

Nuclear magnetic resonance spectroscopy based metabolomics discovers biomarkers of glioblastoma drug response

Thesis for the Master's Degree in Chemistry
60 study points

Ingvild Comfort Hvinden

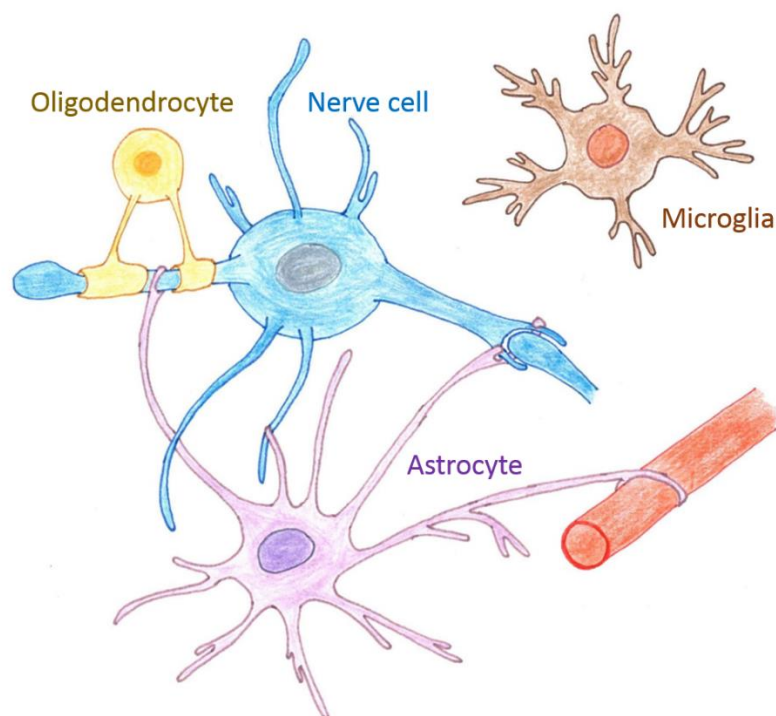


Department of Chemistry
Faculty of Mathematics and Science

UNIVERSITY OF OSLO

May 29th, 2018

Nuclear magnetic resonance spectroscopy based metabolomics discovers biomarkers of glioblastoma drug response



Ingvild Comfort Hvinden

Thesis for the Master's Degree in Chemistry

60 study points

Department of Chemistry

Faculty of Mathematics and Natural Sciences

University of Oslo

May 29th, 2018

© Ingvild Comfort Hvinden

2018

Nuclear magnetic resonance spectroscopy based metabolomics discovers biomarkers of glioblastoma drug response

Ingvild Comfort Hvinden

<http://www.duo.uio.no/>

Printed at: The University Print Centre (Reprosentralen), University of Oslo

Abstract

Glioblastoma is the most common and aggressive form of brain cancer. Even with comprehensive treatment regimes, patients face an expected survival of only 15 months. Currently, methods for assessing treatment response are lacking; it is difficult to accurately determine the efficacy of a treatment. The goal of the present study was to contribute to treatment assessment by scouting for metabolic biomarkers occurring in response to exposure to chemotherapeutic agents temozolomide (TMZ) and sepantronium bromide (YM155).

Untargeted metabolomics of lysate from cultured glioblastoma cells was carried out with liquid state proton nuclear magnetic resonance (NMR) spectroscopy at resonance frequency 800 MHz. Spectral data were analyzed with two different multivariate statistical methods: principal component analysis (PCA) and partial least squares (PLS) regression. For YM155, two biomarker candidates were found: citric and lactic acid. Citric acid appeared to increase most in samples from cell lines less sensitive to YM155. Lactic acid decreased in all cell lines and was considered a more general biomarker of treatment exposure. TMZ-treated samples were not distinguishable from control samples, most likely due to too short exposure time (24 hours). Analyses with nano hydrophilic interaction liquid chromatography coupled with mass spectrometry (MS) corroborated the findings by NMR spectroscopy and statistical analyses.

Both citric acid and lactic acid are biomarker candidates, but a more detailed understanding of their fluctuations in glioblastoma during treatment is needed. Nevertheless, they represent genuine candidates and should be considered for further *in vivo* magnetic resonance spectroscopy (MRS) studies. In the future, the biomarkers could be monitored with MRS, allowing a more unambiguous and personalized assessment of response to treatment in individual patients.

Preface

The present work was performed at the Department of Chemistry at the University of Oslo from August 2016 to May 2018.

I would like to take this opportunity to thank my supervisors associate professor Steven Ray Wilson, professor Frode Rise, PhD candidate Henriette Engen Berg, and professor Elsa Lundanes. Thank you for your support, guidance, and for answering my many questions. The past two years have taught me much about the scientific process, in great part thanks to you.

Furthermore, I would like to thank MD and associate professor Einar Vik-Mo, PhD Cecilie Sandberg, MD and PhD candidate Erlend Skaga, PhD candidate Marit Christensen, and Maria Ewa Walewska from Rikshospitalet, Oslo University Hospital. Thank you to Einar and Cecilie for involving me in such an exciting project. Erlend, Marit and Maria I would like to thank for the many cell samples they cultured for me.

From the NMR lab, I would like to express my gratitude for the help from senior engineer Dirk Petersen, PhD Nils Nyberg from Bruker, senior engineer Per Eugen Kristiansen, and PhD candidates Knut Tormodssønn Hylland and Marte Sofie Holmsen. You have all contributed to making me a better NMR spectroscopist!

I would also like to express my gratitude to PhD Hanne Røberg-Larsen and PhD Ole Kristian Brandtzæg for their invaluable help with the mass spectrometer and to Inge Mikalsen and Marita Clausen for technical support.

The vast and often confusing world of multivariate statistics and programming in R was made understandable by PhD Daniel Sachse. Thank you for taking the time to share you knowledge, I greatly appreciate it.

My two years at the Bioanalytics group were made unforgettable by the wonderful people there, and a special thanks must be given to my office mates Christine Olsen, Tao McQuade, Jenny Skytte af Sätra and Kari Anne Andersen for all the laughter and good times. In addition, I must express my gratitude to Kamilla Bjørseth and Maria Therese Sande for our productive discussions on glioblastoma.

Finally, I would like to express my gratitude to my family, friends, and boyfriend Tor Jan for supporting me throughout my degree.

This work was partly supported by the Research Council of Norway through the Norwegian NMR Platform, NNP (226244/F50).



Ingvid Comfort Hvinden, May 29th, 2018

Table of Contents

Abstract	V
Preface	VI
Table of Contents	VII
1 Abbreviations	1
2 Definitions	4
3 Introduction	7
3.1 Cancer	7
3.2 Glioblastoma.....	8
3.3 Chemotherapeutic agents for treatment of glioblastoma.....	11
3.4 Evaluation of treatment response	15
3.5 Nuclear magnetic resonance spectroscopy	16
3.5.1 General theory of nuclear magnetic resonance	16
3.5.2 Instrumental considerations.....	22
3.5.3 One-dimensional nuclear magnetic resonance spectroscopy	23
3.5.4 Two-dimensional magnetic resonance spectroscopy	24
3.6 Mass spectrometry	26
3.7 Liquid chromatography	28
3.8 Electrospray Ionization.....	31
3.9 Multivariate statistical analysis of metabolomics data	32
3.9.1 Pre-processing of metabolomics data.....	32
3.9.2 Unsupervised and supervised multivariate statistical analysis methods	33
3.10 Aim of Study.....	36
4 Experimental	37
4.1 Instrumentation and equipment	37
4.1.1 Small instruments and equipment	37

4.1.2	Large instruments and equipment	38
4.2	Chemicals	39
4.2.1	Chemicals specific for nuclear magnetic resonance spectroscopy.....	39
4.2.2	Chemicals specific for liquid chromatography based analyses.....	39
4.3	Cell samples.....	40
4.4	Solutions	43
4.4.1	Solutions used in nuclear magnetic resonance spectroscopy	43
4.4.2	Solutions used for liquid chromatography based analysis	43
4.5	Nuclear magnetic resonance spectroscopy analyses	45
4.5.1	Sample preparation.....	45
4.5.2	One-dimensional nuclear magnetic resonance spectroscopy	46
4.5.3	Two-dimensional nuclear magnetic resonance spectroscopy	50
4.5.4	Identification of metabolites: Database searches	52
4.5.5	Statistical analysis of nuclear magnetic resonance data.....	53
4.6	Methods used in liquid chromatography based analysis	55
4.6.1	Making frits and packing capillary columns	55
4.6.2	Testing capillary columns	55
4.6.3	Liquid chromatography-mass spectrometry analyses	56
5	Results and Discussion.....	60
5.1	Choice of sample type, preparation method and nuclear magnetic resonance spectroscopy experiments	61
5.2	Identification of metabolites	63
5.3	Multivariate statistical analysis of nuclear magnetic resonance data	67
5.3.1	Pre-processing of data	67
5.3.2	Principal component analysis revealed two distinct responses to treatment with YM155..	68

5.3.3	Principal component analysis did not reveal any biomarkers for response to temozolomide.....	74
5.3.4	Examination of the second principal component.....	75
5.3.5	Partial least squares regression was congruent with findings in the principal component analyses.....	75
5.4	Liquid chromatography studies.....	78
5.4.1	Preliminary examination of narrow capillary columns.....	78
5.4.2	Targeted metabolomics with liquid chromatography-mass spectrometry.....	80
6	Conclusion.....	83
6.1	Future work.....	83
7	References.....	85
8	Appendix.....	99
8.1	Additional NMR theory.....	99
8.1.1	Frame of reference.....	99
8.1.2	Pulse sequences: solvent suppression.....	100
8.1.3	Pulse sequences: Two-dimensional NMR.....	102
8.2	Description of chemotherapeutic agent candidates and a drug mixture for treatment of glioblastoma.....	105
8.3	Experimental details: Cell culturing.....	111
8.3.1	Chemicals.....	111
8.3.2	Method: cell culturing.....	111
8.3.3	Received cell samples for preliminary nuclear magnetic resonance spectroscopy analyses.....	112
8.4	Experimental details: Preliminary studies in nuclear magnetic resonance spectroscopy.....	114
8.4.1	Instrumentation.....	114
8.4.2	Chemicals.....	115
8.4.3	Solutions.....	115

8.4.4	Methods.....	115
8.5	Results and discussion: preliminary studies	119
8.5.1	Exploration and selection of nuclear magnetic resonance spectroscopy experiments	119
8.5.2	Assessment of narrower sample tubes	128
8.6	Calculation of column efficiency	131
8.7	Chromatographic detection with ultraviolet spectroscopy	132
8.8	Nuclear magnetic resonance spectra.....	133
8.9	Chromatograms from liquid-chromatography-mass spectrometry analyses	138
8.10	Molecule Encyclopedia.....	141
8.11	R scripts	151
8.11.1	R script: pre-processing and PCA of preliminary data	151
8.11.2	R script: pre-processing, PCA and PLS	155
8.11.3	R code for file ‘functions.R’	161
8.12	Nuclear magnetic resonance spectroscopy pulse program scripts	169

1 Abbreviations

Abbreviation	Full name
μL	Microliter
μm	Micrometer
μsec	Microsecond
ACN	Acetonitrile
AIC	5-aminoimidazole-4-carboxamide
APAF-1	Apoptotic protease activating factor 1
AQ	Acquisition time
B_0	External static magnetic field
B_1	Applied magnetic field
Bak	Bcl-2 homologous antagonist killer
Bax	Bcl-2-associated X protein
BBB	Blood-brain barrier
Bcl-2	B-cell lymphoma 2
BER	Base excision repair
bFGF	Basic fibroblast growth factor
Bid	Bcl-2 homology domain 3 (BH3) interacting-domain death agonist
Bim	Bcl-2-like protein 11
BMRB	Biological magnetic resonance bank
CNS	Central nervous system
COLMAR	Complex mixture analysis by NMR
CSC	Cancer stem cell
CUSP9	Coordinated undermining of survival paths
CV	Cross-validation
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's phosphate buffered saline
DS	Number of dummy scans
DW	Dwell time
EC_{50}	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ESI	Electrospray ionization
f_1	Indirect dimension
F16BP	Fructose 1,6-bisphosphate
f_2	Direct dimension

FID	Free induction decay
FIS	Full ion scan
GC	Gas chromatography
GSC	Glioblastoma stem cell
GUI	Graphical user interface
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HILIC	Hydrophilic interaction chromatography
HMDB	Human metabolome database
HSQC	Heteronuclear single quantum coherence spectroscopy
IAP	Inhibitors of apoptosis
ID	Inner diameter
IN_F	Increment for delay
JRES	J-resolved spectroscopy
LB	Line broadening
LC	Liquid chromatography
LOH	Loss of heterozygosity
LPBIN	Number of output points for linear prediction
LV	Latent variable
M	Bulk magnetization vector
MDM2	Murine double minute-2 homolog
ME_mod	Linear prediction for Fourier transformation
MGMT	<i>O</i> 6-methylguanine-DNA methyltransferase
MMR	(DNA) Mismatch repair
MP	Mobile phase
MRI	Magnetic resonance imaging
MRS	<i>In vivo</i> magnetic resonance spectroscopy
MS	Mass spectrometry
MTIC	5-(3-methyltriazene-1-yl)-imidazole-4-carboxamide
NAA	<i>N</i> -acetyl-L-aspartic acid
NAD ⁺	Nicotinamide adenine dinucleotide
NCE	Normalized collision energy
NCOEF	Number of linear prediction coefficients
NIPALS	Non-linear iterative partial least squares
nL	Nanoliter
NMPRTase	Nicotinamide phosphoribosyl transferase
NMR	Nuclear magnetic resonance
NS	Number of scans
NSC	Neural stem cells
NUS	Non-uniform sampling
o1	Frequency transmitter offset
OD	Outer diameter
OMM	Outer mitochondrial membrane

OPC	Oligodendrocyte precursor cells
O-PEA	O-phosphoethanolamine
P	Phosphate
PC	Principal component
PCA	Principal component analysis
PDGF	Platelet-derived growth factor
PEEK	Polyetheretherketone
PEP	Phosphoenolpyruvate
PFG	Pulsed field gradient
PH_mod	Phasing mode
PLS	Partial least squares (regression) Projection to latent structures (regression)
POM	Polyoxymethylene
ppm	Parts per million
PRM	Parallel reaction monitoring
PTEN	Phosphatase and tensin homolog
RG	Receiver gain
RNA	Ribonucleic acid
RP	Reversed phase
SCT	Sample changer temperature
SF	Spectrometer frequency
SI	Size of real spectrum
Smac/Diablo	Second mitochondria-derived activator
SP	Stationary phase
SS	Stainless steel
SW	Sweep width
TCI	Triple resonance cryoprobe
TD	Size of the free induction decay (FID)
TE	Probe temperature
TMZ	Temozolomide
TOCSY	Total correlation spectroscopy
TSP	Trimethylsilyl propionic-2,2,3,3-d ₄ acid
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
WDW	Window function
WHO	World health organization
ZIC	Zwitterionic

2 Definitions

Sources

*Defined by the author.

¹ Biomarkers Definitions Working Group, A. J. Atkinson, W.A. Colburn, V.G. DeGruttola, D. L. DeMets, G.J. Downing, D.F. Hoth, J.A. Oates, C.C. Peck, R. T. Schooley, B. A. Spilker, J. Woodcock, and S. L. Zeger, *Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework*, *Clinical Pharmacology & Therapeutics*, 69 (2001) 89-95.

² Compiled by A. D. McNaught and A. Wilkinson, *IUPAC. Compendium of Chemical Terminology (the "Gold Book")*, 2nd ed., Blackwell Scientific Publications, Oxford, 1997. XML on-line corrected version: <http://goldbook.iupac.org> (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. <https://doi.org/10.1351/goldbook>.

³"*in vitro*", in *Oxford Dictionary of Biochemistry and Molecular Biology*, edited by R. Cammack, T. Atwood, P. Campbell, H. Parish, A. Smith, F. Vella, and J. Stirling, Oxford University Press, Oxford, 2006. Accessed May 26th, 2018:

<http://www.oxfordreference.com/view/10.1093/acref/9780198529170.001.0001/acref-9780198529170-e-10235>

⁴"*in vivo*", in *Oxford Dictionary of Biochemistry and Molecular Biology*, edited by R. Cammack, T. Atwood, P. Campbell, H. Parish, A. Smith, F. Vella, and J. Stirling, Oxford University Press, Oxford, 2006. Accessed May 26th, 2018:

<http://www.oxfordreference.com/view/10.1093/acref/9780198529170.001.0001/acref-9780198529170-e-10239>

^{5a} D.C. Harris, *Quantitative chemical analysis*, 9th ed., W.H Freeman, New York, 2016, pp. GL16 and GL20.

^{5b} C. Wermuth, C. Ganellin, P. Lindberg, L. Mitscher, *Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998)*. *Pure and Applied Chemistry*, 70(5) (2009) 1129-1143.

⁶ D. L. Nelson, A. L. Lehninger, and M. M. Cox, *Lehninger Principles of Biochemistry*, 6th ed., W.H. Freeman, New York, 2013, pp. G-10

⁷ K.K. Murray, R. K. Boyd, M. N. Eberlin, G. J. Langley, L. Li, Y. Naito, *Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013)*. *Pure and Applied Chemistry*. 85(7) (2013) 1515-1609.

⁸ J. Vessman, R. Stefan, J. van Staden, K. Danzer, W. Lindner, D. Thorburn Burns, A. Fajgelj, H. Müller. *Selectivity in analytical chemistry (IUPAC Recommendations 2001)*, *Pure and Applied Chemistry*, 73(8) (2009) 1381-1386.

⁹ A. Klassen, A.T. Faccio, G.A.B. Canuto, P.L.R. da Cruz, H.C. Ribeiro, M.F.M. Tavares, A. Sussulini, *Metabolomics: Definitions and Significance in Systems Biology*, in: A. Sussulini (Ed.) *Metabolomics: From Fundamentals to Clinical Applications*, Springer International Publishing, Cham, **2017**, pp. 3-17

Phrase	Description
Analyte*	Compound(s) of interest in the analysis
Biomarker ¹	“A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.”
Charge number ²	“Ratio of the charge of a particle to the elementary charge.”
Data*	Spectrum or chromatogram associated with one sample
Dataset*	All spectra or chromatograms for a set of samples
In vitro ³	“Latin (of any biological process, reaction, or experiment) occurring or made to occur outside an organism, <i>e.g.</i> in extracts or cultures; literally it means ‘in glass’.”
In vivo ⁴	“Latin (of any biological process, reaction, or experiment) occurring or made to occur within a living organism; literally it means ‘in life’.”
Log p (partition coefficient ^{5a} / lipophilicity ^{5b})	The distribution of neutral or unionized compound (X) in a biphasic system of water and octanol-1-ol: $\log p = \log([X]_{\text{octanol-1-ol}}/[X]_{\text{water}})$.
Metabolite ⁶	“A chemical intermediate in the enzyme-catalyzed reactions of metabolism.”
Metabolome ⁶	“The complete set of small molecule metabolites (metabolic intermediates, signals, secondary metabolites) present in a given cell or tissue under specific conditions.”
Metabolomics ⁶	“The systematic characterization of the metabolome of a cell or tissue.”
m/z ⁷	“Abbreviation representing the dimensionless quantity formed by dividing the ratio of the mass of an ion to the

	unified atomic mass unit, by its charge number (regardless of sign).”
Selectivity ⁸	“... the extent to which the method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behavior.”
Sensitivity ^{5a}	“Capability of responding reliably and measurably to changes in analyte concentration. In quantitative terms, sensitivity is the amount of instrument response per unit change in concentration of analyte.”
Small sample*	Samples ranging in size 10-1000 μ L
Targeted metabolomics ⁹	“...a quantitative analysis (concentrations are determined) or semiquantitative analysis (relative intensities are registered) of a few metabolites and/or substrates of metabolic reactions that might be associated to common chemical classes or linked to selected metabolic pathways.”
Untargeted metabolomics ⁹	“...is based primarily on the qualitative or semiquantitative analysis of the largest possible number of metabolites from a diversity of chemical and biological classes contained in a biological specimen.”

3 Introduction

3.1 Cancer

Humans are made up of approximately 30 trillion cells [1]. Within each cell deoxyribonucleic acid (DNA), the genetic material encoding all genes, is found. A double lipid bilayer compartmentalizes the DNA, separating it from the rest of the cell and protecting it from harm. Furthermore, the DNA is so vital that cells have extensive repair mechanisms to deal with any damage that might occur. If the harm is beyond what the cell can repair it dies on purpose (apoptosis) [2 (p. 1021-1022)], or is removed by the immune system [3 (p. 279-281)].

Cancer originates in surviving defective cells where DNA damage has enabled the cells to grow abnormally and invasively. Two major types of genes contribute to cancer development: *tumor suppressor genes* and *oncogenes*. Tumor suppressor genes encode for proteins that partake in protective processes keeping a cell from turning cancerous. These genes are usually recessive, meaning both alleles¹ have to be lost for the process to cease [5]. On the other hand, the products of oncogenes promote carcinogenesis. Oncogenes are usually dominant and only one allele needs to be mutated for an effect to take place. The products of oncogenes usually give increased protein expression or enzymatic activity [6].

Genetic instability and inflammation pave the way for cells to develop the hallmarks of cancer and evolve into a tumor. Cancer cells are recognized by their ability to take control over proliferative cell signaling pathways and avoiding inhibitory signals from stopping their growth. Furthermore, cancer cells can avoid the mechanisms set in place to stop infinite cell division and even evade apoptosis. Cancer cells also gain the ability to promote the formation of new blood vessels (*angiogenesis*) to feed the tumor with nutrients and oxygen. Finally, cancer cells can become capable of spreading throughout the body, an ability healthy cells lack [7].

For certain cancers, it has also been proposed that only a subset of their cells are capable of endless self-renewal [8-10]. These cells can differentiate and supply the tumor with various cell types, suggesting the tumors are hierarchically organized. The capabilities of these cells are

¹ Allele is a variant of a gene. Humans always have two copies of a gene, which can have the same or different alleles of said gene [4] K.L. Voje, *Alleler*, Store Norske Leksikon, 2014, <https://snl.no/alleler>

similar to that of somatic stem cells (self-renewal and differentiation) and they are aptly named cancer stem cells (CSCs). Often, the CSCs are able to survive initial treatments and pave the way for recurrence because of their ability to produce differentiated cells [3 (p. 178-182), 10]. Exactly how CSCs develop and for which cancers the model applies, is under discussion [10, 11]. Yet, strong evidence for CSCs has been found in *e.g.* leukemia [8], breast [12], colon [13, 14], and brain cancers [15, 16].

Regardless of the organization of cells within a tumor, all the growth and frenzied activity require nutrients for energy and building blocks to establish new cells. Cancer cells have the ability to reprogram their energy metabolism in a variety of ways to achieve further growth. In 1924, Otto Warburg was the first to observe a difference in the metabolism of healthy and cancerous cells [17]. Usually, healthy cells will utilize glucose fully in presence of oxygen and break it down to CO₂. If there is little or no oxygen present, the cells produce lactic acid instead, in part to regenerate an important redox coenzyme (*nicotinamide adenine dinucleotide* (NAD⁺)) [18 (p. 563-565)]. Warburg observed that cancer cells would produce lactic acid even when oxygen was available [17]. The process is known as the Warburg effect or aerobic glycolysis. The exact reason for why cancer cells switch from full glucose oxidation to aerobic glycolysis is uncertain, but there is mounting evidence that it is in part to conserve carbon and funnel it towards building biomass [19]. Furthermore, cancer cells have been shown to reprogram their glutamine, serine/glycine, and acetate metabolism with the same goal in mind [20].

3.2 Glioblastoma

Glioblastoma is an aggressive form of brain cancer. The exact cell type it originates from is not yet determined, but both mature astrocytes [21], oligodendrocyte precursor cells (OPCs) and adult neural stem cells (NSCs) have been suggested [22]. In the World Health Organization's (WHO) Classification of Tumors, glioblastoma is listed as a grade IV (of grades I-IV) central nervous system (CNS) tumor. The higher the grading score, the more malignant the tumor is. Less malignant cancers can also arise in the CNS and are generally called gliomas if thought to arise from glial cells, or astrocytomas if thought to arise from astrocytes [23].

The pathogenesis of glioblastoma is complex and not fully understood. The originator cells undergo several different genetic mutations that together or in part contribute to the development of the cancer [24]. Furthermore, it is important to distinguish between 'primary'

and ‘secondary’ glioblastoma. The former arises *de novo*, while the latter is a malignant progression of a lower grade glioma. Primary glioblastoma is much more common than secondary, accounting for approximately 95% of all cases [25]. Additionally, primary glioblastoma exhibits a different mutational pathway than secondary glioblastoma, indicating their different origins [26]. The various common genetic mutations leading to primary and secondary glioblastoma are illustrated in **Figure 3.1**.

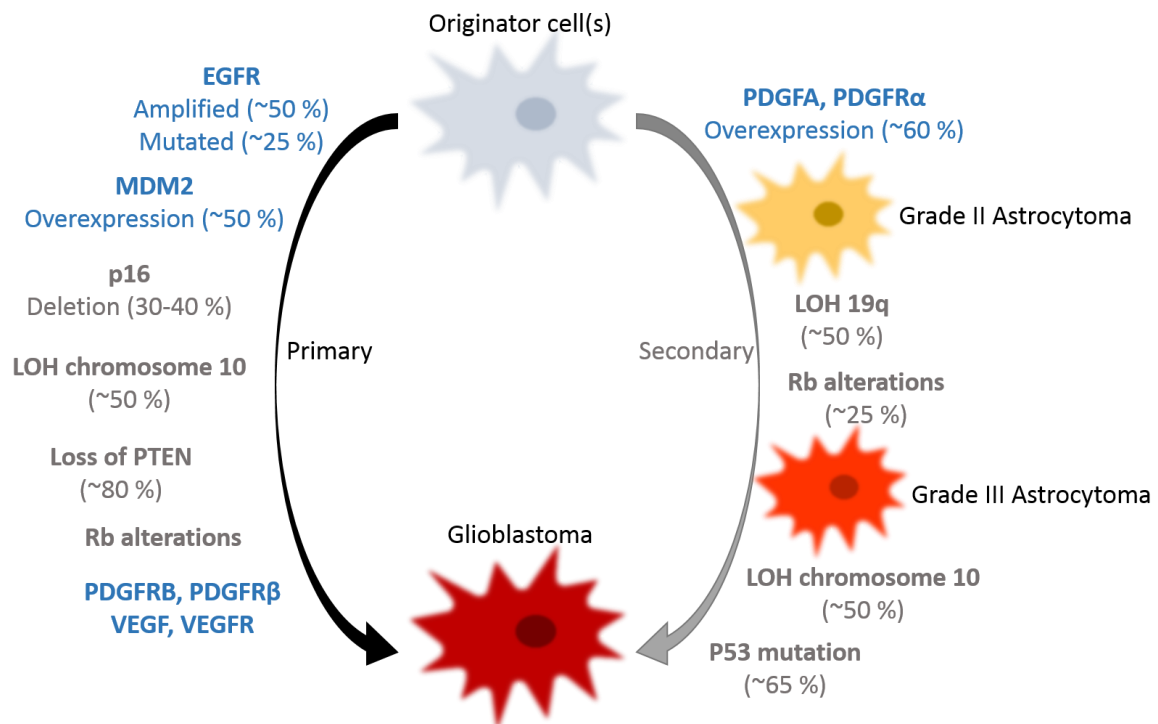


Figure 3.1: The common genetic mutations leading to primary and secondary glioblastoma. Figure adapted from [26]. Oncogenes are written in blue and tumor suppressor genes are written in grey. The percentage of glioblastomas found with the mentioned genetic changes are given for each mutation. Epidermal growth factor receptor (EGFR) takes part in several different signaling pathways related to cancer, *e.g.* apoptosis, proliferation and invasion [27]. Murine double minute-2 homolog (MDM2) is a negative regulator of the p53 gene, an important tumor suppressor gene [28]. The p16 protein inhibits parts of the cell cycle and the deletion of its gene chips away at defense mechanisms against unlimited cell division [2, 26]. Loss of heterozygosity (LOH) is the event where whole genes and surrounding DNA is lost. LOH of regions of chromosome 10 is quite common in both primary and secondary glioblastoma, while loss of regions of chromosome 19 is most common for secondary glioblastoma. Phosphatase and tensin homolog (PTEN) is a part of chromosome 10, and is often a part of LOH of chromosome 10 [24]. Both Rb and p53 alterations and mutations affect the cell’s ability to regulate its cell cycle. Platelet-derived growth factors (PDGF) and vascular endothelial growth factors (VEGF) both contribute to formation of new blood vessels [26].

Glioblastoma is one of several cancers indicated to consist of CSCs and various differentiated cells [15, 29]. More specifically, it appears that most of the cells are sorted into a hierarchy of slow dividing stem cells (GSCs), rapidly cycling progenitor cells, and non-proliferative cells [16].

In epidemiological context, glioblastoma is the most common form of malignant brain cancer, with approximately 300 new cases occurring per year in Norway [30, 31]. It can develop at all ages but is more common among the elderly (average age at diagnosis: 62 years) and men [32, 33]. Median survival is merely ~15 months and only about 5% of patients are still alive 5 years after diagnosis [31, 33].

The current treatment scheme of glioblastoma includes surgical resection or biopsy, followed by combined chemo- and radiation therapy. Surgery alone is not considered curative, as the tumor is too diffuse for complete resection. If resection is not feasible, a biopsy is usually still obtained. Both resected and biopsied tumors are analyzed further to assure that the correct diagnosis has been made and to identify various molecular markers consequential for prognosis and treatment strategies [34]. The chemotherapeutic agent currently in use is temozolomide, described in detail in **section 3.3**. Several studies have shown increased survival rates for patients receiving temozolomide as an adjuvant treatment to radiation therapy, which is why the current treatment scheme includes both [35-37].

In spite of numerous studies into improving treatment and overall survival of glioblastoma patients, little headway has been made [31]. It has been proposed that glioblastoma is challenging to treat because of heterogeneity within the tumor and between patients [38-43]. Intra-tumor heterogeneity poses two main issues. First, analysis of a single biopsy may give inadequate information of the molecular pathology of the cancer since it varies across the tumor. This can influence the choice of treatment and lead to selecting treatment options that possibly only work on a fraction of the cells. Second, certain cells may be resistant to treatment and survive, leading to the patient relapsing [44, 45]. It is especially concerning considering GSCs; several studies have found them resistant to radiation and temozolomide treatment [46-49]. Because of their ability to differentiate into various cancerous cells, their survival sets the stage for recurrence [29]. Finally, concerning heterogeneity between patients, the main issue is the lack of a “one-size-fits-all” treatment option.

3.3 Chemotherapeutic agents for treatment of glioblastoma

Two chemotherapeutic agents are presented in the following section: temozolomide is an established drug and YM155 is a candidate for treatment of glioblastoma. Other drug candidates are described in the **Appendix (section 8.2)**. Chemotherapeutic agents are compounds that are either cytotoxic or cytostatic to cells, *i.e.* they either kill cells or inhibit their growth [50, 51]. Patients receive chemotherapeutic agents either for curative or palliative purposes [52].

Temozolomide (TMZ)

Temozolomide is an alkylating agent; it causes extensive DNA methylation. Moreover, it is small and rather lipophilic and can, therefore, cross the blood-brain barrier (BBB) unaided. The molecule is stable at acidic pH, but becomes labile at pH above seven and spontaneously breaks down to 5-(3-methyltriazen-1-yl)-imidazole-4-carboxamide (MTIC). MTIC can react with water, generating 5-aminoimidazole-4-carboxamide (AIC) and a methyldiazonium cation (as shown in **Figure 3.2**). Once formed, the methyldiazonium cation can methylate DNA, preferably guanine at the *N7* position but other reactions may occur too [53, 54]. In **Figure 3.3**, the three most common methylation products are shown.

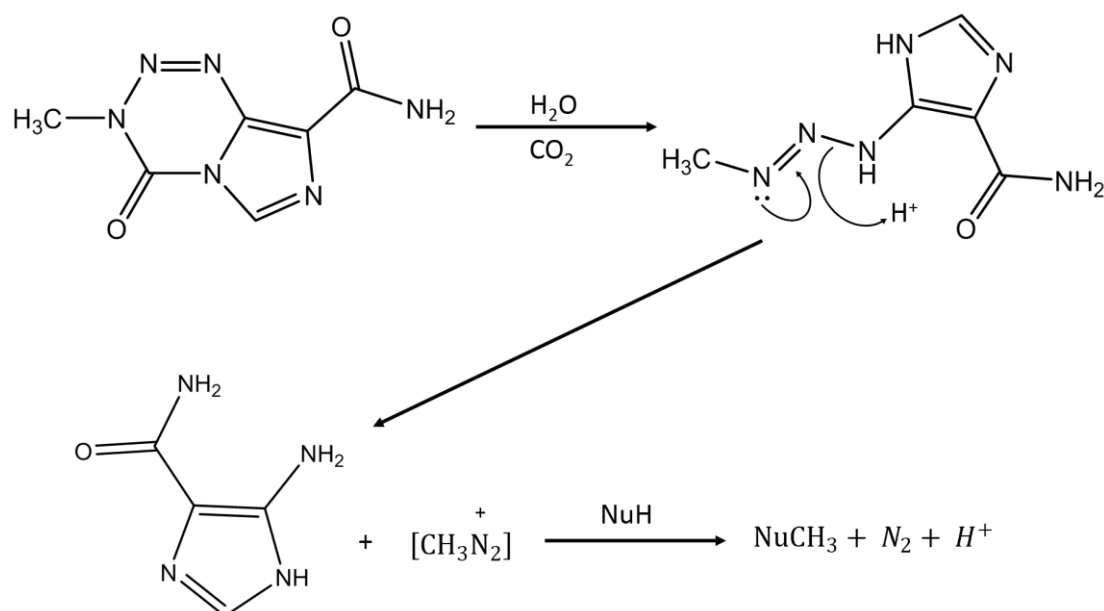


Figure 3.2: Activation of temozolomide. Temozolomide spontaneously forms MTIC at alkaline pH, and MTIC can react with water to release AIC and a methyldiazonium cation. The cation reacts with a nucleophile (NuH) center in the DNA, methylating it.

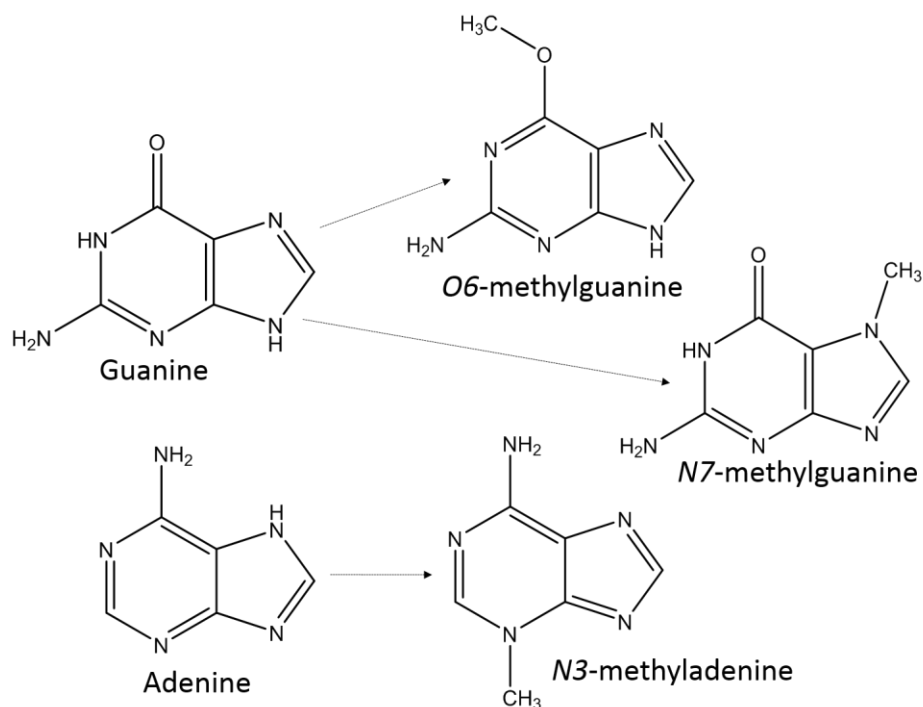


Figure 3.3: The three most common products of DNA methylation by temozolomide. Guanine is most commonly methylated at the *N7* position, but can also be methylated at the *O6* position. Adenine is most commonly methylated on the *N3* position.

The cytotoxicity of temozolomide is dependent on DNA mismatch repair (MMR). *O6*-methylguanine pairs with thymine instead of cytosine. The mismatch is recognized by MMR proteins and they excise the thymine and replace it with cytosine. The proteins are not capable of removing *O6*-methylguanine, thus the mismatch issue persists. For each reparation attempt, DNA strand breaks can occur and these will further on cause replication failure and cell cycle arrest [54, 55].

Cells have various mechanisms for combatting methylated bases, consequently lowering the cytotoxicity of temozolomide. The two most common in glioblastoma are demethylation and base excision. All cells have the enzyme *O6*-methylguanine-DNA methyltransferase (MGMT), which can remove *O6*-methyl groups from guanine. In glioblastoma, the gene for MGMT is often silenced or lost. However, some glioblastomas have a normal or increased expression of MGMT, and they show reduced sensitivity towards temozolomide [54]. Base excision repair (BER) is a pathway in which ‘bulky’ DNA bases are removed and replaced. BER can extract the *N7*-methylguanine and *N3*-methyladenine bases but if it is impaired, temozolomide toxicity is increased [56].

YM155 (Sepantronium bromide)

YM155, structure shown in **Figure 3.4**, is an inhibitor of *survivin* gene transcription [57-60]. *Survivin* is a protein involved in inhibition of apoptosis and regulation of the cell cycle [61]. In other words, by blocking *survivin* transcription apoptosis can proceed. YM155 is a good candidate for treatment of glioblastoma because the cancer has been shown to overexpress *survivin* [62, 63].

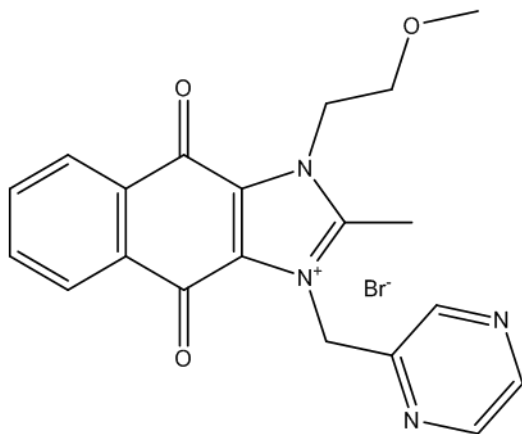


Figure 3.4: Molecular structure of YM155.

Cells have two different pathways for induction of apoptosis: the extrinsic pathway and the intrinsic pathway. The former is initiated by external stimuli, *e.g.* tumor necrosis factors, while the latter is induced by internal stimuli, *e.g.* DNA damage or oxidative stress [3]. They both end with initiating a caspase cascade. Caspases are proteases, enzymes that can cleave proteins. *Survivin* can inhibit certain caspases, as well

as regulate some other participants of the intrinsic apoptotic pathway [61, 64].

In the intrinsic pathway, illustrated in **Figure 3.5**, some of the proteins necessary for initiating the caspase cascade are located in the mitochondria. Their release requires the formation of pores in the outer mitochondrial membrane (OMM). Proteins from the B-cell lymphoma 2 (Bcl-2) protein family (mainly Bax² and Bak³) make up the pores. Other proapoptotic Bcl-2 proteins (Bid⁴ and Bim⁵) can assist Bak and Bax during the necessary conformational changes prior to pore-formation. Once the pores have formed, cytochrome *c* and procaspase 9 can be released to the cytosol. Cytochrome *c* binds to the apoptotic protease activating factor 1 (APAF-1). Next, procaspase 9 can bind to this complex (known as the apoptosome) and become activated. Caspase 9 is an initiator caspase, meaning it activates other caspases and triggers the caspase cascade. *Survivin* can inhibit activation of caspase 9 [65], in addition to inhibiting activated

² Bcl-2-associated X protein.

³ Bcl-2 homologous antagonist killer

⁴ Bcl-2 homology domain 3 (BH3) interacting-domain death agonist

⁵ Bcl-2-like protein 11

downstream caspases. There is some disagreement on which of the downstream caspases *survivin* is capable of inhibiting [66-69].

Cells have mechanisms for stopping the cascade. *Survivin* is a part of the family of inhibitor of apoptosis proteins (IAPs). Like *survivin*, the IAPs are capable of inhibiting the caspases of the cascade. However, when the intrinsic pathway is activated mitochondria also release the protein second mitochondria-derived activator (Smac/Diablo), which can inhibit certain IAPs. *Survivin* can bind to Smac/Diablo and stop it from inactivating the other IAPs [64]. Thus, *survivin* both inhibits caspases and factors that can inhibit other IAPs.

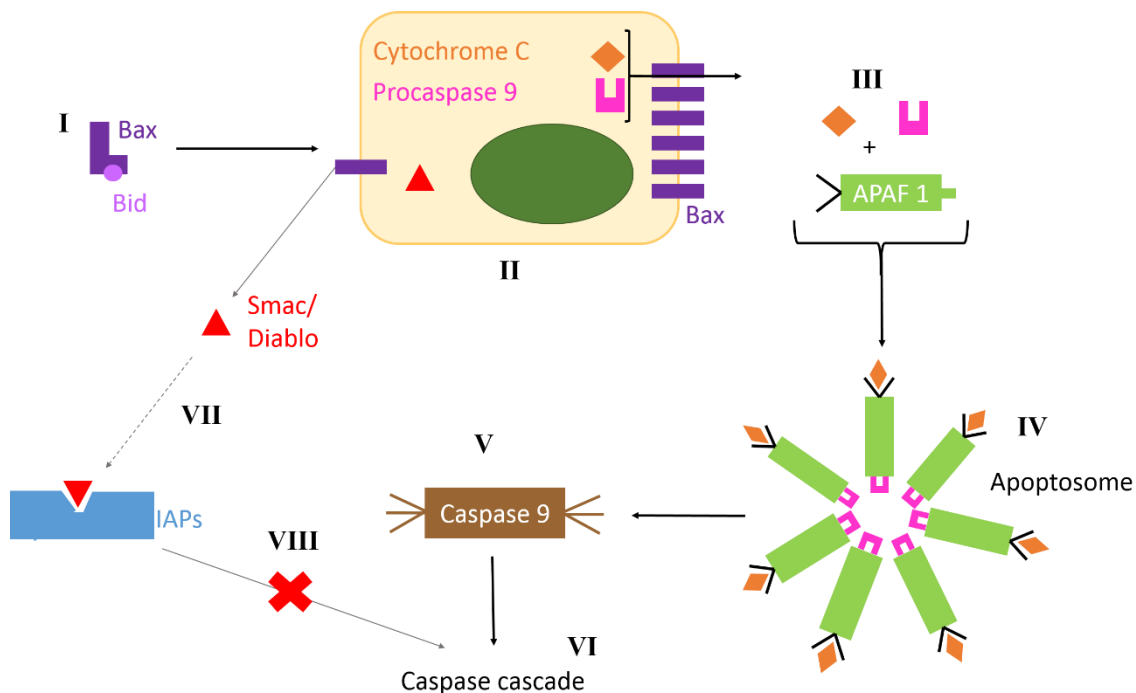


Figure 3.5: The intrinsic apoptotic pathway. It begins with proapoptotic Bcl-2 protein family members (here Bax) being activated by internal signal(s) for cellular stress (I). With help from Bid, another proapoptotic Bcl-2 protein family member, several Bax come together and form pores in the OMM (II). Cytochrome *c*, procaspase 9 and Smac/Diablo are released to the cytosol (III). Cytochrome *c* and APAF-1 form a complex (the apoptosome, IV) that procaspase 9 can bind to and be activated by (V). Once activated, caspase 9 can activate other caspases (e.g. caspase-3, -6 and -7) and initiate a caspase cascade (VI). IAPs can stop the caspases, but Smac/Diablo can inhibit the IAPs (VII), allowing apoptosis to go forth anyways (VIII). *Survivin* can inhibit both caspase 9, some of the caspases in the cascade, as well as Smac/Diablo. Figure adapted from [3 (p. 161)].

3.4 Evaluation of treatment response

Currently, a patient's response to treatment is monitored by assessing the reduction of tumor size with magnetic resonance imaging (MRI), dependence on corticosteroids for symptom alleviation, and overall patient health [70, 71]. The present criteria and established methods have several issues. Most notably, it is difficult to distinguish the tumor from surrounding structures, *e.g.* necrotic tissue and the postoperative cavity. In addition, the MRI measurements are subject to interobserver inconsistencies [70, 72]. Taken together, it is difficult to report treatment response accurately and robustly.

Better tools for assessment are needed, and ideally, they would be as non-invasive and free of side-effects as possible. Furthermore, since glioblastoma has both intratumoral and inter-patient heterogeneity, the new tools should account for this. Biomarkers could increase the specificity of the assessment of a tumor's response to treatment.

The challenge is to find methods or tools that can conveniently provide indirect measurements of the biomarkers or samples containing them. Repeat biopsies of the tumor(s) are not favored because biopsies are invasive and have some serious side-effects, including intracerebral hemorrhage (bleeding in the brain tissue) [73]. Blood tests (liquid biopsies) are another possibility; they are much less invasive than direct biopsies. However, the BBB is not entirely disrupted in glioblastoma [74, 75], and a biomarker found in *in vitro* studies might not be detectable in the blood. Proton magnetic resonance spectroscopy (MRS) is a complementary technique to magnetic resonance imaging (MRI); it allows for detection of chemical compounds in brain tissue. Moreover, MRS is non-invasive and can be carried out *in vivo* [76]. Thus, MRS is the best candidate considering currently available procedures.

MRS can only detect small molecules, such as metabolites (mass < 1500 Dalton) [76, 77], and by refraining from biopsies DNA, ribonucleic acid (RNA), and proteins are not available as biomarker candidates. However, the metabolites provide other information. Since they are the products of a number of metabolic pathways, they represent a broad view and to some extent a summary of the processes occurring in cells. Another advantage of metabolites as biomarkers is that they do not undergo the same type of modifications as DNA, RNA, and proteins, and are thus easier to correlate with phenotype [78].

