amounts. However, to see an effect of an increased amount of enzyme, additional time would also be needed in which the enzyme performs its activity and cells can accumulate increased amounts of the lipid. In light of this reasoning, it is not surprising that no differences in MCFA levels between different treatments was found in my analysis. However, there was a significant increase in C10:0 (decanoic acid) in PBMCs. Seen that PBMCs expressed *OLAH* before treatment, it is possible that the cells were producing the C10:0 using *OLAH*. I think that this is a likely

explanation, though other explanations could be possible, e.g., the liberation of C10:0 by lipases or the breakdown of longer lipids. If I would have had more time, I would have established a more relevant cell culture model that allows for longer time course experiments. Firstly, the knockdown of *OLAH* would have been pivotal in proving or disproving an association between lipid production and *OLAH* expression in immune cells. Secondly, when trying to establish if there is a connection between induction of *OLAH* expression and lipid production, a time course of extending beyond 24hrs would be more likely to yield measurable changes. To this end one could mature monocytes derived from healthy volunteers into adherent macrophages, try different cell lines, e.g., a neutrophil cell line or potentially T-cell lines if shown to express *OLAH*. In order to keep the cells alive while also avoiding a high background, the cells would need to be cultured in standard medium first and then transferred to the medium containing lipid-depleted serum after about 24 to 36 hours. Another option would be to culture the cells in standard medium for the first 24 hours and then switch to a medium containing a labelled glucose or acetate the cells might use for any *de novo* lipid production. This would allow one to distinguish newly produced lipids from background lipids; if the cells however use a different compound in addition to the provided labelled compound to generate MCFAs, this production might be missed. By using a knockdown in combination with the time course, the role of *OLAH* in lipid production by immune cells could be proven or disproven.

In an additional analysis I looked at the expression of *ACADM*, the gene encoding an enzyme (MCAD) pivotal for MCFA catabolism in the mitochondria for energy generation. I reasoned that the presence of MCAD influences the presence of MCFA as well. Unfortunately, the results of this are not clear-cut. The upregulation of *ACADM* in THP-1 cells shows a different pattern to PBMCs, while in neutrophils no clear up- or downregulation is evident. The results in THP-1 cells are in disagreement with results reported in the ebi gene expression atlas for LPS stimulation of THP-1 cells, since the atlas indicates a down-regulation of *ACADM*. In addition, it reports a downregulation of *ACADM* in PBMCs when infected with a bacterium, while in my study *ACADM* was not upregulated in response to LPS. The reported increase in *ACADM* in sepsis according to ebi is also not in agreement with the found downregulation and unchanged expression in neonatal and paediatric sepsis, respectively. It is hence unclear how *ACADM* expression might change in response to different stimuli. This analysis however does provide the information that

Results

ACADM is commonly expressed at a basal level in immune cells. This expression indicates a readiness of the cells to use MCFAs which maybe indicates an availability of MCFAs to the cells consistent enough to warrant a certain baseline level of presence of this enzyme. My analysis of gene expression changes in neonatal sepsis of genes associated with lipid metabolism found a skewing of β -oxidation towards increased utilization of LCFA and decreased utilization of MCFA. This is a preliminary result based on only one transcriptomics dataset and would require further validation by lipidomic and proteomic data and potentially an experimental model. To characterize the genes I used genecards, which is a reviewed resource of scientific information²⁴⁶. Further validation of the roles of the differentially expressed genes by literature research is warranted.

In conclusion, I have shown data supporting the notion that immune cells have the capacity to produce MCFAs and notably could produce C10:0 to increased levels in presence of glucocorticoids. Unfortunately, I was not able to prove or disprove my hypothesis that *OLAH* expression leads to MCFA production in this chapter, but I think that future research using the outlined experiments could elucidate this interesting topic significantly.

8 Discussion

In this thesis I aimed to gain insights into the MCFA-GPR84 signalling axis in sepsis. GPR84, a pro-inflammatory receptor, has been found to be significantly upregulated in sepsis and is an integral innate immune-metabolic biomarker for neonatal sepsis. I hypothesize that MCFA-GPR84 axis is a key regulated host response in sepsis. The first aim of my thesis was to examine if GPR84's putative ligands, the MCFA C10:0 and C12:0, would be present and/or elevated in sepsis. The presence of these lipids in sepsis, or even in inflammatory conditions, has been insufficiently researched. To this end I developed a method to measure these MCFA in blood samples. A range of other lipids was also included in the method to enhance the findings and resulting understanding of lipid changes. The data obtained from lipid measurements with said method in a patient cohort (ILTIS) was combined with metadata to yield insights into lipid profile changes in sepsis with regards to different characteristics, i.e., the type of sepsis or survival. Apart from the presence of MCFA, another important question pertained to their origin. MCFA are relatively understudied compared to SCFA and LCFA, and their origin (and presence) in the human context has been barely studied with exemption of lactation. In addition to GPR84's ligands in sepsis, I was also interested in its expression and transcriptional regulation in sepsis. To this end, I analysed the transcriptional regulation of *GPR84 in silico* and examined its expression in different cell types. This analysis yields some indication into *GPR84* expression pattern in sepsis which should be tied to its function. The main findings of my thesis are summarized in Table 16.

8.1.1 Main findings

The first finding of my thesis that I would like to highlight is the statistically significant increase in C10:0 (decanoic acid) in the plasma of adult sepsis patients. This is an important finding for two reasons: firstly, since C10:0 is one of the putative ligands for GPR84, this means that GPR84 likely will be activated in sepsis. In a subset of patients, the C10:0 concentrations were systemically high enough to activate the receptor according to different *in vitro* studies looking at receptor activation. Secondly, the fact that C10:0 is increased implies that a mechanism associated with sepsis leads to an increase in this lipid. This was not the only lipid significantly changed in sepsis; several lipids were significantly different between sepsis cases and controls in the ILTIS cohort. Among those that I would

Discussion

like to point out is C8:0, since it strongly correlates with C10:0, and was significantly increased in the plasma of sepsis cases in the ILTIS cohort. Furthermore, C12:0, which has been proposed as a ligand for GPR84, was significantly decreased in sepsis of unknown origin and sepsis caused from pneumonia. In the neonatal cohort nSep, C12:0 was significantly decreased in suspected sepsis cases compared to controls. The data hence suggest that in sepsis C10:0 is the more likely GPR84 ligand.

In line with my expectations based on known lipid mediator pathways, the precursors for lipid mediators, C20:4 and C20:5, were significantly decreased in sepsis cases. SCFA- and MCFA- acylcarnitines were significantly increased, a finding that has been reported previously for sepsis in different studies^{50,51,188,192}. An interesting observation in the ILTIS cohort was that urosepsis patients showed fewer changes to their lipid profiles than patients experiencing sepsis caused by pneumonia or undefined aetiology. Fewer interventions, no mortality and shorter hospital stays indicate that these sepsis cases were milder as well. This indicates a connection between severity and lipid profile in sepsis. The nSep cohort was analysed despite the suspected sepsis samples not yet being classified into “sepsis” and “not sepsis” at the time of writing my thesis. These samples showed some significant changes between suspected sepsis cases and healthy controls, including a significant decrease in C12:0 and a significant increase in C6:0 (hexanoic acid), and C16:0-carnitine (palmitoyl-carnitine). Some of the lipids in this cohort show a different alteration than in the ILTIS cohort, however, due to the lack of information on the samples and the different sample type and age of the cohort, this cannot be further interpreted. In addition, and similar to the ILTIS cohort, the heterogeneity in lipid concentrations was high for several of the lipid species. Overall, the preliminary nSep results indicate that it might be possible to use lipid profiles from whole blood as indicators of wellbeing and unwellness. This is also based on the assumption that the majority of suspected sepsis cases did not have sepsis, and hence the observed alterations might be caused by dynamics relating to milder unwellness. Returning to my focus on MCFA and GPR84, I analysed if immune cells might feasibly produce MCFAs. In this analysis I found that *OLAH*, the gene encoding the enzyme required for MCFA production, is upregulated in, paediatric and neonatal sepsis, and my in vitro and ex vivo analyses suggests that its expression is induced by glucocorticoids. Treatment with a glucocorticoid (dexamethasone) did not lead to significantly higher MCFA levels outside of the tested immune cells within 24 hours of

induction compared to cells undergoing no treatment or LPS stimulation. However, a uniform and significant increase in C10:0 when comparing concentration levels at timepoint 24 and 0 was found in PBMCs in all treatment types including the untreated cells. The transcriptional regulation of *GPR84* was analysed as well. Both the transcription factors likely regulating *GPR84*, the promoter architecture and its temporal expression suggest that it is a secondary immune response gene regulated by inflammatory stimuli. Overall, these findings paint a picture of a MCFA-GPR84 axis consisting of a glucocorticoid mediated increase in C10:0 leading to GPR84 activation in sepsis; GPR84 signalling hence likely plays a role in sepsis as part of the secondary immune response and secondary to a hypothalamus-pituitary-adrenal (HPA)-induced cortisol increase. It is likely that the C10:0-GPR84 axis is only active in a subset of patients, as C10:0 was only found increased in a subset of patients.

8.1.2 The increase of C10:0 (and C8:0) in plasma of sepsis patients and the decrease of C12:0

The first key feature to note about the increase of C10:0 (and C8:0) in adult sepsis is that it only appears to occur in a subgroup of sepsis patients. Since C8:0 and C10:0 were correlated, I will discuss C8:0 as well in this context despite it not being a ligand for GPR84. C12:0, which is thought to be a ligand for GPR84, was decreased in most adult sepsis cases to levels likely not able to activate the receptor. There was also no correlation between C10:0 or C8:0 and C12:0, meaning that what leads to increases of C8:0 and C10:0 has no effect on C12:0 levels. Different explanations are possible for the observed C8:0 and C10:0 increase in sepsis. Firstly, the increased expression of *OLAH* in sepsis could lead to an increase in production of C8:0 and C10:0. *OLAH* is a cytosolic enzyme that releases MCFA from the FFA biosynthesis cycle. When both *OLAH* and the enzyme releasing LCFA from the biosynthesis cycle are present, *OLAH* outcompetes the enzyme that would release LCFA, leading to an increased release of MCFA from the synthesis cycle²⁵⁰. Despite *OLAH* being robustly upregulated in neonatal and paediatric sepsis, there are differences apparent in its expression levels between patients. Furthermore, two adult and paediatric studies find differential expression of *OLAH* within sepsis cohorts: Banerjee et al. found that *OLAH* was significantly upregulated in sepsis non-survivors compared to survivors and in patients with a more complicated sepsis course among the survivors in a paediatric sepsis cohort²⁵¹. Basu et al. examined gene expression differences between paediatric sepsis cases that

Discussion

developed acute kidney injury and those that did not²⁵². In this cohort, *OLAH* was among the top differentially expressed genes (up in acute kidney injury). Acute kidney injury in sepsis is associated with higher severity and mortality, e.g., in the cohort examined by Basu et al. the mortality in sepsis-shock associated kidney injury patients was 45% compared to 10% in sepsis patients that did not develop kidney injury. It is worth noting that these studies did not specifically investigate *OLAH* or *GPR84* and neither commented on the possibility of MCFA generation in sepsis.

I wondered if a possible secondary mechanism amplifying the heterogeneity in C8:0 and C10:0 in sepsis could play a role. Two factors potentially affecting plasma lipids concentrations could be the breakdown of these lipids within cells and the release of these lipids into the plasma. For this reason I explored the potential breakdown of these lipids since the mechanism of release of these lipids into the plasma is unknown²⁵³. That is why I investigated *ACADM*, the gene encoding MCAD. MCAD is pivotal for the breakdown of MCFA in the mitochondria after these are imported or have diffused into mitochondria²⁵³. MCAD acts on MCFA-CoA and mediates the first step in β -oxidation of MCFA-CoA. In neonatal sepsis, *ACADM* expression was decreased. Other datasets and data from my *in vitro* experiments gave a mixed picture of *ACADM* gene expression in response to inflammation, glucocorticoids, infection and in sepsis, though in the majority of datasets *ACADM* expression was decreased. A decrease in *ACADM* has also been found in a microarray study of a mouse burn sepsis model²⁵⁴ and in a proteomic study of the monocytes of adult sepsis patients²⁵⁵ MCAD was significantly decreased compared to control cells. It hence seems likely that MCFA will not be metabolized in the mitochondria in sepsis. However, this could lead to an increase in MCFA-acylcarnitines rather than MCFAs. This is further discussed below. A further mechanism that could potentially affect the C8:0 and C10:0 concentrations outside of cells is the export of these lipids out of the cells. It is generally thought that C8:0 and C10:0 can move between compartments using diffusion⁴⁵, however studies by Violante et al. suggest a more nuanced mechanism. In their studies on mitochondrial and peroxisomal function, Violante et al. find that when inhibiting the carnitine shuttle, 50% less C10:0 reached the inside of mitochondria²⁵⁶. This indicates that C10:0 both diffuses and uses the carnitine shuttle. They also tried to establish how acylcarnitines move from the intra- to the extracellular compartment; however, they could not determine which transporter would potentially be involved. Apart from their studies, I

did not find further information on the potential mechanism of acylcarnitine export out of the cells.

The decrease of C12:0 seems to be caused by a different pathway than the hypothetical pathway for C8:0 and C10:0, as these lipids do not correlate at all. In nSep, C12:0 was also the most decreased lipid when comparing suspected sepsis cases to controls. Interestingly, a decrease in C12:0 in the skeletal muscle of a murine sepsis model has been found recently²⁵⁷. The decrease of C12:0 in whole blood and skeletal muscle in addition to in the plasma leads me to hypothesize that the decrease is caused by an intracellular consumption rather than just an increased uptake by cells. Looking at the results of the gene expression in neonatal sepsis, I can generate a hypothesis as to why C12:0 is decreased. The pathway of FA biosynthesis is increased; however, this is not the FA biosynthesis pathway that generates LCFA from acetyl-CoA in the cytoplasm, but the biosynthesis pathway in the endoplasmic reticulum that can lead to increased LC-PUFAs and lipid mediators. This pathway can use acetyl-CoA as substrate, but also other longer lipids. An increase in lipid mediators in sepsis has been found in previous studies. The enzymes in this pathway operate with a relative chain specificity. Among the genes upregulated, it appears that only those are upregulated that would bind fatty acids with a chain length of 12 carbons or longer. I hence think it would be possible that C12:0 is used to generate lipid mediators in the endoplasmic reticulum.

8.1.3 Acylcarnitines in sepsis

Acylcarnitines in sepsis have been identified and have attracted previous attention compared with MCFAs, and they have been measured in several sepsis studies in patient populations. Apart from that, much about the dynamics of acylcarnitines is known from studies into genetic disorders and fasting studies, including in the context of metabolic diseases.