Possible biomarkers should be established *in vitro* prior to studies with MRS. Nuclear magnetic resonance (NMR) spectroscopy is the technique that MRS is developed from [79], and therefore

it could be a suitable choice for *in vitro* metabolomics studies. NMR has several advantages as an analytical technique. First, it is non-destructive, does not require extensive sample preparation or derivatization of analytes, and can handle complex mixtures. Second, it can be used to identify novel compounds and it is relatively easy to quantify compounds of interest. Finally, NMR is robust and reproducible, allowing for straightforward comparison of collected data and standardized databases [80-82]. In a study by Lehtimäki *et al.*, findings in *in vivo* MRS of rat glioma cells generally corresponded to findings in *in vitro* NMR analysis of tissue extracts, showcasing the complementary nature of the two techniques [83]. One major disadvantage of NMR is its lack of sensitivity.

3.5 Nuclear magnetic resonance spectroscopy

3.5.1 General theory of nuclear magnetic resonance spectroscopy

When certain nuclei are placed in a static magnetic field, they can absorb and reemit quantized amounts of energy. The nuclei must have nuclear angular momentum (spin), which in turn gives them a magnetic moment; this is found in nuclei with odd numbers of protons, neutrons or both, *e.g.* ^1H , ^{13}C and ^{19}F . The spin can be further described by its spin quantum number (I), which also depends on the number of protons and neutrons, as shown in **Table 3.1** [84 (p. 561-563)].

Table 3.1: The value of the spin quantum number (I) depends on the number of protons and neutrons in the nucleus.

Spin quantum number	Number of protons/neutrons
Zero	Even/Even
Whole integers	Odd/Odd
Half integers	Odd/Even or Even/Odd

As a result of their magnetic moments, nuclei behave as small magnets and can interact with an external magnetic field. In quantum mechanics, it is stated that a nucleus can have $2I+1$ orientations ('energy levels') when placed in a static magnetic field [84 (p. 561-563)]. Nuclei with $I = \frac{1}{2}$ thus can have two different orientations, commonly referred to as parallel (α) and antiparallel (β) with respect to the external field [85 (p. 11-14)]. Since spin-half nuclei are most often used in NMR the following description will focus on their behavior in static magnetic fields.

The parallel orientation of nucleus spin is associated with lower energy. Because of the energetic advantage, most nuclei will try to align themselves in that manner. From this, it has been popular to describe the nuclei in a sample as all either parallel or antiparallel to the external static magnetic field. However, this is an oversimplification. The random, thermal motion of molecules will disrupt the parallel or antiparallel positions that the nuclei have relative to the external magnetic field. Yet, the slight energetic advantage of being parallel remains and a small net amount of magnetization exists in the sample. The net magnetization is called the *bulk magnetization vector*, and is conventionally drawn along the z -axis. When a sample is placed in a static magnetic field, it takes a certain amount of time for the bulk magnetization vector to appear and stabilize, called the *relaxation time* [85 (p. 11-20), 86 (p. 47-50)]. The process is illustrated schematically in **Figure 3.6**.

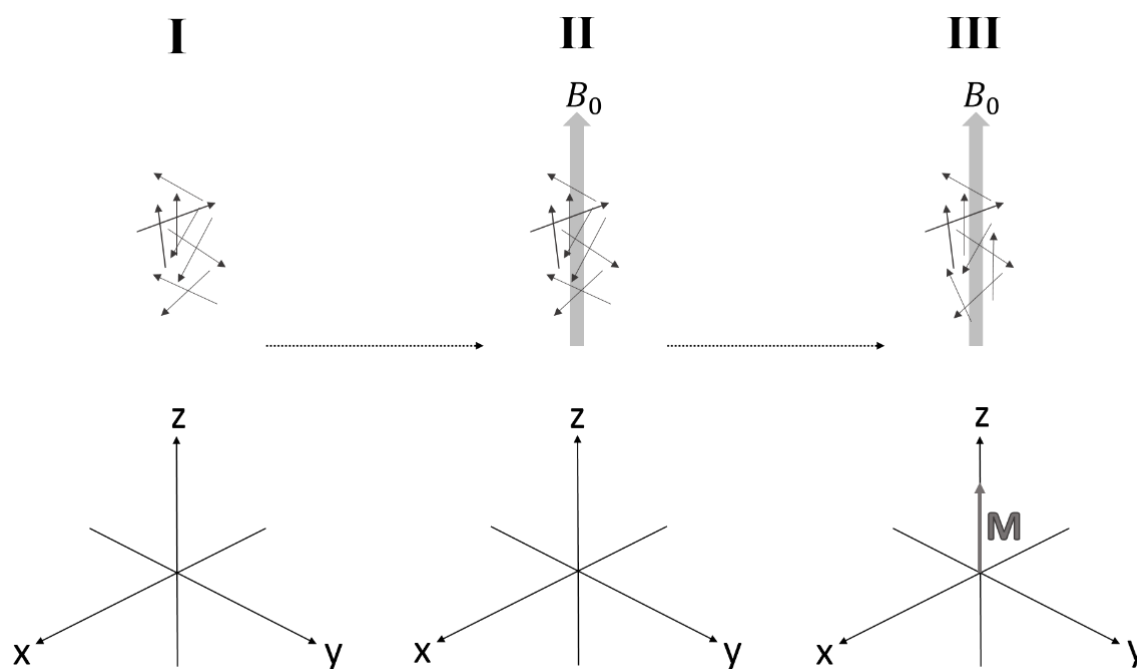


Figure 3.6: Establishing the bulk magnetization vector. (I) When a strong, static magnetic field is absent, nuclei have no net magnetization. (II) When placed in a magnetic field (B_0) it takes some time for the sample to build a net magnetization (M). (III) Once equilibrium is reached, there is a net magnetization across the sample as whole, but each individual nucleus is not necessarily aligned parallel or antiparallel relative to the applied field. The bulk magnetization vector (M) is drawn along the positive z -axis by convention. Figure adapted from [86 (p. 50)].

There is no net magnetization in the xy -plane because there is no energetic advantage in aligning one way or the other; the static magnetic field has only one direction, along the z -axis. Thus, the magnetic moments of the nuclei average out in the xy -plane due to their thermal motion [86 (p. 47-50)].

After the bulk magnetization vector has formed, it does not change in size or direction unless acted upon by a magnetic force. If the vector is tipped away from the z-axis, it will move in a circular motion about the field due to the magnetic field imposing torque on it, as illustrated in **Figure 3.7**. The motion is called Larmor precession and the frequency (ν) of it is dependent on the strength of the static magnetic field (B_0), given by **Equation 3.1**.

$$\nu = -\frac{\gamma B_0}{2\pi} \quad \text{Equation 3.1}$$

Where γ is the magnetogyric ratio of the nucleus in question. The magnetogyric ratio is a constant for given nucleus species and can be negative or positive. The sign of γ determines whether the precession is clockwise or anticlockwise [85 (p. 11-13), 86 (p. 50-51)].

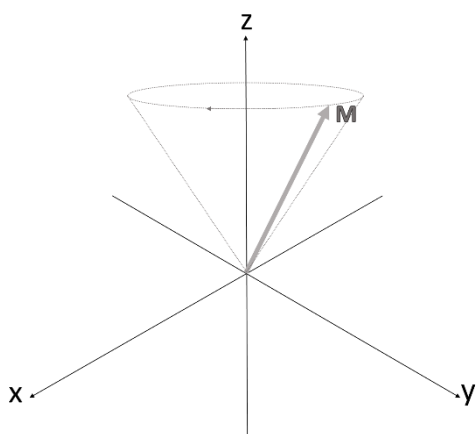


Figure 3.7: Larmor precession of the bulk magnetization vector (M). Figure adapted from [85 (p. 13)]

Individual nuclei also precess at the Larmor frequency when placed in a magnetic field. Nuclear magnetic resonance occurs if a nucleus changes its spin state by absorbing a quantum of energy with the same frequency as the Larmor frequency. The energy required is given by **Equation 3.2**.

$$\Delta E = h\nu = \frac{h\gamma B_0}{2\pi} \quad \text{Equation 3.2}$$

Where h is the Planck constant ($6.62607004 \times 10^{-34} \text{ m}^2 \text{ kg/s}$). The energy difference between

the two spin states is very small. According to the Boltzmann equation, given by **Equation 3.3**, it results in only a small surplus of nuclei in the favored spin state (parallel/ α) and only those nuclei allow observation of resonance [87].

$$\frac{N_\alpha}{N_\beta} = e^{\Delta E/k_B T} \quad \text{Equation 3.3}$$

Where N_α is the number of nuclei in the parallel state, N_β is the number of nuclei in the antiparallel state, T is the temperature, and k_b the Boltzmann constant ($1.38064852 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$). For practical purposes, it means that NMR has low sensitivity compared to many other spectroscopic techniques, *e.g.* ultraviolet (UV) [85 (p. 11-14)].

Pulses

The induction of nuclear magnetic resonance requires radiation with frequency equal to the Larmor frequency of the nuclei in the sample. In NMR spectroscopy, this is accomplished by applying a pulse of oscillating electromagnetic radiation (pulse) through the sample such that the magnetic component (B_1) of electromagnetic radiation is in the transverse plane. Returning to the vector model, the effect of the applied magnetic field is to tip the vector away from the z -axis (by convention drawn moving toward the y -axis) at an angle that depends on the amplitude and duration of the pulse, given by **Equation 3.4**.

$$\theta = \frac{360\gamma B_1 t_p}{2\pi} \quad \text{Equation 3.4}$$

Where t_p is the duration of time the pulse is applied and θ denotes the angle between M and the z -axis.

Pulses are, among other things, named after the degree to which they flip the bulk magnetization vector, *e.g.* a pulse that flips the bulk magnetization vector to the y -axis is known as a 90° pulse. Furthermore, depending on which axis B_1 is applied along, various subscripts such as x or y , and $+$ or $-$ may be added. Additional details on pulses and pulse sequences are given in the **Appendix (section 8.1.2 and 8.1.3)**, including a comment on the choice of frame of reference (**section 8.1.1**).

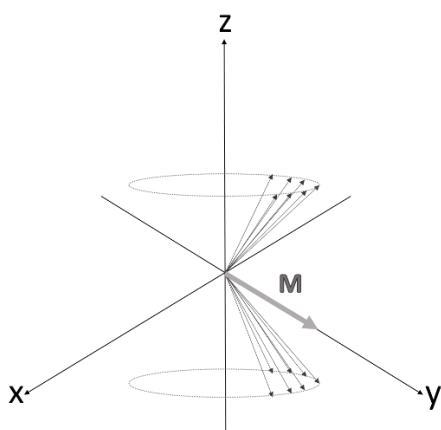


Figure 3.8: Phase coherence. An applied magnetic field (B_1) forces the magnetic moments of the nuclei in a sample to cluster together; they obtain phase coherence. The bulk magnetization vector (M) appears in the xy -plane rather than along the z -axis. Figure adapted from [85 (p. 16)].

It should be kept in mind that the bulk magnetization vector is made up of the magnetic moments of individual nuclei. The effect of *e.g.* a 90° pulse is not to flip the magnetic moments, but “bunch them together” such that they move in phase or have *phase coherence*. The result is that net magnetization no longer occurs along the z -axis, but in the xy -plane, as illustrated in **Figure 3.8** [85 (p. 15-17), 86 (p. 52-55)].

Detection

The electromagnetic pulse used to initiate NMR is generated from a coil surrounding the sample. After the pulse has been stopped, the precessing bulk magnetization vector can induce a current in the coil. The signal is amplified before being sent to a processing computer. Since nothing is forcing the bulk magnetization vector to remain in the xy -plane it returns to the z -axis. This causes the induced signal, called the free induction decay (FID) signal, to die out over time, as illustrated in **Figure 3.9**.



Figure 3.9: Schematic illustration of the free induction decay (FID) signal, which is recorded during an NMR experiment.

The FID is converted to a comprehensible frequency spectrum by Fourier transformation. In short, to find the intensity of a signal at a given frequency, the FID is multiplied with a cosine wave with the same frequency. The area of the

resulting product wave function equals the signal intensity at the chosen frequency. To obtain the whole of the spectrum, the procedure is repeated with incrementally increasing frequencies [86 (p. 77-82)].

Usually more than one FID is collected during the acquisition of an NMR spectrum. The FIDs are added together and only the summed FID is Fourier transformed. Summation of FIDs increases the sensitivity of the measurement: if N FIDs are collected, the signal is N times stronger, while the noise increases only with \sqrt{N} because it is random [85 (p. 81-82)].

Chemical shift

The Larmor frequency of a nucleus depends on its chemical and physical surroundings. Consequently, it is possible to distinguish nuclei of the same species in a molecule, given that they are in different chemical and/or physical environments.

In any given molecule, electrons surround the nuclei to varying degree. The electrons are also affected by the magnetic field utilized in NMR spectroscopy because they also possess spin and a magnetic moment. The valence electrons of each nucleus will begin to circulate around the

nucleus once placed in a static magnetic field, as illustrated in **Figure 3.10**. The movement generates a small magnetic field around the nucleus that counters the main static magnetic field, moderately shielding the nucleus. Thus, because the net magnetic field is diminished, the shielded nucleus will have a lower Larmor frequency than an unshielded one [87 (p. 112-114)]. The higher the electron density around the nucleus, the more shielded it becomes, and the lower its Larmor frequency will be.

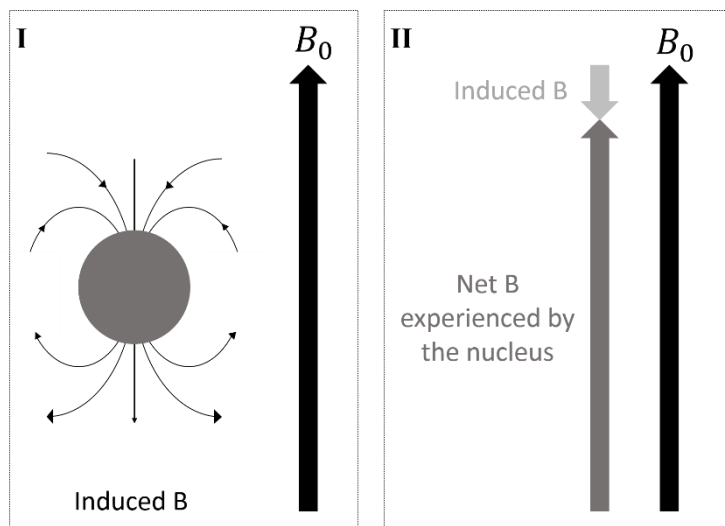


Figure 3.10: The origin of chemical shift. Chemical shift arises due to the valence electrons around the nucleus. They have spin and magnetic moment and when placed in a magnetic field, the electrons will begin to circulate around the nucleus and form a small magnetic field with opposite direction of the external field (I). Thus, the net magnetic field the nucleus experiences is smaller than the external static magnetic field (II). The figure is schematically drawn and chosen sizes should be considered arbitrary. Figure I) adapted from [87 (p. 112)].

chemical shift zero. All other chemical shifts are given relative to the signal from this compound. Furthermore, because the size of the relative shifts also depends on the strength of the static magnetic field, chemical shift is usually listed with a relative value, given by **Equation 3.5**. It ensures easier comparison of spectra from different instruments.

$$\delta = \frac{\nu - \nu_{SC}}{\nu_{SC}}$$

Equation 3.5

Where ν_{SC} is the frequency of the signal from the standard compound and ν is the frequency of the signal from the compound of interest. As mentioned, the difference between two chemical shifts is usually in the kHz range, while the absolute value is in the MHz range. Thus, δ is usually expressed in parts per million (ppm) [86 (p. 6-7), 87 (112-114)].

In the vector model, differences in shift can be depicted as several bulk magnetization vectors, one for each set of chemically distinct nuclei, precessing at different rates. Only the differences in precession frequencies (kHz range) are needed to describe the system, not absolute values (MHz) [85 (p. 17-18)].

In NMR spectroscopy, it is common to add a compound to the sample that has a signal designated as

To excite all nuclei simultaneously in a sample, a range of frequencies must be applied. In order to achieve a simultaneous range of frequencies, the excitation pulse is kept very short (one to ten μsec). Due to Heisenberg's uncertainty principle, the short pulse time does not allow an exact frequency to be established; this ensures that the sample is irradiated with the required range of frequencies. If the pulse is longer, the sample can be irradiated with a more exact frequency [85 (p. 64)].

Coupling

Through chemical bonds, nuclei can 'communicate' with each other. The communication, or coupling, shows up in the splitting patterns found in NMR spectra. Coupling can occur between any nuclei with spin, but it is most common to study the coupling between protons. The combination of information from chemical shift and coupling patterns is in most cases sufficient to identify the molecule in question.

The coupling patterns arise because the chemical shift of one nuclei depends on the spin state of its neighbor(s). Consider a simple model of two coupled spin-half nuclei, A and B. If the spin of B is parallel to the static magnetic field, it adds to the total field A experiences (as illustrated in **Figure 3.10b**), thus increasing its chemical shift. Since B can have two different spin states, A can have two slightly different chemical shifts. Furthermore, A affects B in the same manner [87 (p. 134-136 and 233-247)]. From the vector model point-of-view, the A nuclei give rise to two magnetization vectors with slightly different Larmor frequencies (and therefore chemical shifts), and likewise for the B nuclei [85 (p. 17-18)].

The coupling between two nuclei is constant and independent of the strength of the static magnetic field. Coupling occurring through chemical bonds is called scalar or J coupling, and is measured in Hz. The size of the coupling constant depends largely on the number of bonds between the interacting nuclei. In general, the more bonds between two coupled nuclei, the smaller the coupling constant is [87 (p. 233-247)].

3.5.2 Instrumental considerations

Preceding acquisition of an NMR spectrum, the instrument and sample need to be prepared. In NMR, solvents are usually deuterated for two different reasons. First, if the solvent contains protons the signals of the analytes would be inundated because their concentration would be much lower than that of the solvent. ^2H is NMR active ($I = 1$), but it resonates at different

frequencies than ^1H and therefore does not show up in ^1H spectra. Second, the static magnetic field of NMR instruments has a tendency to drift over time. To correct for the drift, the frequency of the deuterated solvent is measured with a separate coil throughout the experiment. The process of locating the deuterated signal is colloquially called *locking* and is done prior to initiation of experiments [85 (p. 105-107)].

Samples are lowered into a probe inside the static magnetic field. The probe holds the coils needed to conduct NMR experiments. In a process called *tuning and matching*, the circuitry of the probe is optimized for the sample at hand. It ensures that the probe is utilizing the frequencies at which measurements of the nuclei are most sensitive and that the circuitry can operate at the necessary radio frequency energy levels [85 (p. 103-105)].

NMR relies heavily on a perfectly homogenous static magnetic field to acquire high-quality data. The Larmor frequency of a nucleus depends on the magnetic field it is in; any inhomogeneity will change the chemical shift of the nucleus slightly. The consequence is a spectrum with bulging and broad peaks. The main static magnetic field does not provide a sufficiently homogenous field on its own. Therefore, superconducting coils are used to supply an additional small and adjustable magnetic field. The process of homogenizing the net magnetic field is called *shimming* [85 (p. 107-112)].

3.5.3 One-dimensional nuclear magnetic resonance spectroscopy

In NMR-based metabolomics, it is common to use a solvent with a higher proportion of H_2O compared to D_2O to ensure that the signals from labile protons are not lost. The large amount of H_2O complicates the acquisition of one-dimensional spectra. The water peak often completely dominates a spectrum and can shift the dynamic range such that the solutes in low concentrations are no longer detectable. To counter this, various specialized pulse sequences that can suppress or remove the water (solvent) signal have been developed [85 (p. 480-486)]. The following short description highlights different ways of eliminating a solvent peak from a spectrum. The pulse sequences are described in the **Appendix (section 8.1.2)**.

Presaturation of the solvent signal is the simplest and most robust technique. In NMR, *saturating* a signal means stopping the magnetization vector from recovering after the previous pulse, thus making it impossible to observe. In some cases, this is an undesirable effect but in presaturation it is exploited to remove the water signal. While presaturation has the goal of

killing the solvent signal, zero excitation separates the solvent magnetization vector from the others by placing it in the $+z$ -axis prior to detection. Solvent suppression techniques utilizing pulsed field gradients (PFGs) are the most effective at removing the solvent signal. PFGs are time-dependent magnetic fields applied along the z -axis of the sample such that the field no longer is equal throughout the sample. In one variant of PFG-aided solvent suppression, called excitation sculpting, two PFGs, a selective and a short 180° pulse will together ensure that all magnetization vectors except that belonging to the solvent, are in phase and measurable [85 (p. 473-474 and 480-486)].

3.5.4 Two-dimensional magnetic resonance spectroscopy

Human metabolites are estimated to number in the thousands to tens of thousands [88]. Consequently, the most common samples in metabolomics (plasma, urine, cell lysate, tissue extract, and saliva) contain numerous compounds. For one-dimensional NMR spectroscopy, this poses two major issues. First, there is significant overlap between peaks and it can be difficult to discern their true splitting patterns. Second, because of the mixture of compounds, it is nearly impossible to identify which peaks arise from the same molecule [89].

Luckily, NMR is not limited to one-dimensional spectra. A wide variety of two-dimensional experiments is available, which can provide information on splitting patterns and connectivity. In metabolomics, proton total correlation spectroscopy (TOCSY) and J-resolved spectroscopy (JRES) are most commonly used [89]. Descriptions of pulse sequences for JRES and TOCSY are given in the **Appendix (section 8.1.3)**. Observation of ^{13}C is often omitted due to lack of sensitivity because of the combination of the low natural abundance of ^{13}C (ca 1.1% [85 (p. 12)]), lower magnetogyric ratio of ^{13}C compared to ^1H , and the small size of biological samples.

TOCSY reveals which protons come from the same spin system. As mentioned above, neighboring nuclei can couple with each other and that any nuclei with spin are capable of this. If there is an unbroken chain of spin-spin coupled protons, they form a spin system and can relay their coupling from one to the other. In the resulting spectrum, spin systems are revealed by cross peaks that appear on either side of the diagonal peaks. The diagonal peaks are simply a projection of the one-dimensional spectrum onto the two-dimensional plane [85 (p. 220-230)]. See **Figure 3.11** for an illustration of a TOCSY spectrum.

The JRES spectrum provides information of the multiplicity of each peak and the size of its coupling constants. In the simplest JRES experiments, the chemical shifts and couplings appear in the direct dimension (f_2 in **Figure 3.11**), while only the couplings appear in the indirect dimension (f_1 in **Figure 3.11**). More advanced pulse sequences resulting in spectra with only chemical shifts in the direct dimension and the couplings in the indirect dimension have been developed. The advantage with the latter type of experiment is increased resolution of previously overlapping peaks [85 (p. 301-313)]. See figure **3.11** for an illustration of a JRES spectrum.

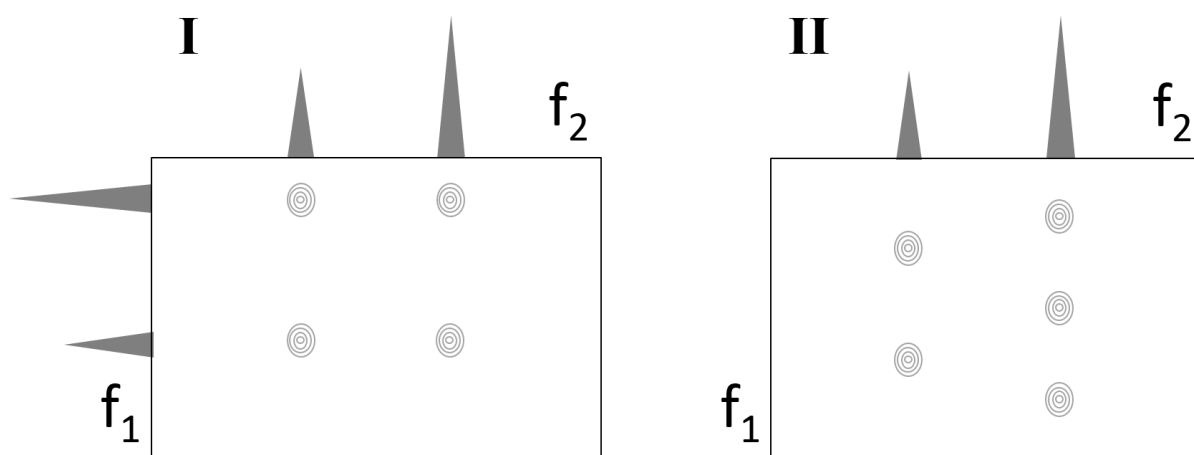


Figure 3.11: (I) Schematic TOCSY spectrum. The cross peaks indicate that the protons giving rise to the two signals are coupled. (II) Schematic JRES spectrum, with coupling pattern and size appearing in the f_1 dimension. f_2 and f_1 are the direct and indirect dimensions, respectively.

Two-dimensional NMR is time-consuming to acquire and this can be especially problematic in metabolomics studies where samples often are small. Fast acquisition methods, in which only parts of the data matrix are actually collected, can reduce experiment time drastically. NMR spectra with two or more dimensions contain few genuine peaks of interest and a lot of noise (“empty” space) [90]. The presence of few peaks makes it possible to reconstruct the whole spectrum from only a small number of measured data points. Non-uniform sampling (NUS), in which a random choice of data points are collected, is one of the more common fast acquisition methods. The number of data points acquired with NUS should not be less than the number of peaks expected in the final spectrum. Too few points can lead to artifacts appearing in the final spectrum [85 (p. 185-186)].

3.6 Mass spectrometry

MS is widely utilized in metabolomics for detection and identification of metabolites [91]. It is a very sensitive and selective technique. While NMR can reach low micromolar detection limits, MS can measure compounds to at least nanomolar concentrations [92, 93]. However, MS is not without drawbacks, which include moderate robustness compared to NMR, being a destructive technique, and that it usually requires more sample preparation than NMR. Most sample types encountered in metabolomics are complex; without sample preparation or prior separation MS often suffers from loss in sensitivity due to matrix effects (*e.g.* ion suppression) [92].

In MS, the ratio of the mass number (molecular mass divided by the unified atomic mass unit) to charge number (m/z) of gaseous ions is determined [87 (p. 418-438), 94]. The mass spectrometer is capable of separating ions with differing m/z values and its ability to do so can be quantified with the parameter mass resolution: $R = m/\Delta m$, where m is the mass of the ion and Δm is the full width of the mass peak at half maximum height [94]. In addition, mass spectrometers are assessed by their accuracy: $E = \frac{\text{Exact } m/z - \text{measured } m/z}{\text{Exact } m/z} \times 10^6 \text{ ppm}$, where the exact mass is calculated based on the molecular formula of the compound [95].

There are several different types of mass spectrometers available. All include an ion source, a mass analyzer and an ion detector. In the ion source, gaseous ions are generated and can also be fragmented, *i.e.* split into smaller pieces. The fragments can aid in identification of analytes. There is a variety of ion sources available; the biggest difference between them is the type of ionization they provide and the degree of fragmentation that occurs. Ion species are generated by gain or loss of electrons or association or disassociation with a charged compound (*e.g.* H^+ , NH_4^+ , CH_4^+ , $\text{C}_4\text{H}_{10}^+$). In the mass analyzer, ions and fragment ions are separated based on their m/z . Finally, the ions are detected either directly or indirectly. In direct detection, the ions induce a current by hitting the detector wall, causing electrons to be emitted. Next, the electrons hit the wall and cause even more electrons to be emitted. The process is repeated several times to amplify the current before the signal is sent to a computer [87 (p. 418-438)]. Indirect detection is used for mass analyzers that keep ions moving in a circular and periodic manner. The ions in motion induce a current in electrodes surrounding the mass analyzer. The detected current, similar to an FID, requires Fourier transformation to obtain the mass spectrum [96 (p. 93-94)].

The Orbitrap™ mass analyzer is popular to use in metabolomics studies because of its high ($R > 100\,000$) resolution and excellent mass accuracy ($E < 5$ ppm), which greatly aid in the identification of metabolites [97-101]. As the name implies, the Orbitrap is a trap; it holds the ions by the help of an electrical field (shown in **Figure 3.12.I**). The ions enter the Orbitrap in concentrated packets and their trajectories are manipulated so they orbit the spindle. After a short while, the ions spread out and form rings that move axially along the spindle. The frequency (ω) of the motion depends on the mass and charge of the ion: $\omega \propto 1/\sqrt{m/z}$. Outer electrodes detect the motion, and the resulting data can be Fourier transformed to reveal which m/z values are present [102].

Most commercially available Orbitraps are hybrids, *i.e.* a combination of two or more mass analyzers. The additional mass analyzer is usually a quadrupole or linear ion trap and it is often used to filter out unwanted ions. Between the first mass analyzer and the Orbitrap there is an additional trapping device to store ions. The device is a type of quadrupole, called a C-trap because of its bent shape, and it can inject packets of ions into the Orbitrap. Finally, since the Orbitrap is unable to fragment ions, a collision cell must be provided if controlled fragmentation is needed [103].

Quadrupole mass analyzers are made up of four parallel rods, as illustrated in **Figure 3.12.II**. Through two of the opposing rods, radio frequency (RF) voltage is applied and through the other two a direct current (DC). The currents are adjusted so that the desired ion(s) have a stable oscillatory trajectory through the poles [87 (p. 430-432)]. The C-trap in Orbitrap instruments uses only RF-voltages when trapping and storing the ions, but a high-voltage DC pulse is applied when the ions are injected into the Orbitrap [98]. If there is a collision cell present, the C-trap can direct the ion package into this for fragmentation. Within the collision cell, an applied voltage accelerates the ions and they collide with neutral gas atoms or molecules (helium, argon or nitrogen). The collision cell is usually an octopole version of the quadrupole and is capable of returning the fragments to the C-trap [104].

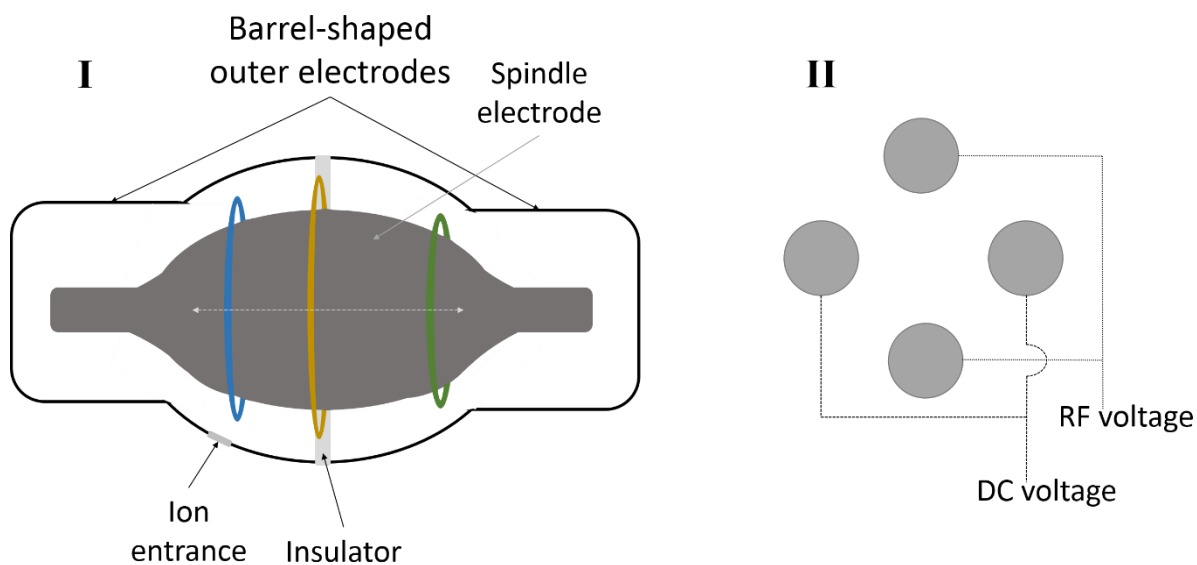


Figure 3.12: Schematic illustration of the Orbitrap mass analyzer (I) and cross section of quadrupole mass analyzer (II, figure adapted from [87 (p. 431)]). In the Orbitrap mass analyzer, the ion rings (blue, yellow and green) move axially along the spindle electrode, *i.e.* back and forth in the direction of the grey dashed arrow.

3.7 Liquid chromatography

In chromatography, compounds in a sample are separated from each other based on their different chemical properties. The sample is carried through the system by a mobile phase (MP), which is either a liquid (liquid chromatography, LC) or a gas (gas chromatography, GC). Separation occurs by interaction with the stationary phase (SP), usually held inside a column, and the MP (only in LC). Generally, in LC the functional groups of the SP (*e.g.* C₁₈-chains) are attached to a support material, but in some cases the support material itself can function as an SP. The most common type of SP “support” is porous silica particles [96 (p. 2-15)].

There are several different types of chromatographic principles, and which one is used depends on the sample composition and the chemical properties of the analyte(s) in question. Reversed phase (RP) chromatography is the most widely used principle and it separates compounds based on differences in their hydrophobicity [96 (p. 64-80)]. Although the vast majority of metabolites found in the human body are ionic or hydrophilic [18 (p. 14-15)], RP chromatography is often employed in metabolomics, but it has issues with co-elution of the most polar and ionic metabolites [105, 106]. Hydrophilic interaction chromatography (HILIC) provides an alternative for separation of hydrophilic and ionic compounds [107, 108].

HILIC columns are available in many different chemistries generally classified as neutral, charged or zwitterionic. Examples of each type are diol, amine (low pH), and sulfobetaine, respectively. Each specific type has different strengths and weaknesses, and choosing the right one is often not a straightforward task [109-111]. Zwitterionic (ZIC) HILIC is appealing due to its ability to separate anionic and cationic species simultaneously, and has been used in a variety of metabolomics studies [112-116]. In **Figure 3.13.I**, the common sulfobetaine ZIC SP is shown.

The MP in HILIC is usually a mixture of a high percentage (60-95%) of organic solvent, most often acetonitrile, and a low percentage of aqueous solvent (40-5%). The aqueous fraction usually has some form of pH control, either an acid, a base or a pH-adjustable salt (*e.g.* ammonium formate/acetate/bicarbonate). Choice of percentage organic phase, pH, and salt can have a large impact on elution time and order [111].

The HILIC materials have a high affinity for water and when the “SP” is exposed to MP, an immobilized layer of water forms as illustrated in **Figure 3.13.II** [117]. It is an important aspect of the retention mechanisms occurring in HILIC; analytes are in part separated by their differing degree of diffusion into the water layer, which is believed to constitute the SP. Electrostatic interactions and hydrogen bonding contribute to the overall retention mechanism as well. Increasing the percentage of water in the MP reduces its difference in polarity compared to the water layer. Consequently, analytes become more soluble in the MP and can elute more easily. Water is therefore considered to have high elution strength in HILIC, while acetonitrile has low (opposite of RP chromatography) [110, 118, 119].

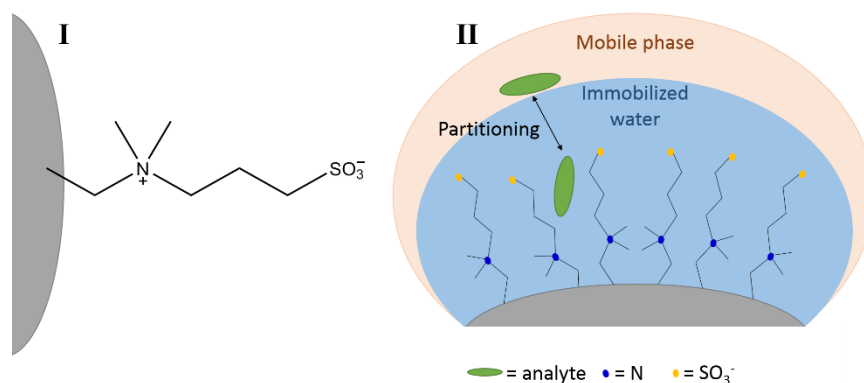


Figure 3.13: (I) The molecular structure of sulfobetaine, commonly used for ZIC-HILIC. (II) The immobilized water layer in HILIC. The grey area in both figures represents the surface of a porous silica or polymeric particle.

Metabolomics samples are usually small in size because their biological sources are often inherently limited. In addition, metabolites are present at low concentrations (nano- to millimolar) [18 (p. 14-15)]. As a consequence, analysis methods must be sensitive to reach adequate detection limits (ng/mL). Mass spectrometers are highly sensitive and can be connected to LC instruments, but additional adjustments of the LC method itself can increase sensitivity.

Most columns in LC have inner diameters (ID) of 4.6 or 2.1 mm and are 50-150 mm long. When a small sample (0.5-5 μL) is introduced into a column of such dimensions, it will be radially diluted due to diffusion, as illustrated in **Figure 3.14**. While the sample components travel through the column they will also be diluted longitudinally. The band of analytes that reaches the detector will have a lower concentration than when it was injected. Even though the MS is mass sensitive, LC-MS with an electrospray ionization (ESI) source is concentration-sensitive, *i.e.* the intensity of the detected signal depends on the concentration of the analyte entering the MS. By reducing column ID radial dilution is decreased and the concentration of analyte in the band reaching the MS will be increased, thus leading to a higher detected signal intensity and higher sensitivity. When the column ID is between 0.01-0.10 mm, the method is prefixed with nano because of the low flow rate of the MP (20-1000 nL/minute) [96 (p. 47-104), 120].

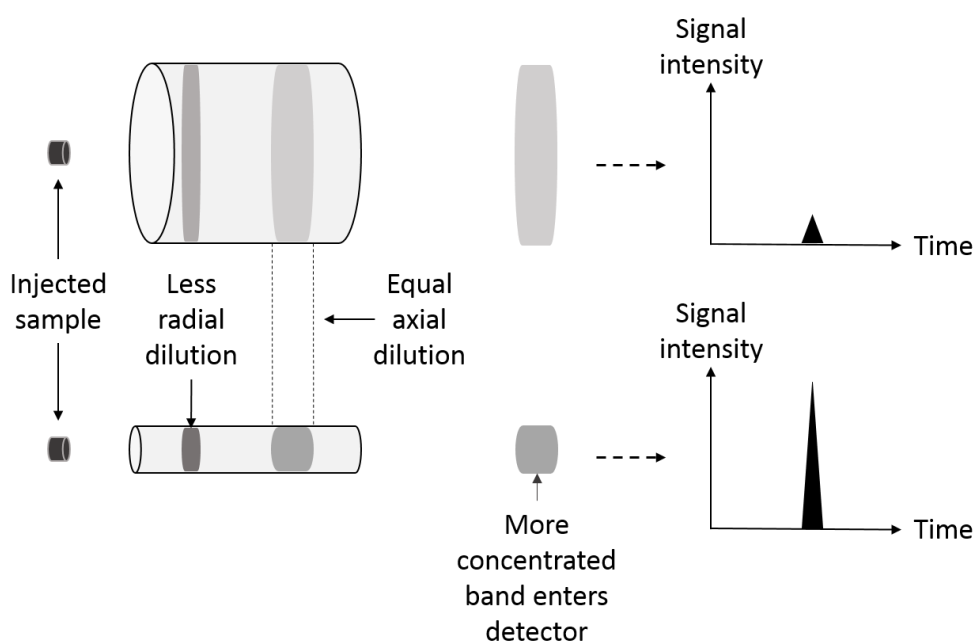


Figure 3.14: Effect of decreasing the inner diameter of the LC column. Longitudinal dilution is not affected by column ID, only length. Figure adapted from [121].

3.8 Electrospray Ionization

MS measures ions in the gas phase, thus connecting LC to MS requires a transition from liquid phase analytes to gaseous ions. There are several different interfaces available, but electrospray ionization (ESI) is popular in metabolomics because it is suitable for charged or easily ionizable compounds [96 (p. 86-88), 122, 123]. In ESI, the MP from the LC column enters a heated capillary with a high voltage (up to ± 5 kV) applied to it; at the outlet of the capillary a Taylor cone forms, from which charged droplets can escape as illustrated in **Figure 3.15**. As solvent evaporates, the charge density in the droplet increases. When the surface tension of the droplet is less than the repulsive electrostatic forces, the droplet explodes. The process is repeated, and in the end, gaseous ions are formed [96 (p. 86-88)]. In addition, ions can be emitted ('kicked out') of a small droplet. Both mechanisms are thought to contribute to formation of gaseous ions [124 (p. 3-26)]. Analytes can carry a charge before reaching the ESI capillary if they are acids or bases and if the MP has a pH suitable for deprotonation/protonation. In addition, at the walls of the capillary the MP can be oxidized or reduced, depending on if the ESI is run in positive or negative mode. The electrochemical reactions can further alter the pH of the MP [125]. When MS is coupled with ESI, the system as a whole becomes a concentration sensitive, because the gas-phase analyte ion current is dependent on the concentration of ions in the ESI spray [126].

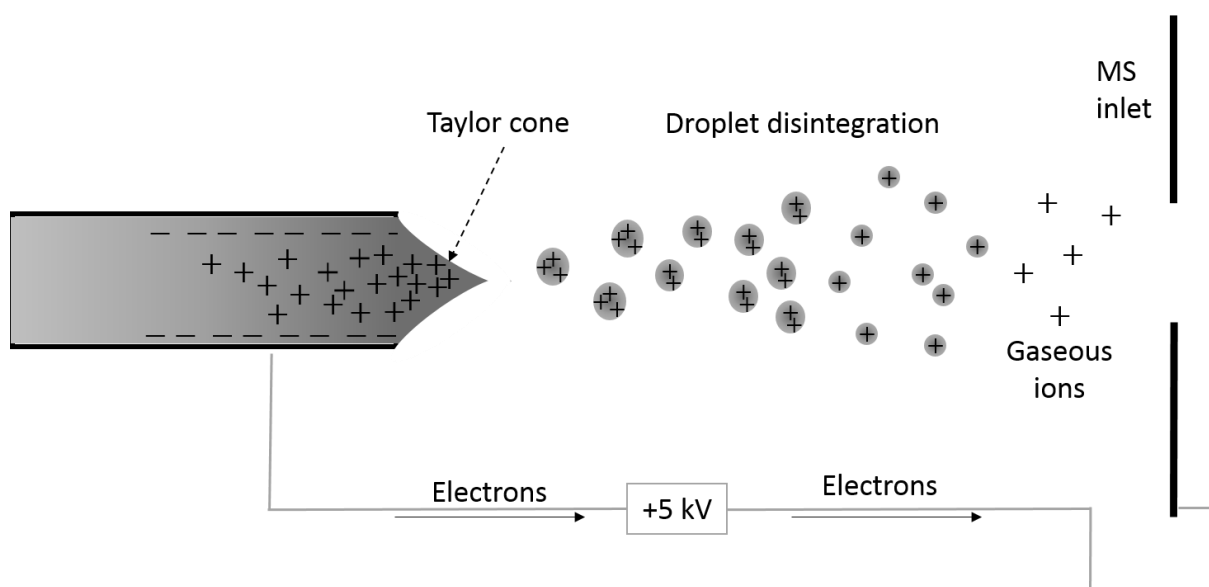


Figure 3.15: Schematic illustration of ESI source and ionization. Adapted from [96 (p. 87)] and [87 (p. 427)]

3.9 Multivariate statistical analysis of metabolomics data

In metabolomics, it is common to measure more than one variable per sample and the collected data is therefore multivariate. An important aspect of metabolomics is finding biomarkers among all the measured variables. The biomarkers should be able to distinguish different sample types, *e.g.* healthy and diseased or treated and untreated [127 (p. 221-250)]. The scouting for biomarkers is greatly aided by multivariate statistical analysis methods capable of extracting possible trends in the data. In some cases, the identified trends can lead to identification of specific compounds.

3.9.1 Pre-processing of metabolomics data

Metabolomics NMR and MS data are complex and require “clean-up” prior to statistical analysis. The clean-up ensures that only biologically relevant data affects the outcome of the analysis. Several different steps must be taken to maintain the quality of the dataset. However, no two datasets are alike and a “one-size-fits-all” method does not exist. Each dataset should be considered in context of how the biological samples were acquired and handled and how the analytical analysis was carried out.

The first pre-processing steps involve removing noise and non-essential data regions to simplify the data matrix and avoid accidentally obscuring meaningful signals. In LC-MS, this includes removing the signals from the MP and de-isotoping the spectrum (combining the isotope peaks corresponding to the same compound into one peak). Sometimes, a signal-to-noise ratio threshold is added as well; any peaks below the threshold are not included in the final data matrix. In NMR data, no signal areas of the spectrum, residual water signal, and any signal from compounds that were added to the sample (*e.g.* internal standard) are removed. A signal-to-noise ratio threshold can be applied as well, if deemed necessary. The data should be visually inspected after processing to ensure that the correct peaks and regions have been removed [128].

The next step is to remove baseline distortions and ensure that all corresponding peaks are aligned [128]. Correction of the baseline is done with processing algorithms and there are several different ones available [129, 130]. It is usually best to try moderate adjustments first to avoid “overcorrecting” or introducing artifacts. Peak alignment is in many cases achieved by simply giving a peak present in all data the same ppm or retention time throughout the dataset. In other cases, certain areas of the data can have large shifts in peak placement; these require

corrective algorithms. Finally, the data should be visually inspected again to catch any badly corrected or faulty data [128].

Two more corrective procedures can be applied to the data: normalization and scaling. The former involves dividing all data by a factor specific for each sample, *e.g.* total spectral intensity, sample weight/volume or intensity of an internal reference compound. In other words, normalization removes systemic differences between samples and ensures that all data is directly comparable. Scaling is similar, but is used to improve comparison of the same features within the dataset. Different types of scaling factors can be implemented, *e.g.* subtraction of the mean value of peak intensity, or multiplying each peak with the inverse of the standard deviation or the inverse of the squared standard deviation. Means and standard deviations are calculated with values for each peak in the dataset, *e.g.* all peaks corresponding to lactate are used to find the mean and standard deviation for lactate in the dataset [128, 131].

Finally, to reduce the number of data points the spectra or chromatograms can be binned. In binning, data is split into small intervals and each interval is given a value corresponding to peak intensity found within [128]. Binning is not as necessary today as it used to be, since computers are more powerful, but can be useful if perfect alignment of peaks is challenging.

3.9.2 Unsupervised and supervised multivariate statistical analysis methods

Multivariate statistical methods can generally be divided into two types: unsupervised and supervised. In the former, the statistical analysis is carried out without knowledge of class membership. Only the spectral or chromatographic data, called the x -data or x -matrix, is utilized. It is considered an exploratory analysis of the data. In the latter method, the class membership is included in the analysis. Usually, supervised methods are used to build a model for identifying the membership of a new sample with only x -data available [132]. Below, two of the more common unsupervised and supervised multivariate analysis methods are briefly explained.