In studies on acylcarnitines in sepsis, an increase in certain acylcarnitines have been found to be associated with non-survival^{50,51,258}. Langley et al. combined transcriptomic data with proteomic data and clinical metadata in a cohort of adult sepsis patients diagnosed based on sepsis-3 criteria⁵¹. Based on their analysis they concluded that the increased acylcarnitines were not a result of increased renal dysfunction. Langley et al. instead hypothesize that the increased acylcarnitine is caused by a defective carnitine shuttle,

Discussion

failing to facilitate β -oxidation in mitochondria. This idea has also been suggested in other publications and the idea that mitochondrial defects in sepsis lead to altered metabolism appears to be widely accepted^{39,183}. In addition, Langley et al. also found that cortisol was increased in the non-survivor group and that lipid transport proteins were decreased. Chung et al. measured acylcarnitines in a targeted approach and correlated the acylcarnitine measurements to clinical parameters and measured cytokines¹⁹². In this study, only C2:0-carnitine was associated with non-survival in sepsis. In addition, they found that acylcarnitine levels (C2:0-, C6:0-, C8:0-, C10:0-carnitine) correlated weakly but significantly with severity as measured by the SOFA score, and they found several SCFA-/MCFA-acylcarnitines to correlate weakly but significantly with different indicators of organ dysfunction. They conclude that these acylcarnitines are increased due to renal dysfunction, hepatobiliary dysfunction, thrombocytopenia and hyperlactatemia. In addition, they found significantly increased C2:0-, C6:0, C8:0 and C10:0-carnitine in patients with acute kidney injury caused by sepsis compared to sepsis patients without acute kidney injury. However, the correlations between acylcarnitines and clinical measurements were relatively weak ($\sim r = 0.3-0.5$), with stronger findings for correlation of C14:0- and C16:0-carnitine with lactate ($\sim r = 0.6$). I would hence regard these findings cautiously. Chung et al. explain their differences in findings to Langley et al. regarding renal dysfunction causing acylcarnitine level increases with the difference in recruitment strategy, as Langley et al. included several patients with chronic kidney disease and cirrhosis, and Chung et al. excluded any patients with a history of kidney disease. Furthermore, Puskarich et al. examined acylcarnitine differences between survivors and non-survivors in a placebo-treated and a carnitine treated patient population¹⁸⁸. In this comparably small study, they do not find significant differences between survivors and non-survivors in comparison with the placebo group. In the treatment group they find similar changes to Langley et al., with increased C2:0, C3:0, C4:0, C5:0, C6:0 and C8:0-carnitine in non-survivors. Based on their analysis of renal creatinine, they exclude an effect of renal dysfunction on acylcarnitine levels. Mickiewicz et al. conducted a metabolomic analysis of sepsis patients vs. SIRS in the ICU and found increased C2:0-carnitine to differentiate between these two groups¹⁶⁶. Liu et al. performed untargeted metabolomics to identify metabolites significantly different between sepsis survivors and non-survivors and found C2:0-carnitine decreased in non-survivors and C4:0-, C6:0, C8:0, C10:0, C12:0, C16:0 and C18:0-carnitine increased in non-

survivors⁵⁰. Their finding regarding C2:0-carnitine differs from the findings of all other groups cited above.

Changes to metabolism in sepsis have been likened to starvation, including the increased release of cortisol³⁹. Several studies have examined the effect of fasting on serum amino acids and acylcarnitines. A study into serum changes of amino acids and acylcarnitines found that healthy volunteers who fasted overnight, i.e., no food intake between dinner and breakfast, had consistently elevated acylcarnitines and decreased amino acids in the morning²⁵⁹. The study authors explained this change with a depletion of glucose overnight and a resulting switch to mitochondrial lipid oxidation leading to increased acylcarnitines. They found the range of acylcarnitines to be significantly higher after overnight fast compared to after breakfast in a very wide range of acylcarnitines, including all even-chain acylcarnitines from C2:0 to C16:0, C18:1-, C18:2, C18:3-carnitine with exemption of C4:0-carnitine and C18:0-carnitine. A similar study with similar findings was conducted in healthy 1 to 7 year old children who fasted for 20 hours²⁶⁰. Increases in acylcarnitines, especially unsaturated LCFA (C14-C18) were found. Saturated even-chain acylcarnitines were also increased, in a measured range of C2:0 to C18:3. C3:0-carnitine was also measured in this cohort but was not found to increase in fasting. Other studies focus on fasting in people having genetic mutations in fatty acid oxidation related genes. Mutations in fatty acid oxidation related genes are relatively rare (1 in 5000 to 10 000 births) and can be diagnosed based on acylcarnitine profiles in the blood, i.e., patients with a mutation in the *ACADM* gene will be diagnosed based on increased MCFA-acylcarnitines in the plasma²⁶¹. These studies could be especially interesting in relation to sepsis, since gene expression of several metabolic genes were changed in sepsis. By comparing the lipid profiles in fatty acid oxidation disorders and those in sepsis, one could find indications how β -oxidation is altered in sepsis. In one study, an infant with LCHAD fasted simultaneously with her twin sister²⁶². Fasting increased the acylcarnitines and FFA (measured as dicarboxylic acids, hydroxy-acids and as total non-esterified fatty acids) in the plasma of both siblings, albeit to a higher degree in the patient than in the sister. The LCHAD patient who was on an MCT-rich diet had increased levels of SCFA and MCFA-acylcarnitines after fasting, which was attributed to her diet. In a different study in adults with varying FAO gene mutations, the authors found that both fasting (overnight) and exercise lead to increased acylcarnitines in the plasma, and that these acylcarnitines were specific to the FAO gene mutations of the

Discussion

participants²⁶³. For example, patients with *CPT2*-mutations showed elevated levels of C18:1-carnitine after fasting and exercise. In addition, total FFA in the plasma were elevated after fasting in all participants, indicating increased lipolysis. Unfortunately, I could not find any fasting studies with participants with *ACADM* mutations. However, a study on muscle cells from healthy volunteers that were non-diabetic, young and lean, provides some relevant insights. Despite the apparent health of the volunteers, the scientists found differences in their respiratory quotient (RQ). The RQ is an indicator of the ratio of carbohydrate to fat oxidation. A higher RQ is thought to indicate a risk of metabolic syndrome, weight gain and is more often found in obese people and people with a family history of type 2 diabetes. The volunteers were divided into a “low RQ” and “high RQ” group after measuring the relevant criteria after an overnight fast. Skeletal muscle cells were sampled from every volunteer and incubated *in vitro* with labelled C16:0 (palmitic acid). After 24 hours of incubation supernatant and cells were sampled. While intracellular acylcarnitine content was similar in both groups, there were significantly increased amounts of labelled and unlabelled MCFA-acylcarnitines in the supernatant of the high RQ muscle cells. In addition, the concentration of extracellular acylcarnitines was several folds higher than the intracellular concentration. An analysis of gene expression found that high RQ muscle cells had decreased expression of *ACADM* compared to the low RQ muscle cells. The authors conclude that muscle cells with decreased *ACADM* expression have incomplete β -oxidation and that the resulting excess acylcarnitines are released into the supernatant to avoid intracellular accumulation. Relating these findings to my analysis of gene expression changes of lipid metabolic pathways in neonatal sepsis, it seems likely that acylcarnitines are increased due to incomplete β -oxidation due to downregulated *ACADM* expression. In this context it is also important to note that peroxisomes can oxidize lipids as well and do this using different enzymes than mitochondria. This could be used as a compensatory mechanism for mitochondrial overload in sepsis. Several peroxisomal genes were upregulated in neonatal sepsis, which could indicate increased peroxisomal activity. *In vitro* studies using KO cells of either the mitochondrial carnitine shuttle (*CPT2*) or a pivotal peroxisomal enzyme (*PEX13*), respectively, and incubation with a chemical mitochondria inhibitor (L-AC) prove that in mitochondrial impairment peroxisomes will convert C12:0 to C12:0-carnitine, C10:0-carnitine, C8:0-carnitine and C6:0-carnitine²⁵³. Peroxisomes do not produce acetyl-carnitine (C2:0-carnitine) as end-product of oxidation.

Despite the hypothesized mitochondrial impairment and increased peroxisomal impairment, C2:0-carnitine is significantly increased in the plasma of sepsis patients in the ILTIS cohort. C2:0-carnitine is also the end product of glucose and amino acid catabolism. In animal models, glucose was quickly depleted after induction of infection/sepsis, which could explain observed increases in C2:0-carnitine in these models^{51,188,192}. Human sepsis patients are usually treated with glucose. The breakdown of proteins and amino acids in sepsis has been observed, evidenced by the measurement of branched chain fatty acids and acylcarnitines with odd-chain length fatty acids, including C3:0-carnitine⁵¹, and the measurement of increased amino acids¹⁶⁶. The accumulation of C2:0-carnitine despite the likely impaired mitochondrial metabolism could also reflect a heterogeneous response within the patient, i.e., some cells having a functional metabolism and others not, or a temporal pattern where increased metabolism is followed by impaired metabolism, leading to the accumulation of acylcarnitines of various chain length up to C10:0-carnitine as observed in the ILTIS cohort and other cohorts^{50,51,192}. The above stated results and conclusions do not support the hypothesis that fatty acid oxidation is impaired at the level of the carnitine shuttle as suggested by Langley et al.⁵¹ or that the acylcarnitine profile is similar to a general fasting response³⁹. Instead, I would propose that the observed acylcarnitine profile in sepsis is in line with impaired mitochondrial fatty acid oxidation at the level of *ACADM*, likely in combination with a fasting response and increased peroxisomal oxidation.

8.1.4 Unsaturated fatty acid changes in sepsis

In sepsis and in animal models of sepsis, increased levels of triacylglycerides and free fatty acids have been found¹⁸⁴. These increases were explained with depletion of glucose and hence increased lipolysis. This would be in line with the increases in circulating triacylglycerides and non-esterified fatty acids observed in fasting studies^{259,263}. In the ILTIS cohort, C18:1 is increased in the plasma of sepsis patients and C18:2 shows a non-significant increased trend. These increases could hence be caused by increased lipid mobilization in response to glucose depletion. In contrast, other long-chain PUFAs are decreased, i.e., C20:4 (arachidonic acid), C20:5 (eicosapentaenoic acid) and C22:6 (docosahexaenoic acid). This finding is in line with previous studies and has been explained by an increased production of lipid mediators¹⁸⁴. In addition, decreased plasma LC-PUFAs with a chain length of >20C has been found to be associated with mortality in human sepsis

patients¹⁸⁴, while eicosanoid and SPM metabolism was one of the pathways with highest fold-changes between survivors and non-survivors in a meta-analysis²⁶⁴. The increased production of lipid mediators has been found in both human and animal studies and makes sense from a biological perspective^{183,264–266}. Looking at the transcriptomic changes in neonatal sepsis, several of the genes involved in conversion of FFA to lipid mediators are upregulated, i.e., *ELOVL5*, and genes involved in the arachidonic acid metabolism are changed as well. This makes sense, both as inflammation induces the conversion of arachidonic acid to eicosanoids, and since it has previously been found that the arachidonic acid metabolism pathway is differently activated in immune cells from sepsis patients compared to those of healthy volunteers¹⁶⁴. The observed decreases in C20:4, C20:5, C22:6 hence likely reflect an increased production of lipid mediators. If the increases in the associated pathways in non-survivors are causative and dysfunctional or merely a reflection of a more severe response in non-survivors remains to be examined.

8.1.5 Urosepsis compared to sepsis OUA and pneumonia sepsis

In the ILTIS cohort, urosepsis cases were found to have a lipid profile more similar to healthy controls than other sepsis patients. This, and the non-existent mortality, fewer interventions and shorter hospital stay in this cohort, indicate a milder disease. I have been assured by clinicians that urosepsis indeed is often less severe than other types of sepsis, and have found in the literature a lower mortality and severity in urosepsis¹⁸⁹. This indicates that changes in the lipid profile might indicate severity of disease, a finding echoed by previous publications in relation to acylcarnitines and lipids not measured in this approach^{168,267}.

8.1.6 Cortisol in sepsis

As mentioned before, Langley et al. found increased cortisol associated with non-survival in sepsis⁵¹. An increase in cortisol has also been found to be associated with non-survival in a study measuring cortisol levels in a patient cohort of severe sepsis and septic shock²¹⁶. Systemic cortisol was measured in samples taken usually in the morning, and as soon as possible and latest within 24 hours after recruitment. All patients were divided into 4 groups based on their cortisol levels. The group with the highest cortisol (serum cortisol >1242 nmol/l) had a significantly increased mortality rate: in this group 81% of patients died, compared to a mortality rate of 51% in the whole cohort. It is important to note that

this group of patients contained more cancer patients and slightly older patients than the other groups (25% compared to 17%, 10% and 10%, mean age 65 compared to 56, 56 and 66). The authors propose that the higher mortality in the high cortisol group might be caused by a disproportionate cortisol response to the infection. In this study no increased mortality was found in the low cortisol group, though a decreased cortisol response to infection has previously been regarded as increasing mortality²¹⁶. The cortisol response in sepsis is of interest to my studies, since *OLAH* is likely a GR-regulated gene and its expression should hence be affected by the cortisol levels. *OLAH* is strongly upregulated in non-survivors and in more severe sepsis cases with a higher mortality rate^{251,252}. The increased expression of *OLAH* in these groups could be caused by increased cortisol. Whether *OLAH* contributes to increased mortality and severity depends on the effects of the likely consequently produced MCFA. This leads me to link *OLAH* to the MCFA-GPR84 signaling axis in sepsis.

8.1.7 The role of GPR84 in sepsis

Based on my analysis of transcriptional regulation, I would suggest that *GPR84* is a secondary response gene to inflammation and is especially involved in innate immune cells. These findings are in line with the previously published literature on GPR84. The receptor has been repeatedly characterized as pro-inflammatory, and it hence seems likely that it plays a pro-inflammatory role in sepsis as well. This could be priming phagocytes for increased phagocytosis, chemotaxis, ROS release and the increased release of pro-inflammatory cytokines by immune cells. Generally, these types of functions could lead to a more effective immune response to the insult but would be detrimental in case of already excessive inflammation. However, it has been previously and is still an ongoing discussion *how* the immune response is dysfunctional in sepsis, and which mechanism within sepsis lead to severity and mortality. Undisputedly though, mortality is caused by organ damage. Hence, rather than discussing the immune response in sepsis, I would like to focus on organ damage in sepsis and GPR84's potential role in it. The lipids likely activating GPR84 have only been found elevated in a subset of patients and were implicated with non-survival, just as *OLAH* has been previously implicated with non-survival and elevated cortisol as well. Could GPR84 signalling contribute a detrimental effect in sepsis? Mouse model studies into kidney and liver fibrosis, among other types of fibrosis, found positive effects from *Gpr84* KO and the use of GPR84 antagonists. In these studies, the use of *Gpr84* KO or GPR84

Discussion

antagonists, led to less immune cell infiltration into the affected organs and less fibrosis. The fact that a GPR84 antagonist reduced immune cell infiltration to the injured liver might indicate that GPR84 was playing a role in the chemotaxis of immune cells to the organ. A study measuring gene expression in human liver biopsies of healthy controls and patients with non-alcoholic fatty liver disease (NAFLD) found increased expression of *GPR84* in NAFLD livers and that *GPR84* expression correlated positively with the inflammation grading of the NAFLD activity score¹²⁴. Chemotaxis and increased local inflammation due to GPR84 could likely also occur in sepsis when kidney or liver injury occurs. Any MCFA-GPR84 signalling potentially occurring in the organ might then exacerbate fibrosis, as in the mouse models. This could be an interesting avenue for further research and has therapeutic potential, as the safety of several antagonists has already been shown¹³¹.

Table 16 shows the main findings of my thesis.

Main findings
- C10:0 is increased in sepsis
- C12:0 is decreased in sepsis OUA (of undefined aetiology) and sepsis caused by pneumonia
- Lipid profiles are different in sepsis, but many lipids do not change
- Acylcarnitines are significantly changed in the plasma of sepsis patients in line with previous findings
- Urosepsis is milder in the ILTIS cohort than pneumonia caused sepsis and sepsis OUA and shows fewer changes to lipid profile
- <i>OLA</i> H gene expression is increased in sepsis
- <i>OLA</i> H gene expression is regulated by glucocorticoids
- <i>GPR84</i> is likely a secondary immune response gene regulated by inflammatory signals, i.e., TLR4 – NFkB, but also others, and a foremost enhancer regulated gene

8.1.8 Pitfalls

In these studies, I have focused on metabolic changes in the plasma and blood and evaluated the scientific literature with a focus on changes as measured in these body fluids. However, sepsis is a condition that affects several organs, and the observed changes in plasma and blood can originate from many sources. When observing metabolomic changes in the blood or plasma, it is essentially impossible to untangle the origin of the metabolomic changes. The reason for this focus is the comparable ease of access to this type of samples for my studies; a convenience that is also reflected in the studies conducted by the wider research community. Studies involving other organs are usually conducted in animal studies. One such study, for example, finds changes in the skeletal muscle tissue in a murine sepsis model on a proteomic and metabolomic level²⁵⁷. Skeletal muscle metabolism could be a (major) contributor to acylcarnitine changes, as it is thought that the majority of the body's carnitine is stored in this tissue, and it has been shown that skeletal muscle can release acylcarnitines into the plasma²⁶¹. In addition, GPR84 has been implicated in a murine model to regulate mitochondrial function in skeletal muscle²⁶⁸. While any dysfunction of immune cells, as at least partially measurable by e.g., transcriptomics and proteomics of the blood, undeniably will play a role in sepsis, organ dysfunction and -failure are what ultimately leads to mortality. It would hence be relevant to research changes to e.g., kidney and liver, as they are metabolically active organs that are often affected in sepsis. In addition, GPR84 signalling might be relevant within these organs due to GPR84's implication in fibrosis as discussed above. The potential presence of MCFA in these organs in sepsis or other conditions has not been previously evaluated.

Another complication in measuring the lipids in sepsis is the heterogeneity of measurements. This has been previously observed in other studies as well, both in the sepsis context and in other fields^{263,269}. This is caused by the several factors that impact on lipid concentrations, including the fasted/fed state, exercise, the diet, and metabolism that could be influenced by comorbidities and other personal factors. In order to achieve valid findings in any patient cohort, it is hence necessary to control any external factors as well as possible and to use adequately sized and hence powered cohorts. In sepsis, the heterogeneity of the condition is a further complicating factor. As a researcher, the recruitment conditions for such a study will hence be a trade-off between inclusivity (bigger cohort, less rules), in order to accurately reflect the general patient population with the

cohort, and the need to exclude conditions that could lead to confounding or outlier results (stricter recruitment rules). Essentially, when putting too many constraints on recruitment, it will be hard to achieve the necessary cohort size for an adequately powered study. Within the ILTIS cohort, the urosepsis patients had a different lipid profile to sepsis OUA and pneumonia-sepsis, which had the potential to confound some results. Simultaneously, if I excluded the urosepsis patients, I would probably not have been able to draw some of the study's conclusions. A good solution to this pitfall might be the diligent documentation of patients' metadata and rigorous analysis of any potential confounding effects. In addition, differing recruitment strategies between studies impact comparisons of results and validation of findings. In sepsis studies, this can be the correct choice of the control cohort. Studies typically recruit patients in the ICU found to not have sepsis, SIRS patients or healthy controls. The use of healthy controls, as done in the ILTIS cohort, might be problematic since this does not evaluate if differences occur due to inflammation (present in SIRS) or unwellness (patient in ICU), or if they are caused by sepsis. This means that any found lipid changes might not be specific to sepsis. This is also illustrated by the fact that suspected sepsis cases had a partially different lipid profile to healthy controls in the nSep cohort, despite the fact that most likely most suspected sepsis cases did not have sepsis. In addition, the recruitment of healthy volunteers could introduce differences in sampling, as samples will be taken (likely) by different staff, in a different location, under different circumstances, i.e., time till cold storage or processing might differ. Differences in sample handling have been found to affect measurement of lactate^{175,176}, and it is possible that this is the case for other compounds as well.

In discussing the potential downstream effects of MCFAs in sepsis I have entirely focused on those effects mediated by GPR84; however, further mechanisms affected by increased MCFA are possible. Firstly, the receptor GPR40 has also been shown to bind MCFA. This receptor can bind both saturated medium to long-chain fatty acids and poly-unsaturated fatty acids²⁷⁰. However, the activation by saturated MCFAs was lower than in response to some of the tested PUFA and saturated LCFA. GPR40 has a very different expression profile to GPR84, since it is foremost expressed in brain and pancreas. With regards to the immune system, it only shows low expression in monocytes in a study, and only negligible to none expression levels in other immune cells²⁷⁰. Activation of this receptor would likely interfere with GPR84-mediated effects, at least, in some tissues. Namely, GPR40 can be activated by

Discussion

some GPR84 antagonists and has been found to have beneficial effects on kidney injury, where antagonism of GPR84 also alleviated kidney damage^{114,271}. Since GPR40 can bind several different fatty acids, some of which are decreased in sepsis, it remains unclear if the total pool of its ligands is increased or decreased in sepsis. Secondly, some studies suggest that MCFAs and other saturated fatty acids could mediate immune regulatory effects, likely by binding to TLR4, leading to increased immune response activation^{272,273}. Several fatty acids are able to bind TLRs, one study finding the highest activation in response to C14:0 and C16:0 in a macrophage cell line²⁷³, another finding the strongest activation in response to C12:0²⁷².

Another pitfall in my studies I would like to point out is the mixing of several different datasets. Using several datasets can confirm robustness of results, if findings are repeatedly found. However, a more robust statistical analysis in lipid changes in connection with transcriptomic changes could have been conducted if the data came from the same set of samples. Instead, there is a certain level of mismatch due to different datasets used, i.e., the lipid measurements were conducted in an adult cohort, but none of the transcriptomic datasets are derived from an adult cohort. It would have been relevant to confirm if any observed changes in transcriptomics in neonatal and paediatric sepsis patients also occur in adult patients. Since most of my findings come from the measurement of samples, *in silico* analysis of different datasets and some *in vitro* experiments, further validation is warranted. While some of these findings align well with previously published findings, further mechanistic studies, i.e., extensive *in vitro* studies or animal models, could confirm the generated hypotheses.

When I conducted my studies into immune cell expression of *OLAH* and subsequent FFA production, I could have chosen my cell model more wisely in order to enable longer time courses. To this end I could have cultured the PBMCs for longer and used labelled glucose or acetate to determine FFA production. This would have omitted the use of charcoal-stripped FCS and any potential background issues, hence leading to better sensitivity. Using adhering cells also means that cells can be washed more easily and hence be sampled for lipid analysis as well.

8.1.9 Future directions

For future studies there are several interesting questions generated in this thesis. Also, additional work to further confirm or refute the hypothesis - that the MCFA-GPR84 axis is

a key regulated host response in sepsis – could be beneficial. One of the generated questions is about the effects of cortisol in affecting metabolism and the immune system in sepsis. To this end, a clinical study linking lipid measurements with transcriptomics and cortisol measurements could be very elucidating. In such an approach, one could look at the role of increased cortisol on gene expression and any resulting lipid changes. If this was coupled with *ex vivo* studies on blood cells obtained from sepsis patients and controls, valuable insights into metabolic dysfunction, gene expression and transcriptional regulation could be obtained. The blood cells could be extracted from the blood similar to this work and incubated for a time course with regular lipid measurements. Then it could be possible to examine if immune cells from sepsis patients with higher severity and/or higher cortisol levels are programmed to produce C8:0, C10:0 or C12:0. By linking transcriptomics and cortisol measurements, it can be evaluated how endogenous cortisol affects gene expression in sepsis to explore potentially detrimental pathways in non-survivors linked to cortisol. Linking lipid measurements to cortisol could confirm if cortisol in sepsis does indeed lead to increased C8:0 and C10:0 levels, as I hypothesized in this thesis. Furthermore, if linking lipid measurements and transcriptomics, it could be evaluated which pathways affect the lipid profile associated with non-survival. In this context the PPAR transcription factor family would also be relevant. Its potential role could be indicated by using ChEA3 to establish enriched transcription factors in different groups of patients. Such a comprehensive approach could evaluate hypotheses produced in this thesis and in the scientific literature.

Using animal models could provide additional insights into the role of GPR84 and OLAH in sepsis. Using a murine sepsis model combined with a GPR84 KO mouse could be used to evaluate if GPR84 affects metabolite levels, severity and survival in the sepsis model. An animal model could also be used to measure lipid levels in different tissues, e.g., in the kidney or liver. If increased levels of C10:0 or C12:0 were present in any of these tissues, then this could lead to chemotaxis and increased inflammation due to GPR84. Naturally, it would be easier to first evaluate if kidney or liver cells are able to produce these fatty acids using *in vitro* incubations of relevant cell lines. Furthermore, the role of OLAH could be evaluated in an animal model; however, it seems indicated that OLAH shows a different expression profile in mice than in humans. The suitability of mice for such a model would

Discussion

need to be tested first, and the use of an alternative animal might be more suitable but could provide additional challenges.

In this approach, I focused on the MCFA C10:0 and C12:0 as potential GPR84 ligands. The most interesting ligand appeared to be C10:0, as it was increased in both the plasma of sepsis patients and in the extracellular medium of incubated PBMCs. However, very recently a paper found 2-cis-decenoic acid and trans-2-decenoic acid as GPR84 ligands⁹⁹. These metabolites are thought to originate from bacteria. While bacteria seemed like a very feasible source for GPR84 ligands, I did not investigate them further in this approach due to time and organizational constraints. In the future, it would be interesting to explore this idea further. Firstly, it would be relevant and technically feasible to include 2-cis-decenoic acid and 3-cis decanoic acid into the list of metabolites measured with the LC-MS approach. Like this, their presence could be detected in samples when using this approach. Furthermore, it would be interesting to use bacterial pathogens obtained from sepsis patients to probe their production of GPR84 ligands. I would like to point out that the potential origin of GPR84 ligands from bacteria does not contradict with a potential production of GPR84 ligands by the body itself. Rather, these could be two additive mechanisms amplifying each other's effect. In other words, if the host increases MCFA production during infection, increasing chemotaxis and phagocytosis towards the infection site, then this would amplify the chemotaxis and phagocytosis induced by GPR84 ligands originating from the invading bacterium. The binding of bacterial ligands would provide an explanation as to the (evolutionary) function of GPR84, as hypothesized by Schulze et al.⁹⁹ and mentioned by Kutschera et al.⁹⁸, who stated that GPR84 might be a bacterial metabolite-sensing receptor. It could be interesting to explore if MCFA are produced by host cells as part of a healthy immune response, or if this is a hallmark of dysfunctional metabolism.

In conclusion, this work found a changed lipid profile to be associated with sepsis in the ILTIS cohort and found indications that this lipid profile might be partially associated with the severity of disease and likelihood of survival. These findings require further validation in more patient cohorts. Since one of the lipids significantly changed is a putative ligand for GPR84 (C10:0), it seems likely that the C10:0-GPR84 axis plays a role in sepsis. It appears feasible that PBMCs produce C10:0, and this production is likely associated with the gene *OLA1*, upregulated in response to glucocorticoids. Both high cortisol and *OLA1* expression

have been associated with non-survival in the literature, and *OLAH* expression with kidney damage in sepsis. Furthermore, GPR84 has been associated with kidney damage. Based on these findings I would like to hypothesise that increased cortisol in sepsis leads to upregulated *OLAH* expression, subsequently increased C10:0, amplifying C10:0-GPR84 signalling which leads to kidney damage. In my studies I did not further investigate the effects of GPR84 signalling in sepsis. As a result, I cannot say via which mechanism C10:0-GPR84 signalling would affect severity and mortality in sepsis. Based on previous studies on GPR84 this might be both by exacerbating organ damage and/or by amplifying the inflammatory immune response. Even an effect on the metabolic response is possible. Regardless of the mechanism, if the link with severity and mortality proves to be true, then GPR84 antagonism could be a treatment option in severe sepsis. Circling back to my hypothesis - MCFA-GPR84 axis is a key regulated host response in sepsis - I would conclude that the data of my thesis supports the notion that the MCFA-GPR84 axis is likely part of the host response in sepsis, with the limitation that this is likely only the case in a subset of patients with increased cortisol. If the MCFA-GPR84 axis is a **key** player in sepsis would require further investigation as outlined in the future directions of this thesis.

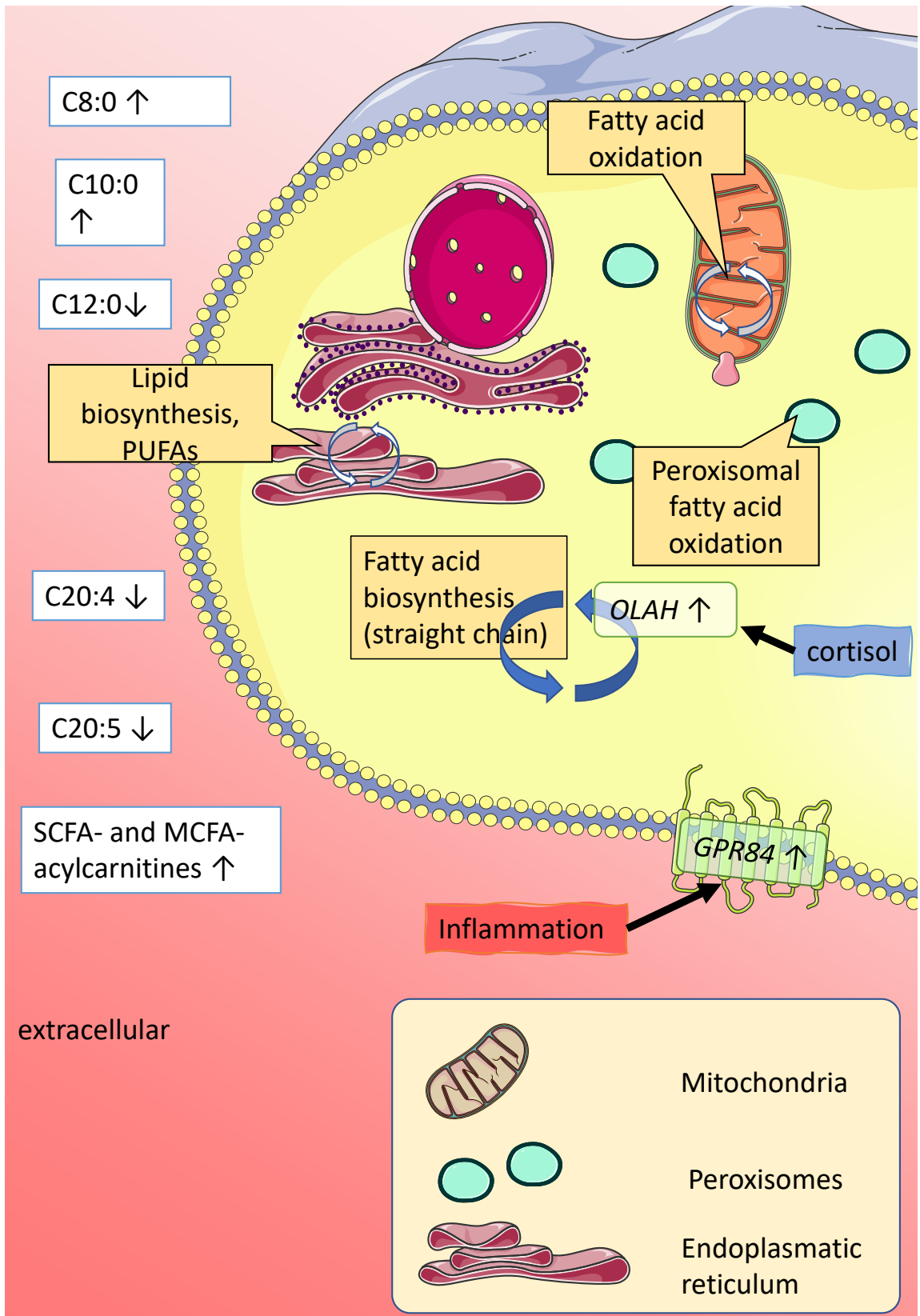


Figure 70 Overview picture of an immune cell depicting several of my main research findings from the thesis. Speech bubbles with orange background indicate biologic processes, green outlined, rounded figures contain genes, blue lined squares indicate lipids.