Unsupervised: Principal component analysis (PCA)

The goal of PCA is to simplify the dataset by reducing dimensionality and see which samples co-vary, *i.e.* have similar spectral/chromatographic data. The spectral/chromatographic data is held in a matrix, with the number of rows equaling the number of samples (n) and the number

of columns equaling the number of variables (p). The reduction of dimensionality is achieved by finding the *principal components* (PCs) of the matrix. Each PC is a linear combination of the n -dimensional vectors found in each column. They are multiplied with a vector of constants (a). The constants are chosen such that the first PC accounts for the most variation found in the dataset, the second PC the second most variation, and so forth. It turns out that the most important variations in the dataset often can be expressed with a few PCs and not all of the measured variables. Plotting the values of the PCs for each sample against each other gives a *scores plot*. In addition, the contribution of each variable to the PCs can be illustrated with a *loadings plot* [127 (p. 224-228), 133].

Supervised analysis: Partial least squares (PLS) regression

In PCA, the variables in the data matrix that contributed most to the variation between samples were identified. In PLS regression, the goal is to find the variables in the data matrix that correlate highly with the class membership of the various samples. It is achieved by finding linear combinations of the original variables (now called latent variables (LV)) that correlate with the response variable(s). For most metabolomics studies, the response variables (control/treatment or healthy/diseased) are often labeled 0 and 1. Variation between the variables is not entirely ignored in PLS, but it is no longer the foundation the model is built on [127 (p. 241-245), 134].

PLS models are evaluated by cross-validation (CV). In one type of CV, the dataset is split into n random and approximately equally sized parts. The PLS regression model is calculated with all but one of the n parts. The final fraction is used to test the predictive ability of the model, and this ability is scored in the following manner [135, 136]. First, the square of the error of prediction for each individual sample assessed by the model is found, and they are subsequently summed (see **Equation 3.6**)

$$PRESS = \sum (y - \hat{y})^2 \quad \text{Equation 3.6}$$

Where, y_i is the true class (e.g. 0 or 1), while \hat{y}_i is the predicted class. The statistic is named PRESS, short for predicted residual error sum of squares. Second, the total sum of squares (TSS) for all n samples is found, which is the sum of the squared difference between the mean value of the response variable and each individual response variable (see **Equation 3.7**).

$$TSS = \sum (y - \bar{y})^2 \quad \text{Equation 3.7}$$

Finally, the score (Q^2) is calculated, given by **Equation 3.8**.

$$Q^2 = 1 - PRESS/TSS \qquad \text{Equation 3.8}$$

The process is repeated n times, and for each turn the model parameters are adjusted until a maximal Q^2 score is obtained. In other words, Q^2 estimates the predictive power of the PLS regression model. A determination coefficient (R^2) between the linear combinations and response variables is also found. Q^2 , R^2 , and Q^2/R^2 can be plotted versus the number of LVs; when the curves flatten or decrease it indicates how many LVs that are necessary to include [132, 134, 137].

Finally, and most importantly, the PLS model should be assessed with a new dataset, called the test set, which cannot have been involved with building the PLS model. It is scored the same way as in cross-validation, except that the score is called V^2 . If the model is overfitted it will perform poorly on the test set. Usually, there are no ways of remedying poor validation and test scores; it either means there are too few samples to build a satisfactory model, or there is no correlation between the measured data and its response variables [132].

3.10 Aim of Study

Glioblastoma is the most common and aggressive form of brain cancer. Patients face dismal survival rates (~ 15 months), even with comprehensive treatment regimes. Current methods for evaluating response to treatment are lacking. Thorough mapping of the metabolic response of the cancer during treatment could lead to the discovery of biomarkers appropriate for detection with *in vivo* with MRS.

*The aim of the study was to identify metabolite biomarkers for response to treatment with the chemotherapeutic agents (TMZ and YM155) in cultured glioblastoma stem cells. Liquid state NMR spectroscopy (^1H resonance frequency: 800 MHz) was used as the principal analysis method. Multivariate statistical analysis methods (PCA and PLS regression) were used to scout for biomarkers and assess metabolic differences between control and treated samples. Targeted HILIC nano-LC-MS was used to support any findings from the statistical analyses. The aim of study is illustrated in **Figure 3.16**.*

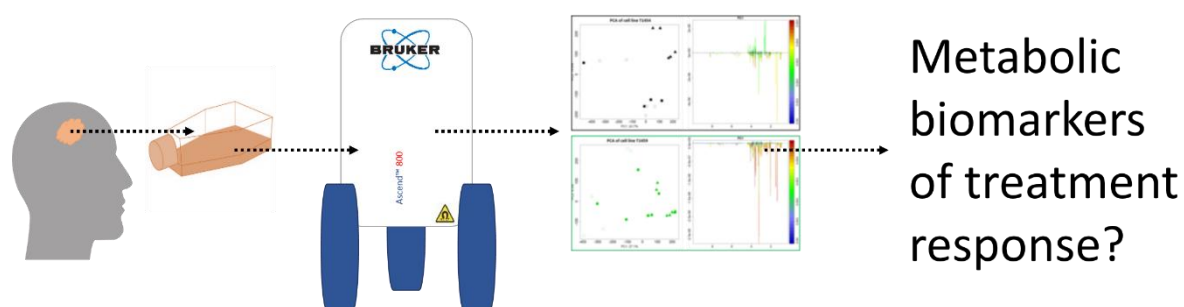


Figure 3.16: Graphical illustration of the aim of study.

4 Experimental

4.1 Instrumentation and equipment

4.1.1 Small instruments and equipment

The following containers were used: protein LoBind 1.5 mL tubes from Eppendorf (Hamburg, Germany), from now on called Eppendorf tubes; Thermo-Tubes (0.5 mL) from Advanced Biotechnologies LTD (Surrey, United Kingdom), from now on called Thermo-tubes; sample vials (0.3 mL Microvials) with screw caps from VWR (Radnor, PA, USA); 15 mL centrifuge tubes (Corning®, Therapak) from VWR; glass vials (Chromatography Autosampler Vials (1.5 mL)) and lids (snap ring caps, 11 mm blue) from VWR; volumetric flasks (MBL™, class A, borosilicate glass) from Thermo Fisher Scientific (Waltham, MA, USA), with sizes that ranged from 5 mL to 500 mL; and MP flasks (round laboratory bottles, Borosilicate 3.3) from VWR, with sizes that ranged from 25 mL to 1000 mL.

Centrifugation was done with two different centrifuges; a BIOFUGE *Fresco* centrifuge from Heraeus (Hanaum, Germany) was used in March and April 2017, and a 5424 R Eppendorf centrifuge was used from May 2017 to April 2018. For quick spinning down of samples, a Ministar Silverline microcentrifuge from VWR was used.

Weighing was done with a Mettler AT200 analytical balance from Mettler Toledo (Greifensee, Switzerland). Ultrasonic treatments were done in a Branson 200 Ultrasonic Cleaner ultrasonic bath from Emerson (Danbury, CT, USA). Sample drying and vacuum concentration were done with a *Concentrator Plus* from Eppendorf. Sample stirring was carried out on an MS2 Minishaker vortex mixer or on a *Topolino* magnet stirrer, both from IKA (Staufen im Breisgau, Germany).

Measurements of pH were done with pH paper or on either one of two pH meters; an 877 Titrino Plus pH meter from Metrohm (Herisau, Switzerland) equipped with a Primatrode pH electrode with an integrated temperature sensor (NTC), also from Metrohm, or a Thermo Orion (model 720) pH meter from Thermo Fischer Scientific equipped with a Schott “BlueLine 11 pH” glass pH electrode from Cole-Parmer (Vernon Hills, IL, USA). The pH paper (MColorpHast™) was from Millipore (Burlington, MA, USA), which is now Merck (Kenilworth, NJ, USA).

NMR analyses were done with 5 mm outer diameter (OD) Boro800-5-7 tubes from Deutero (Kastellaun, Germany). The tubes were placed in polyoxymethylene (POM) standard bore spinners from Bruker (Billerica, MA, USA). The tubes were washed in a Multi-Tube Jet Washer/Dryer (WG-1209-J1) from Wilmad-LabGlass (Vineland, NJ, USA).

Fused silica capillary used as the column body was from Polymicro Technologies (Phoenix, AZ, USA), with an ID of 100 μm and OD of 360 μm . Fused silica capillary used in connections had IDs of either 20 μm or 30 μm and OD of 360 μm , and were also from Polymicro Technologies. Red polyetheretherketone (PEEK) tubing with ID 127 μm and OD 1.6 mm was from IDEX Health and Science (Lake Forest, IL, USA). Magnets used for stirring were three by three mm round magnets from VWR. Manual inspection of capillaries was done with a microscope from Motic (Hong Kong, China).

The following fittings and valves were used, all from VICI AG International (Schenkon, Switzerland): ferrules (Vespel/graphite, 1/16", FS1.3-5), stainless steel (SS) nuts and unions (SS, 1/16", 2.5 mm bore), SS and silicone fill ports (1/16", VISF-1) and a four-port valve with an internal and exchangeable sample loop. Two different loops with volumes 50 nL and 500 nL were used. The ferrules were in-house modified to have a 0.37 mm ID. For MS analyses, the capillary column was connected to an ES 542 SS nanobore emitter (20 μm ID, 40 mm length) from Thermo Scientific. The emitter and capillary were held in separate 360 μm ID "Upchurch Microtight® Tubing Sleeves" and connected together via an "Upchurch PEEK Microtight® Connector Butt" and "MicroFingertight I Fittings", all from Sigma-Aldrich (St. Louis, MO, USA).

The Pump 11 Elite programmable syringe pump was from Harvard Apparatus (Holliston, MA, USA). Glass syringes (5-250 μL) were from SGE Analytical Science, owned by Trajan Scientific and Medical (Ringwood, VIC, Australia).

4.1.2 Large instruments and equipment

An AVIIIHD800 NMR instrument (800 MHz) with a 5 mm Triple Resonance (TCI) cryoprobe and a temperature adjustable Sample Case was used, all from Bruker.

During preparation of capillary column frits, a GC-17A Gas Chromatograph oven from Shimadzu (Kyoto, Japan) was used. Column packing was carried out with an in-house pressure bomb system, described in [138].

LC studies with ultraviolet (UV) detection were done with a Proxeon EASY-nLC pump from Bruker and an Ultimate 3000 RS Variable Wavelength Detector from Thermo Fisher Scientific (formerly Dionex). The flow cell in the UV-detector had a U cell geometry, with a 20 μm ID and 6 mm light path length. Further description of UV detection is given in the **Appendix (section 8.7)**.

The LC-MS study was done with an Agilent 1200 series (G1376) capillary pump with an 1100 series degasser (G1379), both from Agilent (Santa Clara, CA, USA). The Q Exactive™ Hybrid Quadrupole-Orbitrap™ mass spectrometer equipped with a “Nanospray Flex” ion source was from Thermo Fischer Scientific.

4.2 Chemicals

The water used in the project was either type I water purified with a Milli-Q Integral purification system with Q-POD (0.22 μm filter) dispenser from Merck (formerly Merck-Millipore) or HPLC grade water from VWR. Unless specifically noted, “water” refers to type I water.

4.2.1 Chemicals specific for nuclear magnetic resonance spectroscopy

Nitric acid (HNO_3 , AnalaR NORMAPUR® 65%) was obtained from VWR. Potassium hydroxide pellets were from KeboLab (Bohus/Sweden), deuterated water (D_2O , 99% D) was from Cambridge Isotope Laboratories (Andover, MA/USA), potassium dihydrogen phosphate (powder) was from Merck, trimethylsilyl propionic-2,2,3,3- d_4 acid (TSP, D 98%) was from Cambridge Isotope Laboratories, and sodium azide (NaN_3) had unknown purity and unknown producer.

4.2.2 Chemicals specific for liquid chromatography based analyses

A Frit Kit from Next Advance (Troy, NY, USA) was used. It contained formamide, Kasil 1 and Kasil 1624, each stored in separate flasks. Kasil 1 is a solution with a 2.50 weight ratio potassium to silicate (29.1 Wt. %), in addition to water (70.9 Wt. %). Kasil 1624 is a solution with a 1.6 weight ratio of potassium to silicate (24 Wt. %), in addition to water (76 Wt. %).

SeQuant® ZIC®-HILIC 5 µm particles were collected from a column from Merck (formerly Merck-Millipore). Nitrogen gas (>99.99%) was from Praxair (Oslo, Norway).

HPLC quality acetonitrile was from VWR, ammonium formate ($\text{NH}_4\text{CH}_2\text{O}$, $\geq 99.0\%$ purity) and formic acid (reagent grade, $\geq 95\%$) were both from Sigma-Aldrich. Buffer solutions for pH calibration (TITRINORM®, pH 4.00 and 7.00 ± 0.02 (phosphate buffer)) were from VWR.

Tryptophan (chemical reference substance) was from European Pharmacopeia Reference Standard (Strasbourg, France) and toluene was from Rathburn Chemicals Ltd. (Walkerburn, Scotland, UK). Citric acid (ACS reagent, $\geq 99.5\%$) and sodium lactate (~98%) were from Sigma-Aldrich.

4.3 Cell samples

Primary glioblastoma cells were received from the Vilhelm Magnus Laboratory for Neurosurgical Research, Institute of Clinical Medicine, University of Oslo (UiO). Fellow Marit Christensen or research technician Maria Ewa Walewska cultured the cell samples given in **Table 4.1** and **Figure 4.1**. The cells were treated with either YM155 (sepantronium bromide (S1130)) from Selleck Chemicals (Munich, Germany and Houston, TX, USA) or temozolomide/TMZ (T2577) from Sigma-Aldrich. The control samples were “treated” with dimethyl sulfoxide (DMSO, unknown purity) from Sigma-Aldrich.

Details on the chemicals used and how the cells were cultured are given in the **Appendix (section 8.3)**. For details on cell samples used in the preliminary NMR analyses, see the **Appendix (section 8.4)**.

Table 4.1: Overview of samples analyzed with NMR. For each sample the sample number, cell line it originated from, treatment, date collected and identifier are listed. All samples contained approximately 3 million cells. Samples treated with DMSO were considered control samples, and are highlighted with gray. Five different cell lines were analyzed, each with one control group and two different groups treated with either TMZ or YM155. Christensen cultured the cells harvested before March and April 2017. Walewska cultured the cells harvested August-November 2017.

Sample Number	Cell line	Treatment	Date sample collection (day/month/year)	Identifier
45	T1454	DMSO	07/01/2017	D45
46	T1454	DMSO	07/01/2017	D46
21	T1454	DMSO	10/01/2017	D21
50	T1454	DMSO	25/07/2017	D50
51	T1454	DMSO	25/07/2017	D51
41	T1454	TMZ	07/01/2017	T41
42	T1454	TMZ	07/01/2017	T42
19	T1454	TMZ	10/01/2017	T19
52	T1454	TMZ	25/07/2017	T52
43	T1454	YM155	07/01/2017	Y43
44	T1454	YM155	07/01/2017	Y44
20	T1454	YM155	10/01/2017	Y20
53	T1454	YM155	25/07/2017	Y53
54	T1454	YM155	25/07/2017	Y54
55	T1456	DMSO	04/08/2017	D55
56	T1456	DMSO	04/08/2017	D56
57	T1456	TMZ	04/08/2017	T57
58	T1456	TMZ	04/08/2017	T58
62	T1456	TMZ	10/08/2017	T62
59	T1456	YM155	04/08/2017	Y59
60	T1456	YM155	04/08/2017	Y60
61	T1456	YM155	04/08/2017	Y61
63	T1456	YM155	10/08/2017	Y63
28	T1459	DMSO	19/01/2017	D28
24	T1459	DMSO	11/01/2017	D24
25	T1459	DMSO	11/01/2017	D25
39	T1459	DMSO	08/01/2017	D39
40	T1459	DMSO	08/01/2017	D40
64	T1459	TMZ	01/08/2017	T64
65	T1459	TMZ	01/08/2017	T65
22	T1459	TMZ	11/01/2017	T22
26	T1459	TMZ	19/01/2017	T26
35	T1459	TMZ	08/01/2017	T35
36	T1459	TMZ	08/01/2017	T36
23	T1459	YM155	11/01/2017	Y23
27	T1459	YM155	19/01/2017	Y27
37	T1459	YM155	08/01/2017	Y37
38	T1459	YM155	08/01/2017	Y38
102	T1459	YM155	29/09/2017	Y102
104	T1459	YM155	18/10/2017	Y104
105	T1459	YM155	18/10/2017	Y105
106	T1459	YM155	18/10/2017	Y106
66	T1547	DMSO	04/08/2017	D66
67	T1547	DMSO	04/08/2017	D67
73	T1547	DMSO	15/08/2017	D73
98	T1547	DMSO	18/10/2017	D98

99	T1547	DMSO	18/10/2017	D99
68	T1547	TMZ	04/08/2017	T68
74	T1547	TMZ	15/08/2017	T74
75	T1547	TMZ	15/08/2017	T75
71	T1547	TMZ	08/08/2017	T71
72	T1547	YM155	08/08/2017	Y72
69	T1547	YM155	04/08/2017	Y69
70	T1547	YM155	04/08/2017	Y70
100	T1547	YM155	18/10/2017	Y100
101	T1547	YM155	18/10/2017	Y101
103	T1547	YM155	18/10/2017	Y103
76	T1548	DMSO	28/06/2017	D76
77	T1548	DMSO	28/06/2017	D77
78	T1548	DMSO	28/06/2017	D78
85	T1548	DMSO	11/07/2017	D85
86	T1548	DMSO	11/07/2017	D86
79	T1548	TMZ	28/06/2017	T79
80	T1548	TMZ	28/06/2017	T80
81	T1548	TMZ	28/06/2017	T81
87	T1548	TMZ	11/07/2017	T87
88	T1548	TMZ	11/07/2017	T88
82	T1548	YM155	28/06/2017	Y82
83	T1548	YM155	28/06/2017	Y83
84	T1548	YM155	28/06/2017	Y84
89	T1548	YM155	11/07/2017	Y89
90	T1548	YM155	11/07/2017	Y90

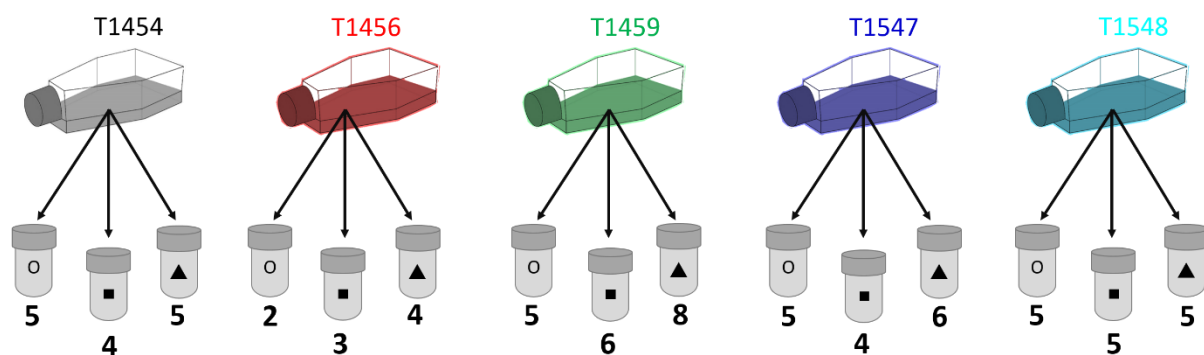


Figure 4.1: Graphical overview of the samples received from Vilhelm Magnus Laboratory for Neurosurgical Research. The shapes on the flasks denote treatment: circle = DMSO (control), square = TMZ, and triangle = YM155. The number beneath each flask represents the number of replicates of each cell line and treatment combination.

4.4 Solutions

4.4.1 Solutions used in nuclear magnetic resonance spectroscopy

NMR buffer

The NMR buffer was a 1.5M potassium phosphate monobasic buffer in D₂O at pH 7.4 with 0.1% TSP, made by first dissolving 20.4 g of KH₂PO₄ in 80 mL of deuterated water (D₂O). Then the pH was adjusted to approximately 5 by dropwise adding a strong (~12M) solution of KOH dissolved in D₂O. Once desired pH was reached, 100 mg of TSP dissolved in 10.0 mL of D₂O was added to the buffer, which was mixed with a vortex mixer. The pH was further adjusted to 7.4 by adding more ~12M KOH in D₂O and the volume was adjusted to 100 mL with D₂O. Finally, approximately 20 mg of NaN₃ was added to the buffer as a bactericide.

The pH was measured with the 877 Plus Titrino pH meter and Primatrode pH electrode. The buffer was stored in a closed volumetric flask at 4 °C between uses.

4.4.2 Solutions used for liquid chromatography based analysis

For preliminary liquid chromatography analysis with UV detection, two different MP buffers were made. Both were 30 mM ammonium formate buffers, but had either pH 3 or 4.5. The buffers were made by transferring 472.8 mg of ammonium formate to 250 mL of water. Then, the solutions were pH adjusted with formic acid to either pH ~3 or ~4.5; measurements were done with pH paper. The two different buffers (pH 3 and pH 4.5) were transferred to individual 25 mL MP flasks. HPLC quality acetonitrile was transferred to a separate 25 mL MP flask. To obtain the chosen MP composition, acetonitrile and buffer was mixed by the Proxeon nLC pump.

Three different sample solutions were made for assessing the capillary columns in the LC-UV analysis; they are described in **Table 4.2**.

Table 4.2: The three different sample solutions made for testing separation abilities of the capillary columns. Both analytes, toluene and tryptophan, were dissolved in 70% acetonitrile and 30% buffer (30 mM ammonium formate, pH 3 or 4.5) (v/v).

Solution name	Toluene concentration ($\mu\text{L/mL}$)	Tryptophan concentration (mg/mL)
S1	18	-
S2	-	1.0
S3	18	1.0

For the LC-MS analyses, the MP buffer consisted of a 30 mM ammonium formate buffer made with HPLC grade water. The pH of the buffer was adjusted to 4.5 with formic acid; measurements were done with the Thermo Orion pH meter and a Schott “BlueLine 11 pH” glass pH electrode. Prior to analysis, the buffer was mixed with acetonitrile. Each component was measured separately with graduated cylinders; the mixture had an acetonitrile/buffer ratio of 65/35 (v/v).

Two different solutions of citric acid and one of lactic acid were made, both are described in **Table 4.3**.

Table 4.3: Citric and lactic acid solutions used in direct injection to MS and LC-MS analyses. The concentration of citric or lactic acid and type of sample solvent are listed. The buffer pH was 4.5.

Solution name	Contained	Concentration ($\mu\text{g/mL}$)	Solvents	Solvent ratio (v/v)	Used for
C0	Citric acid	10	Acetonitrile 30 mM ammonium formate buffer	70/30	Direct injection
L0	Lactic acid	10	Acetonitrile 30 mM ammonium formate buffer	70/30	Direct injection
C1	Citric acid	20	Acetonitrile 30 mM ammonium formate buffer	80/20	LC-MS

4.5 Nuclear magnetic resonance spectroscopy analyses

4.5.1 Sample preparation

The various steps taken during preparation of samples prior to analysis on the NMR instrument are described below.

Cleaning NMR tubes

The 5 mm Boro800-5-7 tubes were washed with the Multi-Tube Jet Washer/Dryer. A schematic illustration of the washer is shown in **Figure 4.2**. The washing acid had a concentration of 6.5% nitric acid in water. The cup was filled three times with washing acid, letting it become entirely empty before refilling. The tubes were rinsed five times with water, by filling the cup five times. Finally, the tubes were left upside down to air dry at room temperature, usually taking from 48 to 72 hours to dry completely.

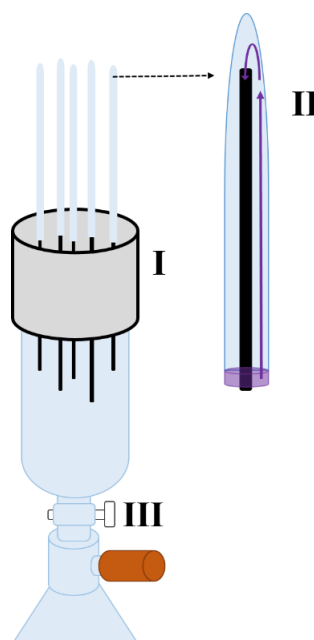


Figure 4.2: Schematic illustration of NMR tube washer. (I) Cup containing acid or water. (II) The NMR tubes were placed over the hollow black tubes, and when the aspirator (III) was opened, liquid traveled from the cup, up along the NMR tube, and out through the hollow black tube.

Preparing metabolomics samples for NMR

The cell samples described in **Section 4.3** were received as pellets. Each pellet was stored in a separate 15 mL centrifuge tube. The tubes were placed on ice and 600 μL water was added. Pellet and liquid were transferred to Eppendorf tubes and then placed in an ultrasonic bath. In the bath, the cells were subjected to ultrasonic treatment for 30 seconds on/30 seconds off 10 times, taking a total of 10 minutes. Afterward the Eppendorf tubes were transferred to a centrifuge cooled to 4 $^{\circ}\text{C}$ and spun for 5 minutes at 12,000 rpm, to spin down cellular debris. Before transfer to NMR tubes, 540 μL of the supernatant was moved to a fresh Eppendorf tube

and mixed with 60 μL NMR buffer. An additional 40 μL of supernatant was set aside in Thermo-Tubes and stored at $-80\text{ }^{\circ}\text{C}$ pending LC-MS analysis.

Before NMR analysis the samples were either stored at $4\text{ }^{\circ}\text{C}$ in a refrigerator or at $7\text{ }^{\circ}\text{C}$ in the Sample Case attached to the 800 MHz NMR instrument.

4.5.2 One-dimensional nuclear magnetic resonance spectroscopy

Optimization procedure prior to acquisition of NMR data

The AVIIIHD800 NMR instrument was run with TopSpin 3.5 patch level 6 (3.5pl6) from Bruker.

Before data acquisition, each sample was left in the magnet for 15 minutes to temperature equilibrate. The probe was kept at 300.8 K with 535 L/hour gas flow, ensuring the samples experienced a temperature of $25\text{ }^{\circ}\text{C}$ during measurements. Then the deuterium signal was locked to D_2O , using the Bruker's lock made for 90% H_2O and 10% D_2O mixtures with salt. After locking, the sample was "tuned and matched" using the command *atma* (automatic tuning and matching), or *atmm* (manual tuning and matching) if *atma* failed.

Once tuning and matching were complete the sample could be shimmed, starting with reading an old shimfile (command *rsh*). The old shimfile would be from an earlier metabolomics sample with same content of water, deuterated water and salt. The shimfile was updated using Bruker's TopShim GUI function, shown in **Figure 4.3**. Under TUNING, in the *before* slot, z-x-y was chosen; in the *after* slot, z-x-y-xz-yz-z was chosen. The shimming was run until all changes of shim directions (z, x, y, xz, yz) were not larger than two digits. The size of changes was given under Report once shimming was completed. The updated shim data was saved to the old shimfile with the command *wsh*.

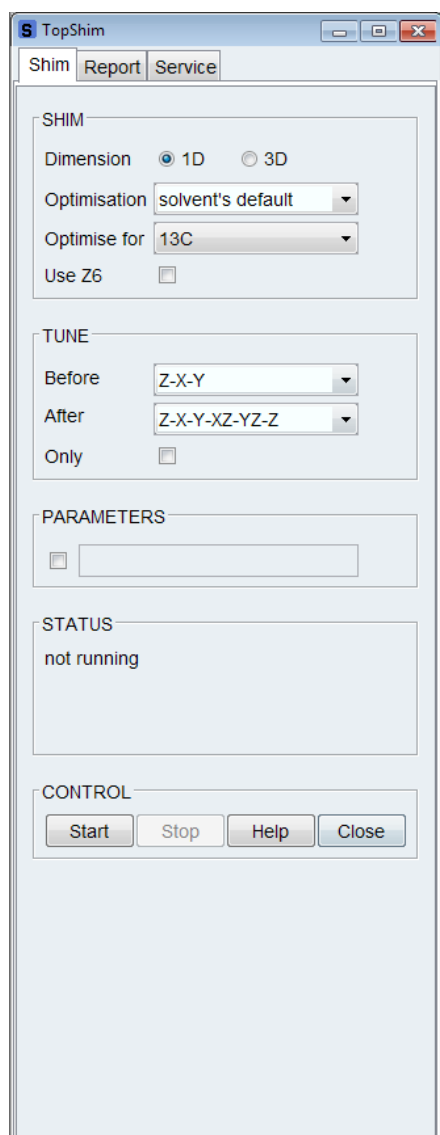


Figure 4.3: The graphical user interface (GUI) in TopSpin for shimming NMR samples. Usually, the samples were optimized for ^1H , unless carbon spectra were being acquired. 3D dimension shimming was only chosen if the sample shimming was especially difficult. The *before* and *after* options for tuning were chosen as Z-X-Y and Z-X-Y-XZ-YZ-Z, respectively.

PROTON experiment was run again, with NS = 1 and DS = 0. Finally, the resulting spectrum was phased 180 degrees, and the o1 value was identified in the spectrum. An example of a final spectrum and where the o1 value was found is shown in **Figure 4.5**.

Samples run with the automated system IconNMR™ had their shimming based on the updated shimfile. The shimfile was updated manually every 2-3 days, unless there was major drift in the magnetic field, if so the shimfile was updated daily.

The one-dimensional proton spectra were acquired with water suppression by excitation sculpting. The pulse program used during acquisition required a specific transmitter frequency offset (o1) value. The program had an automatically set standard o1 value of 3760.14 Hz, but in some cases a self-found o1 value was used instead. To establish an o1 value, the following procedure⁶ was used. First, a ^1H spectrum was recorded with number of scans (NS) 1 and number of dummy scans⁷ (DS) 0 using the experiment PROTON and pulse program zg. Then a pulse calibration was run, with the command *pulsecal*. An example result of the pulse calibration is shown in **Figure 4.4**. Once the calibration was finished, the command *p1* was written into the command line, and the number (length of pulse p1) found during pulse calibration was multiplied by four. The 90-degree pulse was now instead a 360-degree pulse. The

⁶ Taught by senior Engineer Per Eugen Kristiansen (Section for Biochemistry and Molecular Biology, UiO).

⁷ In a dummy scan, the pulsed sequence is run but no FID is acquired.

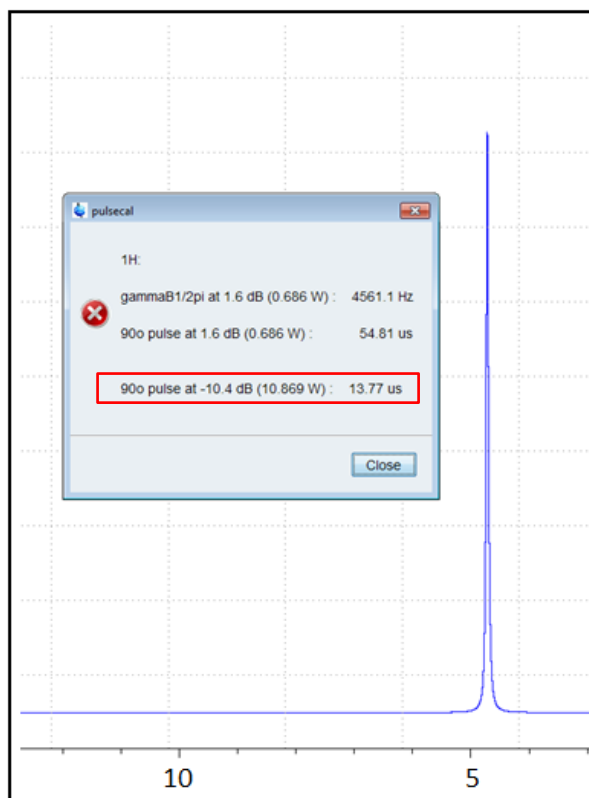


Figure 4.4: Output given by TopSpin after a pulse calibration was run. The length of the 90 degree pulse at -10.4 dB (marked with a red rectangle) was multiplied by four before the next ^1H spectrum was acquired.

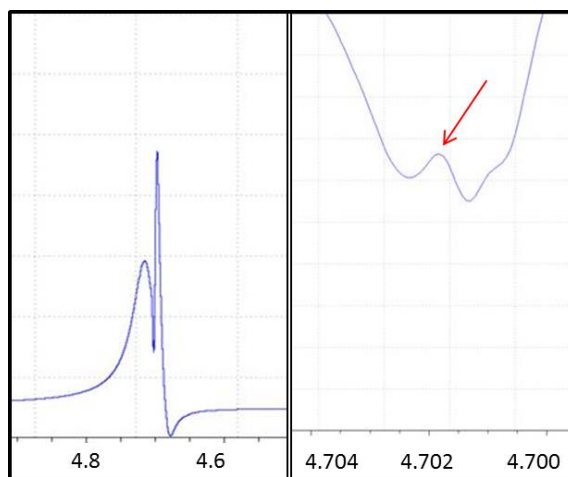


Figure 4.5: An example of a spectrum with long p1 pulse. The picture to the right is zoomed in to show the “bump” where the α_1 value was found. The bump is indicated with a red arrow.

Acquisition of one-dimensional nuclear magnetic resonance spectra

All 1D spectra used for further statistical analysis were acquired with the pulse program *zgesgp.nn*, a program customized by Nils Nyberg from Bruker. The code for the pulse program is given in the **Appendix (section 8.12)**. For each sample, pulses were calibrated prior to acquisition. Suppression of the water signal was obtained with excitation sculpting. Acquisition parameters are given in **Table 4.4**. The individual σ_1 values for each sample are given in **Table 4.5**.

Table 4.4: The general acquisition parameters for all samples analyzed with the pulse program *zgesgp.nn*. Sec is an abbreviation of second and μsec is an abbreviation of microsecond (1×10^{-6} second).

Name of acquisition parameter	Abbreviation	Settings
Size of FID	TD	32,768
Number of dummy scans	DS	4
Number of scans	NS	4096
Spectral width	SW	16.0250 ppm
Acquisition time	AQ	1.278 sec
Receiver gain	RG	203
Dwell time	DW	39.00 μsec
Probe temperature	TE	300.8 K
Sample changer temperature	SCT	280.0 K

Table 4.5: Overview of the ω_1 values used during acquisition of 1D spectrum of the samples listed in Table 4.1. The samples are numbered with the same serial number, but a treatment identifier has also been added (D = DMSO, T = TMZ, Y = YM155).

Sample number and treatment identifier	Value of ω_1 (Hz)	Sample number and treatment identifier	Value of ω_1 (Hz)
D45	3761.58	Y37	3761.91
D46	3761.85	Y38	3761.60
D21	3761.85	D66	3761.59
D50	3761.59	D67	3761.59
D51	3761.59	D73	3761.59
T41	3761.91	T68	3760.14
T42	3761.91	T74	3761.57
T19	3761.85	T75	3761.59
T52	3760.14	T71	3761.59
Y43	3761.85	Y72	3761.59
Y44	3761.85	Y69	3761.59
Y20	3761.85	Y70	3761.00
Y53	3761.59	D76	3761.59
Y54	3760.14	D77	3761.59
D55	3761.57	D78	3761.59
D56	3761.59	D85	3761.57
T57	3761.59	D86	3761.59
T58	3761.59	T79	3761.59
T62	3761.59	T80	3761.59
Y59	3761.57	T81	3760.14
Y60	3761.59	T87	3761.57
Y61	3761.59	T88	3761.59
Y63	3761.59	Y82	3761.59
D28	3761.85	Y83	3761.59
D24	3761.85	Y84	3761.59
D25	3761.85	Y89	3761.57
D39	3761.60	Y90	3761.59
D40	3761.58	D98	3760.14
T64	3761.59	D99	3760.14
T65	3761.59	Y100	3760.14
T22	3761.91	Y101	3760.14
T26	3761.85	Y102	3760.14
T35	3761.60	Y103	3760.14
T36	3761.00	Y104	3760.14
Y23	3761.60	Y105	3760.14
Y27	3761.85	Y106	3760.14

4.5.3 Two-dimensional nuclear magnetic resonance spectroscopy

Optimization procedure prior to acquisition of NMR data

All samples were locked, tuned and matched before acquisition of one-dimensional experiments, and the same parameters were kept in place for all two-dimensional experiments.

Acquisition of two-dimensional nuclear magnetic resonance spectra

The JRES experiments were acquired with the pulse program *jresgppraf* and the TOCSY experiments were acquired with the pulse program *dipsi2esgpph*. The code for the pulse programs are given in the **Appendix (section 8.12)**. General acquisition parameters for the JRES experiment and TOCSY experiment are given in **Table 4.6**. Both experiments were acquired with non-uniform sampling (NUS).

Table 4.6: The general acquisition parameters for all samples analyzed with the pulse program *jresgppraf* and *dipsi2esgpph*. F1 and F2 denote the frequency axes. Sec is an abbreviation of seconds and μsec is an abbreviation of microsecond (1×10^{-6} second).

Name of acquisition parameter	Abbreviation	Settings (JRES)	Settings (TOCSY)
nD acquisition mode	FnTYPE	NUS	NUS
Size of FID	TD	F2: 8192 F1:160	F2: 2048 F1:1024
Number of dummy scans	DS	F2: 16	F2: 32
Number of scans	NS	F2: 32	F2: 32
Spectral width	SW	F2: 16.6216 ppm F1: 0.0975 ppm	F2: 13.9503 ppm F1: 13.9503 ppm
Acquisition time	AQ	F2: 0.3080 seconds F1: 1.0259 seconds	F2: 0.0918 sec F1: 0.0459 sec
Increment for delay	IN_F	12,823.23 μsec	89.60 μsec
Receiver gain	RG	203	203
Dwell time	DW	37.600 μsec	44.800 μsec
Probe temperature	TE	300.8 K	300.8 K
Sample changer temperature	SCT	280.0 K	280.0 K
Amount of sparse sampling	NusAMOUNT	25%	25%
Number of hypercomplex points in indirect dimension	NusPOINTS	40	128

NUS data could not be processed with conventional Fourier transform algorithms. Thus, once acquired, both the JRES and the TOCSY data were processed with the Compressed Sensing algorithm [139]. The JRES spectra all had the size of the real spectrum increased to 16,384 (twice the size of F2's TD) and were calibrated by placing TSP at 0.00 ppm in all spectra. The TOCSY spectra required more processing, and the general parameters applied to all are given in **Table 4.7**.

Table 4.7: The general processing parameters for all samples analyzed with the pulse program *dipsi2esgpph*.

Name of processing parameter	Abbreviation	Settings
Size of real spectrum	SI	F2: 4096 F1: 1024
Spectrometer frequency	SF	F2: 800.03 MHz F1: 800.03 MHz
Window function	WDW	F2: QSINE F1: QSINE
Line broadening	LB	F2: 1.00 Hz F1: 0.30 Hz
Phasing mode	PH_mod	F2: pk F1: pk
Linear prediction for Fourier transformation	ME_mod	F2: None F1: LPfc
Number of linear prediction coefficients	NCOEF	F2: 0 F1: 32
Number of output points for linear prediction	LPBIN	F2: 0 F1: 2048

4.5.4 Identification of metabolites: Database searches

Identification of metabolites in the collected NMR data was carried out with manual peak searches in the following databases: Complex Mixture Analysis by NMR (COLMAR) from Ohio State University [82, 140], the Human Metabolome Database (HMDB) [80], and the Biological Magnetic Resonance Data Bank (BMRB) [141]. The BMRB was mainly used the few times the HMDB was malfunctioning or did not have the necessary NMR spectra.

For each cell line, the spectra from a representative sample of DMSO, TMZ- and YM155-treated cells were selected. A total of 15 samples were examined for metabolite identification. For each sample, all cross peaks in the TOCSY spectrum were found and grouped together according to which peaks were coupled. The multiplicity of each cross peak was identified in the JRES spectra. Each group of identified peaks was run through the COLMAR database for two-dimensional data. If that search gave no conclusive results, the group of peaks was also run through the HMDB. Additional compounds known to occur in glioblastoma cells, but with no correlation peaks, were manually searched for in the different spectra for all 15 samples.

The standard NMR spectra of the suggested molecules were found either on HMDB or BMRB and compared to the observed peaks. Certain or confident identification of a compound required that several known peaks were seen in the one-dimensional NMR spectrum and that multiplicity matched control spectra. Compounds with missing or unclear identification of peaks were categorized as ambiguously determined. If no realistic candidate compound was suggested by the databases, peaks were categorized as having “no good matches”.

4.5.5 Statistical analysis of nuclear magnetic resonance data

Preprocessing and statistical analyses were done with the programming language R, version 3.3.3 from <https://www.r-project.org/> [142]. R was run with R-Studio version 1.0.136 from R-Studio, Inc. (Boston, MA, USA). Daniel Sachse wrote the scripts used for the following statistical analyses. The scripts are presented in the **Appendix (section 8.11)**. The author has made minor changes, but Sachse did the main bulk of coding.

Pre-processing

The one-dimensional spectra were phased and calibrated manually in TopSpin 3.5p6, before being transformed to ASCII files (.txt) with the command *convbin2asc*. Additional general processing parameters and their settings are shown in **Table 4.8**.

Table 4.8: The general processing parameters for all samples analyzed with the pulse program and *zgesgp.nn*.

Name of processing parameter	Abbreviation	Settings (<i>zgesgp.nn</i>)
Size of real spectrum	SI	262,144
Spectrometer frequency	SF	800.03 MHz
Line broadening	LB	0.30 Hz
Phasing mode	PH_mod	Pk
Linear prediction for Fourier transformation	ME_mod	No
Number of linear prediction coefficients	NCOEF	0
Number of output points for linear prediction	LPBIN	0

Certain regions of the NMR spectra were removed before further analysis. This included all regions above 10.0 ppm and below -0.5 ppm, as well as 4.5 to 4.9 ppm (residual water peak), 1.16 to 1.21 ppm and 3.64 to 3.70 ppm (ethanol peaks), and 2.64 to 2.70 ppm (DMSO peak).

Normalization was either to the integral of the TSP peak or to the total integral of the spectrum (not including the TSP peak). Unit variance was used as the scaling method; each variable was divided by its standard deviation. The spectral baselines were corrected (non-polynomial smoothing) prior to multivariate statistical analyses, using an algorithm from the *FTICRMS* R package [143].

PCA

PCA was carried out with a NIPALS (Non-linear Iterative Partial Least Squares) algorithm from the package *pcaMethods* [144], and the TSP peak was excluded during the analysis. Plotting was done with functions from the package *plotrix* [145]. The R script for PCA can be found in the **Appendix (section 8.11.2)**.

PLS regression

PLS regression was carried out with a kernel algorithm (for matrices with many variables compared to samples [146]), which gives similar results as NIPALS, from the package *pls* [147, 148]. The YM155- and TMZ-treated samples were compared separately to the control samples. The PLS regression model for control and YM155-treated samples was calculated and cross-validated with data from cell lines T1454, T1459 and T1547. The dataset was split into four during CV, and a maximum of 10 LVs were calculated. The final model was tested with data from cell lines T1456 and T1548. The PLS regression model for control and TMZ-treated samples was calculated and cross-validated with data from all of the cell lines. The dataset was split into four groups for CV. The R script for PLS regression can be found in the **Appendix (section 8.11.2)**.

4.6 Methods used in liquid chromatography based analysis

LC-UV data were acquired, handled and processed with Chromeleon® 7 version 7.1.0.898 from Dionex (now Thermo Fisher Scientific). LC-MS data were acquired, handled and processed with Xcalibur™ Software from Thermo Fischer Scientific.

4.6.1 Making frits and packing capillary columns

The capillary frits were made with the Frit Kit and in accordance with the method described in [138]. The fritted capillaries were approximately 25 cm long with a 1-2 mm long frit. They were packed with ZIC® HILIC particles (5 µm) following the column packing procedure outlined in [138].

4.6.2 Testing capillary columns

The capillary column (100 µm ID/360 µm OD/length 25 cm, 11 cm with particles) was connected to the Proxeon pump via a four-gated valve with nuts, unions, ferrules and empty capillary (30 µm ID/360 µm OD/length 10 cm). The valve had an inner loop with a volume of 50 nL. The column was connected to the UV flow cell with the same type of capillary (30 µm ID/360 µm OD/length 10 cm), nuts, unions, and ferrules. Detection was carried out at 254 nm. Waste was led away from the valve via red PEEK (127 µm ID/1.6 mm OD/length 5 cm). The MP was led from the pump to the valve via a capillary (30 µm ID/360 µm OD/length 30 cm). The setup is schematically illustrated in **Figure 4.6**.

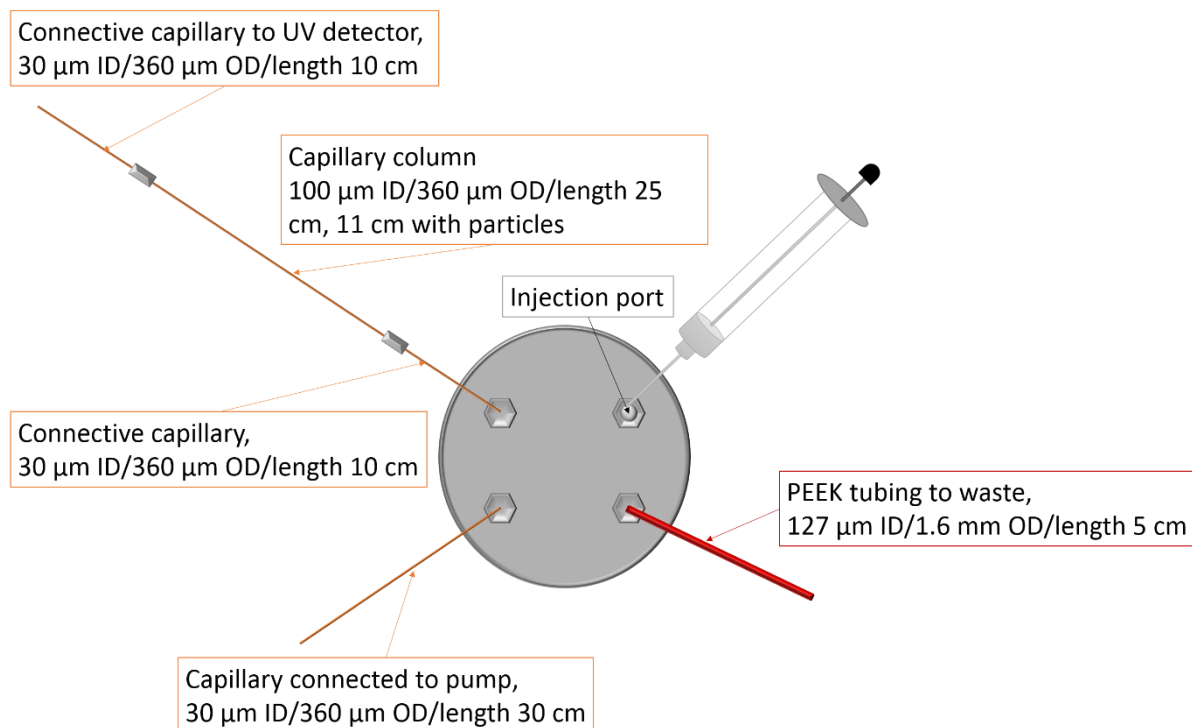


Figure 4.6: Set-up of LC-UV system. The pump was connected to the valve via a capillary (30 µm ID/360 µm OD/length 30 cm). The MP went directly to the column when the valve was in the “load position”, and via the internal loop when in the “inject position”. The 50 nL loop (internal, not shown) was filled manually with a syringe, with excess liquid passing through PEEK tubing. The capillary column (100 µm ID/360 µm OD/length 25 cm, 11 cm with particles) was connected to the valve via empty capillary (30 µm ID/360 µm OD/length 10 cm). From the capillary column, empty capillary (30 µm ID/360 µm OD/length 10 cm) led the MP to the UV flow cell (20 µm ID and 6 mm light path length, U cell geometry).