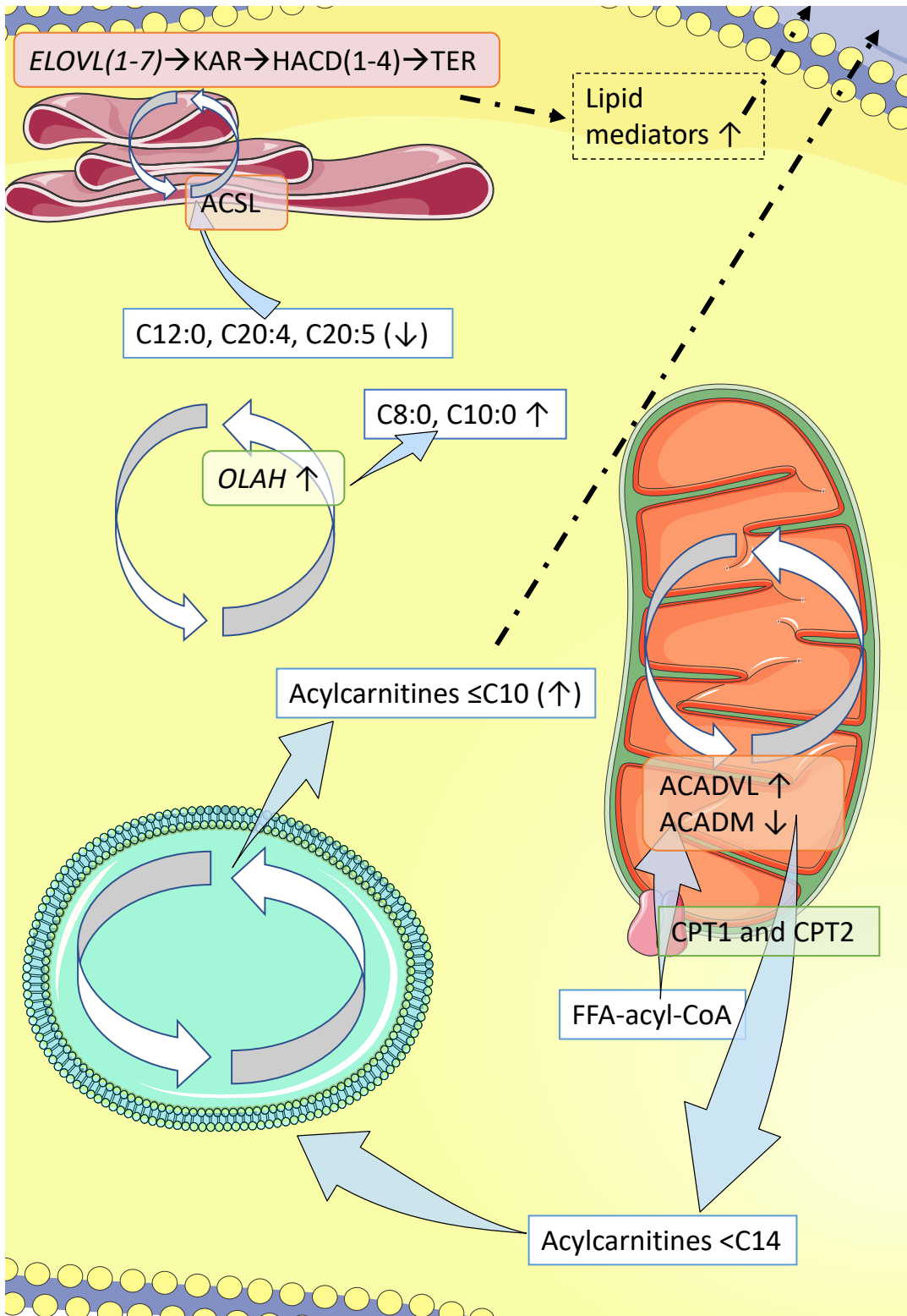


Figure 71 Second overview picture with a more in-depth view of processes within the cell that might lead to the observed lipid concentrations. Blue arrows indicate potential fluxes of lipids within the cell. Two circle arrows indicate a synthesis or catabolic cycle. Lipid mediators and other arrows are marked with dashed lines to indicate a level of uncertainty, as the export mechanisms for these

Discussion

lipids are not quite sure, and the lipid biosynthesis in endoplasmatic reticulum can lead to membrane synthesis and lipid mediators. In this context lipid mediators appear more likely.

9 References

1. Marshall, J. C. Sepsis: rethinking the approach to clinical research. *J. Leukoc. Biol.* **83**, 471–482 (2008).
2. Seventieth World Health Assembly Executive Board. *Improving the prevention, diagnosis and clinical management of sepsis*. World Health Organization vol. 315 (2017).
3. Singer, M. *et al.* The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA - Journal of the American Medical Association* vol. 315 801–810 (2016).
4. The UK SEPSIS TRUST. About - Sepsis Trust. <https://sepsistrust.org/about/about-sepsis/> (2020).
5. Rello, J., Valenzuela-Sánchez, F., Ruiz-Rodríguez, M. & Moyano, S. Sepsis: A Review of Advances in Management. *Adv. Ther.* **34**, 2393–2411 (2017).
6. Mankowski, R. T. *et al.* Older Sepsis Survivors Suffer Persistent Disability Burden and Poor Long-Term Survival. *J. Am. Geriatr. Soc.* (2020) doi:10.1111/jgs.16435.
7. Yende, S. *et al.* Long-term Host Immune Response Trajectories among Hospitalized Patients with Sepsis. *JAMA Netw. Open* **2**, e198686–e198686 (2019).
8. Brakenridge, S. C. *et al.* Current Epidemiology of Surgical Sepsis. *Ann. Surg.* **270**, 502–510 (2019).
9. Fleischmann, C. *et al.* Assessment of global incidence and mortality of hospital-treated sepsis current estimates and limitations. *Am. J. Respir. Crit. Care Med.* **193**, 259–272 (2016).
10. Rudd, K. E. *et al.* Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *Lancet* **395**, 200–211 (2020).
11. Fleischmann-Struzek, C. *et al.* The global burden of paediatric and neonatal sepsis: a systematic review. *The Lancet Respiratory Medicine* vol. 6 223–230 (2018).
12. Naghavi, M. *et al.* Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **385**, 117–171 (2015).
13. Levy, M. M., Evans, L. E. & Rhodes, A. The Surviving Sepsis Campaign Bundle: 2018

References

- update. *Intensive Care Med.* **44**, 925–928 (2018).
14. Nargis, W., Ahamed, B. & Ibrahim, M. Procalcitonin versus C-reactive protein: Usefulness as biomarker of sepsis in ICU patient. *Int. J. Crit. Illn. Inj. Sci.* **4**, 195 (2014).
 15. Kumar, A. *et al.* Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit. Care Med.* **34**, 1589–1596 (2006).
 16. Ferrer, R. *et al.* Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: Results from a guideline-based performance improvement program. *Crit. Care Med.* **42**, 1749–1755 (2014).
 17. Whiles, B. B., Deis, A. S. & Simpson, S. Q. Increased Time to Initial Antimicrobial Administration Is Associated with Progression to Septic Shock in Severe Sepsis Patients. *Crit. Care Med.* **45**, 623–629 (2017).
 18. Sterling, S. A., Miller, W. R., Pryor, J., Puskarich, M. A. & Jones, A. E. The Impact of Timing of Antibiotics on Outcomes in Severe Sepsis and Septic Shock: A Systematic Review and Meta-analysis. *Crit Care Med* **43**, 1907–1915 (2016).
 19. Alam, N. *et al.* Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. *Lancet Respir. Med.* **6**, 40–50 (2018).
 20. Venkatesh, B. *et al.* Adjunctive Glucocorticoid Therapy in Patients with Septic Shock. *N. Engl. J. Med.* **378**, 797–808 (2018).
 21. van der Poll, T., van de Veerdonk, F. L., Scicluna, B. P. & Netea, M. G. The immunopathology of sepsis and potential therapeutic targets. *Nat. Rev. Immunol.* **17**, 407–420 (2017).
 22. Chaplin, D. D. Overview of the immune response. *J. Allergy Clin. Immunol.* **125**, S3 (2010).
 23. Sugimoto, M. A., Sousa, L. P., Pinho, V., Perretti, M. & Teixeira, M. M. Resolution of inflammation: What controls its onset? *Frontiers in Immunology* vol. 7 160 (2016).
 24. Kawai, T. & Akira, S. TLR signaling. *Cell Death and Differentiation* vol. 13 816–825 (2006).
 25. Calder, P. C. Omega-6 and Omega-3 Polyunsaturated Fatty Acids and Inflammatory Bowel Diseases. *Bioact. Food as Diet. Interv. Liver Gastrointest. Dis.* 55–79 (2013) doi:10.1016/B978-0-12-397154-8.00014-2.
 26. Rosales, C. & Uribe-Querol, E. Phagocytosis: A Fundamental Process in Immunity.

BioMed Research International vol. 2017 (2017).

27. Gasteiger, G. *et al.* Cellular Innate Immunity: An Old Game with New Players. *J. Innate Immun.* **9**, 111–125 (2017).
28. Evans, S. S., Repasky, E. A. & Fisher, D. T. Fever and the thermal regulation of immunity: The immune system feels the heat. *Nature Reviews Immunology* vol. 15 335–349 (2015).
29. Delano, M. J. & Ward, P. A. Sepsis-induced immune dysfunction: Can immune therapies reduce mortality? *Journal of Clinical Investigation* vol. 126 23–31 (2016).
30. Boomer, J. S., Green, J. M. & Hotchkiss, R. S. The changing immune system in sepsis: is individualized immuno-modulatory therapy the answer? *Virulence* **5**, 45–56 (2014).
31. Wynn, J. L. *et al.* The influence of developmental age on the early transcriptomic response of children with septic shock. *Mol. Med.* **17**, 1146–56 (2011).
32. Smith, C. L. *et al.* Identification of a human neonatal immune-metabolic network associated with bacterial infection. *Nat. Commun.* **5**, 4649 (2014).
33. Jarczak, D., Kluge, S. & Nierhaus, A. Sepsis—Pathophysiology and Therapeutic Concepts. *Front. Med.* **8**, 628302 (2021).
34. Chang, J. C. Sepsis and septic shock: endothelial molecular pathogenesis associated with vascular microthrombotic disease. *Thromb. J.* 2019 171 **17**, 1–19 (2019).
35. Blanco, A. & Blanco, G. Metabolism in Some Tissues. in *Medical Biochemistry* 447–463 (Elsevier, 2017). doi:10.1016/B978-0-12-803550-4.00020-3.
36. Rui, L. Energy Metabolism in the Liver. *Compr. Physiol.* **4**, 177 (2014).
37. Alsahli, M. & Gerich, J. E. Normal Glucose Physiology. *Encycl. Endocr. Dis.* 72–86 (2014) doi:10.1016/B978-0-12-801238-3.03827-7.
38. O’Neill, L. A. J., Kishton, R. J. & Rathmell, J. A guide to immunometabolism for immunologists. *Nature Reviews Immunology* vol. 16 553–565 (2016).
39. Van Wyngene, L., Vandewalle, J. & Libert, C. Reprogramming of basic metabolic pathways in microbial sepsis: therapeutic targets at last? *EMBO Mol. Med.* **10**, 1–18 (2018).
40. O’Neill, L. A. J. & Pearce, E. J. Immunometabolism governs dendritic cell and macrophage function. *J. Exp. Med.* **213**, 15–23 (2016).
41. Sampath, H. & Ntambi, J. M. The fate and intermediary metabolism of stearic acid.

References

- Lipids* **40**, 1187–1191 (2005).
42. Marion-Letellier, R., Savoye, G. & Ghosh, S. Polyunsaturated fatty acids and inflammation. *IUBMB Life* **67**, 659–667 (2015).
 43. Russo, G. L. Dietary n – 6 and n – 3 polyunsaturated fatty acids: From biochemistry to clinical implications in cardiovascular prevention. *Biochem. Pharmacol.* **77**, 937–946 (2009).
 44. Adeva-Andany, M. M., Carneiro-Freire, N., Seco-Filgueira, M., Fernández-Fernández, C. & Mouriño-Bayolo, D. Mitochondrial β -oxidation of saturated fatty acids in humans. *Mitochondrion* **46**, 73–90 (2019).
 45. Schönfeld, P. & Wojtczak, L. Short- and medium-chain fatty acids in energy metabolism: the cellular perspective. *J. Lipid Res.* (2016) doi:10.1194/jlr.R067629.
 46. Violante, S. *et al.* Peroxisomes contribute to the acylcarnitine production when the carnitine shuttle is deficient. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1831**, 1467–1474 (2013).
 47. Marten, B., Pfeuffer, M. & Schrezenmeir, J. Medium-chain triglycerides. *Int. Dairy J.* **16**, 1374–1382 (2006).
 48. Sharma, N. K. *et al.* Proteomic study revealed cellular assembly and lipid metabolism dysregulation in sepsis secondary to community-acquired pneumonia. *Sci. Reports* **2017 71 7**, 1–13 (2017).
 49. Sharma, N. K. *et al.* Lipid metabolism impairment in patients with sepsis secondary to hospital acquired pneumonia, a proteomic analysis. *Clin. Proteomics* **16**, 1–13 (2019).
 50. Liu, Z., Yin, P., Amathieu, R., Savarin, P. & Xu, G. Application of LC-MS-based metabolomics method in differentiating septic survivors from non-survivors. *Anal. Bioanal. Chem.* **408**, 7641–7649 (2016).
 51. Langley, R. J. *et al.* An Integrated Clinico-Metabolomic Model Improves Prediction of Death in Sepsis. *Sci. Transl. Med.* **5**, 195ra95-195ra95 (2013).
 52. Straub, R. H., Cutolo, M., Buttgereit, F. & Pongratz, G. Energy regulation and neuroendocrine-immune control in chronic inflammatory diseases. *J. Intern. Med.* **267**, 543–560 (2010).
 53. Harbeson, D., Francis, F., Bao, W., Amenyo, N. A. & Kollmann, T. R. Energy Demands of Early Life Drive a Disease Tolerant Phenotype and Dictate Outcome in

- Neonatal Bacterial Sepsis. *Front. Immunol.* **9**, 1918 (2018).
54. Golucci, A. P. B. S., Marson, F. A. L., Ribeiro, A. F. & Nogueira, R. J. N. Lipid profile associated with the systemic inflammatory response syndrome and sepsis in critically ill patients. *Nutrition* **55–56**, 7–14 (2018).
 55. Irahara, T. *et al.* Alterations in energy substrate metabolism in mice with different degrees of sepsis. *J. Surg. Res.* **227**, 44–51 (2018).
 56. Finfer, S. *et al.* Intensive versus Conventional Glucose Control in Critically Ill Patients. *N. Engl. J. Med.* **360**, 1283–1297 (2009).
 57. Savino, P. Knowledge of Constituent Ingredients in Enteral Nutrition Formulas Can Make a Difference in Patient Response to Enteral Feeding. *Nutr. Clin. Pract.* **33**, 90–98 (2018).
 58. O'Donnell, V. *et al.* Failure to apply standard limit-of-detection or limit-of-quantitation criteria to specialized pro-resolving mediator analysis incorrectly characterizes their presence in biological samples. (2021) doi:10.5281/ZENODO.5766267.
 59. Schebb, N. H. *et al.* Formation, Signaling and Occurrence of Specialized Pro-Resolving Lipid Mediators-What is the Evidence so far? *Front. Pharmacol.* **13**, (2022).
 60. Dalli, J., Gomez, E. A. & Serhan, C. N. Evidence for the presence and diagnostic utility of SPM in human peripheral blood. *bioRxiv* 2022.04.28.489064 (2022) doi:10.1101/2022.04.28.489064.
 61. Hayakawa, M. *et al.* Neutrophils biosynthesize leukotoxin, 9, 10-epoxy-12-octadecenoate. *Biochem. Biophys. Res. Commun.* **137**, 424–430 (1986).
 62. López-Vicario, C. *et al.* Targeted lipidomics reveals extensive changes in circulating lipid mediators in patients with acutely decompensated cirrhosis☆. *J. Hepatol.* **73**, 817–828 (2020).
 63. Wasyluk, W. & Zwolak, A. Metabolic Alterations in Sepsis. *J. Clin. Med.* **10**, 2412 (2021).
 64. Misheva, M. *et al.* Oxylipin metabolism is controlled by mitochondrial β -oxidation during bacterial inflammation. *Nat. Commun.* **13**, 1–20 (2022).
 65. Stewart, C. J. *et al.* Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls.