The capillary column was conditioned with 70% acetonitrile and 30% ammonium formate buffer (either pH 3 or pH 4-5) (v/v) for approximately 30 minutes before analysis. The MP composition was kept throughout the assessment of the capillary columns. The MP flow rate was kept at 1.00 µL per minute at all times. The UV lamp was turned on at least 30 minutes prior to any detection. The three different solutions outlined in **Table 4.2** were injected into the column in order.

4.6.3 Liquid chromatography-mass spectrometry analyses

Preparing metabolomics samples for LC-MS

Cell lysate from samples D45, D46, Y20, Y44, Y23, and Y27 (listed in **Table 4.1, section 4.3**) were prepared. The cell lysate (40 µL) was thawed on ice prior to protein precipitation with 80 µL cold acetonitrile. The samples were stirred with a vortex mixer for 15 seconds before waiting 30 minutes on ice. Then the samples were centrifuged for 10 minutes at 12,000 rpm at 4 °C.

From the original sample, 100 μL of the supernatant was transferred to an Eppendorf tube and dried for approximately 1 hour and 45 minutes at 30 $^{\circ}\text{C}$ by vacuum concentration (in the *Concentrator Plus*). Each dry sample was re-dissolved in 20 μL 80/20 (v/v) acetonitrile/ buffer (30 mM ammonium formate pH 4.5).

Direct injection

Connective capillary (30 μm ID/360 μm OD/length 50 cm) was held to the syringe tip via ferrule, nut, union, and fill port. The other end of the capillary was directly connected to a SS emitter (20 μm ID, 40 mm length) with PEEK fittings, sleeves, and butt. A schematic overview of the set-up is given in **Figure 4.7**. The syringe (250 μL) was held in place and pushed by a programmable syringe pump (Pump 11 Elite). The flow rate was kept at 1.00 $\mu\text{L}/\text{minute}$ at all times.

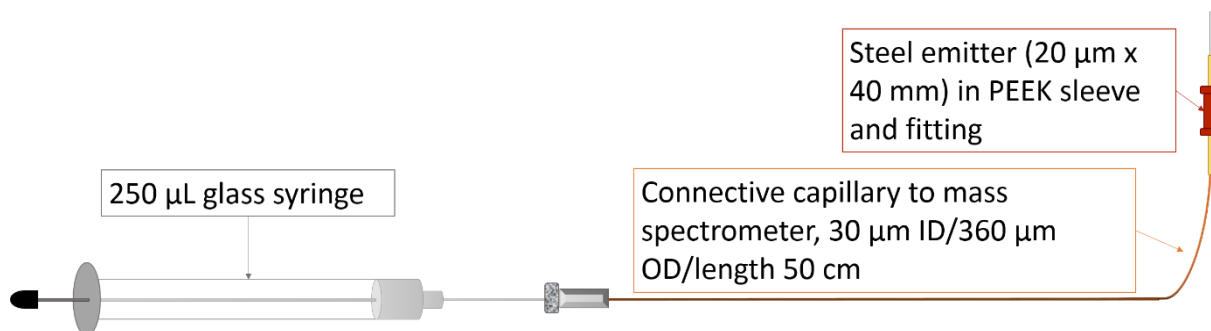


Figure 4.7: Schematic illustration of set-up for direct injection of citric and lactic acid into the MS. Not shown: Programmable syringe pump (holding the syringe in place) and the MS.

Liquid chromatography set up and parameters

The capillary column (100 μm ID/360 μm OD/length 25 cm, 12 cm with particles) was connected to the Agilent 1200 series pump via a four-gated valve with nuts, unions, ferrules and empty capillary (30 μm ID/360 μm OD/length 10 cm). The valve had an inner loop with a volume of 500 nL. The capillary column was connected to a SS emitter (20 μm ID, 40 mm length) with PEEK fittings, sleeves and butt. Waste was led away from the valve via red PEEK (127 μm ID/1.6 mm OD/length 5 cm). The MP was led from the pump to the valve via a capillary (30 μm ID/360 μm OD/length 75 cm). A schematic overview of the set-up is given in **Figure 4.8**. The MP composition was 65/35 (v/v) of acetonitrile and buffer (30 mM ammonium formate, pH 4.5), and the ratio was achieved by mixing the solvents in the same MP flask. The flow rate was kept at 0.60 μL per minute at all times.

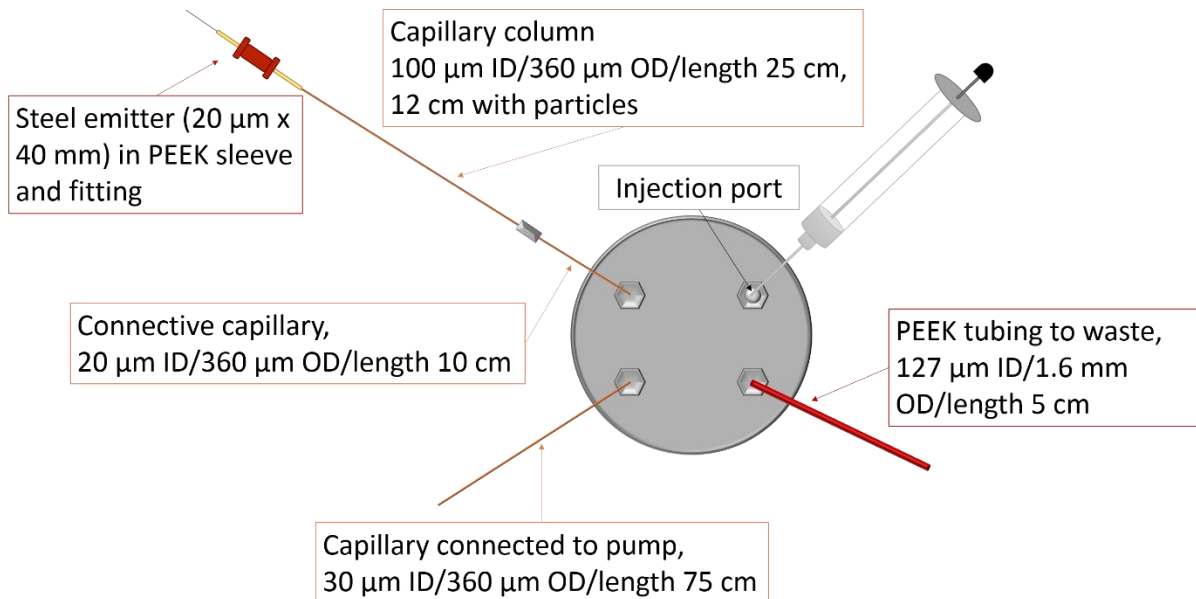


Figure 4.8: Set-up of LC-MS system. The pump was connected to the valve via a capillary (30 μm ID/360 μm OD/length 75 cm). The MP went directly to the column when the valve was in the “load position”, and via the internal loop when in the “inject position”. The 500 nL loop (internal, not shown) was filled manually with a syringe, with excess liquid passing through PEEK tubing (127 μm ID/1.6 mm OD/length 5 cm). The capillary column (100 μm ID/360 μm OD/length 25 cm, 12 cm with particles) was connected to the valve via empty capillary (30 μm ID/360 μm OD/length 10 cm). The capillary column was directly connected to an SS emitter (20 μm x 40 mm) via a PEEK sleeve, fitting and butt.

MS parameters

The mass spectrometer was run in negative mode for all experiments. Further details on MS parameters for both direct injection and when coupled with LC are given in **Table 4.9**.

Table 4.9: MS parameters used during direct injection and when coupled to liquid chromatography.

Parameter	Direct injection	Coupled with LC
Mode	Full ion scan (FIS) Parallel reaction monitoring (PRM)	Parallel reaction monitoring (PRM)
Resolving power	FIS: 70,000 PRM: 70,000 and 17,500	PRM: 70,000 and 17,500
Temperature	225-300 °C	300 °C
Capillary voltage	1.6-2.4 kV	2.2 kV
S-lens radiofrequency level	50.0	50.0
Maximal inject time	120 ms PRM: 100 ms	100 ms
Automatic gain control (AGC) target	FIS: 1 000 000 PRM: 200 000	PRM: 200 000
Inclusion list (PRM)	m/z 191.01975 ($C_6H_7O_7^-$)	m/z 191.01975 ($C_6H_7O_7^-$)
Default charge	-1	-1
Isolation window	PRM: m/z 4.0	PRM: m/z 4.0
Normalized collision energy (NCE)	35.0	35.0
Run time	∞	30 minutes

5 Results and Discussion

In the present study, the main focus was to identify biomarkers for response to treatment with chemotherapeutic agents (TMZ and YM155) in cultured glioblastoma stem cells. Liquid state NMR spectroscopy (^1H resonance frequency: 800 MHz) was used as the principal analysis method and was followed by multivariate statistical analyses (PCA and PLS regression) to scout for biomarkers and assess metabolic differences between control and treated samples. A nano HILIC LC-MS system was subsequently employed to support any findings from the statistical analyses.

In **Figure 5.1**, a graphical presentation of the workflow of the study is given. The preliminary study (red) is presented in the **Appendix (sections 8.4-8.5)**. In **sections 5.1-5.3** the findings from work done with the 800 MHz NMR (blue) is presented, except assessment of 3 mm NMR tubes, which can be found in the **Appendix (section 8.5.1)**. In **section 5.4**, the investigation by nano HILIC LC with UV or MS detection (yellow) is presented. Each section title is colored in correspondence to **Figure 5.1**.

Throughout the sections, short summaries of current conclusions will be given in *italics*.

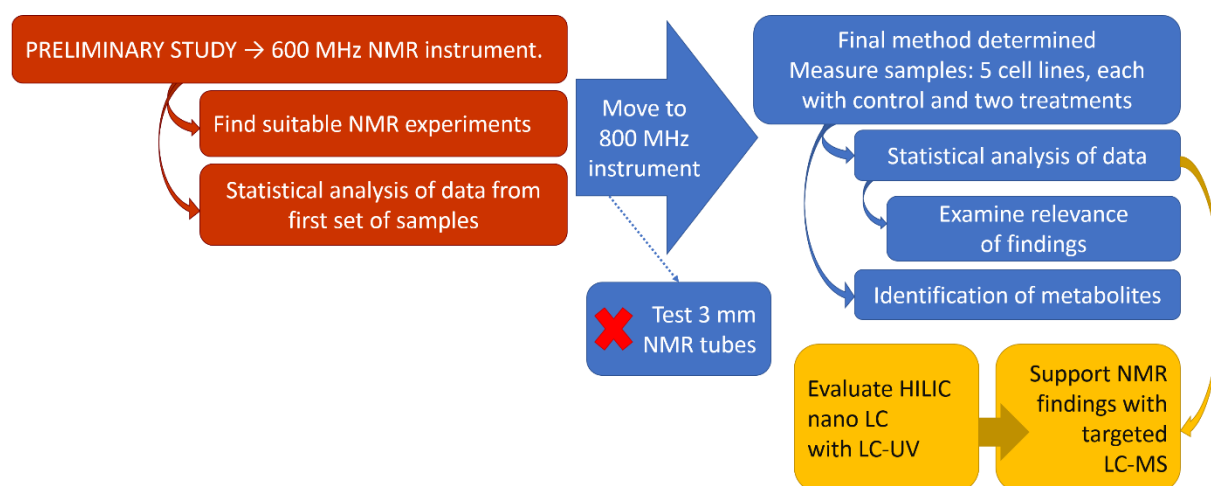


Figure 5.1: Graphical overview of study workflow.

5.1 Choice of sample type, preparation method and nuclear magnetic resonance spectroscopy experiments

An important aspect of metabolomics is the choice of sample type. In the present study, the goal was to find biomarkers of treatment response that later can be detected inside glioblastoma tumors with MRS. It was prudent to begin with surveying the cultured glioblastoma cells response to treatment. The chosen cells provided a simplified model of the cancer [149, 150]. If any likely biomarkers were found, they could be further investigated in more complex systems, *e.g.* tumor biopsies. In other words, cultured cells give findings of limited scope, but they reduce many of the confounding effects biological systems can contribute. The samples received from Rikshospitalet were enriched in glioblastoma stem cells (GSCs), as it was considered a strength to study the most aggressive cells and their response to treatment. In **Figure 5.2**, free-floating tumorspheres, which the cells formed when cultured in suspension, are shown.

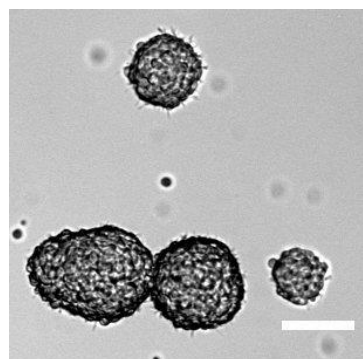


Figure 5.2: Free-floating tumorspheres in suspension culture. The cells are patient-derived and from cell line T1456. The scale bar is 100 μm . Photograph by MD Erlend Skaga.

Prior to NMR spectroscopy analysis, the content of the cells had to be extracted and suspended in a suitable solvent. Lysis and extraction were done with ultrasonic treatment and type I water because it was simple, rapid and reproducible [151]. In addition, there would be limited loss of metabolites or alteration of the sample metabolomes. Other lysis and extraction methods often involve the addition of organic solvent(s), freeze/thaw cycles and ultrasonic treatment [152-155]. These methods are time-consuming and can increase experimental errors and variability [151], and were therefore avoided if possible. The resulting NMR spectra were highly repeatable, exemplified in **Figure 5.3**, in which five control samples from the same cell line are shown. The samples were cultured, prepared, and analyzed on different days. Furthermore, the peak area of the lactate doublet at 1.33 ppm in the five control samples had relative standard deviation of 15.8%, a satisfactory precision for biological analyses [156, 157].

After extraction, a buffer had to be added to ensure equal and stable pH in all samples. Changes or differences in pH can cause peak shifting, which encumbers metabolite identification and convolutes statistical analyses [128, 158]. The buffer chosen in this study was based on a recipe from Bruker, usually used for analysis of urine samples [159]. Earlier work

carried out with this buffer had shown satisfactory repeatability and few issues with tuning, matching or shimming [160]. The same could be said for this study, as exemplified in **Figure 5.3**.

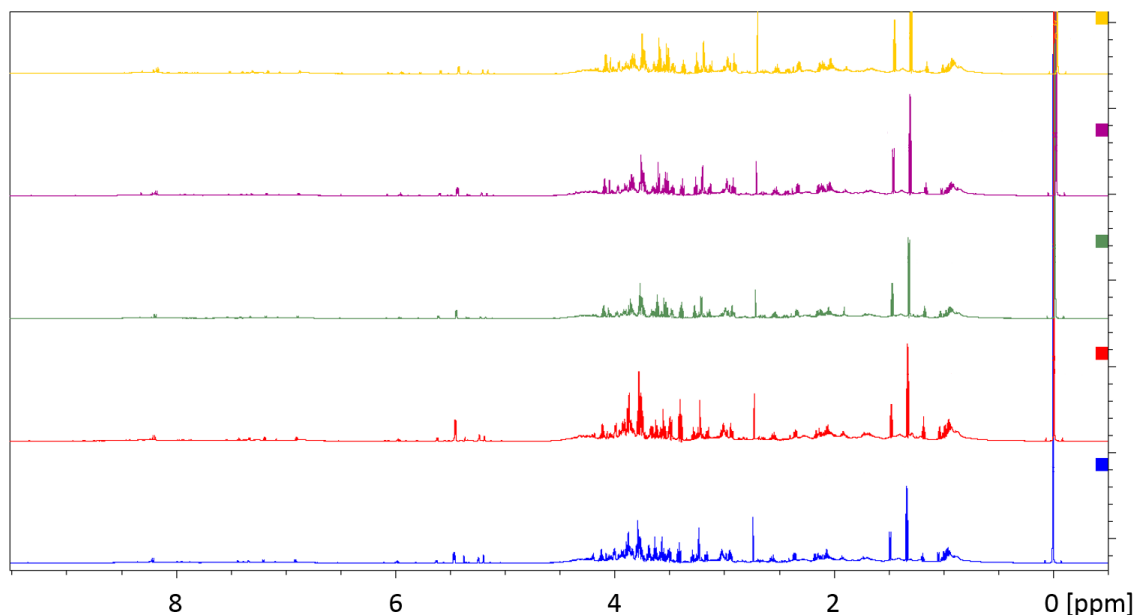


Figure 5.3: One-dimensional ^1H NMR spectra of five different control cell samples, all from the same cell line (T1548). The samples were cultured, prepared, and analyzed on different days. Solvent suppression was by excitation sculpting and NS = 4096.

The NMR experiments were chosen based on a preliminary study carried out during fall 2016, described in detail in the **Appendix (sections 8.4-8.5)**. In short, excitations sculpting was chosen as the solvent suppression method instead of presaturation because it was more effective at removing the water signal. Due to small sample sizes (3 million cells) ^{13}C NMR was not feasible. To aid identification of metabolites, TOCSY and JRES were acquired with 25% NUS coverage. Lower NUS coverage was avoided due to the high number of peaks present in the spectra.

To summarize, cultured glioblastoma cells provide a simple model of the brain cancer, suitable for preliminary and exploratory scouting for biomarkers. Organic solvents and harsh conditions were avoided during sample preparation to minimize the effect on the sample metabolomes. Both the choice of buffer and sample preparation method proved to be suitable and repeatable. Only proton NMR experiments could be carried out because of small samples and lack of sensitivity. Presaturation was deemed unsatisfactory as a solvent suppression method, and excitation sculpting was employed instead.

5.2 Identification of metabolites

Identification of metabolites was mainly aided by the COLMAR database from Ohio State University and the HMDB, as described in **section 4.5.4**. The identification was carried out for two reasons: to establish the identity of biomarkers possibly found in the following statistical analyses and to gain a perspective of the metabolic processes occurring in the cells. In **Table 5.1**, a list of identified metabolites is given. Some of the metabolites could be identified with confidence, while others were more ambiguous due to overlapping or missing peaks. Furthermore, fumaric acid, L-glycine and succinic acid gave rise to only one peak with no coupling, creatine gave rise to two singlets with no coupling, and total choline (free choline, phosphorylcholine, and glycerophosphorylcholine) was identified by only one singlet. Consequently, the identification of all five was considered ambiguous. Additionally, the presumed peak for succinic acid overlapped for some samples, thus the metabolite was not considered even ambiguously detected. Note that the table only states if a metabolite was identified, it does not give the relative intensities of the associated peaks.

All of the samples contained varying amounts of ethanol; it entered the samples during culturing because the workspace where the samples were held was regularly disinfected with ethanol. In addition, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), a common buffering agent added to the cell medium during culturing, was identified. All of the control and TMZ-treated samples also contained DMSO because it was added during culturing.

The NMR peaks and splitting patterns for all the listed compounds are given in the **Appendix (section 8.10)**.

Table 5.1: Identified compounds in glioblastoma cell samples. D = control samples, T = TMZ-treated samples, and Y = YM155-treated samples. The color code indicates the identification and how certain it was: green = confident, orange = ambiguous, grey = none/missing. The following abbreviations were used: F16BP = fructose 1,6-bisphosphate, PEP = phosphoenolpyruvate, O-PEA = O-phosphoethanolamine, P = phosphate group.

Compound	T1454			T1456			T1459			T1547			T1548		
	D	T	Y	D	T	Y	D	T	Y	D	T	Y	D	T	Y
Total choline	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Grey	Orange	Orange	Orange	Orange	Orange	Orange
Citric acid	Grey	Grey	Green	Grey	Grey	Green	Grey	Grey	Green	Grey	Grey	Green	Grey	Grey	Green
Creatine	Orange	Orange	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Grey	Orange	Orange	Orange
Cytidine	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Green	Grey	Grey	Green	Grey	Grey
Cytidine triphosphate	Grey	Grey	Green	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Green
D-Fructose	Grey	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
D-F16BP	Grey	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
D-Glucose	Grey	Grey	Grey	Orange	Grey	Green	Grey	Grey	Orange	Grey	Orange	Green	Grey	Grey	Grey
D-glucose 1-phosphate	Grey	Grey	Grey	Grey	Green	Grey	Grey	Grey	Grey	Grey	Orange	Grey	Grey	Grey	Grey
DMSO	Green	Green	Grey	Green	Green	Grey	Green	Green	Grey	Green	Green	Grey	Green	Green	Grey
Ethanol	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Fumaric acid	Orange	Grey	Grey	Orange	Orange	Grey	Orange	Orange	Grey	Orange	Orange	Grey	Orange	Orange	Orange
Glutathione	Orange	Orange	Orange	Orange	Orange	Grey	Grey	Orange	Grey	Grey	Orange	Grey	Orange	Orange	Grey
HEPES	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Lactic acid	Green	Green	Green	Green	Green	Grey	Green	Green	Grey	Green	Green	Green	Green	Green	Green
L-Alanine	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
L-Glutamate	Green	Green	Green	Green	Green	Grey	Green	Green	Grey	Green	Green	Grey	Green	Green	Green
L-Glutamine	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
L-Glycine	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
L-Isoleucine	Green	Green	Green	Green	Green	Grey	Green	Grey	Green	Green	Green	Green	Green	Green	Green
L-Phenylalanine	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Green	Green	Green	Green
L-Threonine	Green	Green	Green	Green	Green	Orange	Green	Green	Green	Green	Green	Grey	Green	Green	Green
L-Tyrosine	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Green	Green	Green	Green
L-Valine	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Green	Green	Green	Green
Myo-inositol	Green	Green	Green	Orange	Green	Grey	Green	Green	Grey	Green	Green	Grey	Orange	Green	Green
O-PEA	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Green	Grey	Grey	Grey	Grey	Grey	Grey	Grey
3-phosphoglyceric acid	Grey	Grey	Green	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
6-phosphogluconic acid	Grey	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
PEP	Green	Grey	Grey	Grey	Grey	Grey	Orange	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
Succinic acid	Orange	Orange	Grey	Grey	Orange	Grey	Orange	Orange	Orange	Orange	Grey	Orange	Orange	Orange	Grey
Uracil	Green	Green	Green	Grey	Green	Grey	Green	Green	Grey	Green	Grey	Grey	Green	Green	Green

In total, ~30 endogenic compounds could be identified with confidence or slight ambiguity. However, there were many correlated peaks that could not be assigned due to difficulty discerning couplings in TOCSY and splitting patterns in JRES. Mainly, this concerned peaks in the 3-4.5 ppm region, which can be seen in **Figure 5.4**. The peaks in this area usually arise from various carbohydrate molecules, *e.g.* fructose and glucose-6-phosphate. They are known for having many peaks and complicated splitting patterns [161], and it is therefore not surprising that it was an issue to attain reliable identifications. Commercially available programs specialized in identification of metabolites in NMR data (*e.g.* Chenomx [162]) could possibly have increased the number of assigned peaks. However, most of these programs are prohibitively expensive and were therefore not available in this study.

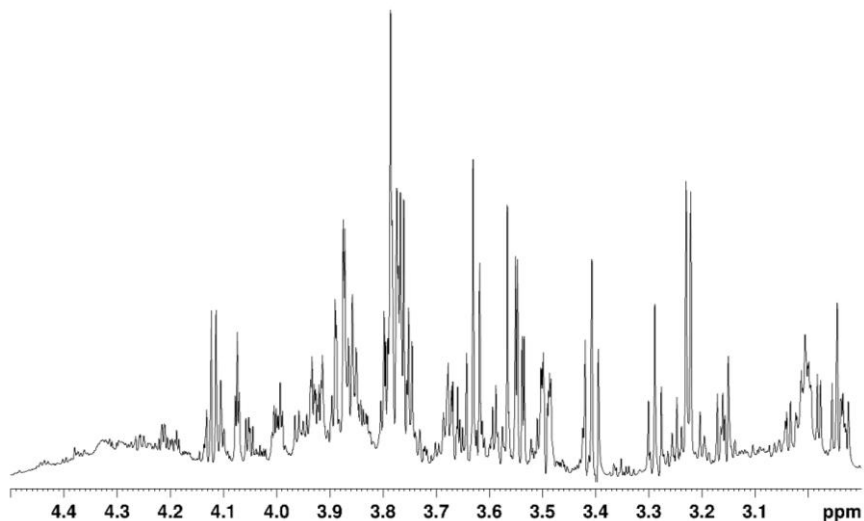


Figure 5.4: The crowded 3-4.5 ppm region of a representative control sample.

Previous studies have identified a number of similar compounds *in vitro*, including glycine, glutamate, glutamine, alanine, lactate, threonine, tyrosine, phenylalanine, *myo*-inositol, creatine, total choline, and citric acid [163, 164]. In addition, *in vivo* studies have identified lactate, alanine, creatine, total choline, *myo*-inositol/glycine, and glutamine/glutamate. The latter two pairs are often difficult to separate in MRS. It is also common for *in vivo* studies to identify *N*-acetyl-L-aspartic acid (NAA) [165-168], but it was not found in the spectral data of the present study. NAA is present at high concentrations in brain tissue, only second to glutamate [169]. Guidoni *et al.* were able to identify it in their *in vitro* study of cultured glioblastoma cells [164], while Cuperlovic-Culf *et al.* only found very low relative concentrations [163, (Supplementary Table 1)]. Perhaps the concentration of NAA was low in the present study as well, making it troublesome to even ambiguously identify.

In the introduction, **section 3.1**, the reprogramming of metabolism that cancers undergo was briefly reviewed. Glioblastoma also undergoes metabolic adaptations; several of the metabolites in **Table 5.1** are indicated to play a role in its metabolism, including glutamine, glutamate, lactate, and alanine [163, 164, 170-172]. The cells in the present study were cultured with both glucose and glutamine present and without restriction of oxygen (see **section 8.3** for details). Both metabolites can be utilized by the cells for energetic and anabolic processes [170].

Lactic acid is present at high levels in all of the control samples, which is indicative that aerobic glycolysis is occurring. Glioblastoma has previously been shown to exhibit the Warburg effect *in vivo* [166, 172]. Furthermore, DeBerardinis *et al.* found that glutamine commonly was converted to glutamate, which was then further converted to α -ketoglutarate [170]. There are three different ways of converting glutamate to α -ketoglutarate: oxidative deamination, transamination of pyruvate (also yields alanine), or transamination of oxaloacetate (also yields aspartic acid) [18 (p. 715-725)]. It was not possible to identify aspartic acid with sufficient confidence in the spectral data of the current study. Thus, the first two conversion strategies seem more likely. The results are congruent with findings by DeBerardinis *et al.*, where high levels of alanine also indicated that alanine transaminase was the most active transaminase enzyme [170].

Compounds from the TCA cycle are often rerouted to anabolic pathways to fuel cancer growth [173]. For example, citric acid can be utilized for the synthesis of fatty acids in both healthy proliferating cells and cancer cells [18 (p. 840), 174], including glioblastoma [170]. Conversion of glutamate to α -ketoglutarate is one way for the cells to replenish the TCA cycle.

Myo-inositol acts as an osmolyte and takes part in the regulation of brain tissue volume [175], and it is found at high levels in astrocytes [176], which are one of three cell types glioblastoma is speculated to originate from. Moreover, *myo*-inositol is found at higher levels in glioblastoma tumor tissue compared to low-grade glioma and healthy (brain) tissue [166, 177]. It is therefore not unreasonable that it would appear in the spectral data of the samples used in the current study.

In summary, ~30 endogenous metabolites could be identified from the spectral data with either confidence or slight ambiguity. Several of the metabolites are indicated to have central roles in glioblastoma metabolism.

5.3 Multivariate statistical analysis of nuclear magnetic resonance data

5.3.1 Pre-processing of data

The one-dimensional ^1H NMR spectra underwent several pre-processing steps (*e.g.* baseline correction) prior to PCA and PLS regression. The removed spectral regions were either empty (above 10 ppm or below -0.5 ppm), artifacts from acquisition (residual water signal), or contained signals from molecules that were not endogenous to the samples (ethanol, DMSO). The signals arising from HEPES did not interfere with the PCA or PLS regression, like ethanol and DMSO did, and were therefore not deemed necessary to remove. Furthermore, the identified HEPES peaks were in the more crowded region of the spectra (from 2.9 ppm and above), and by not having to remove them it was possible to keep as much biologically spectral relevant data as possible.

Normalization to the total spectral integral (excluding the TSP peak) and to the integral of the TSP peak were both assessed. The subsequent distributions of samples in the PCA scores plots were similar, but the loading plots for the dataset normalized to the total integral showed much more influence from noise. The major differences in metabolite concentrations were the same for both normalization methods, thus it was concluded that in this study, normalization to the integral of the TSP peak was preferred. Unit variance scaling was investigated first because it is the standard method [131]. It did not over-enhance the noise in the spectral data and was deemed suitable for the current dataset.

The baseline processing method gave satisfactory smoothing of the spectral data, and this is not surprising as earlier work with urinary NMR metabolomics data has had similar outcomes [160]. However, four spectra consistently failed to have their baselines corrected (sample number 67, 82, 88 and 90 in **Table 4.1**). Visual inspection of the NMR spectra revealed no clues as to why the algorithm failed. The writer of the R code used in the present study, PhD Daniel Sachse, acknowledged that it happened from time to time and that there was no clear explanation why it occurred. PCA without baseline correction provided only mediocre separation of the samples in the scores plot. In the loading plot it became clear that the variations in baseline dominated the PCA, unless removed (plots not shown). Thus, it was concluded that the baseline correction procedure should be carried out regardless of the four failed samples. The spectra that failed baseline correction were left out of subsequent analyses.

5.3.2 Principal component analysis revealed two distinct responses to treatment with YM155

The PCA including all of the samples revealed dense clustering for cell lines T1456, T1459 and T1547 and their sample types (*i.e.* control and TMZ- and YM155-treated), as shown in **Figure 5.5**. The remaining two cell lines, T1454 and T1548, clustered less densely but it was still possible to see grouping by cell line. However, the loading plot, presented in **Figure 5.5**, did not reveal any immediate biomarker candidates. In general, the difference between samples in the PCA scores plot was due to a global decrease in metabolites. The color bar of the loading plot indicates the degree of influence the variation has on the principal component; in **Figure 5.5** it is evident that almost all the metabolites contribute.

Nevertheless, from the PCA of all the cell lines together, separation by cell line was still achieved to some degree. Other studies, by *e.g.* Cuperlovic-Culf *et al.*, were similarly able to find clustering and differences in metabolic profiles between cell lines of glioblastoma [163]. A separate PCA of only the control samples was carried out, but no clear biomarkers separating the different cell lines were found (plot not shown).

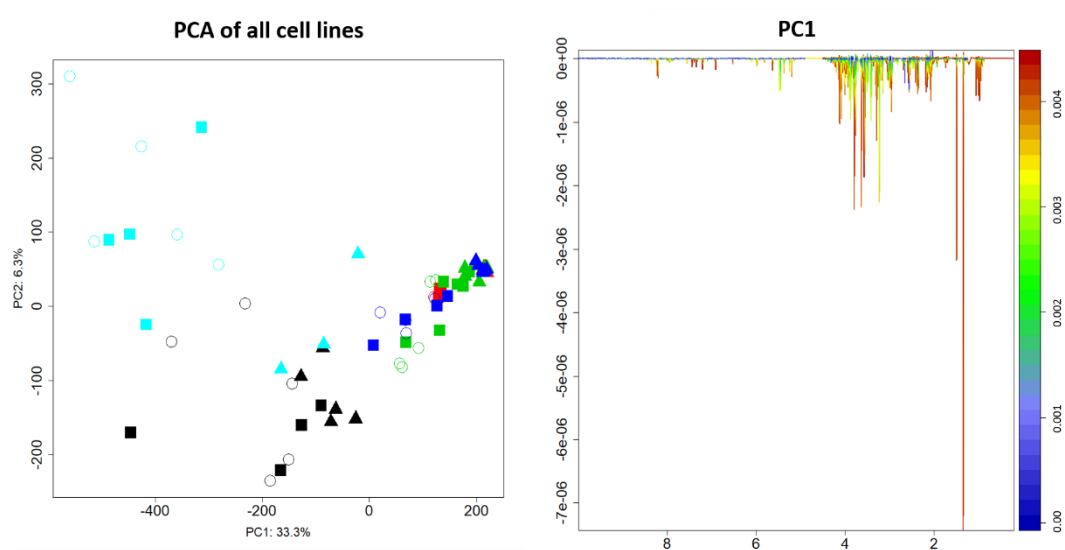


Figure 5.5: PCA scores (left) and loading (right) plots with all cell lines present. Shapes denote treatment: control samples = \circ , TMZ-treated samples = \blacksquare , and YM155-treated samples = \blacktriangle . The colors denote cell lines: black = T1454, red = T1456, green = T1459, blue = T1547, cyan = T1548.

The statistical analysis proceeded with PCAs of control versus treated samples within each cell line. The scores and loading plots of representative samples are shown in **Figure 5.6**. Analogous to the PCA in **Figure 5.5**, the set of PCAs showed two groups of cell lines: T1456/T1459/T1547 and T1454/T1548. Within each group the scores and loading plots had similar appearances.

Moreover, T1454 and T1548 cells treated with YM155 did not exclusively have a global reduction in metabolite levels, but a large increase in citric acid and a small increase in metabolite(s) in the carbohydrate region of the spectrum. Most of the other metabolites were reduced in amount, including lactate and alanine.

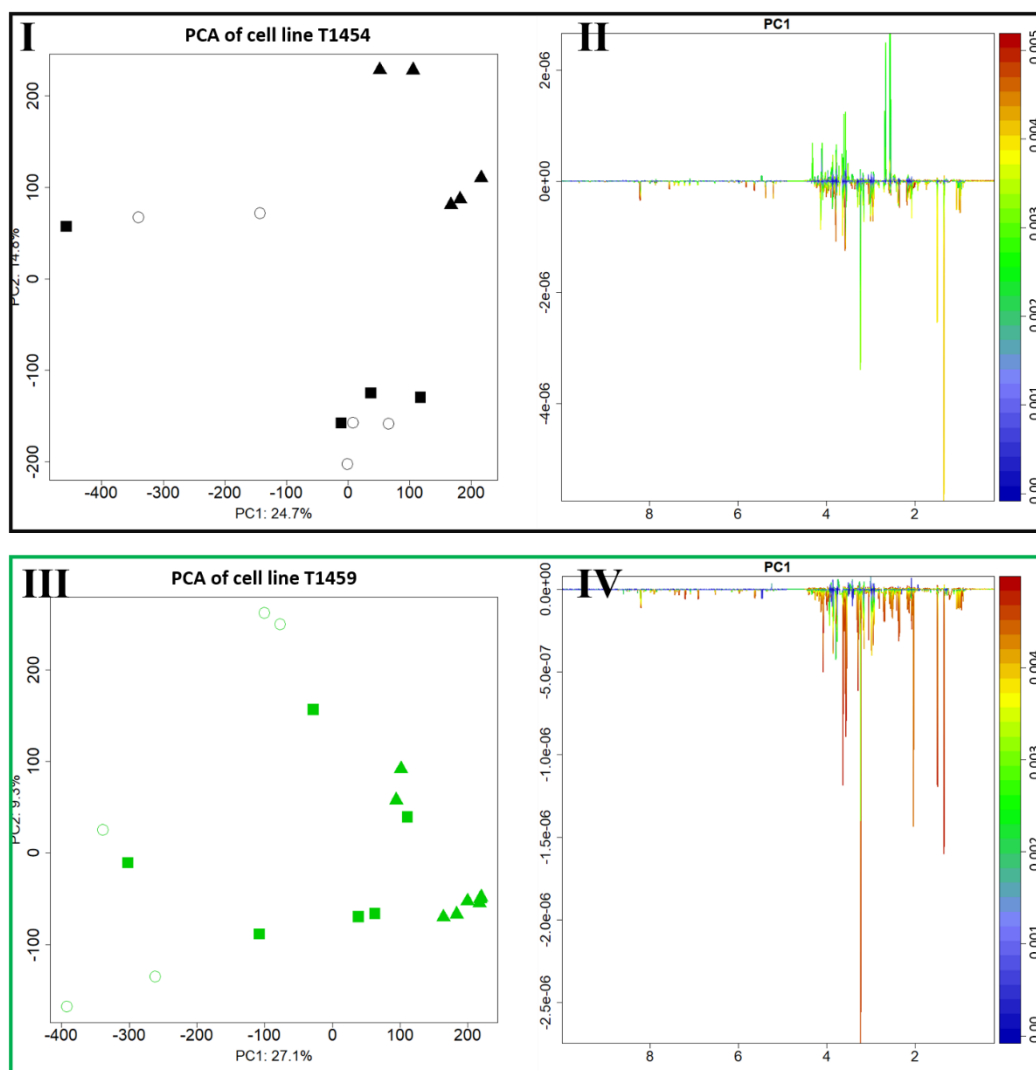


Figure 5.6: PCA scores (I and III) and loading (II and IV) plots for cell line T1454 (black) and T1459 (green). Shapes denote treatment: control samples = ○, TMZ-treated samples = ■, and YM155-treated samples = ▲.

Closer inspection of the spectral regions showing major change in the loading plots confirmed that lactate and alanine were strongly reduced across all cell lines, and citric acid was strongly increased in T1454 and T1548 after treatment with YM155. However, it was difficult to clearly identify which metabolite or metabolites were increased in the carbohydrate region of the YM155-treated for T1454 and T1548 cells. The regions of interest in the spectral data are shown for T1454 and T1459 in **Figure 5.7**. In addition, the integral of the citric acid peak relative to the TSP peak for all samples treated with YM155 is given in **Figure 5.8**.

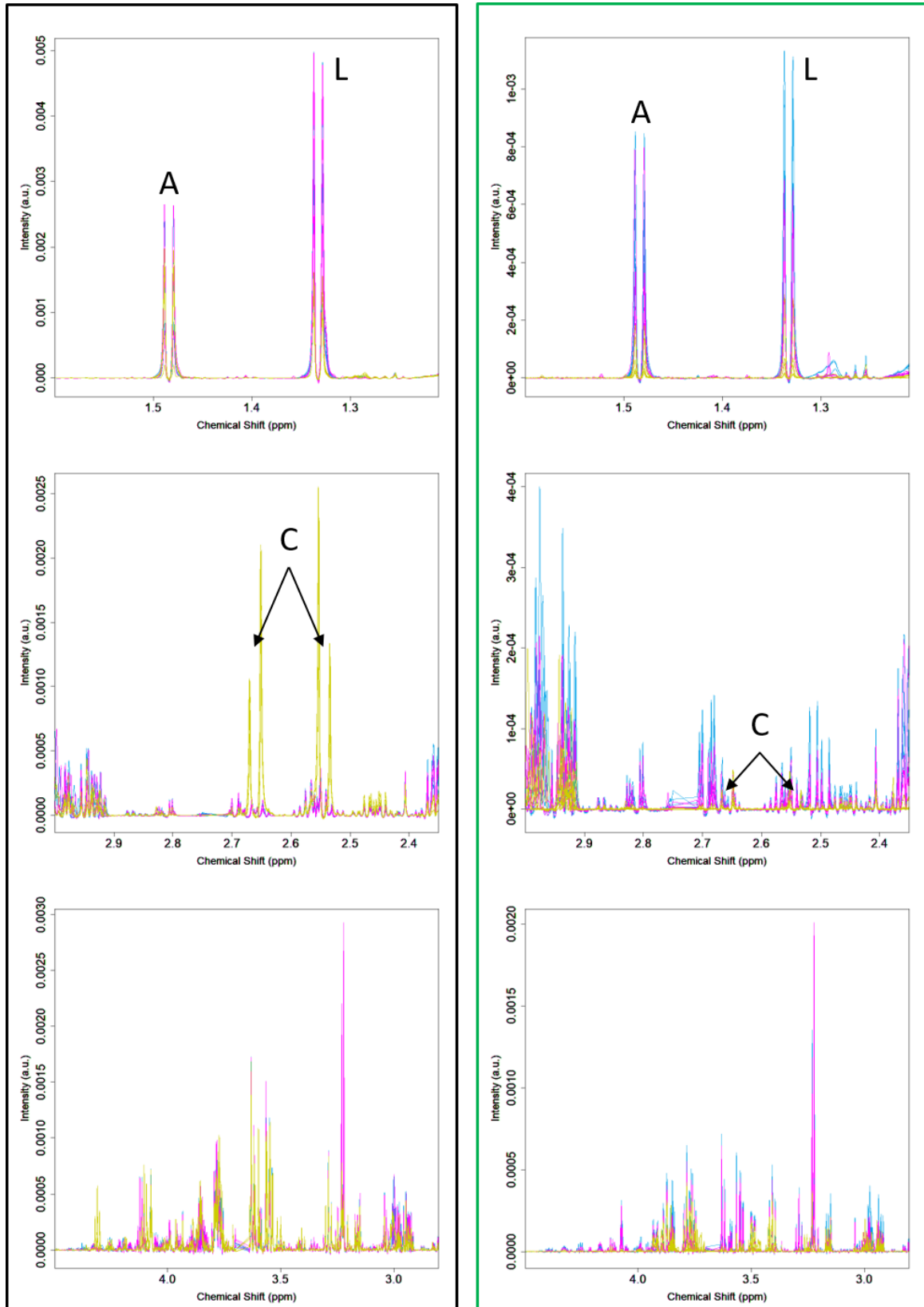


Figure 5.7: Overlaid NMR spectra of three regions of interest: 1.2 to 1.6 ppm (lactate (L) and alanine (A) doublets), 2.35 to 3.0 ppm (citric acid (C)), and 2.8 to 4.5 ppm (carbohydrate region) from T1454 (left) and T1459 (right) cell lines. The spectra are colored by treatment: control = turquoise, TMZ-treated = magenta, YM155-treated = yellow.

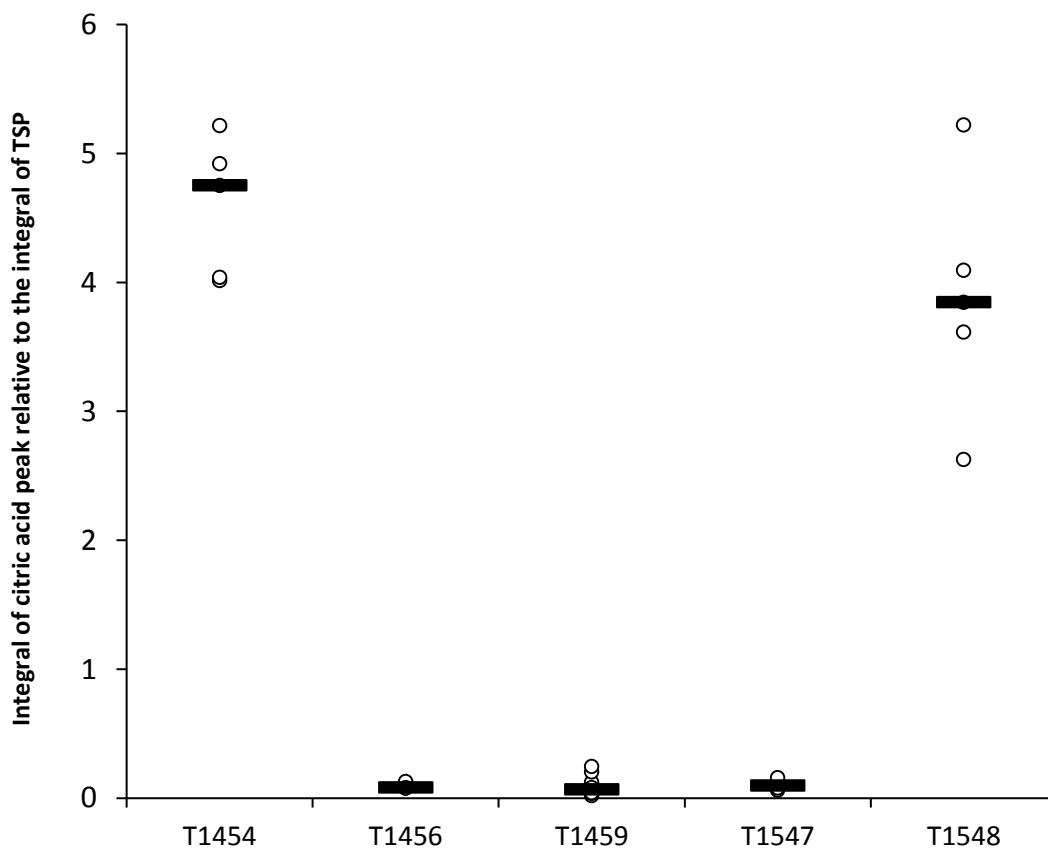


Figure 5.8: Integral of citric acid peak relative to the integral of the TSP peak for all YM155-treated samples. Figure made with Excel template from [178].

The response to YM155 for the five cell lines was also assessed with a dose-response curve, in which the half maximal effective concentration (EC_{50}) can be estimated. The curve was made by comparing the percentage of cells alive after treatment with a certain dosage of drug for a certain amount of time. EC_{50} is the drug dose at which 50% of the cells die. Treated cells were compared to control samples, which were handled exactly the same except for having no drug added to the culture medium [179]. In the present study, the cells were exposed to YM155 for 72 hours before the survival percentage was measured. The measurements were done by MD Erlend Skaga, and the dose-response curve is presented in **Figure 5.9**. Note that the cells received for NMR analysis were treated for 24 hours and not 72 hours.