References

- Microbiome* **5**, 75 (2017).
66. Brown, A. J. *et al.* The orphan G protein-coupled receptors GPR41 and GPR43 are activated by propionate and other short chain carboxylic acids. *J. Biol. Chem.* **278**, 11312–11319 (2003).
 67. Cummings, J. H., Pomare, E. W., Branch, W. J., Naylor, C. P. & Macfarlane, G. T. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* **28**, 1221–7 (1987).
 68. Li, M. *et al.* Pro- and anti-inflammatory effects of short chain fatty acids on immune and endothelial cells. *Eur. J. Pharmacol.* **831**, 52–59 (2018).
 69. Sundqvist, M. *et al.* Similarities and differences between the responses induced in human phagocytes through activation of the medium chain fatty acid receptor GPR84 and the short chain fatty acid receptor FFA2R. *Biochim. Biophys. Acta - Mol. Cell Res.* **1865**, 695–708 (2018).
 70. Wong, H. R. *et al.* The pediatric sepsis biomarker risk model. *Crit. Care* **16**, R174 (2012).
 71. Wong, H. R. *et al.* Improved risk stratification in pediatric septic shock using both protein and mRNA Biomarkers: Persevere-XP. *Am. J. Respir. Crit. Care Med.* **196**, 494–501 (2017).
 72. Sriram, K. & Insel, P. A. G protein-coupled receptors as targets for approved drugs: How many targets and how many drugs? *Molecular Pharmacology* vol. 93 251–258 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5820538/> (2018).
 73. Wettschureck, N. & Offermanns, S. Mammalian G proteins and their cell type specific functions. *Physiological Reviews* vol. 85 1159–1204 (2005).
 74. Wootten, D., Christopoulos, A., Marti-Solano, M., Babu, M. M. & Sexton, P. M. Mechanisms of signalling and biased agonism in G protein-coupled receptors. *Nat. Rev. Mol. Cell Biol.* **19**, 638–653 (2018).
 75. Neves, S. R., Ram, P. T. & Iyengar, R. G protein pathways. *Science* vol. 296 1636–1639 (2002).
 76. Syrovatkina, V., Alegre, K. O., Dey, R. & Huang, X.-Y. Regulation, Signaling, and Physiological Functions of G-Proteins. *J. Mol. Biol.* **428**, 3850–3868 (2016).
 77. Mahmud, Z. Al *et al.* Three classes of ligands each bind to distinct sites on the orphan G protein-coupled receptor GPR84. *Sci. Rep.* **7**, 17953 (2017).

78. Zhang, L. & Shi, G. Gq-coupled receptors in autoimmunity. *J. Immunol. Res.* **2016**, (2016).
79. Sassone-Corsi, P. The Cyclic AMP pathway. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
80. Wen, A. Y., Sakamoto, K. M. & Miller, L. S. The Role of the Transcription Factor CREB in Immune Function. *J. Immunol.* **185**, 6413–6419 (2010).
81. Aronoff, D. M., Canetti, C., Serezani, C. H., Luo, M. & Peters-Golden, M. Cutting Edge: Macrophage Inhibition by Cyclic AMP (cAMP): Differential Roles of Protein Kinase A and Exchange Protein Directly Activated by cAMP-1. *J. Immunol.* **174**, 595–599 (2005).
82. Lewis, A. E., Aesoy, R. & Bakke, M. Role of EPAC in cAMP-Mediated Actions in Adrenocortical Cells. *Front. Endocrinol. (Lausanne)*. **7**, 63 (2016).
83. Aronoff, D. M., Carstens, J. K., Chen, G.-H., Toews, G. B. & Peters-Golden, M. Short Communication : Differences Between Macrophages and Dendritic Cells in the Cyclic AMP-Dependent Regulation of Lipopolysaccharide-Induced Cytokine and Chemokine Synthesis. *J. Interf. Cytokine Res.* **26**, 827–833 (2006).
84. Osaki, M., Oshimura, M. & Ito, H. PI3K-Akt pathway: Its functions and alterations in human cancer. *Apoptosis* **9**, 667–676 (2004).
85. Berridge, M. J., Lipp, P. & Bootman, M. D. The versatility and universality of calcium signalling. *Nature Reviews Molecular Cell Biology* vol. 1 11–21 (2000).
86. Dhanasekaran, N. & Dermott, J. M. Signaling by the G12 class of G proteins. *Cellular Signalling* vol. 8 235–245 (1996).
87. Zeke, A., Misheva, M., Reményi, A. & Bogoyevitch, M. A. JNK Signaling: Regulation and Functions Based on Complex Protein-Protein Partnerships. *Microbiol. Mol. Biol. Rev.* **80**, 793–835 (2016).
88. Ha, J. H., Ward, J. D., Varadarajalu, L., Kim, S. G. & Dhanasekaran, D. N. The gep proto-oncogene Gα12 mediates LPA-stimulated activation of CREB in ovarian cancer cells. *Cell. Signal.* **26**, 122–132 (2014).
89. Khan, S. M., Sung, J. Y. & Hébert, T. E. Gβγ subunits—Different spaces, different faces. *Pharmacol. Res.* **111**, 434–441 (2016).
90. Smrcka, A. V. G protein βγ subunits: Central mediators of G protein-coupled receptor signaling. *Cell. Mol. Life Sci.* **65**, 2191–2214 (2008).
91. Schulman, H. Intracellular Signaling - An Introduction to Cellular and Molecular

References

- Neuroscience. in 335–370 (Academic Press, 2004). doi:10.1016/B978-012148660-0/50013-5.
92. Gaidarov, I. *et al.* Embelin and its derivatives unravel the signaling, proinflammatory and antiatherogenic properties of GPR84 receptor. *Pharmacol. Res.* **131**, 185–198 (2018).
 93. Gundry, J., Glenn, R., Alagesan, P. & Rajagopal, S. A Practical Guide to Approaching Biased Agonism at G Protein Coupled Receptors. *Front. Neurosci.* **11**, 17 (2017).
 94. Wittenberger, T., Schaller, H. C. & Hellebrand, S. An expressed sequence tag (EST) data mining strategy succeeding in the discovery of new G-protein coupled receptors¹¹Edited by J. Thornton. *J. Mol. Biol.* **307**, 799–813 (2001).
 95. Yousefi, S., Cooper, P. R., Potter, S. L., Mueck, B. & Jarai, G. Cloning and expression analysis of a novel G-protein-coupled receptor selectively expressed on granulocytes. *J. Leukoc. Biol.* **69**, 1045–52 (2001).
 96. Wang, J., Wu, X., Simonavicius, N., Tian, H. & Ling, L. Medium-chain fatty acids as ligands for orphan G protein-coupled receptor GPR84. *J. Biol. Chem.* **281**, 34457–64 (2006).
 97. Kaspersen, M. H., Jenkins, L., Dunlop, J., Milligan, G. & Ulven, T. Succinct synthesis of saturated hydroxy fatty acids and in vitro evaluation of all hydroxylauric acids on FFA1, FFA4 and GPR84 † ‡. *Cite this Med. Chem. Commun* **8**, 1360 (2017).
 98. Kutschera, A. *et al.* Bacterial medium-chain 3-hydroxy fatty acid metabolites trigger immunity in Arabidopsis plants. *Science* **364**, 178–181 (2019).
 99. Schulze, A. S. *et al.* Evolutionary analyses reveal immune cell receptor GPR84 as a conserved receptor for bacteria-derived molecules. *iScience* **25**, (2022).
 100. Recio, C. *et al.* Activation of the Immune-Metabolic Receptor GPR84 Enhances Inflammation and Phagocytosis in Macrophages. *Front. Immunol.* **9**, 1419 (2018).
 101. Luscombe, V. B., Lucy, D., Bataille, C. J. R., Russell, A. J. & Greaves, D. R. 20 Years an Orphan: Is GPR84 a Plausible Medium-Chain Fatty Acid-Sensing Receptor? *DNA Cell Biol.* **39**, 1926–1937 (2020).
 102. Amiri Moghaddam, J. *et al.* Cyclopropane-Containing Fatty Acids from the Marine Bacterium *Labrenzia* sp. 011 with Antimicrobial and GPR84 Activity. *Mar. Drugs* **16**, 369 (2018).
 103. Pillaiyar, T. *et al.* Diindolylmethane Derivatives: Potent Agonists of the

- Immunostimulatory Orphan G Protein-Coupled Receptor GPR84. *J. Med. Chem.* **60**, 3636–3655 (2017).
104. Nicol, L. S. C. *et al.* The role of G-protein receptor 84 in experimental neuropathic pain. *J. Neurosci.* **35**, 8959–8969 (2015).
 105. Peters, A. *et al.* Natural biased signaling of hydroxycarboxylic acid receptor 3 and G protein-coupled receptor 84. *Cell Commun. Signal.* **18**, 31 (2020).
 106. Peters, A. *et al.* Hydroxycarboxylic acid receptor 3 and GPR84 – Two metabolite-sensing G protein-coupled receptors with opposing functions in innate immune cells. *Pharmacol. Res.* **176**, 106047 (2022).
 107. Bouchard, C., Pagé, J., Bédard, A., Tremblay, P. & Vallières, L. G protein-coupled receptor 84, a microglia-associated protein expressed in neuroinflammatory conditions. *Glia* **55**, 790–800 (2007).
 108. Cha, S. Bin *et al.* Early transcriptional responses of internalization defective *Brucella abortus* mutants in professional phagocytes, RAW 264.7. *BMC Genomics* **14**, 426 (2013).
 109. Nagasaki, H. *et al.* Inflammatory changes in adipose tissue enhance expression of GPR84, a medium-chain fatty acid receptor: TNF α enhances GPR84 expression in adipocytes. *FEBS Lett.* **586**, 368–372 (2012).
 110. Venkataraman, C. & Kuo, F. The G-protein coupled receptor, GPR84 regulates IL-4 production by T lymphocytes in response to CD3 crosslinking. *Immunol. Lett.* **101**, 144–153 (2005).
 111. Yin, C. *et al.* Regulatory role of Gpr84 in the switch of alveolar macrophages from CD11b^{lo} to CD11b^{hi} status during lung injury process. *Mucosal Immunol.* **13**, 892–907 (2020).
 112. Abdel-Aziz, H. *et al.* GPR84 and TREM-1 signaling contribute to the pathogenesis of reflux esophagitis. *Mol. Med.* **21**, 1011–1024 (2015).
 113. Wei, L., Tokizane, K., Konishi, H., Yu, H. R. & Kiyama, H. Agonists for G-protein-coupled receptor 84 (GPR84) alter cellular morphology and motility but do not induce pro-inflammatory responses in microglia. *J. Neuroinflammation* **14**, (2017).
 114. Gagnon, L. *et al.* A Newly Discovered Antifibrotic Pathway Regulated by Two Fatty Acid Receptors. *Am. J. Pathol.* **188**, 1132–1148 (2018).
 115. Didangelos, A. COVID-19 Hyperinflammation: What about Neutrophils? *mSphere* **5**, 300

References

- (2020).
116. Lu, R. J., Shirvani, P. & Holick, M. F. A Novel Immunomodulatory Mechanism by Which Vitamin D Influences Folate Receptor 3 Expression to Reduce COVID-19 Severity. *Anticancer Res.* **42**, 5043–5048 (2022).
 117. Lai, Y. *et al.* Comprehensive Analysis of Molecular Subtypes and Hub Genes of Sepsis by Gene Expression Profiles. *Front. Genet.* **13**, 1672 (2022).
 118. Ye, N. *et al.* A multi-omic approach reveals utility of CD45 expression in prognosis and novel target discovery. *Front. Genet.* **13**, 2179 (2022).
 119. Madar, I. H. *et al.* Identification of marker genes in Alzheimer's disease using a machine-learning model. *Bioinformatics* **17**, 348 (2021).
 120. Ji, H. *et al.* Novel macrophage-related gene prognostic index for glioblastoma associated with M2 macrophages and T cell dysfunction. *Front. Immunol.* **13**, 5192 (2022).
 121. Widmayer, P., Kusumakshi, S., Hägele, F. A., Boehm, U. & Breer, H. Expression of the Fatty Acid Receptors GPR84 and GPR120 and Cytodifferentiation of Epithelial Cells in the Gastric Mucosa of Mouse Pups in the Course of Dietary Transition. *Front. Physiol.* **8**, 601 (2017).
 122. Suzuki, M. *et al.* Medium-chain Fatty Acid-sensing Receptor, GPR84, Is a Proinflammatory Receptor. *J. Biol. Chem.* **288**, 10684–10691 (2013).
 123. Müller, M. M. *et al.* Global analysis of glycoproteins identifies markers of endotoxin tolerant monocytes and GPR84 as a modulator of TNF α expression. *Sci. Rep.* **7**, 838 (2017).
 124. Puengel, T. *et al.* The Medium-Chain Fatty Acid Receptor GPR84 Mediates Myeloid Cell Infiltration Promoting Steatohepatitis and Fibrosis. *J. Clin. Med.* 2020, Vol. 9, Page 1140 **9**, 1140 (2020).
 125. Wang, M., Zhang, X., Zhang, S. & Liu, Z. Zebrafish fatty acids receptor Gpr84 enhances macrophage phagocytosis. *Fish Shellfish Immunol.* **84**, 1098–1099 (2019).
 126. Fredriksson, J. *et al.* GRK2 selectively attenuates the neutrophil NADPH-oxidase response triggered by β -arrestin recruiting GPR84 agonists. *Biochim. Biophys. Acta - Mol. Cell Res.* **1869**, 119262 (2022).
 127. Nguyen, Q. T. *et al.* PBI-4050 reduces pulmonary hypertension, lung fibrosis, and right ventricular dysfunction in heart failure. *Cardiovasc. Res.* **116**, 171–182 (2020).