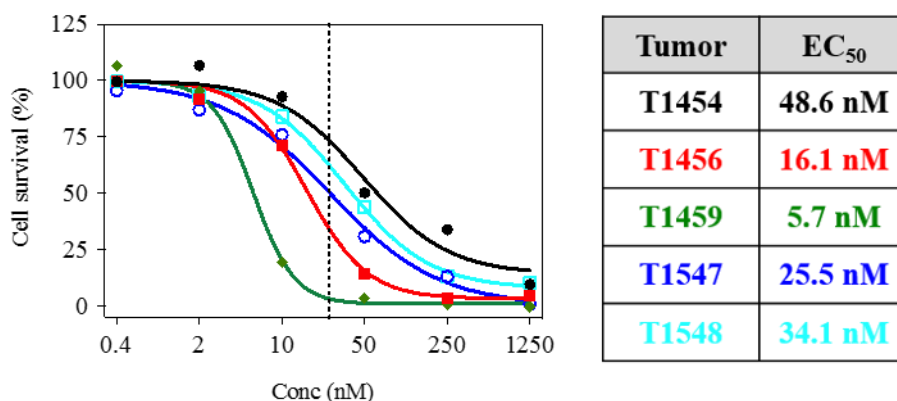


Figure 5.9: EC₅₀ curve of the five cell lines used in the present study. Figure courtesy of MD Erlend Skaga.

From the curve, it is evident that the EC₅₀ values correlate with the two groups of cell lines found in the NMR and statistical analyses. In other words, the largest relative increase in citric acid and the smallest decrease in metabolites occurs in the cell lines that are the least sensitive to YM155. In Figure 5.10, the average integral of the citric acid peak is plotted versus the cell line EC₅₀ value, further illustrating the correlation between sensitivity to YM155 and the amount of citric acid accumulating in the cells during treatment. The two groups of cell lines will from now on be called sensitive (T1456, T1459, T1547) and less sensitive (T1454, T1548) to YM155.

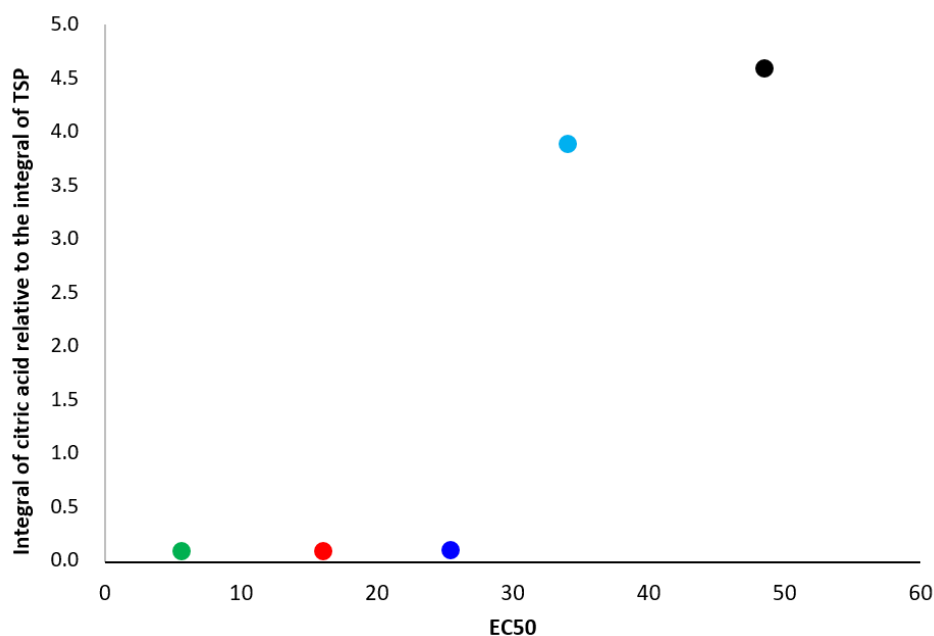


Figure 5.10: Comparison of the average integral of citric acid (relative to the integral of TSP) in YM155-treated samples from each cell line and their respective YM155 EC₅₀ values. The color denotes cell line: green = T1459, red = T1456, blue = T1547, turquoise = T1548, and black = T1454.

At first glance, citric acid could be named the biomarker for lack of sensitivity towards YM155. However, it is important to note that the current data is only for a single time point (24 hours of treatment) and a single, and very high, concentration of drug (1250 nM). Furthermore, citric acid was identified in the one-dimensional spectra of all of the YM155-treated samples. It could be that the amount of citric acid in the sensitive cell lines was so low that it was not caught by the statistical analysis. Yet, even when the data were normalized to the total spectral intensity (instead of the TSP peak), no large citric acid peaks appeared in the sensitive cell lines. Thus, the difference is not solely due to the less sensitive lines having more cells left in the medium after the treatment course.

Due to the limited data, it is difficult to pinpoint an exact reason for the differences in citric acid. In a study by Tiek *et al.*, an increased amount of citric acid was found in the culture medium of TMZ-resistant cells, compared to TMZ-sensitive cells [180]. Perhaps the considerable rise in concentration (relative to the control samples) is a general response for less sensitive or resistant cells to a chemotherapeutic agent.

Citric acid has not been extensively measured *in vivo* in brain tissue because of its low concentration. In a study by Seymour *et al.*, they found an MR signal in pediatric CNS tumors consistent with citric acid, showing that the metabolite can possibly be seen with MRS in brain tissue [181]. Citric acid has been detected with MRS in healthy prostate tissue, where it is found in high concentrations [182, 183].

To summarize, citric acid appears to have an excessively higher concentration in YM155-treated samples from cell lines with high YM155 EC₅₀ values (T1454, T1518). Thus, it is a potential biomarker of drug sensitivity in glioblastoma, and could be an important variable in non-invasive treatment assessment. Citric acid is detectable with MRS, but has not been extensively studied in (healthy) brain tissue because of its low concentration.

Lactic acid: a possible biomarker for treatment exposure?

Unlike citric acid, the lactic acid concentration had a uniform response to YM155 treatment: it decreased across all cell lines. Considering that it is readily detected with MRS [79, 177, 184], it could be considered a suitable biomarker for treatment exposure, *i.e.* that the chemotherapeutic agent has reached the tumor. For the sensitive cell lines, the decrease in lactic acid is likely due to a general loss of cells and thus metabolites. However, in the less sensitive cell lines, where there was both an increase and decrease in metabolite concentrations, the loss

of lactate is more difficult to explain. Possible reasons could be decreased rates of glycolysis or a stronger reliance on oxidative phosphorylation. The former may simply be because the cancer cells no longer are as proliferative as before, the latter as a possible survival adaptation.

In summary, lactic acid decreases across all cell lines after treatment with YM155 and is thus a biomarker candidate for treatment exposure, regardless of sensitivity towards the drug.

5.3.3 Principal component analysis did not reveal any biomarkers for response to temozolomide

The TMZ-treated cell samples had similar distribution in the PCA scores plot as the control samples for all cell lines (see **Figure 5.5**); it was confirmed with an additional PCA comparing only control and TMZ-treated samples, with the scores plot shown in **Figure 5.11**. At first glance all of the five cell lines may appear to be not sensitive to TMZ, since they were similar to the control samples. However, TMZ is a DNA-alkylating agent and it takes some time for its cytotoxic effect to develop. In the recent study by Tiek *et al.*, the glioblastoma cells were cultured with TMZ for 72 hours [180], while in the

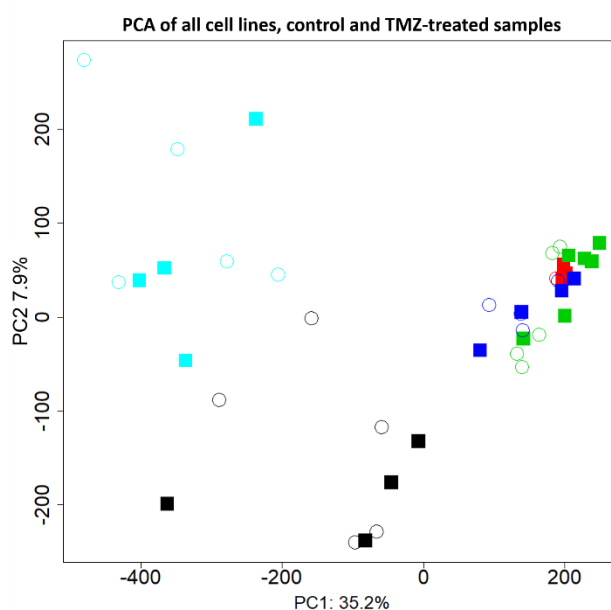


Figure 5.11: PCA scores plot of control and TMZ-treated samples from all cell lines. Shapes denote treatment: control samples = ○ and TMZ-treated samples = ■. The colors denote cell lines: black = T1454, red = T1456, green = T1459, blue = T1547, cyan = T1548.

present study they were cultured for only 24 hours. Thus, the most likely explanation for the similarities between control and TMZ-treated samples is that the cells were not yet experiencing the full cytotoxic effect of the drug. The short exposure period was chosen so that the YM155- and TMZ-treated cells were exposed to their respective chemotherapeutic agents for an equal amount of time, facilitating direct comparison of their corresponding spectral data.

To summarize, the TMZ treated cells did not show any clear difference compared to the control samples, most likely due to too short exposure time to the chemotherapeutic agent.

5.3.4 Examination of the second principal component

The discussion so far has revolved around the loading plots of the first principal component (PC1), which accounts for approximately 25% of the variation seen in the data. The loading plots from PC2 did not lead to identification of any distinct new metabolites, except for cell line T1459. In the PC2 loading plot for T1459, shown in **Figure 5.12**, several peaks are elevated and almost all of the red ones can be assigned to D-glucose. In a study by Rozental *et al.*, glioblastoma tumors were shown

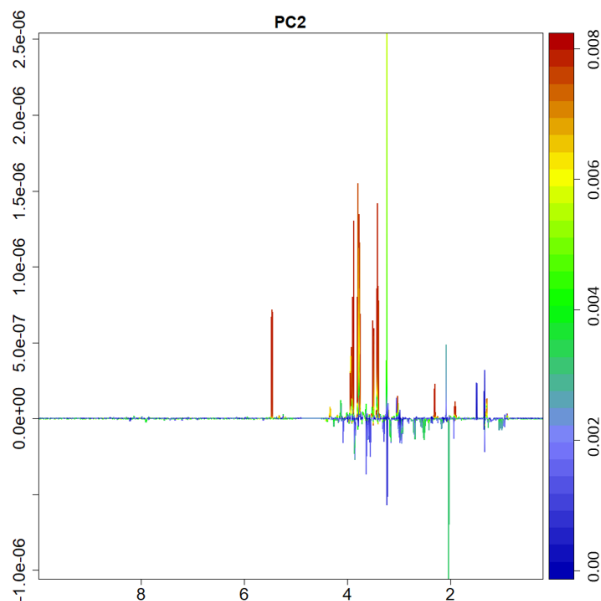


Figure 5.12: Loading plot for the second principal component of the PCA of control and treated samples from cell line T1459.

to increase glucose uptake 24 hours after treatment with carmustine (alkylating agent that can crosslink DNA [185]), and they believed this indicated tumor growth and repair [186]. It is perhaps a similar phenomenon occurring in the YM155-treated T1459 cells.

In summary, the second PC revealed that cell line T1459 had an increase of cellular glucose levels after treatment with YM155. It did not reveal additional biomarker candidates in any of the other cell lines.

5.3.5 Partial least squares regression was congruent with findings in the principal component analyses

Since the PCA could separate the control and YM155-treated samples, a PLS regression was carried out to examine if it gave a congruent assessment of the sample data. The dataset was divided into a training set and a test set. The training set consisted of all of the control and YM155-treated samples from cell lines T1454, T1459 and T1547; the test set consisted of the corresponding samples from cell lines T1456 and T1548. The split was chosen such that the model was both trained and tested with sensitive and less sensitive cell lines. An even more reliable regression model could possibly have been achieved by not splitting along cell lines, but instead semi-randomly assigning the samples to either the training or test set. Both sets

should then have received an approximately equal number of sample data from control/treated and sensitive/less sensitive cell lines, ensuring a reliable regression model. However, due to time constraints, code for the random sorting of the samples was not written and the former sorting method was used instead.

The different assessment parameters of the PLS regression model (R^2 , Q^2 , Q^2/R^2 , and V^2) were plotted versus the number of latent variables (LVs) to determine two properties. First, to examine if the model was capable of predicting the response variable, which usually requires that the ratio between Q^2 and R^2 is equal to or above 0.8. If it is well below 0.8, the model might have incorporated noise and would show poor predictive capabilities [187]. Second, to decide how many LVs should be incorporated to provide a sufficiently accurate prediction without overfitting. Once the Q^2/R^2 curve either leveled off or fell below the maximum ratio, the highest number of necessary LVs had been reached.

The assessment parameters of the control/YM155 PLS regression is shown in **Figure 5.13**. The Q^2/R^2 ratio reached 0.8 (red line) by the first LV, indicating that the model was valid. Considering the assessment parameter of the test set (green, V^2), the model was generally able to predict the correct class of the sample based on its spectral data. Since the dataset was small (training = 33 samples, test = 14 samples), the outcome was satisfactory.

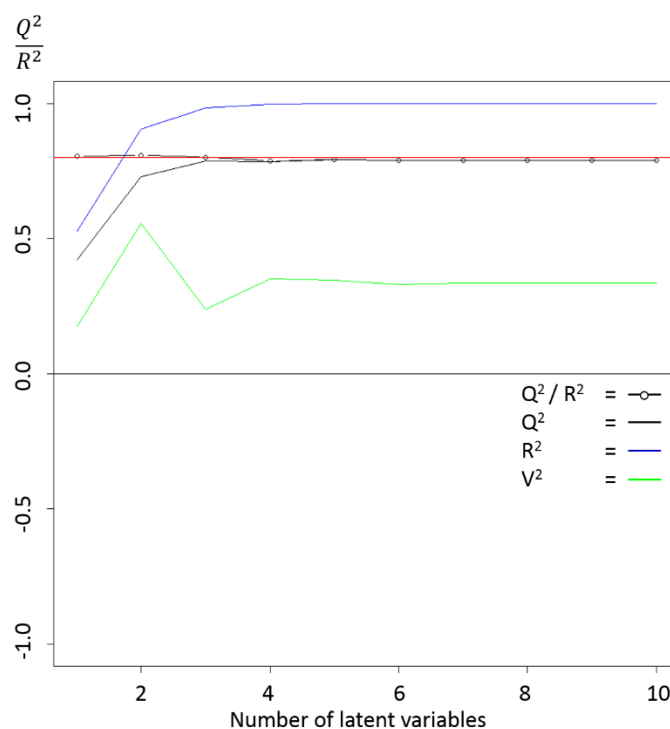


Figure 5.13: The Q^2 , R^2 , Q^2/R^2 , and V^2 values for the PLS regression of control samples versus YM155 treated samples.

The PCA was unable to distinguish control and TMZ samples, and neither could the PLS regression model. The analysis was first carried out with only CV and no test set, and it was immediately clear that the regression had failed; the Q^2/R^2 ratio was very low (<0.5), which can also be seen in **Figure 5.14**. No further test of the model was carried out.

The consistent findings in the PCAs and PLS regressions further demonstrated that the metabolic differences between the control and YM155-treated cells were most likely not random.

To summarize, the outcome of the PLS regression was congruent with the PCA: control and YM155-treated samples could be distinguished, while control and TMZ-treated samples could not.

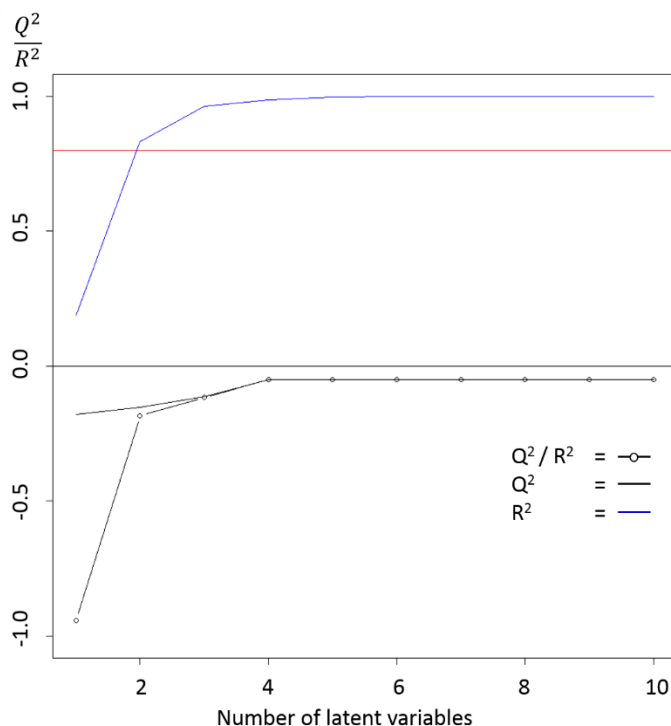


Figure 5.14: The Q^2 , R^2 , and Q^2/R^2 values for the PLS regression of control samples versus TMZ treated samples.

5.4 Liquid chromatography studies

The NMR and statistical analyses revealed at least two metabolites as likely biomarker candidates in response to treatment with YM155: citric and lactic acid. The following liquid chromatography study was carried out to assess whether the findings could be repeated with a different analytical technique. In the two sections below, the assessment of packing silica particles with HILIC SPs in narrow capillaries (ID = 100 μm) and the subsequent coupling to mass spectrometry are described.

5.4.1 Preliminary examination of narrow capillary columns

A description of how column efficiency is calculated and how the UV detector operates is given in the **Appendix** (section 8.6 and 8.7, respectively).

HILIC was chosen instead of RP as the stationary phase because of the high polarity of the two metabolites ($\log p_{\text{lactic acid}} = -0.7$ [188] and $\log p_{\text{citric acid}} = -1.7$ [189]). Furthermore, zwitterionic HILIC was chosen because of its versatility and because it has been utilized in a variety of metabolomics studies [112-116]. Specifically, in a study by Huang *et al.*, ZIC HILIC was shown capable of separating citric and lactic acid [190].

The application of narrow columns was pursued because of the possible gains in sensitivity (as described in **section 3.7**). ESI-MS is concentration sensitive and since the samples at hand were limited (40 μL cell lysate) and with small amounts of analyte, it was sensible to explore nano HILIC LC-MS. Nano-LC columns with HILIC are commercially available, from *e.g.* Merck-Millipore, but they are often quite expensive to purchase. To reduce costs, the columns can be packed in-house, but this had not been done before with HILIC particles with the method described in [138]. Thus, a preliminary assessment of both “ease of packing” and the separation capabilities of HILIC capillary columns with 100 μm ID needed to be completed. In addition, the preliminary assessment would give some indication of which MP(s) worked well for the particular ZIC® HILIC particles that were available.

The packing procedure followed that outlined in [138], the only difference being packing time (two minutes vs one hour). This can be most likely attributed to the difference in capillary ID (50 μm in the cited study vs 100 μm here) and particle diameters. In the study by Berg *et al.*, the capillaries were packed with 2.6 μm diameter silica particles, which will lead to higher back

pressure and longer packing times than the 5 μm diameter silica particles used in the present study.

Two different MP buffers were evaluated: 30 mM ammonium formate at pH 3 and pH 4.5. Earlier work by Johnsen *et al.* ([191]) and Bjørseth ([192]) concluded that ammonium formate buffers were suitable for ZIC® HILIC and the conclusion remains the same in this study. The pH of the buffer component of the MP had a large effect on the efficiency and quality of the chromatographic separation. In the present study, only the MP with buffer pH 4.5 provided satisfactory peak intensity and shape. Furthermore, the MP composition that was found to be most suitable was 70/30 (v/v) acetonitrile/buffer. Increasing the acetonitrile fraction resulted in very long elution times (>30 minutes) and poor peak shape for tryptophan, which was used as a test compound due to its UV absorption. MP gradients can be difficult to use with HILIC, in part because of long equilibration time [193], and it was not further elaborated upon in the present study due to time constraints.

With the selected MP, toluene and tryptophan could be separated within 8 minutes, as shown in **Figure 5.15**. Separation efficiency was acceptable: $N=2200$ plates for the toluene peak with an 11 cm long column, equivalent of plate height $H = 50 \mu\text{m}$. N was calculated at full width, half maximum height of the toluene peak. In comparison, Kawachi *et al.* obtained plate heights of approximately 25 μm with a 2.1 mm x 150 mm ZIC® HILIC column from Merck, also packed with 5 μm diameter particles [194]. The equivalent plate number for the 15 cm long column is 6000.

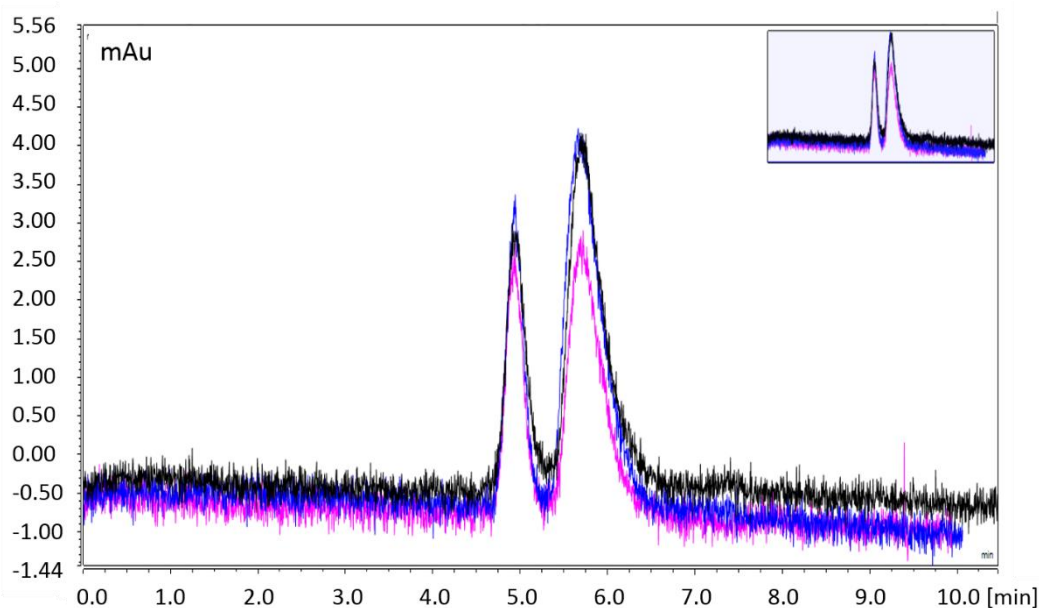


Figure 5.15: Chromatogram of toluene (t_R 4.95 minutes) and tryptophan (t_R 5.70 minutes) separated on a capillary ZIC® HILIC column (100 μm ID/360 μm OD/25 cm long, 11 cm with particles). The MP consisted of 70/30 (v/v) acetonitrile/buffer (30 mM ammonium formate, pH 4.5). The flow rate was kept at 1.00 $\mu\text{L}/\text{minute}$.

In summary, the preliminary assessment of in-house packing of ZIC® HILIC particles in 100 μm ID capillaries showed that the packing resulted in columns with decent efficiency. The packing procedure was straightforward and quick. It was possible to separate toluene and tryptophan within 8 minutes on the column, with an efficiency of $N = 2200$ plates (11 cm column).

5.4.2 Targeted metabolomics with liquid chromatography-mass spectrometry

The mass spectrometry analysis was carried out in negative mode because citric and lactic acid were more likely to form ions via deprotonation than protonation. Prior to coupling the LC system to the MS, the m/z values of the deprotonated ions and fragment ions of citric and lactic acid had to be established by direct injection analysis. Additionally, the direct injection analyses were done to optimize MS parameters, *e.g.* ESI needle temperature and voltage.

Lactic acid has a monoisotopic mass of 90.0317, and when deprotonated its calculated m/z equaled 89.0238. The pK_a of lactic acid is 3.86, and it was assumed that it would be mostly deprotonated in the MP (buffer pH 4.5). Citric acid, on the other hand, has three carboxylic groups with pK_a values of 3.13, 4.76, and 6.39, respectively. It was assumed to only be

singularly deprotonated, but a search for all three corresponding m/z values was carried out. The monoisotopic mass of citric acid is 192.0270, and the following m/z values were calculated: $[M-H]^{-1} = 191.0191$, $[M-2H]^{-2} = 95.0056$, and $[M-3H]^{-3} = 63.0011$. In a study by Huang *et al.*, the main fragment of lactic acid was listed as m/z 43, resulting from loss of the carboxylic acid group. The main fragment of citric acid was listed as m/z 111, resulting from loss one carboxylic acid group and two water molecules [190].

The citric acid $[M-H]^{-1}$ ion and fragment ion with $m/z = 111.007$ were found in the direct injection analysis. However, for lactic acid there was a compound in the MP with very similar m/z . The interfering ion might have been eliminated by switching MPs, but due to time constraints it was not feasible in the present study. The LC-MS analyses thus only involved measurements of citric acid.

The retention time of citric acid was established with a standard solution, and the MP composition was adjusted to 65% acetonitrile and 35% buffer (30 mM ammonium formate, pH 4.5), as it resulted in better peak shape and shorter retention time. All samples, including the cell samples, were dissolved in 80/20 acetonitrile/buffer to promote focusing of the injected sample when it reached the column [195, 196].

Prior to LC-MS analysis, the proteins in the cell lysate were precipitated to avoid clogging of the capillary column and ion suppression in the ion source (ESI). Cold acetonitrile was chosen as the precipitant for several reasons. First, it is miscible with water but also easy to remove by evaporation. Second, it was a constituent of the MP and any small remnants of it from the precipitation would not interfere with the subsequent LC-MS analysis. Finally, salts were avoided, both to avoid clogging the capillary column and because not all salts are compatible with MS. An internal standard was not added because no absolute quantification was performed.

Due to time constraints, only six cell samples were analyzed: two control samples from T1454, two YM155-treated from a sensitive (T1459), and two from a less sensitive (T1454), cell line. In **Figure 5.16**, the relative peak areas for citric acid in the six samples are given. In the **Appendix (section 8.9)**, representative chromatograms are shown. From the figure, it is clear that there was vastly higher concentrations of citric acid in the YM155-treated samples from the less sensitive cell line. Even though the data are limited, they remain congruent with the NMR analyses.

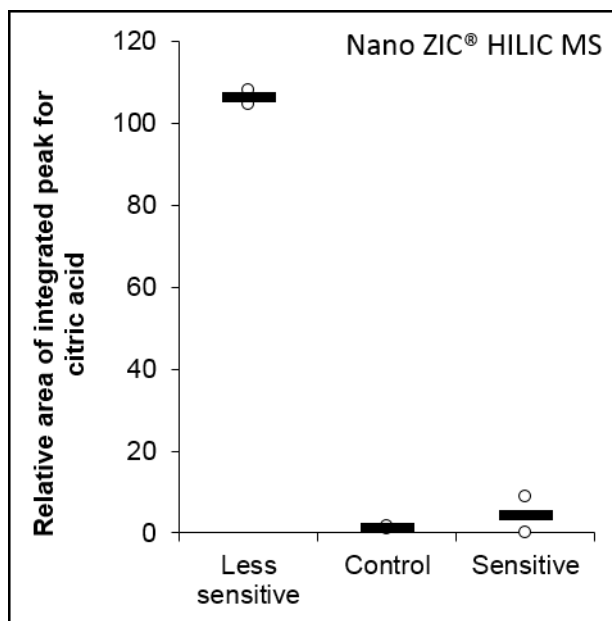


Figure 5.16: The relative peak areas of citric acid in two control samples and YM155-treated samples from a sensitive (T1459) and a less sensitive (T1454) cell line. Figure made with Excel template from [178].

To summarize, only citric acid could be measured by targeted LC-MS, as lactic acid had similar m/z to an ion in the MP. Six cell samples (two control, two sensitive to YM155, two less sensitive to YM155) were measured and the relative peak areas of citric acid were compared. The findings were congruent with the NMR analysis: citric acid was present at high levels in the YM155-treated samples from less sensitive cell lines.

6 Conclusion

In the present study, two biomarker candidates for treatment response were found. Citric acid was identified as a tentative biomarker of sensitivity towards chemotherapeutic agent YM155; its concentration was increased substantially in cell lines less sensitive to YM155 compared to those that were sensitive. On the other hand, the decrease in lactic acid in all cell lines following YM155 treatment indicated that lactic acid is a more general biomarker of exposure to the chemotherapeutic agent. Unfortunately, the TMZ-treated samples were not distinguishable from the control samples, most likely due to the short exposure time to the chemotherapeutic agent (24 hours) prior to NMR analysis.

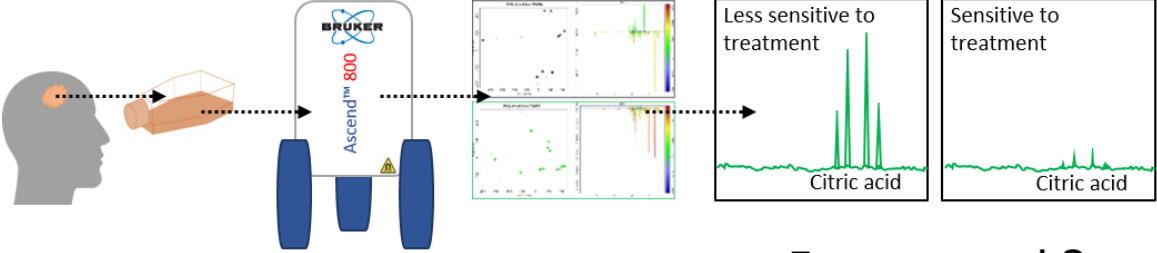
Both citric and lactic acid are measurable with MRS, indicating that the more long-term goal of following treatment in a noninvasive, personalized, and accurate manner could be possible. The clinical application of the current findings would rely on additional experiments to further develop an understanding of the exact fluctuations of citric and lactic acid during treatment. However, it remains clear that the demonstrated methods, especially 800 MHz ^1H NMR with excitation sculpting as solvent suppression, are viable means to scout for biomarkers appropriate for MRS. The biomarkers found in the present study represent genuine candidates.

6.1 Future work

Further investigation of the fluctuations of citric and lactic acid in cultured glioblastoma cells is necessary to be able to more soundly understand their roles as biomarkers in response to YM155 treatment. The study should include additional patient-derived cell lines and different lengths of treatment and drug dosages. Furthermore, a similar study of TMZ-treated cells should be carried out, including cells treated for the appropriate amount of time (≥ 72 hours) with a variety of dosages across several patient-derived cell lines. Finally, biomarker candidates should be examined in patients being treated with TMZ, and YM155 if it is approved for clinical use. If successful, the scope could be broadened even further and other drugs could be investigated as well.

The conclusion and suggestions for future work are illustrated in **Figure 6.1**.

Conclusion



Future work?

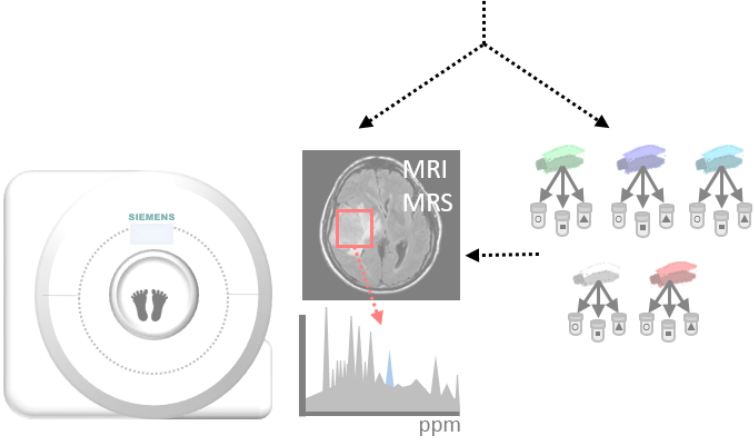


Figure 6.1: Graphical illustration of conclusion and future work. The glioblastoma MRI is a case courtesy of Professor Frank Gaillard, Radiopaedia.org, rID: 8939.

7 References

- [1] R. Sender, S. Fuchs, R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body*, PLOS Biology, 14 (2016) 1-14.
- [2] B. Alberts, J.H. Wilson, T. Hunt, *Molecular biology of the cell*, 6th ed. ed., Garland Science, New York, 2015.
- [3] L. Pecorino, *Molecular biology of cancer: mechanisms, targets, and therapeutics*, 4th ed., Oxford University Press, Oxford, 2016.
- [4] K.L. Voje, *Alleler*, Store Norske Leksikon, 2014, <https://snl.no/alleler>
- [5] C. Caldas, A.R. Venkitaraman, *Tumor Suppressor Genes*, Brenner's Encyclopedia of Genetics, 2001, <https://www.sciencedirect.com/science/article/pii/B9780123749840015953>
- [6] M. Hartl, M.K. Bister, *Oncogenes*, Brenner's Encyclopedia of Genetics, 2013,
- [7] D. Hanahan, R. Weinberg, *Hallmarks of Cancer: The Next Generation*, Cell, 144 (2011) 646-674.
- [8] T. Lapidot, C. Sirard, J. Vormoor, B. Murdoch, T. Hoang, J. Caceres-Cortes, M. Minden, B. Paterson, M.A. Caligiuri, J.E. Dick, *A cell initiating human acute myeloid leukaemia after transplantation into SCID mice*, Nature, 367 (1994) 645-648.
- [9] J.E. Dick, *Stem cell concepts renew cancer research*, Blood, 112 (2008) 4793-4807.
- [10] M. Shackleton, E. Quintana, E.R. Fearon, S.J. Morrison, *Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution*, Cell, 138 (2009) 822-829.
- [11] M. Greaves, C.C. Maley, *Clonal evolution in cancer*, Nature, 481 (2012) 306-313.
- [12] M. Al-Hajj, M.S. Wicha, A. Benito-Hernandez, S.J. Morrison, M.F. Clarke, *Prospective identification of tumorigenic breast cancer cells*, Proceedings of the National Academy of Sciences, 100 (2003) 3983-3988.
- [13] C.A. O'Brien, A. Pollett, S. Gallinger, J.E. Dick, *A human colon cancer cell capable of initiating tumour growth in immunodeficient mice*, Nature, 445 (2006) 106-110.
- [14] L. Ricci-Vitiani, D.G. Lombardi, E. Pilozzi, M. Biffoni, M. Todaro, C. Peschle, R. De Maria, *Identification and expansion of human colon-cancer-initiating cells*, Nature, 445 (2006) 111-115.
- [15] S.K. Singh, C. Hawkins, I.D. Clarke, J.A. Squire, J. Bayani, T. Hide, R.M. Henkelman, M.D. Cusimano, P.B. Dirks, *Identification of human brain tumour initiating cells*, Nature, 432 (2004) 396-401.
- [16] X. Lan, D.J. Jörg, F.M.G. Cavalli, L.M. Richards, L.V. Nguyen, R.J. Vanner, P. Guilhamon, L. Lee, M.M. Kushida, D. Pellacani, N.I. Park, F.J. Coutinho, H. Whetstone, H.J. Selvadurai, C. Che, B. Luu, A. Carles, M. Moksa, N. Rastegar, R. Head, S. Dolma, P. Prinos, M.D. Cusimano, S. Das, M. Bernstein, C.H. Arrowsmith, A.J. Mungall, R.A. Moore, Y. Ma, M. Gallo, M. Lupien, T.J. Pugh, M.D. Taylor, M. Hirst, C.J. Eaves, B.D. Simons, P.B. Dirks, *Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy*, Nature, 549 (2017) 227-232.
- [17] O. Warburg, *The metabolism of tumors in the body*, The Journal of General Physiology, 8 (1927) 519-530.
- [18] D.L. Nelson, A.L. Lehninger, M.M. Cox, A.L. Lehninger, *Lehninger Principles of Biochemistry*, 6th ed. ed., W.H. Freeman, New York, 2013.
- [19] M.G. Vander Heiden, L.C. Cantley, C.B. Thompson, *Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation*, Science, 324 (2009) 1029-1033.
- [20] W. Dong, M.A. Keibler, G. Stephanopoulos, *Review of metabolic pathways activated in cancer cells as determined through isotopic labeling and network analysis*, Metabolic Engineering, 43 (2017) 113-124.

- [21] S.K. Ray, *Glioblastoma: Molecular Mechanisms of Pathogenesis and Current Therapeutic Strategies*, New York, NY, US: Springer New York, New York, **2009**.
- [22] F. Shao, C. Liu, *Revisit the Candidacy of Brain Cell Types as the Cell(s) of Origin for Human High-Grade Glioma*, *Frontiers in Molecular Neuroscience*, 11 (**2018**) 1-10.
- [23] D.N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W.K. Cavenee, H. Ohgaki, O.D. Wiestler, P. Kleihues, D.W. Ellison, *The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary*, *Acta Neuropathologica*, 131 (**2016**) 803-820.
- [24] K.G. Abdullah, C. Adamson, S. Brem, *Chapter 3 - The Molecular Pathogenesis of Glioblastoma*, in: *Glioblastoma*, Elsevier, **2016**, pp. 21-31.
- [25] H. Ohgaki, P. Dessen, B. Jourde, S. Horstmann, T. Nishikawa, P.-L. Di Patre, C. Burkhard, D. Schüler, N.M. Probst-Hensch, P.C. Maiorka, N. Baeza, P. Pisani, Y. Yonekawa, M.G. Yasargil, U.M. Lütolf, P. Kleihues, *Genetic Pathways to Glioblastoma - A Population-Based Study*, *Cancer research*, 64 (**2004**) 6892-6899.
- [26] G. Carrabba, D. Mukhopadhyay, A. Guha, *Aberrant Signalling Complexes in GBMs: Prognostic and Therapeutic Implications*, in: S.K. Ray (Ed.) *Glioblastoma: Molecular Mechanisms of Pathogenesis and Current Therapeutic Strategies*, Springer New York, New York, NY, **2010**, pp. 95-129.
- [27] H. Xu, H. Zong, C. Ma, X. Ming, M. Shang, K. Li, X. He, H. Du, L. Cao, *Epidermal growth factor receptor in glioblastoma*, *Oncology Letters*, 14 (**2017**) 512-516.
- [28] B. Costa, S. Bendinelli, P. Gabelloni, E. Da Pozzo, S. Daniele, F. Scatena, R. Vanacore, P. Campiglia, A. Bertamino, I. Gomez-Monterrey, D. Sorriento, C. Del Giudice, G. Iaccarino, E. Novellino, C. Martini, *Human Glioblastoma Multiforme: p53 Reactivation by a Novel MDM2 Inhibitor*, *PLOS ONE*, 8 (**2013**) 1-19.
- [29] J. Chen, Y. Li, T.-S. Yu, R.M. McKay, D.K. Burns, S.G. Kernie, L.F. Parada, *A restricted cell population propagates glioblastoma growth after chemotherapy*, *Nature*, 488 (**2012**) 522-526.
- [30] O. Klepp, *Glioblastom*, *Store medisinske leksikon*, **2016**, <https://sml.snl.no/glioblastom>
- [31] Q.T. Ostrom, P. Liao, L.C. Stetson, J.S. Barnholtz-Sloan, *Chapter 2 - Epidemiology of Glioblastoma and Trends in Glioblastoma Survivorship*, in: *Glioblastoma*, Elsevier, **2016**, pp. 11-19.
- [32] Anette Storstein, Eirik Helseth, Tom Børge Johannesen, Till Schellhorn, Sverre Mørk, R.v. Helvoirt, *Høygradige gliomer hos voksne*, *Tidsskriftet Den Norske Legeforening*, 131 (**2011**) 238-241.
- [33] B.M. IK Larsen, TB Johannesen, S Larønningen, TE Robsahm, TK Grimsrud, G Ursin, *Cancer in Norway 2016*, in: I.K. Larsen (Ed.), *Cancer Registry of Norway*, **2017**.
- [34] A. Omuro, L.M. DeAngelis, *Glioblastoma and other malignant gliomas: A clinical review*, *The Journal of the American Medical Association (JAMA)*, 310 (**2013**) 1842-1850.
- [35] R. Stupp, W.P. Mason, M.J. van den Bent, M. Weller, B. Fisher, M.J.B. Taphoorn, K. Belanger, A.A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R.C. Janzer, S.K. Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J.G. Cairncross, E. Eisenhauer, R.O. Mirimanoff, *Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma*, *New England Journal of Medicine*, 352 (**2005**) 987-996.
- [36] R. Stupp, M.E. Hegi, W.P. Mason, M.J. van den Bent, M.J.B. Taphoorn, R.C. Janzer, S.K. Ludwin, A. Allgeier, B. Fisher, K. Belanger, P. Hau, A.A. Brandes, J. Gijtenbeek, C. Marosi, C.J. Vecht, K. Mokhtari, P. Wesseling, S. Villa, E. Eisenhauer, T. Gorlia, M. Weller, D. Lacombe, J.G. Cairncross, R.-O. Mirimanoff, *Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial*, *The Lancet Oncology*, 10 (**2009**) 459-466.

- [37] A.F. Hottinger, K.G. Abdullah, R. Stupp, *Chapter 6 - Current Standards of Care in Glioblastoma Therapy*, in: *Glioblastoma*, Elsevier, **2016**, pp. 73-80.
- [38] M.-d.-M. Inda, R. Bonavia, A. Mukasa, Y. Narita, D.W.Y. Sah, S. Vandenberg, C. Brennan, T.G. Johns, R. Bachoo, P. Hadwiger, P. Tan, R.A. Depinho, W. Cavenee, F. Furnari, *Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma*, *Genes & Development*, 24 (**2010**) 1731-1745.
- [39] R.G.W. Verhaak, K.A. Hoadley, E. Purdom, V. Wang, Y. Qi, M.D. Wilkerson, C.R. Miller, L. Ding, T. Golub, J.P. Mesirov, G. Alexe, M. Lawrence, M. O'Kelly, P. Tamayo, B.A. Weir, S. Gabriel, W. Winckler, S. Gupta, L. Jakkula, H.S. Feiler, J.G. Hodgson, C.D. James, J.N. Sarkaria, C. Brennan, A. Kahn, P.T. Spellman, R.K. Wilson, T.P. Speed, J.W. Gray, M. Meyerson, G. Getz, C.M. Perou, D.N. Hayes, *Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1*, *Cancer Cell*, 17 (**2010**) 98-110.
- [40] R. Bonavia, M.-d.-M. Inda, W.K. Cavenee, F.B. Furnari, *Heterogeneity Maintenance in Glioblastoma: A Social Network*, *Cancer research*, 71 (**2011**) 4055-4060.
- [41] V. Frattini, V. Trifonov, J.M. Chan, A. Castano, M. Lia, F. Abate, S.T. Keir, A.X. Ji, P. Zoppoli, F. Niola, C. Danussi, I. Dolgalev, P. Porrati, S. Pellegatta, A. Heguy, G. Gupta, D.J. Pisapia, P. Canoll, J.N. Bruce, R.E. McLendon, H. Yan, K. Aldape, G. Finocchiaro, T. Mikkelsen, G.G. Privé, D.D. Bigner, A. Lasorella, R. Rabadan, A. Iavarone, *The integrated landscape of driver genomic alterations in glioblastoma*, *Nature Genetics*, 45 (**2013**) 1141-1149.
- [42] A. Sottoriva, I. Spiteri, S.G.M. Piccirillo, A. Touloumis, V.P. Collins, J.C. Marioni, C. Curtis, C. Watts, S. Tavaré, *Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics*, *Proceedings of the National Academy of Sciences*, 110 (**2013**) 4009-4014.
- [43] A.P. Patel, I. Tirosh, J.J. Trombetta, A.K. Shalek, S.M. Gillespie, H. Wakimoto, D.P. Cahill, B.V. Nahed, W.T. Curry, R.L. Martuza, D.N. Louis, O. Rozenblatt-Rosen, M.L. Suvà, A. Regev, B.E. Bernstein, *Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma*, *Science*, 344 (**2014**) 1396-1401.
- [44] J.K. Park, T. Hodges, L. Arko, M. Shen, D. Dello Iacono, A. McNabb, N. Olsen Bailey, T.N. Kreisl, F.M. Iwamoto, J. Sul, S. Auh, G.E. Park, H.A. Fine, P.M. Black, *Scale to Predict Survival After Surgery for Recurrent Glioblastoma Multiforme*, *Journal of Clinical Oncology*, 28 (**2010**) 3838-3843.
- [45] N.R. Parker, P. Khong, J.F. Parkinson, V.M. Howell, H.R. Wheeler, *Molecular Heterogeneity in Glioblastoma: Potential Clinical Implications*, *Frontiers in Oncology*, 5 (**2015**) 1-9.
- [46] S. Bao, Q. Wu, R.E. McLendon, Y. Hao, Q. Shi, A.B. Hjelmeland, M.W. Dewhirst, D.D. Bigner, J.N. Rich, *Glioma stem cells promote radioresistance by preferential activation of the DNA damage response*, *Nature*, 444 (**2006**) 756-760.
- [47] A. Eramo, L. Ricci-Vitiani, A. Zeuner, R. Pallini, F. Lotti, G. Sette, E. Pilozzi, L.M. Larocca, C. Peschle, R. De Maria, *Chemotherapy resistance of glioblastoma stem cells*, *Cell Death And Differentiation*, 13 (**2006**) 1238-1241.
- [48] G. Liu, X. Yuan, Z. Zeng, P. Tunici, H. Ng, I.R. Abdulkadir, L. Lu, D. Irvin, K.L. Black, J.S. Yu, *Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in glioblastoma*, *Molecular Cancer*, 5 (**2006**) 67-79.
- [49] A. Murat, E. Migliavacca, T. Gorlia, W.L. Lambiv, T. Shay, M.-F. Hamou, N. de Tribolet, L. Regli, W. Wick, M.C.M. Kouwenhoven, J.A. Hainfellner, F.L. Heppner, P.-Y. Dietrich, Y. Zimmer, J.G. Cairncross, R.-C. Janzer, E. Domany, M. Delorenzi, R. Stupp, M.E. Hegi, *Stem Cell-Related "Self-Renewal" Signature and High Epidermal Growth Factor Receptor*

Expression Associated With Resistance to Concomitant Chemoradiotherapy in Glioblastoma, Journal of Clinical Oncology, 26 (2008) 3015-3024.