128. Gagnon, L. *et al.* FP266PBI-4050 REDUCES SYSTEMIC INFLAMMATION, ELECTROLYTE DISTURBANCES, AND RENAL INJURY IN MICE WITH SEPSIS-INDUCED ACUTE KIDNEY INJURY; ROLE OF GPR84. *Nephrol. Dial. Transplant.* **34**, (2019).
129. Simard, J.-C. *et al.* Fatty acid mimetic PBI-4547 restores metabolic homeostasis via GPR84 in mice with non-alcoholic fatty liver disease. *Sci. Rep.* **10**, 12778 (2020).
130. Labéguère, F. *et al.* Discovery of 9-Cyclopropylethynyl-2-((S)-1-[1,4]dioxan-2-ylmethoxy)-6,7-dihydropyrimido[6,1- a]isoquinolin-4-one (GLPG1205), a Unique GPR84 Negative Allosteric Modulator Undergoing Evaluation in a Phase II Clinical Trial. *J. Med. Chem.* **63**, 13526–13545 (2020).
131. Chen, L. H., Zhang, Q., Xie, X. & Nan, F. J. Modulation of the G-Protein-Coupled Receptor 84 (GPR84) by Agonists and Antagonists. *Journal of Medicinal Chemistry* vol. 63 15399–15409 (2020).
132. Zhang, Q. *et al.* GPR84 signaling promotes intestinal mucosal inflammation via enhancing NLRP3 inflammasome activation in macrophages. *Acta Pharmacol. Sin.* **43**, 2042 (2022).
133. Lambert, S. A. *et al.* The Human Transcription Factors. *Cell* **172**, 650–665 (2018).
134. Mitsis, T. *et al.* Transcription factors and evolution: An integral part of gene expression (Review). *World Acad. Sci. J.* **2**, 3–8 (2020).
135. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* **167**, 1170–1187 (2016).
136. Yoshida, H. *et al.* The cis-Regulatory Atlas of the Mouse Immune System. *Cell* **176**, 897-912.e20 (2019).
137. Enhancer Networks. <http://rstats.immgen.org/EnhancerControl/>.
138. Dass, C. *Fundamentals of Contemporary Mass Spectrometry. Fundamentals of Contemporary Mass Spectrometry* (John Wiley & Sons, Inc., 2007). doi:10.1002/0470118490.
139. Han, J., Gagnon, S., Eckle, T. & Borchers, C. H. Metabolomic analysis of key central carbon metabolism carboxylic acids as their 3-nitrophenylhydrazones by UPLC/ESI-MS. *Electrophoresis* **34**, n/a-n/a (2013).
140. Bligh, E.G. and Dyer, W. J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, (1959).
141. Wishart, D. S. *et al.* HMDB 4.0: The human metabolome database for 2018. *Nucleic*

References

- Acids Res.* **46**, D608–D617 (2018).
142. Chong, J. *et al.* MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **46**, W486–W494 (2018).
143. Toni, L. S. *et al.* Optimization of phenol-chloroform RNA extraction. *MethodsX* **5**, 599–608 (2018).
144. Sun, Y. & O’Riordan, M. X. D. Regulation of bacterial pathogenesis by intestinal short-chain Fatty acids. *Adv. Appl. Microbiol.* **85**, 93–118 (2013).
145. Maslowski, K. M. *et al.* Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* **461**, 1282–1286 (2009).
146. Kelly, C. J. *et al.* Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function. *Cell Host Microbe* **17**, 662–671 (2015).
147. Psichas, A. *et al.* The short chain fatty acid propionate stimulates GLP-1 and PYY secretion via free fatty acid receptor 2 in rodents. *Int. J. Obes.* **39**, 424–429 (2015).
148. Tan, J. K., McKenzie, C., Mariño, E., Macia, L. & Mackay, C. R. Metabolite-Sensing G Protein–Coupled Receptors—Facilitators of Diet-Related Immune Regulation. *Annu. Rev. Immunol.* **35**, 371–402 (2017).
149. Vici, C. D., Bachmann, C., Gradwoht, M. & Colombo, J. P. Determination of medium chain fatty acids in serum. *Clin. Chim. Acta* **172**, 233–238 (1988).
150. Shrestha, R. *et al.* Plasma capric acid concentrations in healthy subjects determined by high-performance liquid chromatography. *Ann. Clin. Biochem.* **52**, 588–596 (2015).
151. Sills, M. A. & Forsythe, W. I. Role of octanoic and decanoic acids in the control of seizures. *Arch. Dis. Child.* **61**, 1173–1177 (1986).
152. Staggars, J. E., Fernando-Warnakulasuriya, G. J. & Wells, M. A. Studies on fat digestion, absorption, and transport in the suckling rat. II. Triacylglycerols: molecular species, stereospecific analysis, and specificity of hydrolysis by lingual lipase. *J. Lipid Res.* **22**, 675–9 (1981).
153. Dehaven, C. D., Evans, A. M., Dai, H. & Lawton, K. A. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J. Cheminform.* **2**, 1–12 (2010).
154. Alder, L., Greulich, K., Kempe, G. & Vieth, B. Residue analysis of 500 high priority pesticides: Better by GC–MS or LC–MS/MS? *Mass Spectrom. Rev.* **25**, 838–865

(2006).

155. Perez, E. R. *et al.* Comparison of LC–MS–MS and GC–MS Analysis of Benzodiazepine Compounds Included in the Drug Demand Reduction Urinalysis Program. *J. Anal. Toxicol.* **40**, 201 (2016).
156. Santa, T. Derivatization in LC-MS Bioanalysis. in *Handbook of LC-MS Bioanalysis* 239–248 (John Wiley & Sons Inc., 2013). doi:10.1002/9781118671276.ch19.
157. Han, J., Lin, K., Sequeira, C. & Borchers, C. H. An isotope-labeled chemical derivatization method for the quantitation of short-chain fatty acids in human feces by liquid chromatography–tandem mass spectrometry. *Anal. Chim. Acta* **854**, 86–94 (2015).
158. Chan, J. C. Y., Kioh, D. Y. Q., Yap, G. C., Lee, B. W. & Chan, E. C. Y. A novel LCMSMS method for quantitative measurement of short-chain fatty acids in human stool derivatized with ¹²C- and ¹³C-labelled aniline. *J. Pharm. Biomed. Anal.* **138**, 43–53 (2017).
159. Zeng, M. & Cao, H. Fast quantification of short chain fatty acids and ketone bodies by liquid chromatography-tandem mass spectrometry after facile derivatization coupled with liquid-liquid extraction. *J. Chromatogr. B* **1083**, 137–145 (2018).
160. Meierhofer, D. Acylcarnitine profiling by low-resolution LC-MS. *PLoS One* **14**, e0221342 (2019).
161. Stone, J. A practical guide to sample preparation for liquid chromatography-tandem mass spectrometry in clinical research and toxicology. *Spectrosc. Eur.* **30**, 15–21 (2018).
162. Han, J. *et al.* Isotope-labeling derivatization with 3-nitrophenylhydrazine for LC/multiple-reaction monitoring-mass-spectrometry-based quantitation of carnitines in dried blood spots. *Anal. Chim. Acta* **1037**, 177–187 (2018).
163. Liebisch, G. *et al.* Quantification of fecal short chain fatty acids by liquid chromatography tandem mass spectrometry—investigation of pre-analytic stability. *Biomolecules* **9**, (2019).
164. Bruegel, M. *et al.* Sepsis-associated changes of the arachidonic acid metabolism and their diagnostic potential in septic patients. *Crit. Care Med.* **40**, 1478–1486 (2012).
165. Jaurila, H. *et al.* 1H NMR Based Metabolomics in Human Sepsis and Healthy Serum. *Metab. 2020, Vol. 10, Page 70* **10**, 70 (2020).

References

166. Mickiewicz, B. *et al.* Integration of metabolic and inflammatory mediator profiles as a potential prognostic approach for septic shock in the intensive care unit. *Crit. Care* **19**, (2015).
167. Grauslys, A. *et al.* Title NMR-based metabolic profiling provides diagnostic and prognostic information in critically ill children with suspected infection. *Sci. Rep.* **10**, (2020).
168. Rogers, A. J. *et al.* Metabolomic Derangements Are Associated with Mortality in Critically Ill Adult Patients. *PLoS One* **9**, e87538 (2014).
169. Reinke, S. N., Chaleckis, R. & Wheelock, C. E. Metabolomics in pulmonary medicine - extracting the most from your data. *Eur. Respir. J.* 2200102 (2022) doi:10.1183/13993003.00102-2022.
170. Piehowski, P. D. *et al.* Sources of Technical Variability in Quantitative LC-MS Proteomics: Human Brain Tissue Sample Analysis. *J. Proteome Res.* **12**, 2128 (2013).
171. Sveshnikova, N., Yuan, T., Warren, J. M. & Piercey-Normore, M. D. Development and validation of a reliable LC-MS/MS method for quantitative analysis of usnic acid in *Cladonia uncialis*. *BMC Res. Notes* **12**, 1–6 (2019).
172. Mendes, V. M., Coelho, M., Tomé, A. R., Cunha, R. A. & Manadas, B. Validation of an LC-MS/MS Method for the quantification of caffeine and theobromine using non-matched matrix calibration curve. *Molecules* **24**, (2019).
173. O'Donnell, V. B. *et al.* Steps Toward Minimal Reporting Standards for Lipidomics Mass Spectrometry in Biomedical Research Publications. *Circ. Genomic Precis. Med.* **13**, e003019 (2020).
174. Liebisch, G. *et al.* Lipidomics needs more standardization. *Nat. Metab.* 2019 18 **1**, 745–747 (2019).
175. Seymour, C. W. *et al.* Temperature and time stability of whole blood lactate: Implications for feasibility of pre-hospital measurement. *BMC Res. Notes* **4**, 1–4 (2011).
176. Andersen, O. *et al.* Preanalytical handling of samples for measurement of plasma lactate in HIV patients. *Scand. J. Clin. Lab. Invest.* **63**, 449–454 (2003).
177. Hashim, I. A., Mohamed, M., Cox, A., Fernandez, F. & Kutscher, P. Plasma lactate measurement as an example of encountered gaps between routine clinical laboratory processes and manufactures' sample-handling instructions. *Pract. Lab.*

- Med.* **12**, e00109 (2018).
178. Crotti, S. *et al.* Altered plasma levels of decanoic acid in colorectal cancer as a new diagnostic biomarker. *Anal. Bioanal. Chem.* **408**, 6321–6328 (2016).
 179. Ruiz-Perez, D., Guan, H., Madhivanan, P., Mathee, K. & Narasimhan, G. So you think you can PLS-DA? *BMC Bioinformatics* **21**, 1–10 (2020).
 180. Köse, M. *et al.* An Agonist Radioligand for the Proinflammatory Lipid-Activated G Protein-Coupled Receptor GPR84 Providing Structural Insights. *J. Med. Chem.* **63**, 2391–2410 (2020).
 181. Southern, C. *et al.* Screening β -Arrestin Recruitment for the Identification of Natural Ligands for Orphan G-Protein–Coupled Receptors. *J. Biomol. Screen.* **18**, 599–609 (2013).
 182. Nikaido, Y., Koyama, Y., Yoshikawa, Y., Furuya, T. & Takeda, S. Mutation analysis and molecular modeling for the investigation of ligand-binding modes of GPR84. *J. Biochem.* **157**, 311–320 (2015).
 183. Amunugama, K., Pike, D. P. & Ford, D. A. The lipid biology of sepsis. *J. Lipid Res.* **62**, (2021).
 184. Rival, T. *et al.* Alteration of plasma phospholipid fatty acid profile in patients with septic shock. *Biochimie* **95**, 2177–2181 (2013).
 185. Wang, C., Han, D., Feng, X. & Wu, J. Omega-3 fatty acid supplementation is associated with favorable outcomes in patients with sepsis: an updated meta-analysis. *J. Int. Med. Res.* **48**, (2020).
 186. Svahn, S. L. *et al.* Dietary Polyunsaturated Fatty Acids Increase Survival and Decrease Bacterial Load during Septic Staphylococcus aureus Infection and Improve Neutrophil Function in Mice. *Infect. Immun.* **83**, 514 (2015).
 187. Schooneman, M. G., Achterkamp, N., Argmann, C. A., Soeters, M. R. & Houten, S. M. Plasma acylcarnitines inadequately reflect tissue acylcarnitine metabolism. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1841**, 987–994 (2014).
 188. Puskarich, M. A. *et al.* Septic Shock Nonsurvivors Have Persistently Elevated Acylcarnitines Following Carnitine Supplementation. *Shock* **49**, 412–419 (2018).
 189. Chou, E. H. *et al.* Incidence, trends, and outcomes of infection sites among hospitalizations of sepsis: A nationwide study. *PLoS One* **15**, e0227752 (2020).
 190. Chakraborty, M. *et al.* nSeP: immune and metabolic biomarkers for early detection

References

- of neonatal sepsis—protocol for a prospective multicohort study. *BMJ Open* **11**, e050100 (2021).
191. Kauppi, A. M. *et al.* Metabolites in Blood for Prediction of Bacteremic Sepsis in the Emergency Room. *PLoS One* **11**, e0147670 (2016).
 192. Chung, K. P. *et al.* Increased Plasma Acetylcarnitine in Sepsis Is Associated With Multiple Organ Dysfunction and Mortality: A Multicenter Cohort Study. *Crit. Care Med.* **47**, 210–218 (2019).
 193. Stringer, K. A. *et al.* Whole Blood Reveals More Metabolic Detail of the Human Metabolome than Serum as Measured by ¹H-NMR Spectroscopy: Implications for Sepsis Metabolomics. *Shock* **44**, 200–208 (2015).
 194. Peng, Y. *et al.* Fatty acid composition of diet, cord blood and breast milk in Chinese mothers with different dietary habits. *Prostaglandins Leukot. Essent. Fat. Acids* **81**, 325–330 (2009).
 195. Tullai, J. W. *et al.* Immediate-Early and Delayed Primary Response Genes Are Distinct in Function and Genomic Architecture*. *J. Biol. Chem.* **282**, 23981–23995 (2007).
 196. Ochsner, S. A. *et al.* The Signaling Pathways Project, an integrated ‘omics knowledgebase for mammalian cellular signaling pathways. *Sci. Data* **6**, 1–21 (2019).
 197. Myers, R. M. *et al.* A User’s Guide to the Encyclopedia of DNA Elements (ENCODE). *PLoS Biol.* **9**, e1001046 (2011).
 198. Rouillard, A. D. *et al.* The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* **2016**, baw100 (2016).
 199. Chèneby, J. *et al.* ReMap 2020: A database of regulatory regions from an integrative analysis of Human and Arabidopsis DNA-binding sequencing experiments. *Nucleic Acids Res.* **48**, D180–D188 (2020).
 200. Alexandra B Keenan, Denis Torre, Alexander Lachmann, Ariel K Leong, Megan L Wojciechowicz, Vivian Utti, Kathleen M Jagodnik, Eryk Kropiwnicki, Zichen Wang, A. M. ChEA3: transcription factor enrichment analysis by orthogonal omics integration | Nucleic Acids Research | Oxford Academic. <https://academic.oup.com/nar/article/47/W1/W212/5494769>.
 201. Lachmann, A. *et al.* Massive mining of publicly available RNA-seq data from human and mouse. *Nat. Commun.* **9**, 1–10 (2018).

202. JASPAR - a database of transcription factor binding profiles. <http://jaspar.genereg.net/>.
203. Fornes, O. *et al.* JASPAR 2020: Update of the open-Access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
204. Jayaram, N., Usvyat, D. & Martin, A. C. Evaluating tools for transcription factor binding site prediction. *BMC Bioinformatics* **17**, 547 (2016).
205. Kim-Hellmuth, S. *et al.* Cell type–specific genetic regulation of gene expression across human tissues. *Science (80-)*. **369**, eaaz8528 (2020).
206. Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn - A web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488 (2008).
207. Szklarczyk, D. *et al.* STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
208. Navarro Gonzalez, J. *et al.* The UCSC genome browser database: 2021 update. *Nucleic Acids Res.* **49**, D1046–D1057 (2021).
209. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
210. Ramirez-Carrozzi, V. R. *et al.* A Unifying Model for the Selective Regulation of Inducible Transcription by CpG Islands and Nucleosome Remodeling. *Cell* **138**, 114–128 (2009).
211. UCSC Genome Browser Home. <http://genome.ucsc.edu/>.
212. Farré, D. *et al.* Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res.* **31**, 3651–3653 (2003).
213. Al Ait, L., Yamak, Z. & Morgenstern, B. DIALIGN at GOBICS--multiple sequence alignment using various sources of external information. *Nucleic Acids Res.* **41**, W3–W7 (2013).
214. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
215. Vandewalle, J. & Libert, C. Glucocorticoids in Sepsis: To Be or Not to Be. *Frontiers in Immunology* vol. 11 1318 (2020).