[50] A. Brøgger, *Cytotoksisk*, in: Store Medisinske Leksikon, 2018.

[51] O. Klepp, *Cytostatika*, in: Store Medisinske Leksikon, 2018.

[52] V.R. Archer, L.J. Billingham, M.H. Cullen, *Palliative Chemotherapy: No Longer a Contradiction in Terms*, The Oncologist, 4 (1999) 470-477.

[53] B.J. Denny, R.T. Wheelhouse, M.F. Stevens, L.L. Tsang, J.A. Slack, *NMR and molecular modeling investigation of the mechanism of activation of the antitumor drug temozolomide and its interaction with DNA*, Biochemistry, 33 (1994) 9045-9051.

[54] J. Zhang, M. F.G. Stevens, T. D. Bradshaw, *Temozolomide: Mechanisms of Action, Repair and Resistance*, Current Molecular Pharmacology, 5 (2012) 102-114.

[55] N. Mojas, M. Lopes, J. Jiricny, *Mismatch repair-dependent processing of methylation damage gives rise to persistent single-stranded gaps in newly replicated DNA*, Genes & Development, 21 (2007) 3342-3355.

[56] E.M. Goellner, B. Grimme, A.R. Brown, Y.-C. Lin, X.-H. Wang, K.F. Sugrue, L. Mitchell, R.N. Trivedi, J.-b. Tang, R.W. Sobol, *Overcoming Temozolomide Resistance in Glioblastoma via Dual Inhibition of NAD⁺ Biosynthesis and Base Excision Repair*, Cancer research, 71 (2011) 2308-2317.

[57] T. Nakahara, M. Takeuchi, I. Kinoyama, T. Minematsu, K. Shirasuna, A. Matsuhisa, A. Kita, F. Tominaga, K. Yamanaka, M. Kudoh, M. Sasamata, *YM155, a Novel Small-Molecule Survivin Suppressant, Induces Regression of Established Human Hormone-Refractory Prostate Tumor Xenografts*, Cancer research, 67 (2007) 8014-8021.

[58] A.W. Tolcher, A. Mita, L.D. Lewis, C.R. Garrett, E. Till, A.I. Daud, A. Patnaik, K. Papadopoulos, C. Takimoto, P. Bartels, A. Keating, S. Antonia, *Phase I and Pharmacokinetic Study of YM155, a Small-Molecule Inhibitor of Survivin*, Journal of Clinical Oncology, 26 (2008) 5198-5203.

[59] P.C. Lai, S.H. Chen, S.H. Yang, C.C. Cheng, T.H. Chiu, Y.T. Huang, *Novel Survivin Inhibitor YM155 elicits Cytotoxicity in Glioblastoma Cell Lines with Normal or Deficiency DNA-Dependent Protein Kinase Activity*, Pediatrics & Neonatology, 53 (2012) 199-204.

[60] H. Guo, Y. Wang, T. Song, T. Xin, Z. Zheng, P. Zhong, X. Zhang, *Silencing of Survivin Using YM155 Inhibits Invasion and Suppresses Proliferation in Glioma Cells*, Cell Biochemistry and Biophysics, 71 (2015) 587-593.

[61] N.K. Sah, Z. Khan, G.J. Khan, P.S. Bisen, *Structural, functional and therapeutic biology of survivin*, Cancer Letters, 244 (2006) 164-171.

[62] A. Das, W.-L. Tan, J. Teo, D.R. Smith, *Expression of survivin in primary glioblastomas*, Journal of Cancer Research and Clinical Oncology, 128 (2002) 302-306.

[63] D. Xie, Y.X. Zeng, H.J. Wang, J.M. Wen, Y. Tao, J.S.T. Sham, X.Y. Guan, *Expression of cytoplasmic and nuclear Survivin in primary and secondary human glioblastoma*, British Journal Of Cancer, 94 (2005) 108-114.

[64] Z. Song, X. Yao, M. Wu, *Direct interaction between survivin and Smac/DIABLO is essential for the anti-apoptotic activity of survivin during taxol-induced apoptosis*, Journal of Biological Chemistry, 278 (2003) 23130-23140.

[65] T. Dohi, E. Beltrami, N.R. Wall, J. Plescia, D.C. Altieri, *Mitochondrial survivin inhibits apoptosis and promotes tumorigenesis*, The Journal of Clinical Investigation, 114 (2004) 1117-1127.

[66] I. Tamm, Y. Wang, E. Sausville, D.A. Scudiero, N. Vigna, T. Oltersdorf, J.C. Reed, *IAP-Family Protein Survivin Inhibits Caspase Activity and Apoptosis Induced by Fas (CD95), Bax, Caspases, and Anticancer Drugs*, Cancer research, 58 (1998) 5315-5320.

[67] D.P. Banks, J. Plescia, D.C. Altieri, J. Chen, S.H. Rosenberg, H. Zhang, S.-C. Ng, *Survivin does not inhibit caspase-3 activity*, Blood, 96 (2000) 4002-4003.

- [68] M.A. Verdecia, H.-k. Huang, E. Dutil, D.A. Kaiser, T. Hunter, J.P. Noel, *Structure of the human anti-apoptotic protein survivin reveals a dimeric arrangement*, *Nature Structural Biology*, 7 (2000) 602-608.
- [69] D.C. Altieri, *Survivin, cancer networks and pathway-directed drug discovery*, *Nature Reviews Cancer*, 8 (2008) 61-70.
- [70] P.Y. Wen, D.R. Macdonald, D.A. Reardon, T.F. Cloughesy, A.G. Sorensen, E. Galanis, J. Degroot, W. Wick, M.R. Gilbert, A.B. Lassman, C. Tsien, T. Mikkelsen, E.T. Wong, M.C. Chamberlain, R. Stupp, K.R. Lamborn, M.A. Vogelbaum, M.J. van den Bent, S.M. Chang, *Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group*, *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 28 (2010) 1963-1972.
- [71] M. Bazerbashi, D. Gomez-Hassan, *Response Assessment in Treated Brain Tumors: The Fundamentals*, *Seminars in Roentgenology*, 53 (2018) 37-44.
- [72] A.I. Mehta, C.W. Kanaly, A.H. Friedman, D.D. Bigner, J.H. Sampson, *Monitoring Radiographic Brain Tumor Progression*, *Toxins*, 3 (2011) 191-200.
- [73] H. Malone, J. Yang, D.L. Hershman, J.D. Wright, J.N. Bruce, A.I. Neugut, *Complications Following Stereotactic Needle Biopsy of Intracranial Tumors*, *World Neurosurgery*, 84 (2015) 1084-1089.
- [74] O. van Tellingen, B. Yetkin-Arik, M.C. de Gooijer, P. Wesseling, T. Wurdinger, H.E. de Vries, *Overcoming the blood-brain tumor barrier for effective glioblastoma treatment*, *Drug Resistance Updates*, 19 (2015) 1-12.
- [75] J.N. Sarkaria, L.S. Hu, I.F. Parney, D.H. Pafundi, D.H. Brinkmann, N.N. Laack, C. Giannini, T.C. Burns, S.H. Kizilbash, J.K. Laramy, K.R. Swanson, T.J. Kaufmann, P.D. Brown, N.Y.R. Agar, E. Galanis, J.C. Buckner, W.F. Elmquist, *Is the blood-brain barrier really disrupted in all glioblastomas? A critical assessment of existing clinical data*, *Neuro-Oncology*, 20 (2018) 184-191.
- [76] S. Currie, M. Hadjivassiliou, I.J. Craven, I.D. Wilkinson, P.D. Griffiths, N. Hoggard, *Magnetic resonance spectroscopy of the brain*, *Postgraduate Medical Journal*, (2013) 94-106.
- [77] M.C. Preul, Z. Caramanos, D.L. Collins, J.-G. Villemure, R. Leblanc, A. Olivier, R. Pokrupa, D.L. Arnold, *Accurate, noninvasive diagnosis of human brain tumors by using proton magnetic resonance spectroscopy*, *Nature Medicine*, 2 (1996) 323-325.
- [78] G.J. Patti, O. Yanes, G. Siuzdak, *Innovation: Metabolomics: the apogee of the omics trilogy*, *Nature Reviews Molecular Cell Biology*, 13 (2012) 263-269.
- [79] M.C. Muñoz-Hernández, M.L. García-Martín, *In Vivo 1H Magnetic Resonance Spectroscopy*, in: M.L. García Martín, P. López Larrubia (Eds.) *Preclinical MRI: Methods and Protocols*, Springer New York, New York, NY, 2018, pp. 151-167.
- [80] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G.E. Duggan, G.D. Macinnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B.D. Sykes, H.J. Vogel, L. Querengesser, *HMDB: the Human Metabolome Database*, *Nucleic acids research*, 35 (2007) D521-526.
- [81] D.S. Wishart, *Quantitative metabolomics using NMR*, *TrAC Trends in Analytical Chemistry*, 27 (2008) 228-237.
- [82] K. Bingol, L. Bruschweiler-Li, D.-W. Li, R. Bruschweiler, *Customized Metabolomics Database for the Analysis of NMR 1H-1H TOCSY and 13C-1H HSQC-TOCSY Spectra of Complex Mixtures*, *Analytical Chemistry*, 86 (2014) 5494-5501.
- [83] K.K. Lehtimäki, P.K. Valonen, J.L. Griffin, T.H. Väisänen, O.H.J. Gröhn, M.I. Kettunen, J. Vepsäläinen, S. Ylä-Herttuala, J. Nicholson, R.A. Kauppinen, *Metabolite Changes in BT4C*

Rat Gliomas Undergoing Ganciclovir-Thymidine Kinase Gene Therapy-induced Programmed Cell Death as Studied by 1H NMR Spectroscopy in Vivo, ex Vivo, and in Vitro, Journal of Biological Chemistry, 278 (2003) 45915-45923.

[84] P.W. Atkins, *Atkins' physical chemistry*, 10th ed., Oxford University Press, Oxford, 2014.

[85] T.D.W. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, 3 ed., Elsevier, Amsterdam, 2016.

[86] J. Keeler, *Understanding NMR spectroscopy*, 2nd ed., John Wiley and Sons, Chichester, U.K., 2010.

[87] D.L. Pavia, *Introduction to spectroscopy*, 4th ed., Brooks/Cole, Belmont, Calif, 2009.

[88] J.L. Spratlin, N.J. Serkova, S.G. Eckhardt, *Clinical Applications of Metabolomics in Oncology: A Review*, Clinical Cancer Research, 15 (2009) 431-440.

[89] O. Beckonert, H.C. Keun, T.M.D. Ebbels, J. Bundy, E. Holmes, J.C. Lindon, J.K. Nicholson, *Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts*, Nature Protocols, 2 (2007) 2692-2703.

[90] K. Kazimierczuk, V. Orekhov, *Non-uniform sampling: post-Fourier era of NMR data collection and processing*, Magnetic Resonance in Chemistry, 53 (2015) 921-926.

[91] F. Tugizimana, A. Steenkamp Paul, A. Piater Lizelle, A. Dubery Ian, *Mass spectrometry in untargeted liquid chromatography/mass spectrometry metabolomics: Electrospray ionisation parameters and global coverage of the metabolome*, Rapid Communications in Mass Spectrometry, 32 (2017) 121-132.

[92] A.-H.M. Emwas, *The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research*, in: J.T. Bjerrum (Ed.) *Metabonomics: Methods and Protocols*, Springer New York, New York, NY, 2015, pp. 161-193.

[93] T. A. Pertinhez, E. Casali, L. Zambianchi, A. Spisni, R. Baricchi, *Statistical validation of 1H NMR protocol vs standard biochemical assay in quality control of RBC packed units*, Journal of Pharmaceutical and Biomedical Analysis, 147 (2018) 485-492.

[94] K. Murray Kermit, K. Boyd Robert, N. Eberlin Marcos, G.J. Langley, L. Li, Y. Naito, *Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013)*, in: *Pure and Applied Chemistry*, 2013, pp. 1515-1609.

[95] A.G. Brenton, A.R. Godfrey, *Accurate Mass Measurement: Terminology and Treatment of Data*, Journal of the American Society for Mass Spectrometry, 21 (2010) 1821-1835.

[96] E. Lundanes, L.T. Reubsaet, *Chromatography : Basic Principles, Sample Preparations and Related Methods*, Somerset, NJ, USA: John Wiley & Sons, Somerset, 2013.

[97] J.V. Olsen, L.M.F. de Godoy, G. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning, M. Mann, *Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap*, Molecular & Cellular Proteomics, 4 (2005) 2010-2021.

[98] A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning, *Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer*, Analytical Chemistry, 78 (2006) 2113-2120.

[99] L. Cui, H. Lu, H. Lee Yie, *Challenges and emergent solutions for LC-MS/MS based untargeted metabolomics in diseases*, Mass Spectrometry Reviews, (2018) 1-21.

[100] I. Gertsman, B.A. Barshop, *Promises and pitfalls of untargeted metabolomics*, Journal of Inherited Metabolic Disease, (2018) 355-366.

[101] M. Yan, G. Xu, *Current and future perspectives of functional metabolomics in disease studies—A review*, Analytica Chimica Acta, (2018) 1-14.

[102] A. Makarov, *Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis*, Analytical Chemistry, 72 (2000) 1156-1162.

- [103] S. Eliuk, A. Makarov, *Evolution of Orbitrap Mass Spectrometry Instrumentation*, Annual Review of Analytical Chemistry, 8 (2015) 61-80.
- [104] J.V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, *Higher-energy C-trap dissociation for peptide modification analysis*, Nature Methods, 4 (2007) 709-712.
- [105] W. Lu, E. Kimball, J.D. Rabinowitz, *A High-Performance Liquid Chromatography-Tandem Mass Spectrometry Method for Quantitation of Nitrogen-Containing Intracellular Metabolites*, Journal of the American Society for Mass Spectrometry, 17 (2006) 37-50.
- [106] K. Spagou, I.D. Wilson, P. Masson, G. Theodoridis, N. Raikos, M. Coen, E. Holmes, J.C. Lindon, R.S. Plumb, J.K. Nicholson, E.J. Want, *HILIC-UPLC-MS for Exploratory Urinary Metabolic Profiling in Toxicological Studies*, Analytical Chemistry, 83 (2011) 382-390.
- [107] K. Spagou, H. Tsoukali, N. Raikos, H. Gika, D. Wilson Ian, G. Theodoridis, *Hydrophilic interaction chromatography coupled to MS for metabonomic/metabolomic studies*, Journal of Separation Science, 33 (2010) 716-727.
- [108] D.Q. Tang, L. Zou, X.X. Yin, N. Ong Choon, *HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS*, Mass Spectrometry Reviews, 35 (2016) 574-600.
- [109] T. Ikegami, K. Tomomatsu, H. Takubo, K. Horie, N. Tanaka, *Separation efficiencies in hydrophilic interaction chromatography*, Journal of Chromatography A, 1184 (2008) 474-503.
- [110] D.V. McCalley, *Study of the selectivity, retention mechanisms and performance of alternative silica-based stationary phases for separation of ionised solutes in hydrophilic interaction chromatography*, Journal of Chromatography A, 1217 (2010) 3408-3417.
- [111] P. Jandera, *Stationary and mobile phases in hydrophilic interaction chromatography: a review*, Analytica Chimica Acta, 692 (2011) 1-25.
- [112] H. Idborg, L. Zamani, P.-O. Edlund, I. Schuppe-Koistinen, S.P. Jacobsson, *Metabolic fingerprinting of rat urine by LC/MS: Part 1. Analysis by hydrophilic interaction liquid chromatography–electrospray ionization mass spectrometry*, Journal of Chromatography B, 828 (2005) 9-13.
- [113] A. Kamleh, M.P. Barrett, D. Wildridge, R.J.S. Burchmore, R.A. Scheltema, D.G. Watson, *Metabolomic profiling using Orbitrap Fourier transform mass spectrometry with hydrophilic interaction chromatography: a method with wide applicability to analysis of biomolecules*, Rapid Communications in Mass Spectrometry, 22 (2008) 1912-1918.
- [114] K. Contrepois, L. Jiang, M. Snyder, *Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography (RPLC)–Mass Spectrometry*, Molecular & Cellular Proteomics, 14 (2015) 1684-1695.
- [115] I. Roci, H. Gallart-Ayala, A. Schmidt, J. Watrous, M. Jain, C.E. Wheelock, R. Nilsson, *Metabolite Profiling and Stable Isotope Tracing in Sorted Subpopulations of Mammalian Cells*, Analytical Chemistry, 88 (2016) 2707-2713.
- [116] H. Gallart-Ayala, I. Konz, F. Mehl, T. Teav, A. Oikonomidi, G. Peyratout, V. van der Velpen, J. Popp, J. Ivanisevic, *A global HILIC-MS approach to measure polar human cerebrospinal fluid metabolome: Exploring gender-associated variation in a cohort of elderly cognitively healthy subjects*, Analytica Chimica Acta, (2018) 1-11.
- [117] E. Wikberg, T. Sparrman, C. Viklund, T. Jonsson, K. Irgum, *A ²H nuclear magnetic resonance study of the state of water in neat silica and zwitterionic stationary phases and its influence on the chromatographic retention characteristics in hydrophilic interaction high-performance liquid chromatography*, Journal of Chromatography A, 1218 (2011) 6630-6638.
- [118] P. Hemström, K. Irgum, *Hydrophilic interaction chromatography*, Journal of Separation Science, 29 (2006) 1784-1821.
- [119] G. Greco, T. Letzel, *Main Interactions and Influences of the Chromatographic Parameters in HILIC Separations*, Journal of Chromatographic Science, 51 (2013) 684-693.

- [120] J.F. Banks, *High-sensitivity peptide mapping using packed-capillary liquid chromatography and electrospray ionization mass spectrometry*, Journal of Chromatography A, 743 (1996) 99-104.
- [121] S.R. Wilson, T. Vehus, H.S. Berg, E. Lundanes, *Nano-LC in proteomics: recent advances and approaches*, Bioanalysis, 7 (2015) 1799-1815.
- [122] M. Bedair, L.W. Sumner, *Current and emerging mass-spectrometry technologies for metabolomics*, TrAC Trends in Analytical Chemistry, 27 (2008) 238-250.
- [123] Z. Lei, D.V. Huhman, L.W. Sumner, *Mass Spectrometry Strategies in Metabolomics*, Journal of Biological Chemistry, 286 (2011) 25435-25442.
- [124] R.B. Cole, *Electrospray ionization mass spectrometry: fundamentals, instrumentation, and applications*, Wiley, New York, 1997.
- [125] G.J. Van Berkel, F. Zhou, J.T. Aronson, *Changes in bulk solution pH caused by the inherent controlled-current electrolytic process of an electrospray ion source*, International Journal of Mass Spectrometry and Ion Processes, 162 (1997) 55-67.
- [126] L. Tang, P. Kebarle, *Effect of the conductivity of the electrosprayed solution on the electrospray current. Factors determining analyte sensitivity in electrospray mass spectrometry*, Analytical Chemistry, 63 (1991) 2709-2715.
- [127] J.N. Miller, *Statistics and chemometrics for analytical chemistry*, 6th ed., Pearson Prentice Hall, Harlow, 2010.
- [128] R. Vettukattil, *Preprocessing of Raw Metabonomic Data*, in: J.T. Bjerrum (Ed.) *Metabonomics: Methods and Protocols*, Springer New York, New York, NY, 2015, pp. 123-136.
- [129] D. Chang, C.D. Banack, S.L. Shah, *Robust baseline correction algorithm for signal dense NMR spectra*, Journal of Magnetic Resonance, 187 (2007) 288-292.
- [130] M. Katajamaa, M. Orešič, *Data processing for mass spectrometry-based metabolomics*, Journal of Chromatography A, 1158 (2007) 318-328.
- [131] R.A. van den Berg, H.C.J. Hoefsloot, J.A. Westerhuis, A.K. Smilde, M.J. van der Werf, *Centering, scaling, and transformations: improving the biological information content of metabolomics data*, BMC Genomics, 7 (2006) 142-142.
- [132] D. Broadhurst, D. Kell, *Statistical strategies for avoiding false discoveries in metabolomics and related experiments*, Metabolomics, 2 (2006) 171-196.
- [133] I.T. Jolliffe, J. Cadima, *Principal component analysis: a review and recent developments*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374 (2016).
- [134] S. Wold, M. Sjöström, L. Eriksson, *PLS-regression: a basic tool of chemometrics*, Chemometrics and Intelligent Laboratory Systems, 58 (2001) 109-130.
- [135] J.A. Westerhuis, E.J.J. van Velzen, H.C.J. Hoefsloot, A.K. Smilde, *Discriminant Q2 (DQ2) for improved discrimination in PLS-DA models*, Metabolomics, 4 (2008) 293-296.
- [136] D.M. Allen, *The Relationship between Variable Selection and Data Augmentation and a Method for Prediction*, Technometrics, 16 (1974) 125-127.
- [137] A. Golbraikh, A. Tropsha, *Beware of $q^2!$* , Journal of Molecular Graphics and Modelling, 20 (2002) 269-276.
- [138] H.S. Berg, K.E. Seterdal, T. Smetop, R. Rozenvalds, O.K. Brandtzaeg, T. Vehus, E. Lundanes, S.R. Wilson, *Self-packed core shell nano liquid chromatography columns and silica-based monolithic trap columns for targeted proteomics*, Journal of Chromatography A, 1498 (2017) 111-119.
- [139] K. Kazimierczuk, Y. Orekhov Vladislav, *Accelerated NMR Spectroscopy by Using Compressed Sensing*, Angewandte Chemie International Edition, 50 (2011) 5556-5559.
- [140] S.L. Robinette, F. Zhang, L. Brüschweiler-Li, R. Brüschweiler, *Web Server Based Complex Mixture Analysis by NMR*, Analytical Chemistry, 80 (2008) 3606-3611.

- [141] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C.F. Schulte, D.E. Tolmie, R. Kent Wenger, H. Yao, J.L. Markley, *BioMagResBank*, Nucleic acids research, 36 (2008) D402-D408.
- [142] R Core Team, *R: A language and environment for statistical computing*, in, R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [143] D. Barkauskas, *FTICRMS: Programs for Analyzing Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry Data*, in, 2012.
- [144] W. Stacklies, H. Redestig, M. Scholz, D. Walther, J. Selbig, *pcaMethods—a bioconductor package providing PCA methods for incomplete data*, Bioinformatics, 23 (2007) 1164-1167.
- [145] L. J., *Plotrix: a package in the red light district of R*, R-News, 6 (2006) 8-12.
- [146] S. Dayal Bhupinder, F. MacGregor John, *Improved PLS algorithms*, Journal of Chemometrics, 11 (1997) 73-85.
- [147] B.-H. Mevik, R. Wehrens, *The pls Package: Principal Component and Partial Least Squares Regression in R*, Journal of Statistical Software, 18 (2007) 1-23.
- [148] B.-H. Mevik, R. Wehrens, K.H. Liland, *pls: Partial Least Squares and Principal Component Regression*, in, 2016.
- [149] S.S. Gill, R.K. Small, D.G.T. Thomas, P. Patel, R. Porteous, N. Van Bruggen, D.G. Gadian, R.A. Kauppinen, S.R. Williams, *Brain metabolites as ¹H NMR markers of neuronal and glial disorders*, NMR in Biomedicine, 2 (1989) 196-200.
- [150] C.L. Florian, N.E. Preece, K.K. Bhakoo, S.R. Williams, M.D. Noble, *Cell Type-specific Fingerprinting of Meningioma and Meningeal Cells by Proton Nuclear Magnetic Resonance Spectroscopy*, Cancer research, 55 (1995) 420-427.
- [151] N. Matheus, S. Hansen, E. Rozet, P. Peixoto, E. Maquoi, V. Lambert, A. Noel, M. Frederich, D. Mottet, P. de Tullio, *An easy, convenient cell and tissue extraction protocol for nuclear magnetic resonance metabolomics*, Phytochemical Analysis, 25 (2014) 342-349.
- [152] J.B. Ritter, Y. Genzel, U. Reichl, *Simultaneous extraction of several metabolites of energy metabolism and related substances in mammalian cells: Optimization using experimental design*, Analytical Biochemistry, 373 (2008) 349-369.
- [153] S. Dietmair, N.E. Timmins, P.P. Gray, L.K. Nielsen, J.O. Krömer, *Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol*, Analytical Biochemistry, 404 (2010) 155-164.
- [154] K. Dettmer, N. Nürnberger, H. Kaspar, M.A. Gruber, M.F. Almstetter, P.J. Oefner, *Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols*, Analytical and Bioanalytical Chemistry, 399 (2011) 1127-1139.
- [155] M.A. Lorenz, C.F. Burant, R.T. Kennedy, *Reducing Time and Increasing Sensitivity in Sample Preparation for Adherent Mammalian Cell Metabolomics*, Analytical Chemistry, 83 (2011) 3406-3414.
- [156] US Food & Drug Administration, *Guidelines for the Validation of Chemical Methods for the FDA FVM Program (2nd edition)*, Department of health & human services, 2015
- [157] S. Bansal, A. DeStefano, *Key elements of bioanalytical method validation for small molecules*, The AAPS Journal, 9 (2007) E109-E114.
- [158] G.F. Giskeødegård, T.G. Bloemberg, G. Postma, B. Sitter, M.-B. Tessem, I.S. Gribbestad, T.F. Bathen, L.M.C. Buydens, *Alignment of high resolution magic angle spinning magnetic resonance spectra using warping methods*, Analytica Chimica Acta, 683 (2010) 1-11.
- [159] A.C. Dona, B. Jiménez, H. Schäfer, E. Humpfer, M. Spraul, M.R. Lewis, J.T.M. Pearce, E. Holmes, J.C. Lindon, J.K. Nicholson, *Precision High-Throughput Proton NMR Spectroscopy of Human Urine, Serum, and Plasma for Large-Scale Metabolic Phenotyping*, Analytical Chemistry, 86 (2014) 9887-9894.

- [160] D. Sachse, *The metabolic profile at the crossroads of pregnancy and infancy: exploratory biomarker research using NMR-based urinary metabolomics*, Faculty of Medicine, University of Oslo, PhD (2014).
- [161] A. Bubb William, *NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity*, Concepts in Magnetic Resonance Part A, 19A (2003) 1-19.
- [162] A.M. Weljie, J. Newton, P. Mercier, E. Carlson, C.M. Slupsky, *Targeted profiling: Quantitative analysis of ¹H NMR metabolomics data*, Analytical Chemistry, 78 (2006) 4430-4442.
- [163] M. Cuperlovic-Culf, D. Ferguson, A. Culf, P. Morin, M. Touaibia, *¹H NMR metabolomics analysis of glioblastoma subtypes: correlation between metabolomics and gene expression characteristics*, The Journal of biological chemistry, 287 (2012) 20164-20175.
- [164] L. Guidoni, L. Ricci-Vitiani, A. Rosi, A. Palma, S. Grande, M. Luciani Anna, F. Pelacchi, S. Martino, C. Colosimo, M. Biffoni, R. De Maria, R. Pallini, V. Viti, *¹H NMR detects different metabolic profiles in glioblastoma stem-like cells*, NMR in Biomedicine, 27 (2013) 129-145.
- [165] F.A. Howe, S.J. Barton, S.A. Cudlip, M. Stubbs, D.E. Saunders, M. Murphy, P. Wilkins, K.S. Opstad, V.L. Doyle, M.A. McLean, B.A. Bell, J.R. Griffiths, *Metabolic profiles of human brain tumors using quantitative in vivo ¹H magnetic resonance spectroscopy*, Magnetic Resonance in Medicine, 49 (2003) 223-232.
- [166] K. Kallenberg, H.C. Bock, G. Helms, K. Jung, A. Wrede, J.-H. Buhk, A. Giese, J. Frahm, H. Strik, P. Dechent, M. Knauth, *Untreated Glioblastoma Multiforme: Increased Myo-inositol and Glutamine Levels in the Contralateral Cerebral Hemisphere at Proton MR Spectroscopy*, Radiology, 253 (2009) 805-812.
- [167] A. Candiota, C. Majós, M. Julià-Sapé, *Non-invasive grading of astrocytic tumours from the relative contents of myo-inositol and glycine measured by in vivo MRS*, Journal of the Belgian Society of Radiology, 94 (2011) 319-329.
- [168] A.A. Maudsley, R.K. Gupta, R. Stoyanova, N.A. Parra, B. Roy, S. Sheriff, N. Hussain, S. Behari, *Mapping of Glycine Distributions in Gliomas*, American Journal of Neuroradiology, 35 (2014) S31-S36.
- [169] V. Govindaraju, K. Young, A. Maudsley Andrew, *Proton NMR chemical shifts and coupling constants for brain metabolites*, NMR in Biomedicine, 13 (2000) 129-153.
- [170] R.J. DeBerardinis, A. Mancuso, E. Daikhin, I. Nissim, M. Yudkoff, S. Wehrli, C.B. Thompson, *Beyond aerobic glycolysis: Transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis*, Proceedings of the National Academy of Sciences, 104 (2007) 19345-19350.
- [171] C. Yang, J. Sudderth, T. Dang, R.G. Bachoo, J.G. McDonald, R.J. DeBerardinis, *Glioblastoma Cells Require Glutamate Dehydrogenase to Survive Impairments of Glucose Metabolism or Akt Signaling*, Cancer research, 69 (2009) 7986-7993.
- [172] E.A. Maher, I. Marin-Valencia, R.M. Bachoo, T. Mashimo, J. Raisanen, K.J. Hatanpaa, A. Jindal, F.M. Jeffrey, C. Choi, C. Madden, D. Mathews, J.M. Pascual, B.E. Mickey, C.R. Malloy, R.J. DeBerardinis, *Metabolism of [¹³C]glucose in human brain tumors in vivo*, NMR in Biomedicine, 25 (2012) 1234-1244.
- [173] R.J. DeBerardinis, J.J. Lum, G. Hatzivassiliou, C.B. Thompson, *The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation*, Cell Metabolism, 7 (2008) 11-20.
- [174] N. Zaidi, J.V. Swinnen, K. Smans, *ATP-Citrate Lyase: A Key Player in Cancer Metabolism*, Cancer research, 72 (2012) 3709-3714.
- [175] R.E. Isaacks, A.S. Bender, C.Y. Kim, N.M. Prieto, M.D. Norenberg, *Osmotic regulation of myo-inositol uptake in primary astrocyte cultures*, Neurochemical Research, 19 (1994) 331-338.

- [176] A. Brand, C. Richter-Landsberg, D. Leibfritz, *Multinuclear NMR Studies on the Energy Metabolism of Glial and Neuronal Cells*, *Developmental Neuroscience*, 15 (1993) 289-298.
- [177] H. Nagashima, T. Sasayama, K. Tanaka, K. Kyotani, N. Sato, M. Maeyama, M. Kohta, J. Sakata, Y. Yamamoto, K. Hosoda, T. Itoh, R. Sasaki, E. Kohmura, *Myo-inositol concentration in MR spectroscopy for differentiating high grade glioma from primary central nervous system lymphoma*, *Journal of neuro-oncology*, 136 (2018) 317-326.
- [178] T.L. Weissgerber, N.M. Milic, S.J. Winham, V.D. Garovic, *Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm*, *PLOS Biology*, 13 (2015) 1-10.
- [179] T. Pemovska, M. Kontro, B. Yadav, H. Edgren, S. Eldfors, A. Sz wajda, H. Almusa, M.M. Bepalov, P. Ellonen, E. Elonen, B.T. Gjertsen, R. Karjalainen, E. Kuleskiy, S. Lagström, A. Lehto, M. Lepistö, T. Lundán, M.M. Majumder, J.M.L. Marti, P. Mattila, A. Murumägi, S. Mustjoki, A. Palva, A. Parsons, T. Pirttinen, M.E. Rämetsä, M. Suvela, L. Turunen, I. Väström, M. Wolf, J. Knowles, T. Aittokallio, C.A. Heckman, K. Porkka, O. Kallioniemi, K. Wennerberg, *Individualized Systems Medicine Strategy to Tailor Treatments for Patients with Chemorefractory Acute Myeloid Leukemia*, *Cancer Discovery*, 3 (2013) 1416-1429.
- [180] D.M. Tiek, J.D. Rone, G.T. Graham, E.L. Pannkuk, B.R. Haddad, R.B. Riggins, *Alterations in Cell Motility, Proliferation, and Metabolism in Novel Models of Acquired Temozolomide Resistant Glioblastoma*, *Scientific Reports*, 8 (2018) 1-11.
- [181] Z.A. Seymour, A. Panigrahy, J.L. Finlay, M.D. Nelson, S. Blüml, *Citrate in Pediatric CNS Tumors?*, *American Journal of Neuroradiology*, 29 (2008) 1006-1011.
- [182] G.P. Liney, L.W. Turnbull, M. Lowry, L.S. Turnbull, A.J. Knowles, A. Horsman, *In vivo quantification of citrate concentration and water T2 relaxation time of the pathologic prostate gland using 1H MRS and MRI*, *Magnetic Resonance Imaging*, 15 (1997) 1177-1186.
- [183] V. Kumar, G.S. Bora, R. Kumar, N.R. Jagannathan, *Multiparametric (mp) MRI of prostate cancer*, *Progress in Nuclear Magnetic Resonance Spectroscopy*, 105 (2018) 23-40.
- [184] K. Herholz, W. Heindel, P.R. Luyten, J.A. denHollander, U. Pietrzyk, J. Voges, H. Kugel, G. Friedmann, W.D. Heiss, *In vivo imaging of glucose consumption and lactate concentration in human gliomas*, *Annals of Neurology*, 31 (1992) 319-327.
- [185] K.W. Kohn, *Interstrand Cross-linking of DNA by 1,3-Bis(2-chloroethyl)-1-nitrosourea and Other 1-(2-haloethyl)-1-nitrosoureas*, *Cancer research*, 37 (1977) 1450-1454.
- [186] J.M. Rozental, J.D. Cohen, M.P. Mehta, R.L. Levine, J.M. Hanson, R.J. Nickles, *Acute changes in glucose uptake after treatment: the effects of carmustine (BCNU) on human glioblastoma multiforme*, *Journal of neuro-oncology*, 15 (1993) 57-66.
- [187] D. Sachse, L. Sletner, K. Mørkrid, A.K. Jenum, K.I. Birkeland, F. Rise, A.P. Piehler, J.P. Berg, *Metabolic Changes in Urine during and after Pregnancy in a Large, Multiethnic Population-Based Cohort Study of Gestational Diabetes*, *PLOS ONE*, 7 (2012) 1-12.
- [188] National Center for Biotechnology Information, *Compound identifier = 612*, PubChem Compound Database, accessed May 26, 2018, <https://pubchem.ncbi.nlm.nih.gov/compound/612>
- [189] National Center for Biotechnology Information, *Compound identifier = 311*, PubChem Compound Database, accessed May 26, 2018, <https://pubchem.ncbi.nlm.nih.gov/compound/311>
- [190] Y. Huang, Y. Tian, Z. Zhang, C. Peng, *A HILIC-MS/MS method for the simultaneous determination of seven organic acids in rat urine as biomarkers of exposure to realgar*, *Journal of Chromatography B*, 905 (2012) 37-42.
- [191] E. Johnsen, S. Leknes, S.R. Wilson, E. Lundanes, *Liquid chromatography-mass spectrometry platform for both small neurotransmitters and neuropeptides in blood, with automatic and robust solid phase extraction*, *Scientific Reports*, 5 (2015) 9308-9316.

- [192] K. Bjørseth, *Evaluation of exosome isolation techniques for hydrophilic interaction liquid chromatography-mass spectrometry based metabolomics*, Department of Chemistry, University of Oslo, MSc (2017).
- [193] D.V. McCalley, *A study of column equilibration time in hydrophilic interaction chromatography*, Journal of Chromatography A, 1554 (2018) 61-70.
- [194] Y. Kawachi, T. Ikegami, H. Takubo, Y. Ikegami, M. Miyamoto, N. Tanaka, *Chromatographic characterization of hydrophilic interaction liquid chromatography stationary phases: Hydrophilicity, charge effects, structural selectivity, and separation efficiency*, Journal of Chromatography A, 1218 (2011) 5903-5919.
- [195] J.P. Chervet, M. Ursem, J.P. Salzmann, *Instrumental Requirements for Nanoscale Liquid Chromatography*, Analytical Chemistry, 68 (1996) 1507-1512.
- [196] J. Layne, T. Farcas, I. Rustamov, F. Ahmed, *Volume-load capacity in fast-gradient liquid chromatography: Effect of sample solvent composition and injection volume on chromatographic performance*, Journal of Chromatography A, 913 (2001) 233-242.
- [197] P. Plateau, M. Gueron, *Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences*, Journal of the American Chemical Society, 104 (1982) 7310-7311.
- [198] M. Piotto, V. Saudek, V. Sklenář, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*, Journal of Biomolecular NMR, 2 (1992) 661-665.
- [199] K. Stott, J. Stonehouse, J. Keeler, T.-L. Hwang, A.J. Shaka, *Excitation Sculpting in High-Resolution Nuclear Magnetic Resonance Spectroscopy: Application to Selective NOE Experiments*, Journal of the American Chemical Society, 117 (1995) 4199-4200.
- [200] C. Qian, C.-J. Lai, R. Bao, D.-G. Wang, J. Wang, G.-X. Xu, R. Atoyian, H. Qu, L. Yin, M. Samson, B. Zifcak, A.W.S. Ma, S. DellaRocca, M. Borek, H.-X. Zhai, X. Cai, M. Voi, *Cancer Network Disruption by a Single Molecule Inhibitor Targeting Both Histone Deacetylase Activity and Phosphatidylinositol 3-Kinase Signaling*, Clinical Cancer Research, 18 (2012) 4104-4113.
- [201] N. The Cancer Genome Atlas Research, *Comprehensive genomic characterization defines human glioblastoma genes and core pathways*, Nature, 455 (2008) 1061-1068.
- [202] A. Sami, M. Karsy, *Targeting the PI3K/AKT/mTOR signaling pathway in glioblastoma: novel therapeutic agents and advances in understanding*, Tumor Biology, 34 (2013) 1991-2002.
- [203] K.-i. Ozaki, M. Kosugi, N. Baba, K. Fujio, T. Sakamoto, S. Kimura, S. Tanimura, M. Kohno, *Blockade of the ERK or PI3K–Akt signaling pathway enhances the cytotoxicity of histone deacetylase inhibitors in tumor cells resistant to gefitinib or imatinib*, Biochemical and Biophysical Research Communications, 391 (2010) 1610-1615.
- [204] V.M. Richon, S. Emiliani, E. Verdin, Y. Webb, R. Breslow, R.A. Rifkind, P.A. Marks, *A class of hybrid polar inducers of transformed cell differentiation inhibits histone deacetylases*, Proceedings of the National Academy of Sciences, 95 (1998) 3003-3007.
- [205] C.E. Quartararo, E. Reznik, A.C. deCarvalho, T. Mikkelsen, B.R. Stockwell, *High-Throughput Screening of Patient-Derived Cultures Reveals Potential for Precision Medicine in Glioblastoma*, ACS Medicinal Chemistry Letters, 6 (2015) 948-952.
- [206] M. Hasmann, I. Schemainda, *FK866, a Highly Specific Noncompetitive Inhibitor of Nicotinamide Phosphoribosyltransferase, Represents a Novel Mechanism for Induction of Tumor Cell Apoptosis*, Cancer research, 63 (2003) 7436-7442.
- [207] G. Magni, A. Amici, M. Emanuelli, G. Orsomando, N. Raffaelli, S. Ruggieri, *Enzymology of NAD⁺ homeostasis in man*, Cellular and Molecular Life Sciences CMLS, 61 (2004) 19-34.
- [208] A.D. Gujar, S. Le, D.D. Mao, D.Y.A. Dadey, A. Turski, Y. Sasaki, D. Aum, J. Luo, S. Dahiya, L. Yuan, K.M. Rich, J. Milbrandt, D.E. Hallahan, H. Yano, D.D. Tran, A.H. Kim, *An NAD⁺-dependent transcriptional program governs self-renewal and radiation resistance in glioblastoma*, Proceedings of the National Academy of Sciences, 113 (2016) E8247-E8256.

- [209] R.E. Kast, J.A. Boockvar, A. Brüning, F. Cappello, W.-W. Chang, B. Cvek, Q.P. Dou, A. Duenas-Gonzalez, T. Efferth, D. Focosi, S.H. Ghaffari, G. Karpel-Massler, K. Ketola, A. Khoshnevisan, D. Keizman, N. Magné, C. Marosi, K. McDonald, M. Muñoz, A. Paranjpe, M.H. Pourgholami, I. Sardi, A. Sella, K.S. Srivenugopal, M. Tuccori, W. Wang, C.R. Wirtz, M.-E. Halatsch, *A conceptually new treatment approach for relapsed glioblastoma: Coordinated undermining of survival paths with nine repurposed drugs (CUSP9) by the International Initiative for Accelerated Improvement of Glioblastoma Care*, *Oncotarget*, 4 (2013) 502-530.
- [210] R.E. Kast, G. Karpel-Massler, M.-E. Halatsch, *CUSP9* treatment protocol for recurrent glioblastoma: aprepitant, artesunate, auranofin, captopril, celecoxib, disulfiram, itraconazole, ritonavir, sertraline augmenting continuous low dose temozolomide*, *Oncotarget*, 5 (2014) 8052-8082.
- [211] T. Akazawa, G. Kwatra Shawn, E. Goldsmith Laura, D. Richardson Mark, A. Cox Elizabeth, H. Sampson John, M. Kwatra Madan, *A constitutively active form of neurokinin 1 receptor and neurokinin 1 receptor-mediated apoptosis in glioblastomas*, *Journal of Neurochemistry*, 109 (2009) 1079-1086.
- [212] X. Wang, Y. Chen, S. Zhang, L. Zhang, X. Liu, L. Zhang, X. Li, D. Chen, *Co-expression of COX-2 and 5-LO in primary glioblastoma is associated with poor prognosis*, *Journal of neuro-oncology*, 125 (2015) 277-285.
- [213] H. Bow, L.S. Hwang, N. Schildhaus, J. Xing, L. Murray, Q. Salditch, X. Ye, Y. Zhang, J. Weingart, H. Brem, B. Tyler, *Local delivery of angiogenesis-inhibitor minocycline combined with radiotherapy and oral temozolomide chemotherapy in 9L glioma*, *Journal of Neurosurgery*, 120 (2013) 662-669.
- [214] N. Garrido-Mesa, A. Zarzuelo, J. Gálvez, *Minocycline: far beyond an antibiotic*, *British Journal of Pharmacology*, 169 (2013) 337-352.
- [215] Y. Wang, N. Huang, H. Li, S. Liu, X. Chen, S. Yu, N. Wu, X.W. Bian, H.Y. Shen, C. Li, L. Xiao, *Promoting oligodendroglial-oriented differentiation of glioma stem cell: a repurposing of quetiapine for the treatment of malignant glioma*, *Oncotarget*, 8 (2017) 37511-37524.
- [216] S.K. Tan, A. Jermakowicz, A.K. Mookhtiar, C.B. Nemeroff, S.C. Schürer, N.G. Ayad, *Drug Repositioning in Glioblastoma: A Pathway Perspective*, *Frontiers in Pharmacology*, 9 (2018) 1-19.
- [217] R.B. Gil, A. Ortiz, M.D. Sanchez-Niño, K. Markoska, E. Schepers, R. Vanholder, G. Glorieux, P. Schmitt-Kopplin, S.S. Heinzmann, *Increased urinary osmolyte excretion indicates chronic kidney disease severity and progression rate*, *Nephrology Dialysis Transplantation*, (2018) 1-9.
- [218] R.O. Bahado-Singh, R. Akolekar, R. Mandal, E. Dong, J. Xia, M. Kruger, D.S. Wishart, K. Nicolaides, *Metabolomics and first-trimester prediction of early-onset preeclampsia*, *The Journal of Maternal-Fetal & Neonatal Medicine*, 25 (2012) 1840-1847.
- [219] E. Locci, A. Noto, M. Puddu, G. Pomero, R. Demontis, C. Dalmazzo, A. Delogu, V. Fanos, E. d'Aloja, P. Gancia, *A longitudinal ¹H-NMR metabolomics analysis of urine from newborns with hypoxic-ischemic encephalopathy undergoing hypothermia therapy*. *Clinical and medical legal insights*, *PLOS ONE*, 13 (2018) 1-19.
- [220] T.W.M. Fan, A.N. Lane, *NMR-based stable isotope resolved metabolomics in systems biochemistry*, *Journal of Biomolecular NMR*, 49 (2011) 267-280.
- [221] M. Bylesjö, *Extracting meaningful information from metabonomic data using multivariate statistics*, *Methods in Molecular Biology*, 1277 (2015) 137-146.
- [222] J. Feng, P.-F. Yan, H.-y. Zhao, F.-C. Zhang, W.-H. Zhao, M. Feng, *Inhibitor of Nicotinamide Phosphoribosyltransferase Sensitizes Glioblastoma Cells to Temozolomide via Activating ROS/JNK Signaling Pathway*, *BioMed Research International*, 2016 (2016) 1-11.

- [223] E.P. Jane, D.R. Premkumar, J.D. DiDomenico, B. Hu, S.-Y. Cheng, I.F. Pollack, *YM-155 Potentiates the Effect of ABT-737 in Malignant Human Glioma Cells via Survivin and Mcl-1 Downregulation in an EGFR-Dependent Context*, *Molecular Cancer Therapeutics*, 12 (2013) 326-338.
- [224] D.C. Harris, *Quantitative chemical analysis*, 9th ed. ed., W.H. Freeman, New York, 2016.

8 Appendix

8.1 Additional NMR theory

8.1.1 Frame of reference

It is common to describe NMR and the effect of pulses on the bulk magnetization vector from a rotating frame of reference rather than the laboratory frame of reference. If NMR is considered from a laboratory frame of reference, both B_1 and the bulk magnetization vector are moving simultaneously, which is difficult to consider visually. In the rotating frame of reference, the x , y , z coordinate system is rotating with the same frequency as the bulk magnetization vector and B_1 , *i.e.* the Larmor frequency.

In the rotating frame of reference, it is simpler to consider B_1 as the sum of two vectors moving opposite of each other, as illustrated in **Figure 8.1**. The first vector remains static since it is moving with the same frequency as the coordinate system. The other vector moves with twice the frequency and in the opposite direction of the rotating frame; it has no significant interaction with the bulk magnetization vector and is ignored. Thus, the system has been simplified to two stationary vectors, M and the first vector of B_1 , and is much easier to study [85 (p. 14-15), 86 (p. 53-55)]. In **Figure 8.2**, a schematic illustration of laboratory frame vs rotating frame is shown.

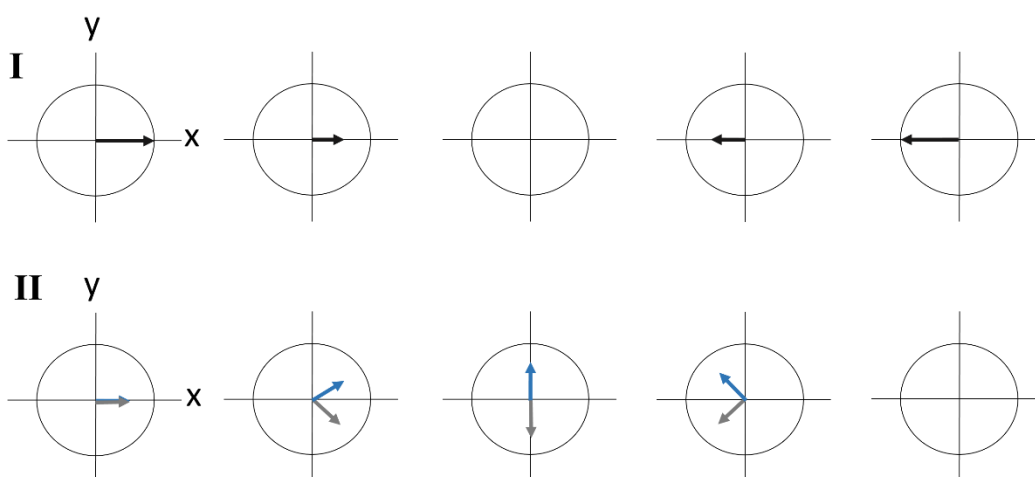


Figure 8.1: The applied magnetic field B_1 can be considered as one vector or the sum of two vectors. (I) B_1 oscillation can be illustrated as a single vector growing and shrinking along the x -axis. (II) B_1 is made up of two vectors moving in the opposite directions, with the sum of the blue and grey vector equaling the black vector in (I). Figure adapted from [86 (p. 53)]

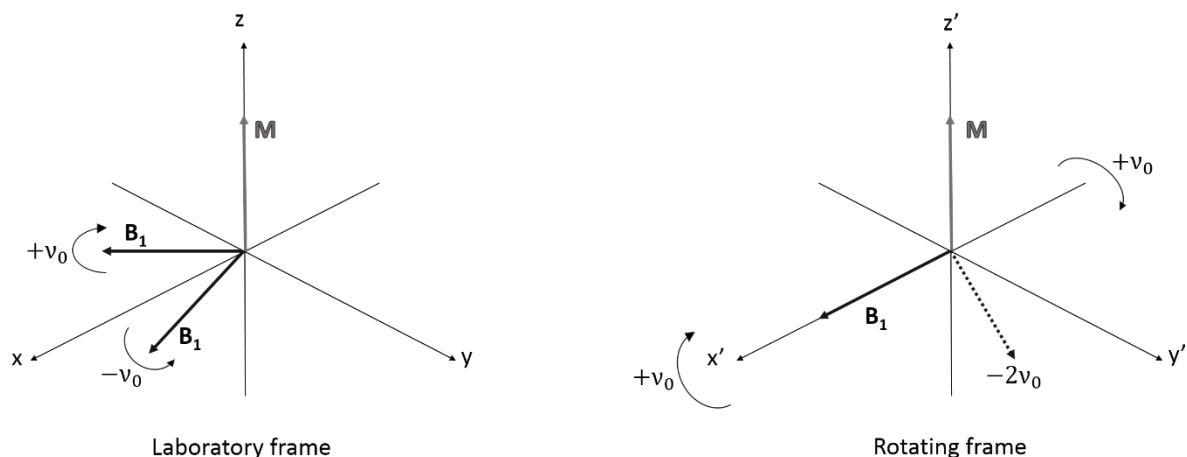


Figure 8.2: Laboratory frame of reference versus rotating frame of reference. When observing M and B_1 in the laboratory frame of reference, the coordinate system is stationary while both B_1 and M can move. This can be simplified by observing them in the rotating frame of reference, because B_1 and M become stationary while the coordinate system rotates at the Larmor frequency. B_1 is observed as the combination of two vectors, but only one of them is considered in the rotating frame of reference. While the first vector remains stationary, the other moves at twice the frequency, has no significant interaction with M and is therefore ignored. Figure adapted from [85 (p. 15)].

A common and illustrative analogy is to picture the laboratory frame of reference as observing the passengers on a merry-go-round from the ground. The merry-go-round is revolving and the passengers are moving up and down; their motion is a complex combination of both processes. In a rotating frame of reference, observation of the passengers would occur *on* the merry-go-round. Then the motion of the passengers is much easier to understand: they are simply moving up and down.

8.1.2 Pulse sequences: solvent suppression

In **Figure 8.3.I**, the simplest one-dimensional pulse sequence is illustrated. The pulse sequence for solvent suppression by presaturation starts with a long, weak pulse that continuously irradiates the water signal, as shown in **Figure 8.3.II**. Since the pulse is long, the frequency is exact. After 1-3 seconds, the sample is irradiated with a short 90° pulse to acquire the signals from the other protons in the sample. The suppression of the water signal can be improved by adding various additional pulses with short delays (0-10 ms). Because of the total length of each repetition of the pulse program (ca 2 seconds), the signals from exchangeable protons are usually lost [85 (p. 480-482)].

Zero excitation can be achieved with a variety of pulse sequences. The simplest one begins by moving the bulk magnetization vector into the xy -plane with a 90° pulse, as shown in **Figure 8.3.III**. As the various vectors fan out due to the differences in chemical shift of the nuclei, the solvent peak remains along the x -axis because it is on-resonance. A second 90° pulse, applied in the opposite direction, moves the solvent magnetization vector back to the z -axis. The other magnetization vectors do not perfectly align with the z -axis and the remaining x -component of their magnetization vectors enables detection [85 (p. 482-483), 197]. Improved pulse programs that remove more of the solvent signal have also been developed [85 (p. 482-483)].

Recall that PFGs are magnetic fields applied along the z -axis of the sample such that the field no longer is equal throughout the sample. Additionally, recall that the Larmor frequency of nuclei depend on the static magnetic field they are placed in. If a PFG is applied after the initial excitation pulse, the nuclei in the sample are no longer precessing in an equal field and will fan out at varying rates, becoming *dephased* or *defocused*. The magnetization vectors will cancel each other out and no NMR signals can be detected. If a PFG of equal magnitude but opposite direction is applied to the sample, the magnetization vectors will be refocused [85 (p. 191-195 and 473-485)].

In excitation sculpting, a variant of solvent suppression with PFGs, the solvent nuclei are dephased and made unmeasurable by the combined action of PFGs and selective 180° pulses. In the same set of pulses, the other magnetization vectors are *refocused* by hard (short) 180° pulses. The process can be repeated prior to acquisition for an increased suppression, but the second set of PFGs should have different strength to avoid refocusing any of the previously defocused elements. The name excitation sculpting was chosen in part because the solvent signal is “chipped” away for each repetition of the PFGs and selective 180° pulses [85 (p. 191-195 and 473-485), 198, 199]. In **Figure 8.3.IV**, an illustration of an example of an excitation sculpting pulse program is shown.

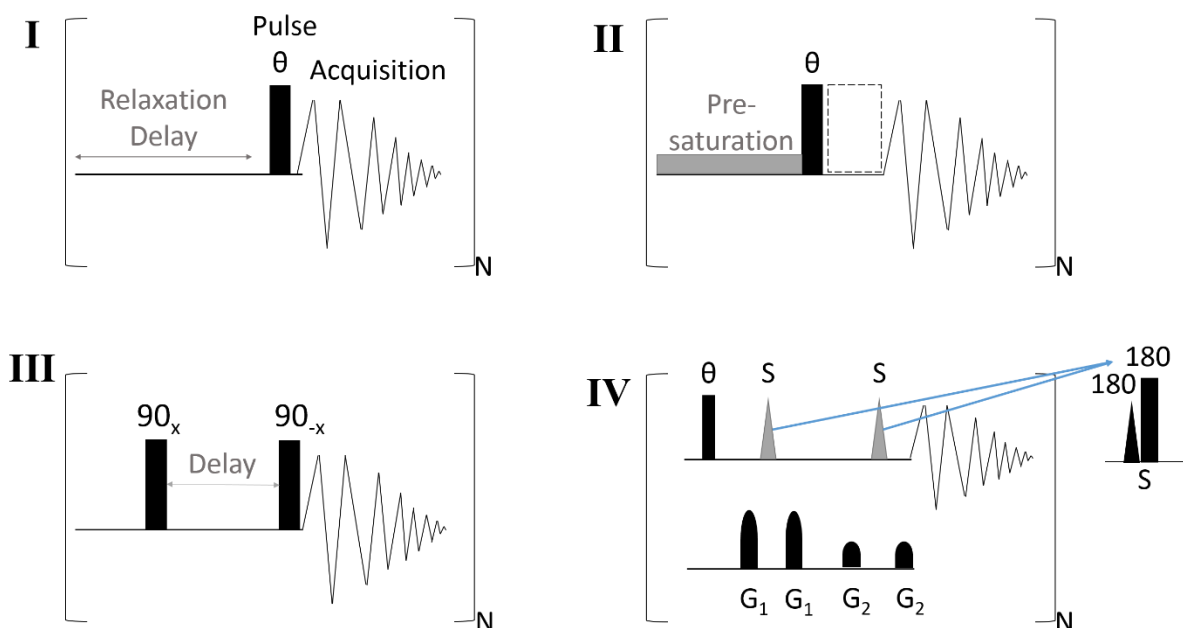


Figure 8.3: One-dimensional pulse sequences. (I) Simplest one-dimensional pulse sequence with relaxation time prior to a 90° excitation pulse. (II) Pulse program for solvent suppression by presaturation. The dashed rectangle illustrates where additional pulses would be added. (III) Simple pulse program for solvent suppression by zero excitation. (IV) Pulse program for solvent suppression by excitation sculpting. The selective pulses are shown as grey triangles, but each actually consists of a selective and a hard 180° pulse. The PFGs are the rounded pulses on the second line. Figures adapted from [85 (p. 134, 481, and 484)].

8.1.3 Pulse sequences: Two-dimensional NMR

All two-dimensional NMR pulse programs are based on the same general set up, schematically illustrated in **Figure 8.4.I**. The pulse sequence starts with a preparation pulse that excites the nuclei in the sample. It is followed by an evolution period (t_1) that is incrementally increased for each repetition of the pulse sequence. Then a mixing pulse is applied to ensure that the desired signals are observable. The sequence ends with acquisition of the FID, which gives rise to the signals in the first dimension of the spectrum. The second dimension is acquired indirectly. For each increase of the evolution period, the magnetization vectors can precess for longer and the resulting frequency spectrum reflects that. If all of the first dimension frequency spectra are placed in a row, the changes in intensity can be traced and a second FID will emerge, as shown in **Figure 8.4.II** [85 (p. 171-176)]. In NUS, a random selection of evolution periods (t_1) are carried out and acquired, the remaining data is calculated based on the few actual measurements made [85 (p. 185-186)].

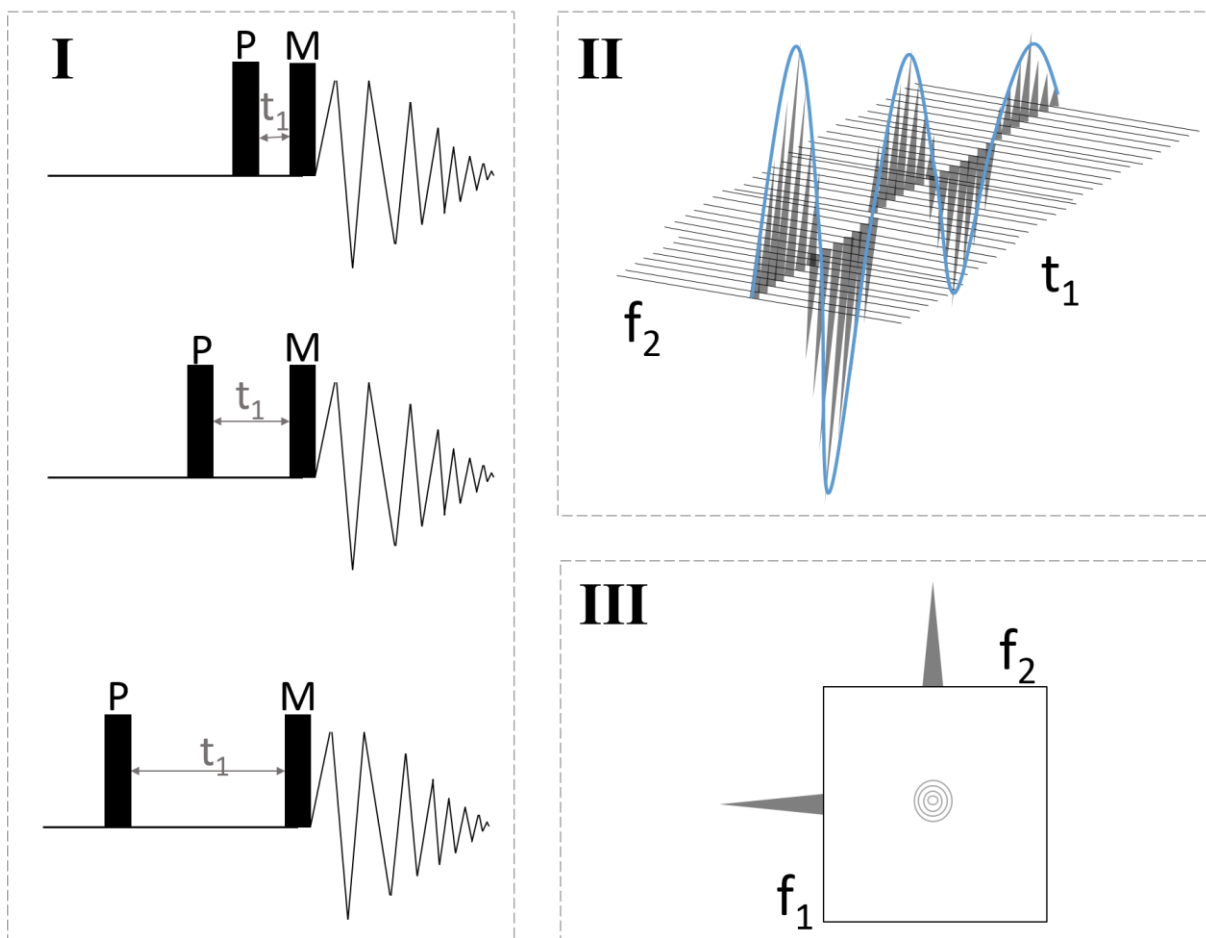


Figure 8.4: General theory of two-dimensional NMR. (I) The pulse sequence of two-dimensional NMR experiments are all based on the same general set up. A preparation pulse (P) excites all nuclei in the sample, followed by an evolution time (t_1) that is increased incrementally for each repetition of the pulse sequence. Then the mixing pulse (M) ensures that the desired magnetization is observable prior to detection. (II) The FID for the second dimension (t_1) is found by following the modulation of signal intensity of peaks in the first dimension (f_2). (III) The two-dimensional spectrum is usually displayed with one-dimensional spectra on its top and left side. The direct dimension is named f_2 because it is acquired after the indirect dimension (f_1). Figures adapted from [85 (p-171-176)].

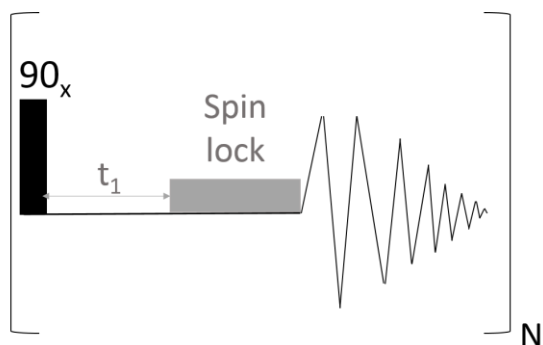


Figure 8.5: TOCSY pulse sequence. Figure adapted from [85 (p. 221)].

continuously refocused and no chemical shift evolution occurs. However, evolution due to spin-spin coupling is not affected by the spin lock and continues as if there were no pulse irradiating the sample [85 (p. 220-229)].

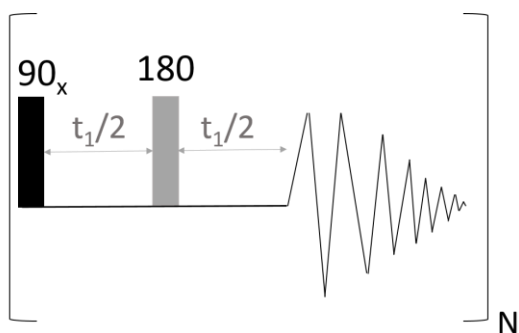


Figure 8.6: Simple JRES pulse sequence. Figure adapted from [85 (p. 301)].

The coupling patterns in the indirect dimension are not vertical; they have a 45° tilt. The tilted peaks can cause overlap and confusion during examination of the spectrum. Processing packages are usually available in the NMR software [85 (p. 301-313)].

Heteronuclear Single Quantum Coherence (HSQC) spectroscopy

The HSQC spectrum reveals which protons and carbons are directly coupled with each other. Unlike TOCSY, it only has cross peaks and no diagonal line [85 (p. 246-256)]. Because of the very limited use of HSQC, no detailed description of its pulse sequence is given.

The TOCSY pulse sequence, illustrated in **Figure 8.5**, consists of a 90° excitation pulse, the variable evolution time and a *spin lock* prior to detection. The spin lock is a continuous low-frequency pulse applied along the y-axis, *i.e.* where the magnetization vectors of the nuclei are first placed. While the spin lock is in place the magnetization vectors are

The simplest JRES pulse sequence begins with a 90° excitation pulse and is followed by a 180° pulse applied in the middle of the evolution time, as shown in **Figure 8.6**. The effect of the 180° pulse is similar to the spin lock in TOCSY: the chemical shifts are refocused but not the spin-spin couplings. After acquisition, the spectrum requires some processing because the

8.2 Description of chemotherapeutic agent candidates and a drug mixture for treatment of glioblastoma

CUDC-907

CUDC-907 is a dual inhibitor of class I and II histone deacetylases (HDAC) and class I phosphoinositide 3-kinases (PI3K). The molecule is a combination of the active groups known to inhibit HDAC and PI3K, held together by a linker [200]. The various parts are highlighted in **Figure 8.7**, in which the molecular structure is shown.

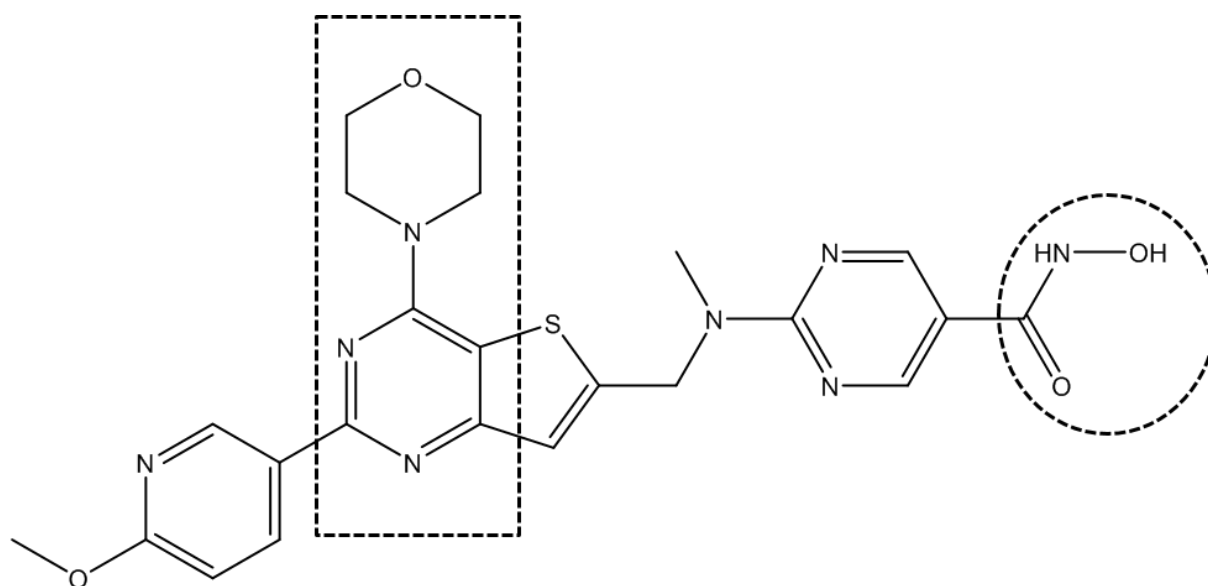


Figure 8.7: The molecular structure of CUDC-907. The HDAC inhibitor functionality site is encircled, while the PI3K inhibitor functionality site is enclosed in a rectangle, with the linker in between. The 2-methoxyphenyl connected to the PI3K inhibitor functionality had no specified function in the original article presenting CUDC-907, where it was simply shown as an R-group [200].

PI3Ks are involved in cell signaling pathways stimulating cell growth and survival [2], and have been found to undergo genetic alterations in glioblastoma [201]. Current chemotherapeutic agents targeting the cell signaling pathways involving PI3Ks often fail. It is speculated that cancer cells are capable of reducing treatment efficacy by reprogramming their cell signaling pathways [202]. A suggested approach to sensitize cancer cells to PI3K inhibitors is to block HDACs simultaneously [203]. HDACs are involved in deacetylation of histones, proteins that take part in packing DNA into tighter structures. The packing of DNA affects transcription, and HDACs *generally* decreases the cells ability to transcribe genes [2 (p. 196-197), 3 (p. 62-63)]. By inhibiting HDACs, cell differentiation and apoptosis can be induced [204]. The exact mechanisms involved are not clear. It has been shown that by inhibiting both

HDAC and PI3K simultaneously increases the efficacy of both treatments [200, 203]. CUDC-907 has been tested on patient-derived glioblastoma cells and was found to be potent [205].

FK866 (Daporinad)

FK866 inhibits synthesis of nicotinamide adenine dinucleotide (NAD⁺) [206], an important co-enzyme in several metabolic reactions occurring in cells [18 (p. 306-309)]. The depletion of NAD⁺ leads to apoptosis [206]. Specifically, FK866 is a noncompetitive inhibitor of the enzyme nicotinamide phosphoribosyl transferase (NMPRTase, also abbreviated to NAMPT), which catalyzes the conversion of nicotinamide and phosphoribosyl pyrophosphate to nicotinamide mononucleotide and pyrophosphate (P₂O₇⁻⁴), the rate-limiting step of NAD⁺ synthesis [207]. Several cancers have elevated expression of NMPRTase and are highly dependent of NAD⁺, including glioblastoma [208]. The molecular structure of FK866 is shown in **Figure 8.8**.

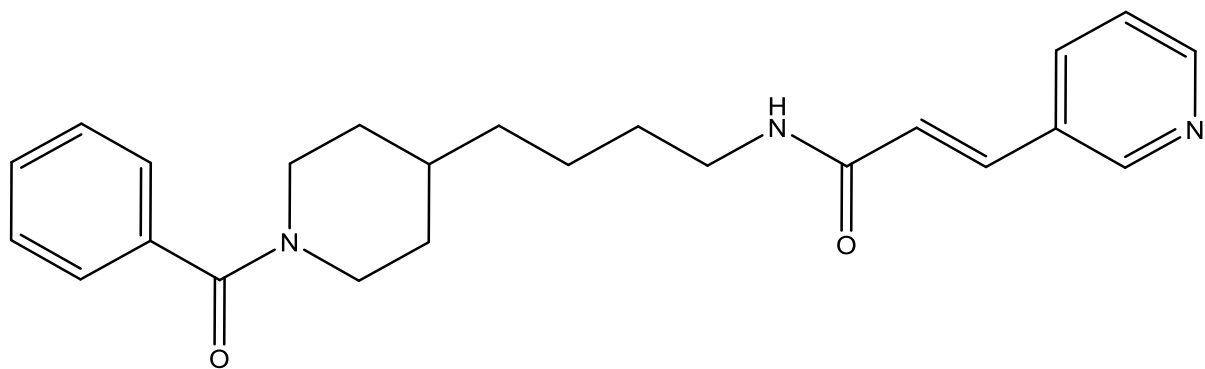


Figure 8.8: The molecular structure of the novel chemotherapeutic agent FK866.

CUSP9

CUSP9 (Coordinated Undermining of Survival Paths) started as a mixture of eight different drugs given in addition to temozolomide to patients suffering from recurrent glioblastoma. All of the drugs were already established and well tolerated in treatment of other diseases. They were chosen because they targeted vulnerable pathways in glioblastoma or were thought to give a general increase in overall survival. Five of the drugs were given with the goal of impairing or inhibiting growth factors and growth signaling pathways: aprepitant, artesunate, auranofin, disulfiram with copper gluconate, and nelfinavir. The three remaining drugs had published results of increasing overall survival: captopril, sertraline, and ketoconazole [209].

A year later, the authors published an article describing an updated mixture, CUSP9*. The main differences between the two mixtures were changing drugs that were no longer available

(ketoconazole to itraconazole, nelfinavir to ritonavir), no longer adding an exogenous copper source for disulfiram to chelate with as it was deemed unnecessary, and adding celecoxib because it also targeted survival signaling pathways in glioblastoma [210].

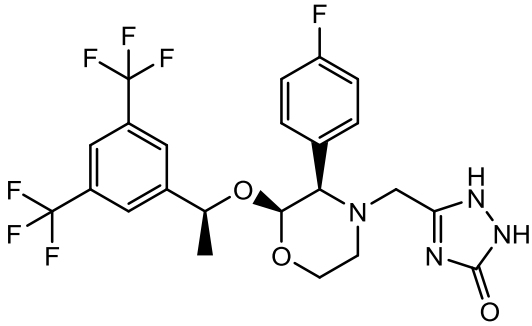
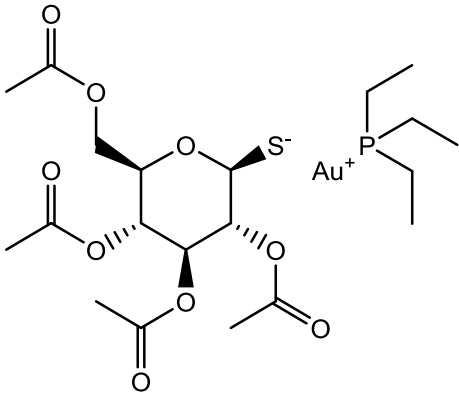
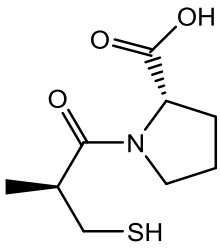
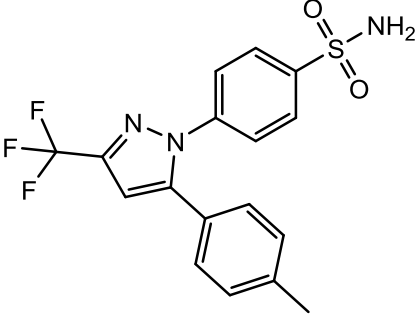
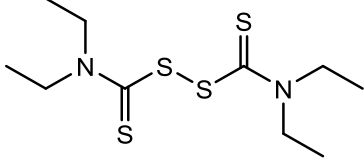
The authors of the first article carefully evaluated the possibility of interactions between the nine drugs and concluded that the only two with potential harmful interaction was artesunate and auranofin. However, side-effects did turn out to be an issue, even with CUSP9*, and further adjustments of dose and type of drugs has been made. The mixture used in this project (CUSP9v4) contains aprepitant, auranofin, captopril, celecoxib, disulfiram, itraconazole, minocycline, quetiapine, and sertraline, in addition to temozolomide. The expected benefit of the drugs and their molecular structures are given in **Table 8.1** and **Table 8.2** respectively. Temozolomide is not included because it is described in detail in **section 3.3**.

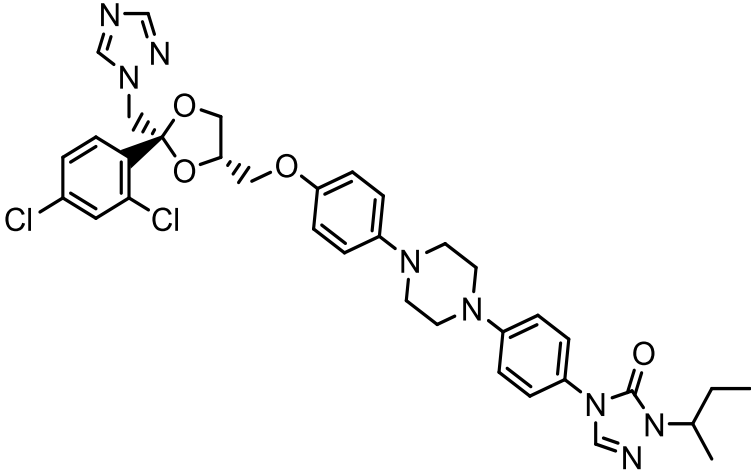
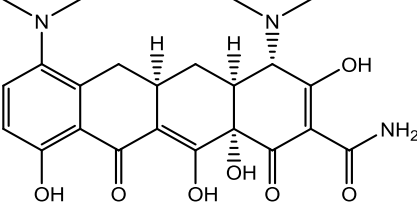
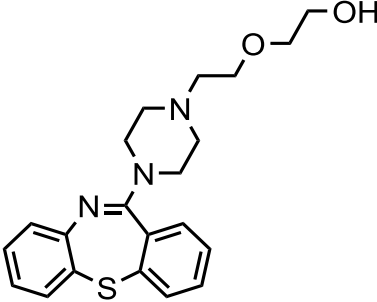
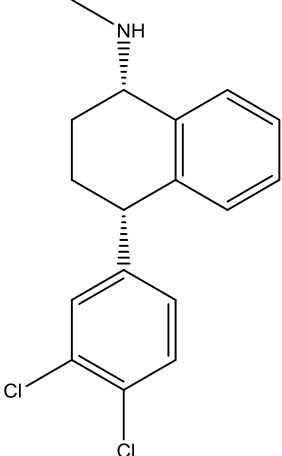
Table 8.1: The expected benefit of the drugs found in CUSP9v4, not including temozolomide.

Drug	Expected benefit	Source
Aprepitant	Commonly used to treat nausea. Inhibitor of neurokinin-1 receptor (NK1R). NK1R is involved in cell proliferation and inhibition of it has been suggested to cause glioblastoma cells to undergo apoptosis.	[210, 211]
Auranofin	Commonly used to treat rheumatoid arthritis. In glioblastoma, auranofin has three possible effects. First, it can cause increased generation of reactive oxygen species because of its ability to inhibit thioredoxin reductase. Second, it weakly inhibits cathepsin B, a protein that is upregulated and contributes to glioblastoma growth. Finally, auranofin can inhibit the protein 5-lipoxygenase (5-LO), yet another important enzyme in glioblastoma survival pathways.	[210]
Captopril	Commonly used to treat hypertension and congestive heart failure because of its ability to inhibit the angiotensin conversion enzyme (ACE). In glioblastoma, it has been shown to reduce the need of steroids, inhibit angiogenesis, and inhibit growth in cancer cell models.	[209, 210]
Celecoxib	Commonly used for treatment of pain. Celecoxib inhibits cyclooxygenase (COX). COX, together with 5-LO, are suggested to have roles in the tumorigenesis and proliferation of glioblastoma.	[210, 212]
Disulfiram	Commonly used to treat alcoholism. Disulfiram makes consumption of ethanol highly unpleasant by inhibiting aldehyde dehydrogenase, which leads to build-up of acetaldehyde.	[209, 210]

	Disulfiram has several possible anti-cancer capabilities, among them inhibition of MGMT (thus sensitizing glioblastoma to temozolomide). In addition, it can inhibit a drug exporter found in the brain (P-glycoprotein).	
Itraconazole	Commonly used as an anti-fungal agent. The anti-tumor effect of itraconazole includes anti-angiogenic abilities and inhibition of VEGF.	[210]
Minocycline	Commonly used as an antibiotic in treatment of acne vulgaris. In glioblastoma, it has been shown to have antiangiogenic effects in mouse models and is being considered as an adjuvant therapy to temozolomide.	[213, 214]
Quetiapine	Commonly used as an antipsychotic, <i>e.g.</i> for treatment of schizophrenia, bipolar disorder, major depressive disorder, and generalized anxiety. In glioblastoma, quetiapine has shown an ability to suppress cell growth, downregulate PI3K, and sensitize glioblastoma cells to temozolomide, among other things.	[215, 216]
Sertraline	Commonly used as an anti-depressant and for treatment of excessive anxiety. Sertraline is also used as an anti-depressant in CUSP9 because it has been shown to have a positive effect on glioblastoma patients. Additionally, the compound has shown to have some anti-cancer abilities. It inhibits translationally controlled tumor protein (TCTP), which in turn is an inhibitor of the protein p53. By reestablishing the pro-apoptotic signaling from p53, cancer cells can be sensitized to other cytotoxic compounds.	[210]

Table 8.2: The molecular structure of the drugs found in the mixture CUSP9v4, not including temozolomide.

Drug	Molecular structure
Aprepitant	 <p>The structure of Aprepitant consists of a piperazine ring system. One nitrogen of the piperazine is substituted with a 4-fluorophenyl group. The other nitrogen is substituted with a propylidene hydantoin group. The 2-position of the piperazine ring is linked via an oxygen atom to a chiral center (marked with a wedge bond) that is also bonded to a methyl group (marked with a dash bond) and a 3,5-bis(trifluoromethyl)phenyl group.</p>
Auranofin	 <p>The structure of Auranofin features a central gold atom (Au⁺) coordinated to a triethylphosphine ligand (P(Et)₃) and a thiolate group (S⁻) from a thiozinc sugar derivative. The sugar moiety is a six-membered ring with four acetoxy groups attached at various positions and a methyl group at the 2-position.</p>
Captopril	 <p>The structure of Captopril is a proline derivative. It features a proline ring with a methyl group at the 2-position and a propylsulfhydryl group at the 3-position. A carboxylic acid group is attached to the 4-position of the proline ring.</p>
Celecoxib	 <p>The structure of Celecoxib is a pyrazole derivative. It features a pyrazole ring substituted with a trifluoromethyl group at the 3-position, a 4-methylphenyl group at the 5-position, and a 4-sulfamoylphenyl group at the 4-position.</p>
Disulfiram	 <p>The structure of Disulfiram is a dithiocarbamate derivative. It consists of two diethylthiocarbamate groups linked together by a disulfide bridge (S-S) between their sulfur atoms.</p>

Itraconazole	 <p>The structure of Itraconazole features a spirocyclic core consisting of a benzimidazole ring fused to a cyclohexane ring. The benzimidazole ring has two chlorine atoms at the 2 and 6 positions. The cyclohexane ring is substituted with a propyl chain at the 4-position and a 1,2,4-triazole ring at the 5-position. The triazole ring is further substituted with an isopropyl group. The cyclohexane ring is also linked via an ether bridge to a piperazine ring, which is in turn connected to a para-substituted benzene ring. This benzene ring is further substituted with a 1,2,4-triazole ring, which is finally substituted with an isopropyl group.</p>
Minocycline	 <p>The structure of Minocycline is a tetracycline derivative. It consists of a tetracyclic core with a dimethylamino group at the 7-position, a hydroxyl group at the 8-position, and a dimethylamino group at the 12-position. The 12-position also has a hydroxyl group and an amino group (NH₂) at the 13-position. The 14-position has a hydroxyl group, and the 15-position has a hydroxyl group and a carbonyl group.</p>
Quetiapine	 <p>The structure of Quetiapine features a benzothiazine core. The benzothiazine ring is substituted with a piperazine ring at the 4-position. The piperazine ring is further substituted with a propyl chain at the 3-position, which is terminated by a hydroxyl group (OH).</p>
Sertraline	 <p>The structure of Sertraline consists of a tetrahydroindole ring system. The indole ring is substituted with a dimethylamino group at the 2-position and a 3,4-dichlorophenyl group at the 3-position. The tetrahydroindole ring is further substituted with a methyl group at the 1-position and a hydroxyl group at the 4-position.</p>

8.3 Experimental details: Cell culturing

MD and PhD candidate Erlend Skaga cultured the cell samples received during fall 2016, and fellow Marit Christensen cultured samples received from December 2016 to April 2017. Lab technician Maria Ewa Walewska cultured the samples received from May to November 2017. All three carried out their work at the facilities at the Vilhelm Magnus Laboratory for Neurosurgical Research, Institute of Clinical Medicine, University of Oslo (UiO).

8.3.1 Chemicals

The following chemicals were used during cell culturing: trypsin-EDTA (ethylenediaminetetraacetic acid) from Invitrogen (Carlsbad, CA, USA), owned by Thermo Fisher; human albumin from Octapharma Pharmazeutika Produktiones (Lachen, Switzerland); L-15 cell culture medium from Lonza (Basel, Switzerland); serum free Dulbecco's Modified Eagle's Medium (DMEM/F-12 GlutaMAX™) from Gibco/Thermo Fisher Scientific; 10 ng/mL basic fibroblast growth factor (bFGF) and 20 ng/mL epidermal growth factor (EGF), both from R&D Systems (Minneapolis, MN, USA); penicillin/streptomycin 100 U/mL from Lonza; heparin 1 ng/mL from Leo Pharma (Ballerup, Denmark); 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) 8 mM from Lonza; 1:50 B27-supplement without retinoic acid from Gibco/Thermo Fisher Scientific.

The following drugs (in addition to DMSO, temozolomide, and YM155 listed in **section 5.3**) were used to treat the cultured cells: CUDC-907 (ID S2759) from Selleck Chemicals, daporinad (also known as FK866, ID Axon 1546) from Axon MedChem (Reston, VA, USA), and CUSP9 which contains auranofin (ID sc-202476) from Santa Cruz Biotechnology (Dallas, TX, USA), temozolomide, and aprepitant (ID S1189), captopril (ID S2051), celecoxib (ID S1261), disulfiram (ID S1680), itraconazole (ID S2476), minocycline (ID S4226), quetiapine fumarate (ID S1763), sertraline (ID S4052) all from Selleck Chemicals.

8.3.2 Method: cell culturing

Skaga wrote the following description of how the samples were obtained and the cells cultured.

Technical information: Spinning procedure: 300xg in 5 minutes.

Glioblastoma biopsies were obtained from seven informed and consenting patients undergoing surgery for glioblastoma at Oslo University Hospital, Norway, approved by The Norwegian Regional Committee for Medical Research Ethics. Histopathological diagnostics were performed according to the WHO classification. Cell cultures were established both from the tumor biopsy and ultrasonic aspirate generated during surgery. Single cells were isolated mechanically and enzymatically with trypsin-EDTA, blocked by 2 mg/mL human albumin and washed in L-15 cell culture medium. Cells were cultured under sphere forming conditions containing serum free Dulbecco's Modified Eagle's Medium (DMEM), 10 ng/mL bFGF, 20 ng/mL EGF, penicillin/streptomycin 100 U/mL, heparin 1 ng/mL (Leo Pharma), HEPES 8 mM and 1:50 B27-supplement without retinoic acid.

Cell treatment and sample preparations

Cells were plated at 300,000 cells/mL under sphere-forming conditions. YM155 (final concentration 1250 nM), temozolomide (final concentration 250 μ M) and negative control (0.5% DMSO) were added in three different flasks, respectively. Following incubation for 24 h cells were washed two times in Dulbecco's Phosphate Buffered Saline (DPBS) and snap frozen in liquid nitrogen before storage in -80 °C. The samples were transported to the NMR facility with dry ice in a Styrofoam box with a lid.

8.3.3 Received cell samples for preliminary nuclear magnetic resonance spectroscopy analyses

The samples received fall 2016 and January 2017 were used for preliminary NMR analyses and testing of 3 ID mm NMR tubes, respectively. In **Table 8.3** an overview of the samples received from Skaga are shown, while in **Table 8.4** an overview of Christensen's samples are given. The remaining samples received are listed in **Table 4.1**.

Table 8.3: Overview of samples analyzed in the preliminary NMR analysis fall 2016. For each sample the sample number, cell line it originated from, the approximate number of cells in the sample, treatment and identifier are listed. Samples with treatment cell medium and 0.5% DMSO were control samples for the samples treated with temozolomide. The samples only treated with cell medium were control samples for the samples treated with YM155 or daporinad. The sample treated with cell medium and 0.1% DMSO was the control sample for the sample treated with CUDC-907. Finally, the samples treated with cell medium, 0.1% DMSO and 20 μ M Cu(ii)Cl were control samples for samples treated with CUSP9. Skaga cultured the cells.

Sample number	Cell line	Number of cells (millions)	Treatment	Identifier
1	T1456	3	Temozolomide	pT1
2	T1456	3	Cell medium + 0.5% DMSO	pD2
3	T1456	3	YM155	pY3
4	T1456	3	Only cell medium	pM4
5	T1456	3	Cell medium + 0.1% DMSO	pD5
6	T1456	3	CUDC-907	pC6
7	T1456	3	CUSP9	pCU7
8	T1456	3	Cell medium + 0.1% DMSO + 20 μ M Cu(ii)Cl	pDCu8
9	T1502	2.5	Temozolomide	pT9
10	T1502	2.5	Cell medium + 0.5% DMSO	pD10
11	T0965	1.5	Only cell medium	pM11
12	T0965	1.5	YM155	pY12
13	T0965	3	Cell medium + 0.5% DMSO	pD13
14	T0965	3	Temozolomide	pT14
15	T1459	2	Temozolomide	pT15
16	T1459	2	Cell medium + 0.5% DMSO	pD16
17	T1459	2	YM155	pY17
18	T1459	2	Only cell medium	pM18
19	T1454	3	Temozolomide	pT19
20	T1454	3	Cell medium + 0.5% DMSO	pD20
21	T1454	3	CUSP9	pCU21
22	T1454	3	Cell medium + 0.1% DMSO + 20 μ M Cu(ii)Cl	pDCu22
23	T1454	3	YM155	pY23
24	T1454	3	Only cell medium	pM24
25	T1459	3	Daporinad	pDa25
26	T1459	3	Daporinad	pDa26
27	T1459	4	Daporinad	pDa27
28	T1459	3	Cell medium	pM28
29	T1459	3	Cell medium	pM29
30	T1459	4	Cell medium	pM30

Table 8.4: Overview of samples analyzed when 3 mm NMR tubes were tested, January 2017. For each sample the sample number, cell line it originated from, treatment and identifier are listed. Three different cell lines were analyzed, each with one control group and two different groups treated with either temozolomide or YM155. All the samples contained approximately 3 million cells. Samples treated with DMSO were considered control samples, and are highlighted with gray. Christensen cultured the cells.

Sample Number	Cell line	Treatment	Date sample collection (day/month/year)	Identifier
1	T1456	Temozolomide	23/12/2016	T1
2	T1456	Temozolomide	23/12/2016	T2
3	T1456	YM155	23/12/2016	Y3
4	T1456	YM155	23/12/2016	Y4
5	T1456	DMSO	23/12/2016	D5
6	T1456	DMSO	23/12/2016	D6
7	T1456	Temozolomide	06/01/2016	T7
8	T1456	YM155	06/01/2016	Y8
9	T1456	DMSO	06/01/2016	D9
10	T1459	DMSO	23/12/2016	D10
11	T1459	DMSO	23/12/2016	D11
12	T1459	DMSO	23/12/2016	D12
13	T1454	Temozolomide	23/12/2016	T13
14	T1454	Temozolomide	23/12/2016	T14
15	T1454	YM155	23/12/2016	Y15
16	T1454	YM155	23/12/2016	Y16
17	T1454	DMSO	23/12/2016	D17
18	T1454	DMSO	23/12/2016	D18

8.4 Experimental details: Preliminary studies in nuclear magnetic resonance spectroscopy

8.4.1 Instrumentation

Preliminary NMR analyses were done with 5 mm OD Boroeco-5-7 tubes from Deutero. The tubes were placed in POM standard bore spinners from Bruker. NMR tubes with 3 outer mm diameter, model 341-PP-7 from Wilmad-LabGlass, were assessed as well. The tubes had 5X3INS-B Optimizer Inserts™ from Norell (Morganton, NC, USA) in addition, to fit into the standard bore POM spinners.

Centrifugation was done with A 5415 R Eppendorf centrifuge.

The following NMR instrumentation from Bruker was used: an AVI600 (600 MHz) NMR instrument with a 5 mm Triple Resonance (TCI) cryoprobe and Bruker Automated Control

System (B-ACS) 60-sample changer was used. The 600 MHz instrument was run with the program TopSpin versions 2.1 patch level 6 (2.1pl6).

8.4.2 Chemicals

The same chemicals as listed in **section 4.2.1** were used, in addition to acetone (GPR RECTAPUR®, $\geq 99.5\%$) from VWR.

8.4.3 Solutions

The NMR buffer described in **section 4.4.1** was used.

8.4.4 Methods

NMR sample preparation

The Boroeco-5-7 NMR tubes, which were used for cancer samples from September to December 2016, were rinsed manually once with acetone and three times with type 1 water. The 3 mm NMR tubes were washed with the Multi-Tube Jet Washer/Dryer, as described in **Section 4.5.1**.

The cell samples were prepared as described in **section 4.5.1** in the preliminary NMR study. For the test of 3 mm NMR tubes, only 220 μL of type 1 water was added before cell lysis by ultrasound, and only 180 μL of sample and 20 μL of buffer was transferred to the NMR tubes.

Setup of preliminary NMR analyses

Probe temperature was 299.7 K and gas flow was 800 L/hour, which ensured the samples experienced a temperature of 25 °C during measurements. Once in the magnet, the sample was locked to the deuterium signal of D_2O , using Bruker's lock made for 90% H_2O and 10% D_2O mixtures. After locking, the sample was tuned and matched using the command *atma* (automatic tuning and matching). The samples were shimmed in the same way as described in **section 4.5.2**.

The 3 mm tubes were set up in as described in section 4.5.2. Only one-dimensional spectra with excitation sculpting as solvent suppression was carried out, also as described in **section 4.5.2**.