References

216. Sam, S., Corbridge, T. C., Mokhlesi, B., Comellas, A. P. & Molitch, M. E. Cortisol levels and mortality in severe sepsis. *Clin. Endocrinol. (Oxf)*. **60**, 29–35 (2004).
217. Hurst, L. D., Sachenkova, O., Daub, C., Forrest, A. R. R. & Huminiecki, L. A simple metric of promoter architecture robustly predicts expression breadth of human genes suggesting that most transcription factors are positive regulators. *Genome Biol.* **15**, 413 (2014).
218. Wingender, E., Schoeps, T., Haubrock, M., Krull, M. & Dönitz, J. TFClass: Expanding the classification of human transcription factors to their mammalian orthologs. *Nucleic Acids Res.* **46**, D343–D347 (2018).
219. Suryamohan, K. & Halfon, M. S. Identifying transcriptional cis -regulatory modules in animal genomes. *Wiley Interdiscip. Rev. Dev. Biol.* **4**, 59–84 (2015).
220. Fishilevich, S. *et al.* GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*. **2017**, (2017).
221. Escoubet-Lozach, L. *et al.* Mechanisms Establishing TLR4-Responsive Activation States of Inflammatory Response Genes. *PLoS Genet.* **7**, e1002401 (2011).
222. Yang, C., Bolotin, E., Jiang, T., Sladek, F. M. & Martinez, E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* **389**, 52–65 (2007).
223. Ghisletti, S. *et al.* Identification and Characterization of Enhancers Controlling the Inflammatory Gene Expression Program in Macrophages. *Immunity* **32**, 317–328 (2010).
224. Kuznetsova, T. *et al.* Glucocorticoid receptor and nuclear factor kappa-b affect three-dimensional chromatin organization. *Genome Biol.* **16**, 264 (2015).
225. Hermoso, M. A. & Cidlowski, J. A. Putting the Brake on Inflammatory Responses: The Role of Glucocorticoids. *IUBMB Life* **55**, 497–504 (2003).
226. Fernando-Warnakulasuriya, G. J., Staggers, J. E., Frost, S. C. & Wells, M. A. Studies on fat digestion, absorption, and transport in the suckling rat. I. Fatty acid composition and concentrations of major lipid components. *J. Lipid Res.* **22**, 668–74 (1981).
227. Libertini, L. J., Chu Yuan Lin, Abraham, S., Hilf, R. & Smith, S. Medium chain fatty acid synthesis in rodent mammary adenocarcinomas in vitro. *Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab.* **618**, 185–191 (1980).
228. Spear, M. L. *et al.* Human mammary gland function at the onset of lactation:

- medium-chain fatty acid synthesis. *Lipids* **27**, 908–911 (1992).
229. Hachey, D. L., Silber, G. H., Wong, W. W. & Garza, C. Human Lactation II: Endogenous Fatty Acid Synthesis by the Mammary Gland. *Pediatr. Res.* **25**, 63–68 (1989).
230. Al-Tamer, Y. Y. & Mahmood, A. Fatty-acid composition of the colostrum and serum of fullterm and preterm delivering Iraqi mothers. *Eur. J. Clin. Nutr.* **58**, 1119–1124 (2004).
231. Mohammad, M. A., Sunehag, A. L. & Haymond, M. W. De novo synthesis of milk triglycerides in humans. *Am. J. Physiol. Metab.* **306**, E838–E847 (2014).
232. Thompson, B. J. & Smith, S. Biosynthesis of fatty acids by lactating human breast epithelial cells: An evaluation of the contribution to the overall composition of human milk fat. *Pediatr. Res.* **19**, 139–143 (1985).
233. Mosley, E. E., Wright, A. L., McGuire, M. K. & McGuire, M. A. trans Fatty acids in milk produced by women in the United States. *Am. J. Clin. Nutr.* **82**, 1292–1297 (2005).
234. Matsumoto, M. *et al.* Impact of intestinal microbiota on intestinal luminal metabolome. *Sci. Rep.* **2**, (2012).
235. Steimle, A. & Autenrieth, I. B. Structure and function: Lipid A modifications in commensals and pathogens. *Int. J. Med. Microbiol.* **306**, 290–301 (2016).
236. Scarborough, M. J., Lawson, C. E., Hamilton, J. J., Donohue, T. J. & Noguera, D. R. Metatranscriptomic and Thermodynamic Insights into Medium-Chain Fatty Acid Production Using an Anaerobic Microbiome. *mSystems* **3**, (2018).
237. Chandler, C. E. & Ernst, R. K. Bacterial lipids: powerful modifiers of the innate immune response. *F1000Research* **6**, (2017).
238. Lass, A., Zimmermann, R., Oberer, M. & Zechner, R. Lipolysis – A highly regulated multi-enzyme complex mediates the catabolism of cellular fat stores. *Prog. Lipid Res.* **50**, 14–27 (2011).
239. Papatheodorou, I. *et al.* Expression Atlas update: From tissues to single cells. *Nucleic Acids Res.* **48**, D77–D83 (2020).
240. Varga, T., Czimmerer, Z. & Nagy, L. PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation. *Biochimica et Biophysica Acta - Molecular Basis of Disease* vol. 1812 1007–1022 (2011).
241. Jiang, D. *et al.* Potential biomarkers screening to predict side effects of dexamethasone in different cancers. *Mol. Genet. Genomic Med.* (2020)

References

- doi:10.1002/mgg3.1160.
242. Wong, H. R. *et al.* Corticosteroids Are Associated with Repression of Adaptive Immunity Gene Programs in Pediatric Septic Shock. *Am. J. Respir. Crit. Care Med.* **189**, 940–946 (2014).
 243. Xue, J. *et al.* Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity* **40**, 274–288 (2014).
 244. Waddell, S. J. *et al.* Dissecting interferon-induced transcriptional programs in human peripheral blood cells. *PLoS One* **5**, (2010).
 245. Bosshart, H. & Heinzelmann, M. THP-1 cells as a model for human monocytes. *Annals of Translational Medicine* vol. 4 (2016).
 246. Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **54**, 1.30.1-1.30.33 (2016).
 247. Sawai, M. *et al.* The 3-hydroxyacyl-CoA dehydratases HACD1 and HACD2 exhibit functional redundancy and are active in a wide range of fatty acid elongation pathways. *J. Biol. Chem.* **292**, 15538 (2017).
 248. Sassa, T. & Kihara, A. Metabolism of Very Long-Chain Fatty Acids: Genes and Pathophysiology. *Biomol. Ther. (Seoul)*. **22**, 83 (2014).
 249. Libertini, L. J. & Smith, S. Purification and properties of a thioesterase from lactating rat mammary gland which modifies the product specificity of fatty acid synthetase. *J. Biol. Chem.* **253**, 1393–1401 (1978).
 250. Ritchie, M. K. *et al.* Crystal Structure and Substrate Specificity of Human Thioesterase 2: INSIGHTS INTO THE MOLECULAR BASIS FOR THE MODULATION OF FATTY ACID SYNTHASE. *J. Biol. Chem.* **291**, 3520–3530 (2016).
 251. Banerjee, S., Mohammed, A., Wong, H. R., Palaniyar, N. & Kamaleswaran, R. Machine Learning Identifies Complicated Sepsis Course and Subsequent Mortality Based on 20 Genes in Peripheral Blood Immune Cells at 24 H Post-ICU Admission. *Front. Immunol.* **12**, (2021).
 252. Basu, R. K. *et al.* Identification of candidate serum biomarkers for severe septic shock-associated kidney injury via microarray. *Crit. Care* **15**, (2011).
 253. Violante, S. *et al.* Peroxisomes can oxidize medium- and long-chain fatty acids through a pathway involving ABCD3 and HSD17B4. *FASEB J.* **33**, 4355–4364 (2019).
 254. Xu, X. *et al.* Identification of differentially expressed genes associated with burn

- sepsis using microarray. *Int. J. Mol. Med.* **36**, 1623–1629 (2015).
255. de Azambuja Rodrigues, P. M. *et al.* Proteomics reveals disturbances in the immune response and energy metabolism of monocytes from patients with septic shock. *Sci. Reports* **2021 111** **11**, 1–10 (2021).
256. Violante, S. *et al.* Carnitine palmitoyltransferase 2 and carnitine/acylcarnitine translocase are involved in the mitochondrial synthesis and export of acylcarnitines. *FASEB J.* **27**, 2039–2044 (2013).
257. Jiang, Y. *et al.* Exploring the Muscle Metabolomics in the Mouse Model of Sepsis-Induced Acquired Weakness. *Evid. Based. Complement. Alternat. Med.* **2022**, (2022).
258. Chung, K.-P. *et al.* Increased Plasma Acetylcarnitine in Sepsis Is Associated With Multiple Organ Dysfunction and Mortality. *Crit. Care Med.* **47**, 210–218 (2019).
259. Thompson, D. K. *et al.* Daily Variation of Serum Acylcarnitines and Amino Acids. *Metabolomics* **8**, 556 (2012).
260. Costa, C. C. G., Tavares De Almeida, I., Jakobs, C., Poll-The, B. T. & Duran, M. Dynamic Changes of Plasma Acylcarnitine Levels Induced by Fasting and Sunflower Oil Challenge Test in Children. *Pediatr. Res.* **1999 464** **46**, 440–440 (1999).
261. J. Lawrence Merritt, I., Norris, M. & Kanungo, S. Fatty acid oxidation disorders. *Ann. Transl. Med.* **6**, 181–183 (2018).
262. Halldin, M. U., Forslund, A., von Döbeln, U., Eklund, C. & Gustafsson, J. Increased lipolysis in LCHAD deficiency. *J. Inherit. Metab. Dis.* **30**, 39–46 (2007).
263. Elizondo, G., Matern, D., Vockley, J., Harding, C. O. & Gillingham, M. B. Effects of fasting, feeding and exercise on plasma acylcarnitines among subjects with CPT2D, VLCADD and LCHADD/TFPD. *Mol. Genet. Metab.* **131**, 90–97 (2020).
264. Wang, J., Sun, Y., Teng, S. & Li, K. Prediction of sepsis mortality using metabolite biomarkers in the blood: a meta-analysis of death-related pathways and prospective validation. *BMC Med.* **18**, 83 (2020).
265. Willenberg, I. *et al.* Characterization of changes in plasma and tissue oxylipin levels in LPS and CLP induced murine sepsis. *Inflamm. Res.* **65**, 133–142 (2016).
266. Ahmad, N. S., Tan, T. L., Arifin, K. T., Ngah, W. Z. W. & Yusof, Y. A. M. High sPLA2-IIA level is associated with eicosanoid metabolism in patients with bacterial sepsis syndrome. *PLoS One* **15**, e0230285 (2020).
267. Schmerler, D. *et al.* Targeted metabolomics for discrimination of systemic

References

- inflammatory disorders in critically ill patients. *J. Lipid Res.* **53**, 1369–75 (2012).
268. Montgomery, M. K. *et al.* Regulation of mitochondrial metabolism in murine skeletal muscle by the medium-chain fatty acid receptor Gpr84. *FASEB J.* **33**, 12264–12276 (2019).
269. Yaméogo, C. W. *et al.* Correlates of whole-blood polyunsaturated fatty acids among young children with moderate acute malnutrition. *Nutr. J.* **16**, 1–11 (2017).
270. Briscoe, C. P. *et al.* The orphan G protein-coupled receptor GPR40 is activated by medium and long chain fatty acids. *J. Biol. Chem.* **278**, 11303–11 (2003).
271. Cheff, V. *et al.* P0710DUAL GPR40/GPR84 FATTY ACID RECEPTOR DELETION IMPROVES ADENINE-INDUCED RENAL INJURY IN MICE. *Nephrol. Dial. Transplant.* **35**, (2020).
272. Lee, J. Y., Sohn, K. H., Rhee, S. H. & Hwang, D. Saturated Fatty Acids, but Not Unsaturated Fatty Acids, Induce the Expression of Cyclooxygenase-2 Mediated through Toll-like Receptor 4. *J. Biol. Chem.* **276**, 16683–16689 (2001).
273. Shi, H. *et al.* TLR4 links innate immunity and fatty acid-induced insulin resistance. *J. Clin. Invest.* **116**, 3015–3025 (2006).

10 Supplementary materials

Supplementary table 1 depicts the ChEA3 TF enrichment hits for paediatric sepsis and neonatal sepsis filtered for hits with a hit for GPR84 in one of the databases. The right column indicates the overlap in TF hits between the two datasets.

TF enriched paediatric sepsis with hit for GPR84	TF enriched neonatal sepsis with hit for GPR84	overlap
SPI1	SPI1	SPI1
HLX	MTF1	MTF1
MTF1	CEBPB	CEBPB
NFE2	HLX	HLX
CEBPB	BATF2	BATF2
ZNF467	ZNF467	ZNF467
ZNF438	NFE2	NFE2
PLSCR1	CEBPE	CEBPE
CEBPE	PLSCR1	PLSCR1
TFEC	TFEC	TFEC
ZNF267	RELB	RELB
BATF2	IRF2	IRF2
BATF	ZNF438	ZNF438
RELB	BATF	BATF
IKZF1	NFKB2	NFKB2
RUNX2	ZNF267	ZNF267
ATF3	IKZF1	IKZF1
RUNX1	CSRNP1	CSRNP1
CSRNP1	GFI1B	GFI1B
IRF2	GATA1	GATA1
FLI1	ATF3	ATF3
NFE2L2	STAT1	STAT1
GFI1B	NFKB1	NFKB1
GATA1	ETS2	ETS2
NFKB1	ZBTB17	RUNX1
NFKB2	RUNX1	NFE2L2

References

ZEB2	NFE2L2	PBX2
FOSL1	PBX2	ELF1
ZNF697	ELF1	TCF7
RFX8	EGR2	VDR
VDR	TCF7	RUNX2
STAT1	VDR	MSC
AHR	RUNX2	AHR
MSC	EGR1	SNAI1
ELK3	IRF3	FOSL1
SNAI1	MSC	ETV3L
PBX2	MAFB	KLF9
ETS2	AHR	RFX8
MITF	SNAI1	SP1
NR1H3	FOSL1	MITF
ELF1	ETV3L	KLF4
MITF	RFX5	ZEB2
NCOA1	KLF9	LEF1
TCF7	RFX8	CREB1
SP1	SP1	SPIC
ETV3L	MITF	ZNF697
KLF9	ZNF143	NR2F2
CUX1	KLF4	HOXB7
SPIC	ZEB2	
NR2F2	LEF1	
KLF3	MYC	
CREB1	NFYB	
SOX9	CREB1	
CREM	SPIC	
NR4A3	ZNF697	
MEIS1	NR2F2	
LEF1	SMAD1	
KLF4	HOXB7	
HOXB7	ESRRA	

TWIST1	NANOG	
FOXP1		

Supplementary table 2 TFs binding to the GPR84 regulatory region according to Chip hits from the ENCODE and Signalling Pathways Project (SPP) databases.

ENCODE hits for GPR84	hits for SPP hits all cells	SPP hits immune cells	Overlap ENCODE and all SPP hits
ATF2	FOS 4HT	CTCF	ATF2
BCLAF1	JUN DEX	RUNX1	CEBPB
CEBPB	CTCF	NFYB	CTCF
CHD4	E2F4 4HT	MYB	EBF1
CTCF	IRF1	STAT1 IFNG	FOXM1
EBF1	RUNX1	GABPA	IKZF1
EZH2	GATA3	ETS1	MAX
FOXM1	MAX	SPI1	MEF2A
H2AFZ	CEBPB	CEBPA	MTA3
HDAC2	NFYB	USF1	MYC
IKZF1	MYB	MAX	NFYB
MAFK	STAT1 IFNG	SP1	SPI1
MAX	GABPA	GATA1	STAT3
MEF2A	TCF12	STAT3	STAT5A
MTA3	CREB1	FOXM1	TBP
MYC	ETS1	STAT5A	ZNF384
MYOG	SPI1	E2F4	MAFK
NFIC	CEBPA	BHLHE40	
NFYB	HEYL	PAX5	
PML	ELF3	FOSL1	
POLR2A	USF1	YY1	

References

RBBP5	GATA4	IRF1 IFNG	
RUNX3	FOXA1	RELA BAY082	
SPI1	FOXA1 CREB1	TFAP4	
STAT1	JUND	MYC IFNA1	
STAT3	FOSL2	MAZ	
STAT5A	SP1	MXI1	
TAF1	TCF21	BCL3	
TBL1XR1	RELA	RELA TNF	
TBP	MXI1	SREBF2	
WRNIP1	MEF2A	HMG3	
ZNF384	GATA1	GATA2	
	ATF2	EBF1	
	STAT3	FOS	
	SREBF1 22RHC; INS	CEBPB	
	FOXM1	CREB1	
	STAT5A	TAL1	
	HIF1A	IKZF1	
	YY1 ATRA	SPI1 ROSI	
	E2F4	SMAD1	
	MYC	ELF1	
	MAZ	SRF	
	BHLHE40	POU2F2	
	ZBTB17	TCF12	
	ZNF280D	JUN IFNG	
	E2F1 BICALU	LYL1	
	PAX5	MEF2A	
	FOSL1	PBX3	
	FOS	TCF3	
	ZNF83	NFKB1	
	ELF1	ATF3	
	YY1	E2F6	

	JUN	MYC	
	SRF	IRF1 IFNA1	
	NANOG	ERG	
	IRF1 IFNG	JUN	
	RELA BAY082	NFYA	
	KLF11	REST	
	ZBTB7A	ATF1	
	NFYA	ZNF143	
	REST	RELA	
	TFAP4	TBP	
	MYC IFNA1	IRF4	
	E2F1 HES6	ZBTB7A	
	FOXH1	MYC IFNG	
	TEAD4	ATF2	
	ATF3	JUND	
	BCL3	UBTF	
	RELA TNF	ZNF384	
	TCF7L2	JUN IFNA1	
	SOX2	IRF3	
	SREBF2	HEY1	
	EGR1	SP2	
	MAFK	MTA3	
	USF1 ATRA	EGR1	
	FOXA1 17BE2	TEAD4	
	HMGN3	RUNX1T1	
	E2F1		
	GATA2		
	ARNT		
	EBF1		
	E2F6		
	GABPA R1881		

References

	TAL1		
	CREB1 FOXA1		
	KLF5		
	PBX3		
	SOX9		
	KLF11 ROSI		
	CREB1 DEX		
	FOXA2		
	IKZF1		
	TBP		
	SPI1 ROSI		
	FOXA1 DEX		
	SIX2		
	SMAD1		
	POU2F2		
	CTCF DEX		
	SP2		
	JUN IFNG		
	LYL1		
	FOXA1 R1881		
	E2F1 BICALU HES6		
	MYC DHT		
	TP53		
	ERG DHT		
	TCF3		
	JUNB		
	ZNF143		
	NFKB1		
	SMARCC1		
	FOXA1 DHT		
	GTF2B		

	PDX1		
	IRF1 IFNA1		
	ERG		
	OTX2		
	CDX2 125DVD3		
	GLI2		
	FOXP1 DHT		
	ZNF263		
	ATF1		
	ARID3A		
	IRF4		
	IRF3		
	STAT2		
	MYC 4HT		
	GTF2F1		
	FOXP1		
	HEY1		
	ZEB1		
	CEBPB FORSK		
	SIX5		
	MYC IFNG		
	ETV5		
	FOXP2		
	E2F3		
	UBTF		
	STAT3 4HT		
	ZNF384		
	BARX1		
	JUN IFNA1		
	MYBL2		
	MTA3		

References

	USF1 DEX		
	NFYC		
	CDX2		
	ZKSCAN1		
	HOXA6		
	MEIS1		
	RUNX1T1		
	HOXA4		

Supplementary table 3 List of all TFs identified in at least one analysis and columns indicating if a TF was found in this analysis or not.

Transcription factor	Chea3_all_hits	Chea3_GPR8 4hits	Cistromics_GPR 84hits	upregulated_TF_in_sepsis
SPI1	✓	✓	✓	✓
HLX	✓	✓		
MTF1	✓	✓		✓
NFE2	✓	✓		✓
CREB5	✓			
CEBPB	✓	✓	✓	✓
NFE4	✓			✓
ZNF467	✓	✓		
ZNF438	✓	✓		✓
PLSCR1	✓	✓		✓
CEBPE	✓	✓		
TFEC	✓	✓		
MXD1	✓			
ZNF746	✓			
NFIL3	✓			✓
ZNF267	✓	✓		
BCL6	✓			✓
JDP2	✓			
LYL1	✓		✓	
ELF4	✓			

BATF2	✓	✓		
TET2	✓			
CEBPA	✓		✓	✓
BATF	✓	✓		
STAT5A	✓		✓	
LTF	✓			✓
SP110	✓			
BACH1	✓			
FOXP2	✓			
BORCS8MEF2B	✓			
IRF7	✓			
IRF9	✓			
TRAFD1	✓			
STAT3	✓		✓	
ZNF641	✓			
TFE3	✓			
TFEB	✓			
IRF5	✓			
RELB	✓	✓		
IRF1	✓		✓	
TAL1	✓		✓	
ARID5A	✓			
IKZF1	✓	✓	✓	
CEBPD	✓			✓
FOSL2	✓		✓	
STAT5B	✓			
DNTTIP1	✓			
ETV6	✓			
RUNX2	✓	✓		
SNAI3	✓			
GTF2B	✓		✓	
ATF3	✓	✓	✓	
ATF6	✓			
RUNX1	✓	✓	✓	
HIF1A	✓		✓	
CSRNP1	✓	✓		

References

ZBTB34	✓		
JUNB	✓		✓
IRF2	✓	✓	
FLI1	✓	✓	
MAFG	✓		
NFE2L2	✓	✓	
GFI1B	✓	✓	
JAZF1	✓		
GLMP	✓		
GATA1	✓	✓	✓
PPARG	✓		
KLF2	✓		
NFKB1	✓	✓	✓
GFI1	✓		
TIGD3	✓		
PRDM1	✓		
HHEX	✓		
SP100	✓		
NFKB2	✓	✓	
RARA	✓		
ZNF319	✓		
USF1	✓		✓
BATF3	✓		
KLF1	✓		
RXRA	✓		
STAT4	✓		✓
AKNA	✓		
ZEB2	✓	✓	
IRF8	✓		
PHF21A	✓		
FOSL1	✓	✓	✓
MLX	✓		
ZNF697	✓	✓	
MAX	✓		✓
KLF6	✓		
RFX8	✓	✓	

VDR	✓	✓		
STAT1	✓	✓	✓	
MNT	✓			
TBX21	✓			
AHR	✓	✓		
NR4A1	✓			
MSC	✓	✓		
MXI1	✓		✓	
NR1H2	✓			
ELK3	✓	✓		
EGR1	✓		✓	
JUN	✓		✓	
POU2F2	✓		✓	
SNAI1	✓	✓		
CREBL2	✓			
TCF7L2	✓		✓	
ASCL2	✓			
TRERF1	✓			
ZNF394	✓			
MAFK	✓		✓	
PBX2	✓	✓		
FOSB	✓			
VENTX	✓			
FOXP3	✓			
ARNTL	✓			
MEF2A	✓		✓	
CREB3	✓			
ETS1	✓		✓	✓
REL	✓			
ETS2	✓	✓		✓
FOXO4	✓			
SMAD3	✓			
MAFF	✓			
ZNF217	✓			
MAFB	✓			
ARID3B	✓			

References

NR1H3	✓	✓	
ELF1	✓	✓	✓
DDIT3	✓		
ARID3A	✓		✓
BHLHE40	✓		✓
PPARD	✓		
MBD2	✓		
ZNF341	✓		
ZBTB7B	✓		
ARNTL2	✓		
RELA	✓		✓
ZBED1	✓		
MBNL2	✓		
RUNX3	✓		
TGIF1	✓		
ATF5	✓		
NAIF1	✓		
EPAS1	✓		
MITF	✓	✓	
L3MBTL3	✓		
ARID5B	✓		
STAT2	✓		✓
ZNF710	✓		
NR3C1	✓		
NFATC1	✓		
ZBTB48	✓		
SP140	✓		
C11ORF95	✓		
NCOA1	✓	✓	
STAT6	✓		
ZNF346	✓		
FOXO3	✓		
FOS	✓		✓
EHF	✓		
MYB	✓		✓
ZNF792	✓		

TSHZ3	✓			
CREB3L2	✓			
TCF7	✓	✓		✓
TWIST2	✓			
SP140L	✓			
MEF2B	✓			
ZNF276	✓			
ZNF281	✓			
MBD4	✓			
SPIB	✓			
ERG	✓		✓	
SP1	✓	✓	✓	
JUND	✓		✓	
NKX23	✓			
ETV3L	✓	✓		
KLF9	✓	✓		
NRF1	✓			
IRF4	✓		✓	
RAG1	✓			
MTERF2	✓			
KLF10	✓			
CUX1	✓	✓		
LIN54	✓			
ZBTB10	✓			
PBX1	✓			
ZEB1	✓		✓	
SPIC	✓	✓		
NR2F2	✓	✓		
KLF13	✓			
ZBED2	✓			
ZNF18	✓			
SP2	✓		✓	
GATA2	✓		✓	
KLF3	✓	✓		
ZBTB49	✓			
ZNF331	✓			

References

CREB1	✓	✓	✓	
HBP1	✓			
SRF	✓		✓	
ZNF586	✓			
RBPJ	✓			
ZNF75D	✓			
USF2	✓			
NFYC	✓		✓	
ZNF117	✓			
ZNF469	✓			
ZNF672	✓			
SOX9	✓	✓	✓	
CREM	✓	✓		
NR4A3	✓	✓		
SREBF1	✓		✓	
ZHX1	✓			
KLF11	✓		✓	
FOXJ2	✓			
ZNF274	✓			
MEIS1	✓	✓	✓	
ELF3	✓		✓	
LEF1	✓	✓		✓
ZNF516	✓			
GZF1	✓			
SATB1	✓			✓
SKIL	✓			
KLF4	✓	✓		
HOXB7	✓	✓		
ZBTB22	✓			
ZNF296	✓			
SNAI2	✓			
HMGN3	✓		✓	✓
TWIST1	✓	✓		
ZNF462	✓			
NPAS2	✓			
ZNF860	✓			

SNAPC2	✓			
ZNF524	✓			
RLF	✓			
ZNF410	✓			
FOXO1	✓			
TCF12	✓		✓	
MEF2C	✓			✓
ZNF608	✓			
AR	✓			
FOXP1	✓	✓	✓	
ZNF443	✓			
GMEB2	✓			
CREB3L1	✓			
GLIS3	✓			
CTCF			✓	
E2F4			✓	
GATA3			✓	✓
NFYB			✓	
GABPA			✓	
HEYL			✓	
GATA4			✓	
FOXA1			✓	
TCF21			✓	
ATF2			✓	
FOXM1			✓	
YY1			✓	
MYC			✓	
MAZ			✓	
ZBTB17			✓	
ZNF280D			✓	
E2F1			✓	
PAX5			✓	
ZNF83			✓	✓
NANOG			✓	
ZBTB7A			✓	
NFYA			✓	

References

REST	✓
TFAP4	✓
FOXH1	✓
TEAD4	✓
BCL3	✓
SOX2	✓
SREBF2	✓
ARNT	✓
EBF1	✓
E2F6	✓
KLF5	✓
PBX3	✓
FOXA2	✓
TBP	✓
SIX2	✓
SMAD1	✓
TP53	✓
TCF3	✓
ZNF143	✓
SMARCC1	✓
PDX1	✓
OTX2	✓
CDX2	✓
GLI2	✓
ZNF263	✓
ATF1	✓
IRF3	✓
GTF2F1	✓
HEY1	✓
SIX5	✓
ETV5	✓
FOXP2	✓
E2F3	✓
UBTF	✓
ZNF384	✓
BARX1	✓

MYBL2	✓	
MTA3	✓	
ZKSCAN1	✓	
HOXA6	✓	
RUNX1T1	✓	
HOXA4	✓	
IFNG	✓	
IFNA1	✓	
HES6	✓	
BACH2		✓
BCL11A		✓
BCL11B		✓
CXXC5		✓
DACH1		✓
DNMT1		✓
E2F5		✓
EOMES		✓
KDM2B		✓
MYBL1		✓
NR1D2		✓
PRDM5		✓
RORA		✓
TSHZ1		✓
ZBTB41		✓
ZFP3		✓
ZFP62		✓
ZNF121		✓
ZNF260		✓
ZNF266		✓
ZNF329		✓
ZNF420		✓
ZNF439		✓
ZNF573		✓
ZNF600		✓

References

Supplementary table 4" Knowledge" on GPR84 expression according to Enrichr

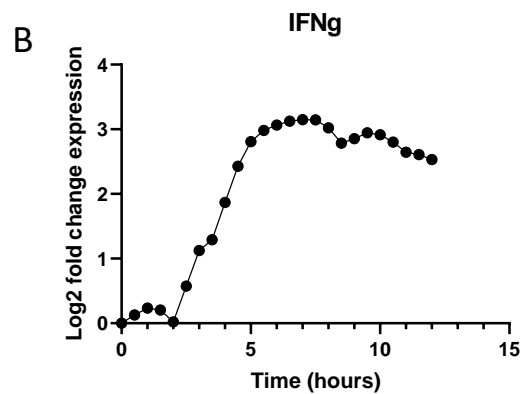
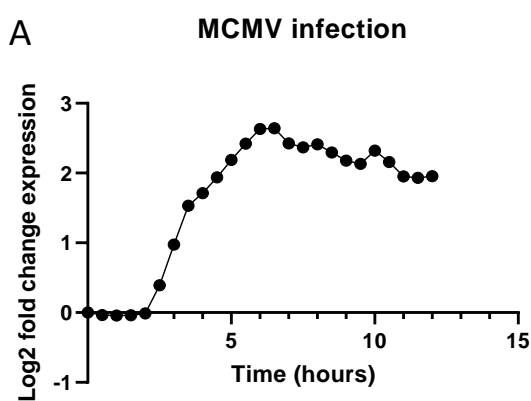
Transcription Factor	Source
NR1H3	ARCHS4 Co-expression
FOSL2	ARCHS4 Co-expression
NFKB2	ARCHS4 Co-expression
CREM	ARCHS4 Co-expression
RELB	ARCHS4 Co-expression
TFEC	ARCHS4 Co-expression
NFKB1	ARCHS4 Co-expression
VDR	ARCHS4 Co-expression
PPARD	ARCHS4 Co-expression
NFE2L2	ARCHS4 Co-expression
ZBTB17	ARCHS4 Co-expression
ETV3	ARCHS4 Co-expression
NR4A3	ARCHS4 Co-expression

EGR2	ARCHS4	Co-expression
SNAI1	ARCHS4	Co-expression
HIVEP3	ARCHS4	Co-expression
IRF5	ARCHS4	Co-expression
HLX	ARCHS4	Co-expression
BATF	ARCHS4	Co-expression
ETS2	ARCHS4	Co-expression
ZNF385A	ARCHS4	Co-expression
ATF3	ARCHS4	Co-expression
BATF2	ARCHS4	Co-expression
ATF5	ARCHS4	Co-expression
CEBPB	ARCHS4	Co-expression
MSC	ARCHS4	Co-expression
ZNF267	ARCHS4	Co-expression
MTF1	ARCHS4	Co-expression

References

MAFG	ARCHS4 Co-expression
MAFB	ARCHS4 Co-expression
JDP2	ARCHS4 Co-expression
ZNF697	ARCHS4 Co-expression
ETV3L	ARCHS4 Co-expression
ZNF438	ARCHS4 Co-expression
PLEK	ARCHS4 Co-expression
SMAD7	ARCHS4 Co-expression
STAT1	KO Mouse, GPR84 up
Myc	KD Human cell line, GPR84 up
Bach2	KO Mouse, GPR84 up
XBP1	KO Mouse, GPR84 up
IRF8	KO Mouse, GPR84 up

SATB2	Overexpression Mouse, GPR84 down
VAX2	KO Mouse, GPR84 down
ZFX	KO Mouse, GPR84 down
ZBTB7A	KO Mouse, GPR84 down
ESRRA	KO Mouse, GPR84 down
ARID3B	KO Mouse, GPR84 down
ZXDC	KD/ depletion Human, GPR84 down



Supplemental figure 1 Gpr84 expression in BMDM infected with MCMV or B) treated with IFN- γ . The y-axis depicts the Log₂ fold change in expression relative to expression at t=0, the x-axis the hours after infection of the cells with the virus (n=1) or stimulated by IFN- γ