One-dimensional NMR analyses

Two different one-dimensional spectra were acquired: proton spectrum with NOESY presaturation solvent suppression (pulse program *noesygprr1d.comp*) and proton spectrum with excitation sculpting solvent suppression (pulse program *zgesgp.dp*). The code for each pulse program is given in **section 8.12**. Details concerning acquisition parameters are given in **Table 8.5** and details concerning processing parameters are given in **Table 8.6**.

Table 8.5: The general acquisition parameters for all samples analyzed with the pulse programs *noesygprr1d.comp* and *zgesgp.dp*. Sec is an abbreviation of seconds and μsec is an abbreviation of microsecond (1×10^{-6} second).

Name of acquisition parameter	Abbreviation	Settings (NOESY presaturation)	Settings (Excitation sculpting)
Size of FID	TD	65536	65536
Number of dummy scans	DS	4	2
Number of scans	NS	32	128
Spectral width	SW	20.6225 ppm	16.0221 ppm
Acquisition time	AQ	2.6476543 sec	3.4078720 sec
Receiver gain	RG	512	256 or 512
Dwell time	DW	40.400 μsec	52.000
Probe temperature	TE	302.5 K	300.0
Transmitter frequency offset	o1	2816.20 ppm	2816.20 ppm

Table 8.6: The general processing parameters for all samples analyzed with the pulse programs *noesygprr1d.comp* and *zgesgp.dp*.

Name of processing parameter	Abbreviation	Settings (NOESY presaturation)	Settings (Excitation sculpting)
Size of real spectrum	SI	65536	65536
Spectrometer frequency	SF	600.13 MHz	600.13 MHz
Line broadening	LB	0.30 Hz	1.00 Hz
Phasing mode	PH_mod	pk	No
Linear prediction for Fourier transformation	ME_mod	No	No
Number of linear prediction coefficients	NCOEF	0	0
Number of output points for linear prediction	LPBIN	0	0

Two-dimensional NMR analyses

The JRES experiments were acquired with the pulse program *jresgppraf*, the TOCSY experiments were acquired with the pulse program *dipsi2esgpph*, and the HSQC experiment was acquired with pulse program *hsqcdietgpsisp.2*. Details concerning acquisition parameters are given in **Table 8.7** and details concerning processing parameters for TOCSY are given in **Table 8.8**. The JRES spectra all had the size of the real spectrum increased to 16,384 (twice the size of F2's TD) and were calibrated by placing TSP at 0.00 ppm in all spectra. The HSQC spectrum was not further processed once acquired.

The code for each pulse program is given in **section 8.12**.

Table 8.7: The general acquisition parameters for all samples analyzed with the pulse programs *jresgppraf*, *dipsi2esgpph*, and *hsqcdietgpsisp.2*. F1 and F2 denote the frequency axes. Sec is an abbreviation of seconds and μsec is an abbreviation of microsecond (1×10^{-6} second).

Name of acquisition parameter	Abbreviation	Settings (JRES)	Settings (TOCSY)	Settings (HSQC)
nD acquisition mode	FnTYPE	Traditional(planes)	Traditional(planes)	Traditional (planes)
Size of FID	TD	F2: 8192 F1: 40	F2: 2048 F1: 512	F2: 1024 F1: 512
Number of dummy scans	DS	F2: 16	F2: 32	F2: 64
Number of scans	NS	F2: 128	F2: 8	F2: 16
Spectral width	SW	F2: 16.6630 ppm F1: 0.1320 ppm	F2: 10.2102 ppm F1: 10.2016 ppm	F2: 10.0138 ppm F1: 160.0000 ppm
Acquisition time	AQ	F2: 0.40960 sec F1: 0.25600 sec	F2: 0.1671168 sec F1: 0.0418143 sec	F2: 0.0851968 sec F1: 0.0106020 sec
Increment for delay	IN_F	F1: 12800.00 μsec	F1: 163.34 μsec	F1: 41.41 μsec
Receiver gain	RG	512	256	11585.2
Dwell time	DW	50.000 μsec	81.600 μsec	83.200 μsec
Probe temperature	TE	3002.5 K	300.0 K	300.0 K

Table 8.8: The general processing parameters for all samples analyzed with the pulse program *dipsi2esgpph*.

Name of processing parameter	Abbreviation	Settings
Size of real spectrum	SI	F2: 2048 F1: 1024
Spectrometer frequency	SF	F2: 600.13 MHz F1: 600.13 MHz
Window function	WDW	F2: QSINE F1: QSINE
Line broadening	LB	F2: 1.00 Hz F1: 0.30 Hz
Phasing mode	PH_mod	F2: pk F1: pk
Linear prediction for Fourier transformation	ME_mod	F2: None F1: LPfc
Number of linear prediction coefficients	NCOEF	F2: 0 F1: 32
Number of output points for linear prediction	LPBIN	F2: 0 F1: 2048

Statistical analysis: pre-processing

The one-dimensional spectra were phased and calibrated manually in TopSpin, before being transformed to ASCII files (.txt) with the command *convbin2asc*.

The NMR data were pre-processed as described in **section 4.5.5**, except normalization was to the integral of the total spectrum (not including the TSP peak), not to the integral of the TSP peak. Pareto was used as the scaling method; each variable was divided by its squared standard deviation. The spectral baselines were corrected (non-polynomial smoothing) prior to multivariate statistical analyses, using an algorithm from the *FTICRMS* R package [143].

PCA with R

PCA was carried out the same way as described **section 4.5.5**. The script used for the PCA of the preliminary data is given in **section 8.11.1**.

8.5 Results and discussion: preliminary studies

8.5.1 Exploration and selection of nuclear magnetic resonance spectroscopy experiments

The goal of the preliminary studies was to decide what kind of solvent suppression method best suited the samples, which type of NMR experiments to continue with, and perform a preliminary statistical analysis of data to hone in on which chemotherapies would be the focus for the remainder of the project.

Assessment of solvent suppression methods

The cell samples used in the preliminary NMR analysis contained 90% water ($^1\text{H}_2\text{O}$). The cell lysate was not diluted in pure deuterated water (D_2O) because it would result in loss of signal from exchangeable protons on acids and bases. Since biological samples contain a variety of acids and bases, valuable information could have been lost by diluting the lysate with D_2O only. However, without the use of solvent suppression during the acquisition of spectra, the overpowering water ($^1\text{H}_2\text{O}$) signal would disguise most of the other signals coming from the sample. [85 (p. 480-486)]. In **Figure 8.9**, a comparison of a sample before and after solvent suppression is given. It is clear that collection of useful data was dependent on acquisition of spectra with successful suppression of the solvent signal.

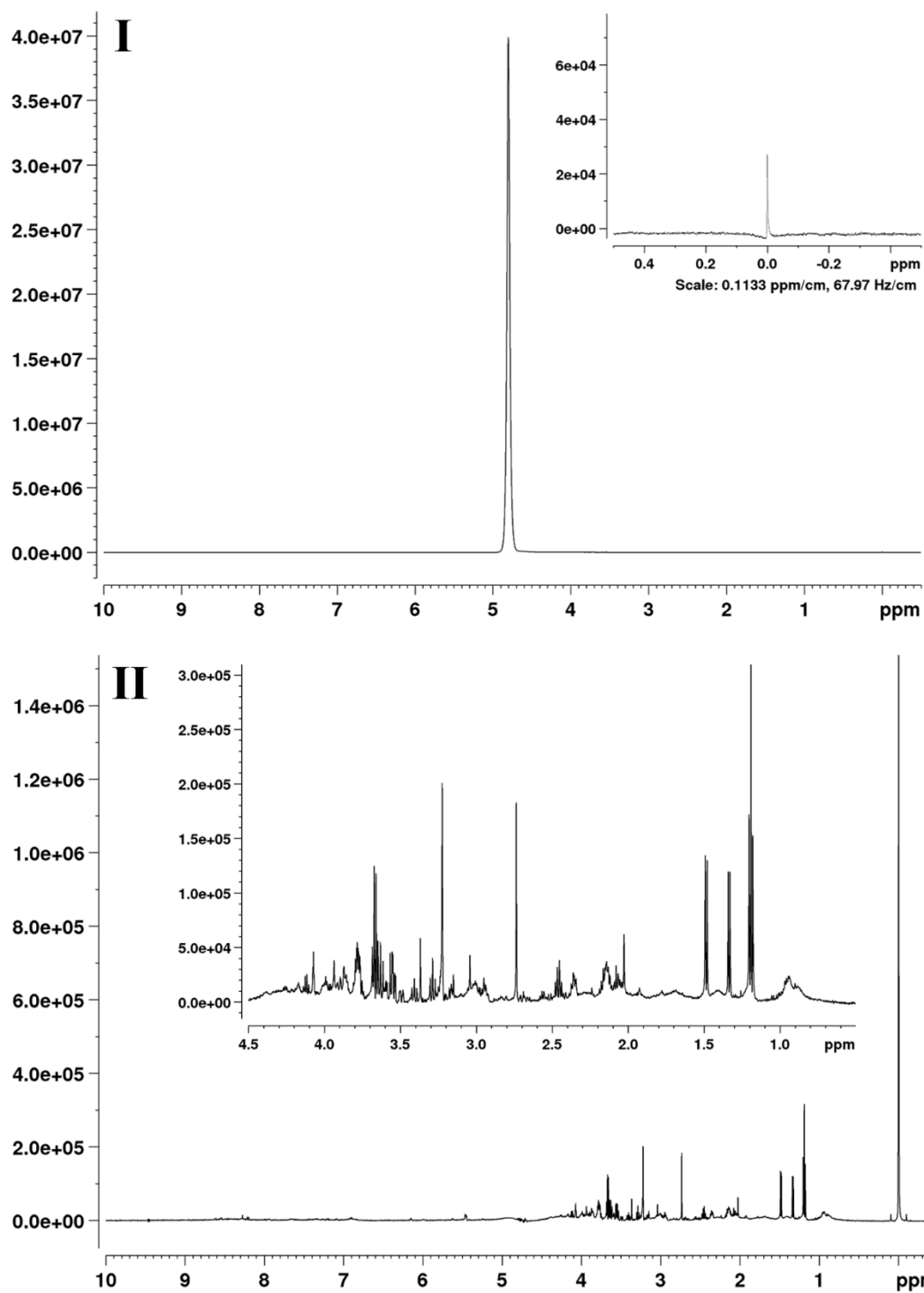


FIGURE 8.9: Difference between acquiring one-dimensional NMR spectra without (I) and with (II) solvent suppression. The same sample was used to acquire both spectra. In (I), NS = 16 and DS = 2. The highlighted area of the spectrum is the TSP peak. In (II), NS = 128 and DS = 2. The solvent was suppressed with excitation sculpting.

In **section 3.5.3**, three different methods for suppression of solvent signals were described: presaturation, zero excitation and excitation sculpting. Of these three, presaturation and excitation sculpting were chosen to be examined. The presaturation pulse sequence included pulses usually found in Nuclear Overhauser effect spectroscopy (NOESY) sequences. The additional pulses are added to further suppress the solvent signal. NOESY presaturation was chosen because it is a simple and robust technique, while excitation sculpting was chosen because it has fewer issues with baseline distortions and has improved phase properties [85 (p. 480-486)]. Of the two, NOESY presaturation is the most common to use, as shown in recent metabolomics studies of urine [217], and serum [218, 219]. Excitation sculpting usually has more efficient solvent suppression, and can be employed if deemed necessary [89].

A representative spectrum of solvent suppression with NOESY presaturation is shown in **Figure 8.10** and a representative spectrum with excitation sculpting is shown in **Figure 8.11**. It is clear from the spectra shown in the two following figures that excitation sculpting was the better solvent suppression method because the residual water signal was three factors smaller than in the NOESY presaturation spectrum.

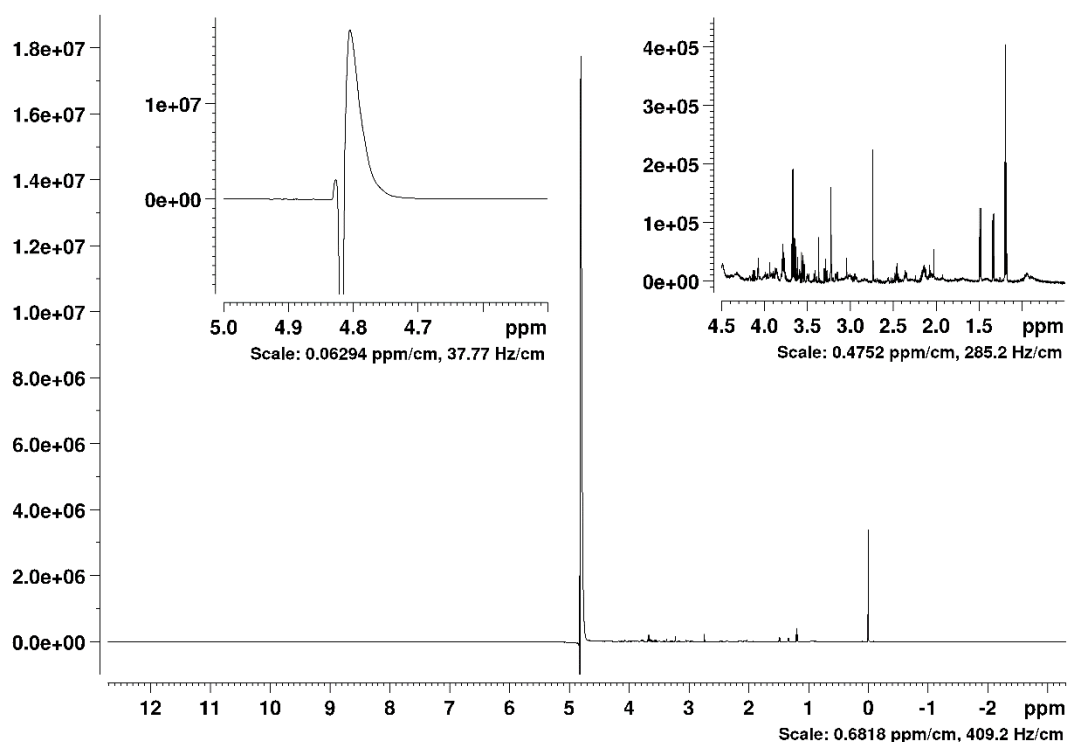


Figure 8.10: Solvent suppression by NOESY presaturation. Notice the large, residual water peak at 4.5 to 5 ppm and the increased intensity of the constituent molecules of the sample. NS = 32 and DS = 2.

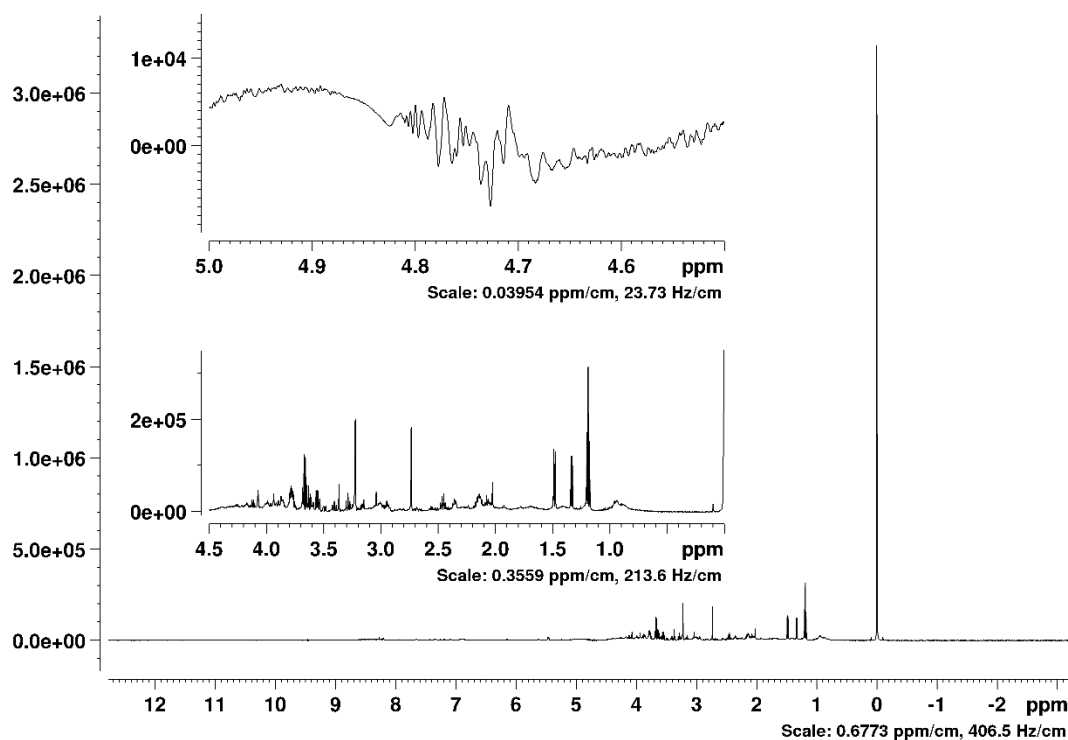


Figure 8.11: Solvent suppression by excitation sculpting. Notice how much smaller the residual water peak at 4.5 to 5 ppm is. Furthermore, the constituent molecules of the sample show increased signal intensity. NS = 128 and DS = 2.

Thus, in summary, further work with one-dimensional experiments was carried out with excitation sculpting as the solvent suppression method.

Assessment of two-dimensional spectra

In section 3.5.4, the advantages of two-dimensional spectra for metabolite identification was highlighted. Briefly, two-dimensional spectra can provide information on connectivity between nuclei and coupling patterns otherwise overlapped in the one-dimensional spectra. In the preliminary study, coupling between ^{13}C and ^1H was attempted examined by acquiring HSQC spectra. The additional information about carbon would be a great aid in identification of the constituent molecules [89].

The two-dimensional spectra had to be possible to acquire in a reasonable amount of time and with satisfactory signal intensity. Considering several dozens of samples needed to be measured, a reasonable amount of time was set as maximum 8 hours total per sample for all spectral measurements. Within that time, a satisfactory signal intensity was considered detection of peaks (signal-to-noise ratio > 3) from approximately zero to eight ppm.

The low natural abundance of ^{13}C (1.1%) and small magnetogyric ratio (γ) severely impacts the sensitivity of acquisition, but it can be alleviated to some degree. Three common approaches are to acquire several scans, increasing the sample concentration [85 (p. 97-103)], or enriching the sample with ^{13}C [170, 220]. There are drawbacks with all three solutions: additional scans mean longer experiments, growing many cells is time-consuming and often technically difficult, and ^{13}C enrichment is prohibitively expensive. Thus, if ^{13}C NMR spectra could not be acquired with the samples on hand (three million cultured cells per sample, unenriched) within a reasonable amount of time (< 8 hours), it would be not be carried out.

Three different two-dimensional NMR experiments were assessed: ^1H JRES, ^1H TOCSY, and HSQC, as they are the most commonly used in metabolomics [89]. Representative spectra are shown in **Figures 8.12** (JRES), **8.13** (TOCSY), and **8.14** (HSQC). Both JRES and TOCSY filled the criteria set: adding overall useful additional information within a practical time frame and with satisfactory signal intensity.

The JRES spectrum untangled the coupling patterns of the various peaks in the one-dimensional proton spectrum. Many of the peaks were overlapping and difficult to discern properly in the one-dimensional spectrum alone. The experimental time on the AVI600 was 2 hours and 25 minutes, but this was lowered to 1 hour and 4 minutes on the AVIIIHD800 because of intrinsic higher sensitivity and use of NUS (25% coverage).

The TOCSY spectrum gave valuable information on the coupling of the various peaks in the one-dimensional proton spectrum. By acquiring TOCSY spectra, it was possible to know which peaks were derived from the same molecule – knowledge that would be exceedingly difficult to unravel from the one-dimensional spectrum only. Experimental time was 2 hours and 4 minutes on the AVI600, but it was increased to 4 hours on the AVIIIHD800, in addition to using NUS (25% coverage), to maximize the amount of information gathered.

The HSQC experiment was not feasible to do. Even with 4 hours experiment time, few to no peaks were seen. It was briefly considered to do HSQC on one replicate sample per cell line and treatment scheme. However, this could have possibly added weeks of measuring time, even on the AVIIIHD800 instrument, and it was decided that the JRES and TOCSY experiments would suffice.

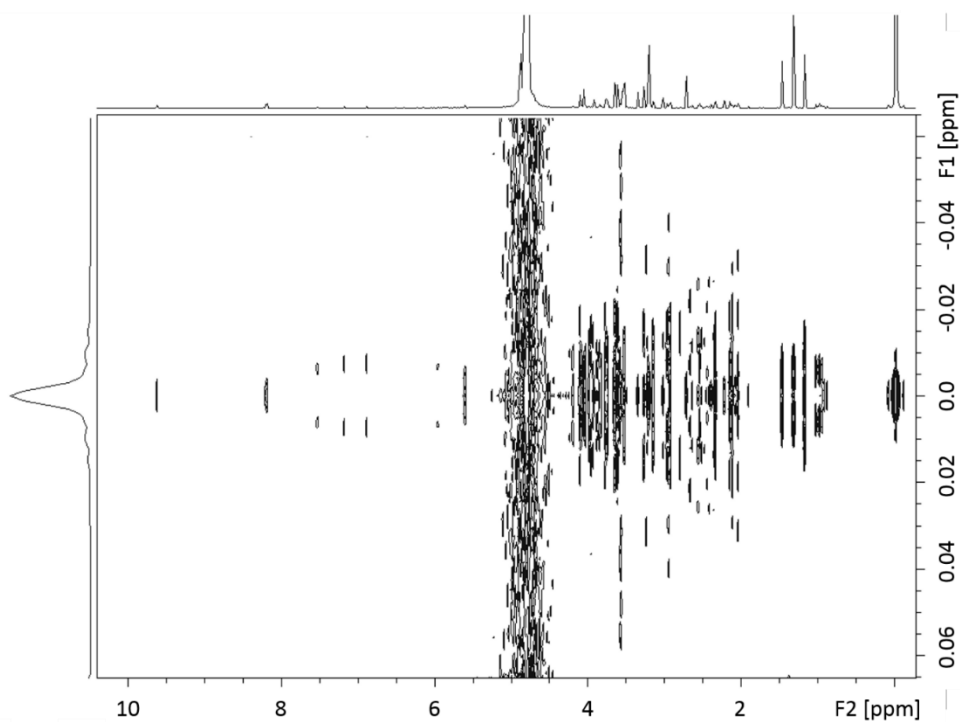


Figure 8.12: Representative JRES spectrum, acquired on the AVI600 instrument. NS = 128 and number of t_1 increments was 40. Experiment time was 2 hours and 25 minutes.

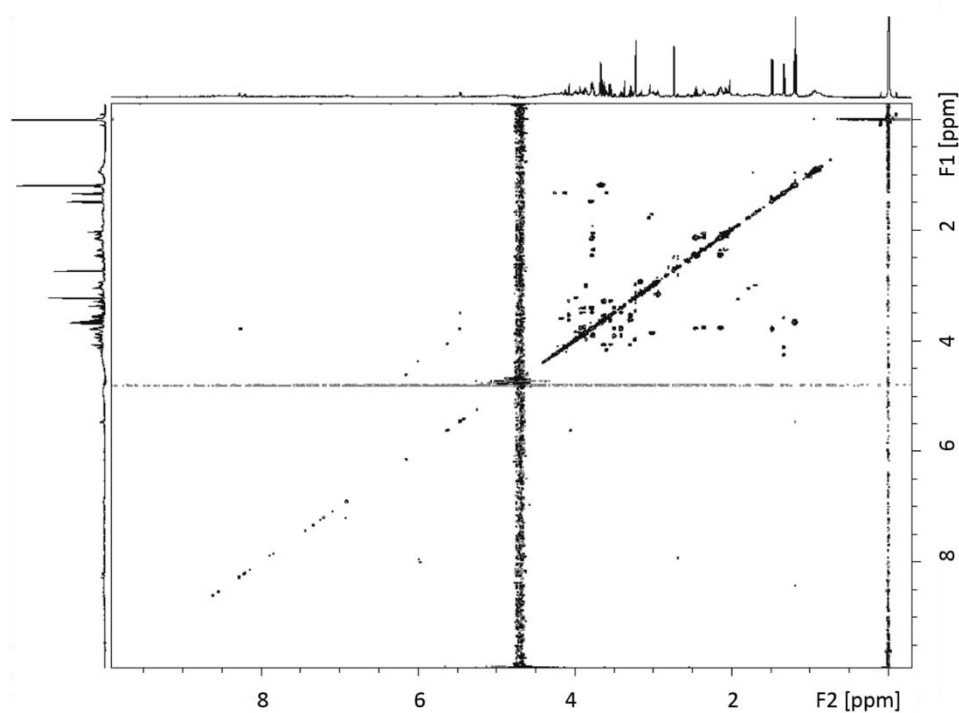


Figure 8.13: Representative TOCSY spectrum, acquired on the AVI600 instrument. NS = 8 and number of t_1 increments was 512. Experiment time was 2 hours and 4 minutes.

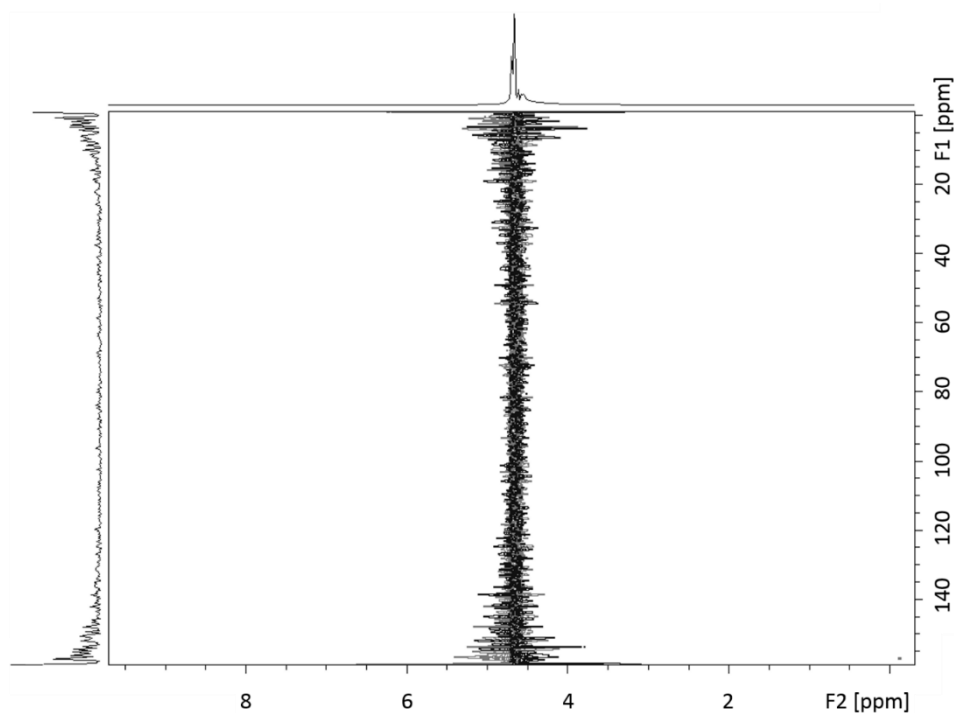


Figure 8.14: Representative HSQC spectrum, acquired on the AVI600 instrument. NS = 16 and number of t_1 increments was 512. Experiment time was 4 hours.

In summary, three different two-dimensional NMR experiments were assessed: ^1H JRES, ^1H TOCSY, and HSQC. Due to sensitivity issues, only proton two-dimensional spectra could be acquired. JRES and TOCSY experiments proved to be both rich in information and possible to acquire within a reasonable time frame.

Initial statistical analysis

The cultured cells had been treated with five different drugs or drug mixtures: temozolomide, YM155, CUSP9, CUDC-907, and daporinad. It would not be feasible to measure adequate numbers of replicates (≥ 3) for all treatments on a variety of cell lines (≥ 3) within the time constraints of the project. Instead, it was decided to focus on temozolomide and one additional chemotherapeutic agent. Temozolomide is an adjuvant chemotherapy given together with radiation therapy in treatment of glioblastoma patients [35, 36], and was therefore of interest to study further.

To facilitate the choice of one of the other four chemotherapies, PCA of the data listed in **Table 8.3** was carried out. In metabolomics, it is common to use PCA in the initial exploratory analysis of data [221]; it simplifies complex datasets and can reveal similarities and differences between samples. Ideally, it would point out which of the treated samples were dissimilar to the control

samples, *i.e.* show in which samples metabolic change had occurred. Samples with ID 9, 10, 13 and 14 were not included because they did not have satisfactory one-dimensional spectra. PCA with and without baseline correction was carried out, and it was clear that it affected the scores plot of the PCA quite heavily, further illustrated in **Figure 8.15** and **8.16**. Three samples failed to have their baselines corrected: 3, 6, and 8. Scaling with unit variance was deemed too harsh and introduced too much noise. Instead, pareto scaling, division by the squared standard deviation, was used.

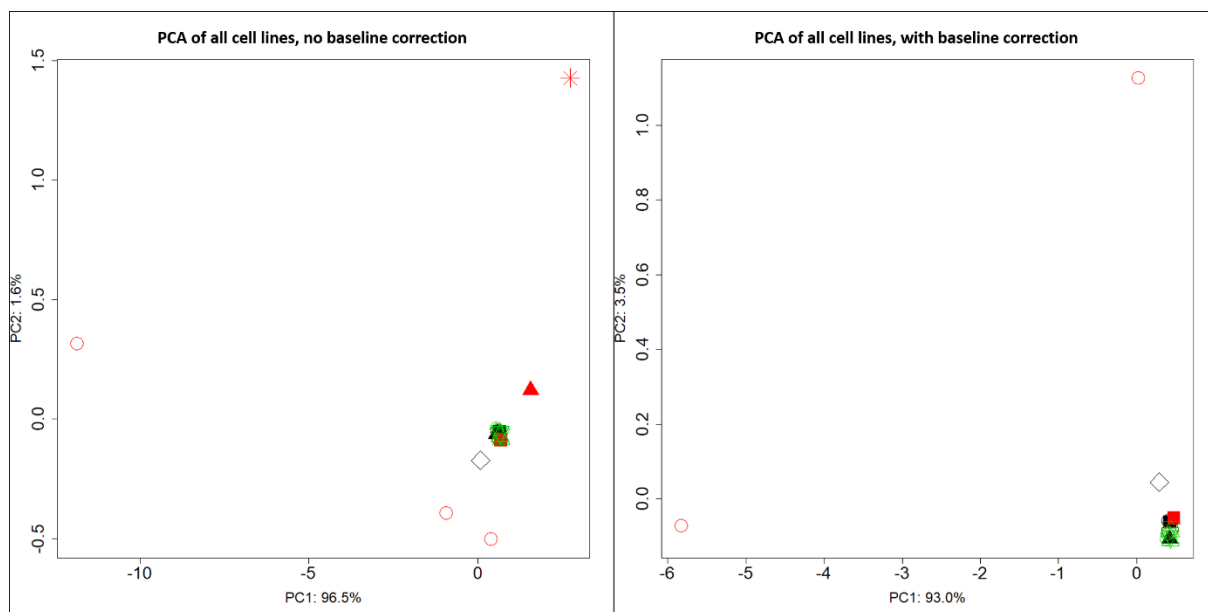


Figure 8.15: PCA scores plots of all cell lines without (left) and with (right) baseline correction. Color denotes cell line: black = T1454, red = T1456, and green = T1459. Shape denotes treatment: control = \circ , TMZ = \blacksquare , YM155 = \blacktriangle , CUSP9 = \diamond , FK866/daporinad = \star , CUDC-907 = $*$.

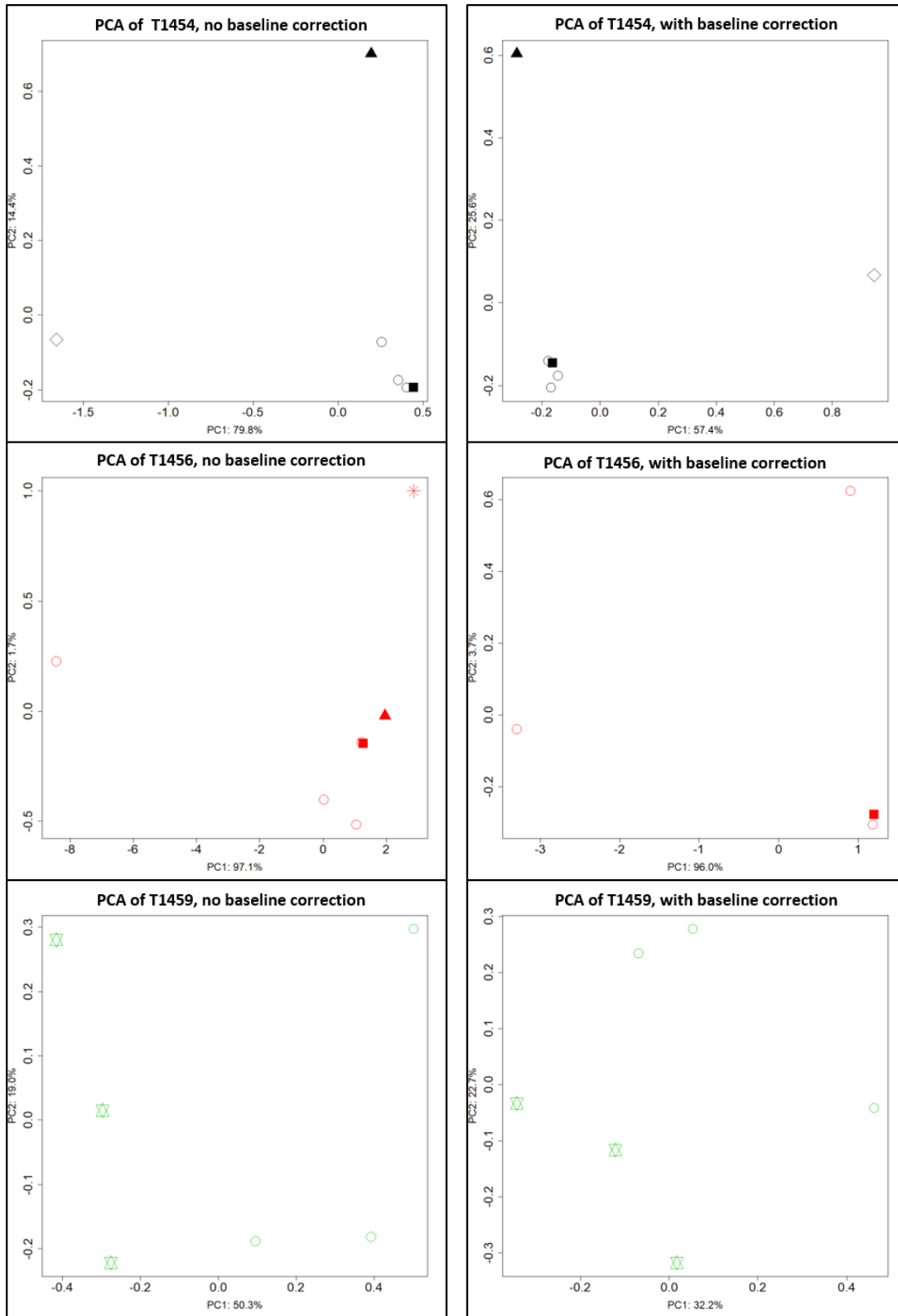


Figure 8.16: PCA scores plots of cell lines T1454, T1456, and T1459 without (left) and with (right) baseline correction. In the baseline corrected PCA of T1456, three samples were removed because they failed correction (sample ID 3, 6, and 8). Color denotes cell line: black = T1454, red = T1456, and green = T1459. Shape denotes treatment: control = \circ , TMZ = \blacksquare , YM155 = \blacktriangle , CUSP9 = \diamond , FK866/daporinad = \star , CUDC-907 = $*$.

The most noticeable effect seen in the PCA scores plots is how much baseline correction affected the outcome of the PCA. In addition, there were very few replicates of each sample type (cell line and treatment). Thus, the outcome of the PCA could not be considered without caution.

Samples treated with YM155, FK866/daporinad, CUDC-907 and CUSP9 were all separated from their respective control samples. CUSP9 was rejected as an option because it involved many different drugs and would probably not be available for widespread clinical use in the near future. FK866/daporinad showed consistent separation from control samples, but its benefit is largely in sensitizing cells to TMZ and not as a standalone chemotherapeutic agent [56, 222]. The present study was interested in examining the individual differences of response to chemotherapeutic agents, and mixing could confound the metabolic findings. CUDC-907 was at the time (autumn 2016) in early clinical trials. It was of greater interest to the surgeons at Rikshospitalet to pursue YM155, which had shown promising effect on glioblastoma cell lines [59, 60, 223]. The decision was supported by the fact that YM155 was clearly separated from control samples in two different cell lines (T1454 and T1456), as shown in **Figure 8.16**.

Two different chemotherapies were chosen as candidates for further study: temozolomide and YM155. The former because of its widespread use for glioblastoma patients and the latter because of its promising clinical usefulness.

8.5.2 Assessment of narrower sample tubes

After the preliminary study was concluded, there was a desire to increase the sensitivity of the measurements, even though the 800 MHz instrument had become available. In short, the sensitivity of detection in NMR is described by **Equation 8.1**.

$$\frac{\text{Signal}}{\text{Noise}} \propto N_m A T_s^{-1} B_0^{3/2} \gamma^{5/2} T_2^* (N_s)^{1/2} \quad \text{Equation 8.1}$$

Where N_M denotes the number of molecules in the observed sample volume, A is the abundance of the observed nucleus, T_s is the temperature of the sample and the surrounding coil, B_0 is the strength of the magnetic field, γ is the magnetogyric ratio of the observed nucleus, T_2^* is the effective time it takes for magnetization to disappear from the transverse plane (depends on field homogeneity and the molecule being measured), and N_s is the number of scans collected [85 (p. 97-99)].

T_2^* , γ and A are intrinsic and cannot be changed, except for ensuring as homogenous magnetic field as possible considering T_2^* . The probes used in both the AVI 600 MHz and AVIIIHD 800 MHz were cryogenically cooled, thus the S/N ratio was already increased in that aspect. The remaining factors were number of scans or concentration of analytes. Since additional scans are time-consuming, it was decided to attempt to increase the concentration by using narrower sample tubes: 3 mm outer diameter instead of 5 mm. In NMR, the sample must fill the tube enough to have solution above the coils, *i.e.* the sample height in the tube must be ≥ 5 cm. For a 5 mm tube, that requires at least 500 μL (preferably 600 μL) of sample, while in a 3 mm tube it only requires 180 μL sample. Thus, theoretically, a three-fold gain of sensitivity could be achieved.

Disappointingly, it turned out that working with 3 mm tubes was fraught with difficulty. Not only was it very challenging to manually fill the tubes with sample, but sufficient suppression of the solvent signal turned into an unnecessarily laborious task. It could take three to four attempts to obtain a spectrum without a large residual solvent peak. In **Figure 8.17**, an example of a typical bad suppression is shown.

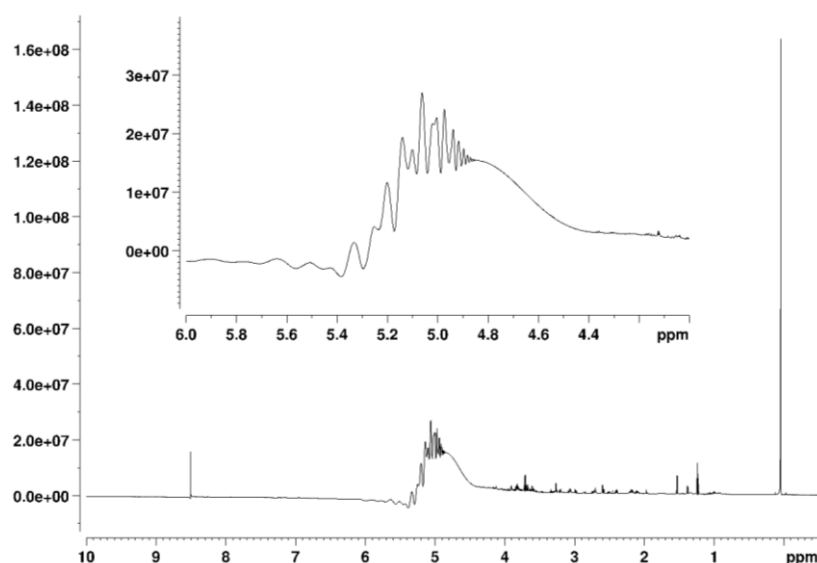


Figure 8.17: A representative spectrum of bad solvent suppression often occurring when using 3 mm NMR tubes instead of 5 mm NMR tubes. Compared to Figure 8.11, the residual water peak is a factor of three greater in intensity and four times wider on the x -axis.

Several different explanations were considered, among them contamination of the sample, instability of the sample within the magnet (wobbling of the tube), or issues with the tubes themselves. To address the first issue, contamination, the NMR buffer was mixed with pure water in a 3 mm NMR tube; solvent suppression was feasible without too many issues.

Contamination was therefore not the cause of the bad solvent suppression. The second issue was addressed by adding small weights to the spinners holding the 3 mm NMR tubes or using a heavier spinner. This had little to no effect on the spectral quality, and wobbling was ruled out as a cause. Finally, moving samples from 3 mm tubes to 5 mm tubes seemed to fix the issue, *i.e.* satisfactory suppression was usually obtained on the first or second try. It was concluded that it was the 3 mm tubes themselves that were causing issues with water signal suppression.

Nils Nyberg from Bruker was also consulted on the issue, and he came to a similar conclusion: 5 mm tubes were easier to work with than 3 mm tubes. It turned out that shimming with 5 mm tubes was much less troublesome than with 3 mm tubes. If the shimming was not adequate the water peak was not narrow enough and thus could not be successfully suppressed. The 3 mm tubes were therefore replaced with 5 mm tubes to reduce issues with shimming. New 5 mm NMR tubes of higher quality (thinner and straighter glass walls) than those used in the preliminary studies, were purchased to further reduce external effects on shimming.

However, when returning to 5 mm tubes, the issue with solvent suppression persisted. Something else was spoiling the shimming of the samples, independently of the NMR tubes. The main difference between the AVI600 and the AVIIIHD800 was the automatic sample changer; the latter one was possible to cool and was kept at 7 °C. The probe temperature was 25 °C and the samples must have experienced solvent convection after being placed in the probe. This would reduce the quality of the shimming. If the samples were given a minimum of 15 minutes to temperature equilibrate in the magnet prior to shimming, satisfactory solvent suppression was achieved in almost every single spectrum. In hindsight, this probably applied to the 3 mm tubes as well. However, due to the difficulty in handling and filling them with samples they were not used again.

3 mm tubes were tested with the goal of increasing sensitivity by increasing sample concentration. Issues with solvent suppression encumbered the work. No explanation other than that 3 mm tubes were more difficult to shim with than 5 mm tubes, could be found. However, when returning to 5 mm tubes, the issues persisted and it was found that the samples required a minimum of 15 minutes to reach probe temperatures when coming from a cooled sample holder. Temperature fluctuations and accompanying solvent movements would encumber shimming, and this was the cause of the issues with solvent suppression, not the NMR tube geometry. However, the 3 mm tubes were difficult to use in other respects too, and were therefore not reconsidered.

8.6 Calculation of column efficiency

Column efficiency expresses the extent of band broadening, *i.e.* how much a band of solutes is expanded as it travels through a column. It can be expressed by plate number, which is defined by **Equation 8.2**.

$$N = (t_R/\sigma)^2 \quad \text{Equation 8.2}$$

Where t_R is the retention time and σ its standard deviation. If the peaks are assumed to have a Gaussian distribution, peak width can be used to measure σ . There are three different approaches commonly used to measure N , given by **Equation 8.3**, **8.4**, and **8.5**, respectively.

$$N = 5.54 \left(t_R / w_{\frac{1}{2}h} \right)^2 \quad \text{Equation 8.3}$$

$$N = 16(t_R/w_{base})^2 \quad \text{Equation 8.4}$$

$$N = 25(t_R/w_{4.4\%h})^2 \quad \text{Equation 8.5}$$

Measuring peak width at half maximum height (**Equation 8.3**) is common if the peak is symmetrical, while the measuring at the base or 4.4% of peak height should be employed if the peak has a lot of tailing. Furthermore, column efficiency must be determined in experiments run without a MP composition gradient or temperature gradient (isocratic and isothermal elution). The injected sample must be dissolved in the MP [96 (p. 5-10)]. In **Figure 8.18**, the different ways of measuring peak width are shown.

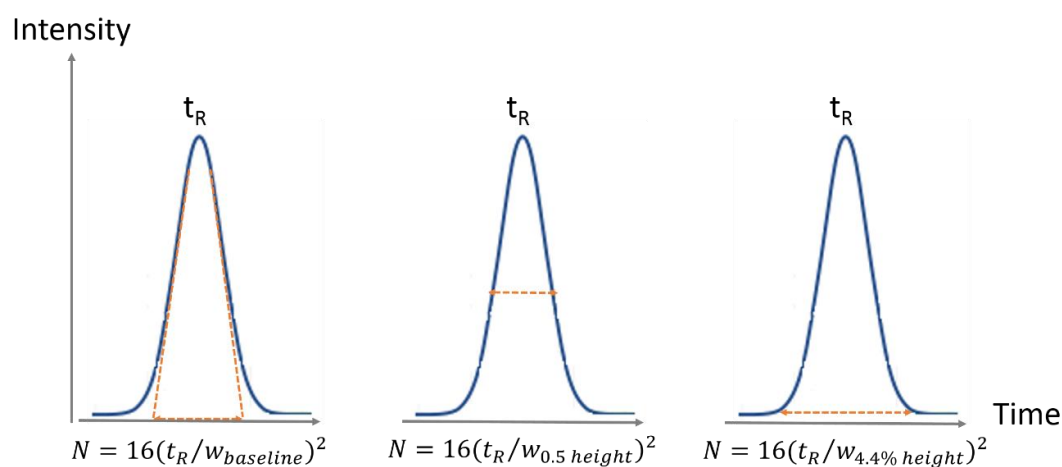


Figure 8.18: The three different approaches of calculating plate number (N). Figure adapted from [96 (p. 10)].

For comparison of columns with different length, plate height (H) can be calculated instead, given by **Equation 8.6**.

$$H = L/N \quad \text{Equation 8.6}$$

Where L is the length of the column [96 (p. 10-11)].

8.7 Chromatographic detection with ultraviolet spectroscopy

Detection of compounds based on their UV absorption at specific wavelengths is used extensively in liquid chromatography analyses because it is simple and robust [96 (p. 80-84)]. It is based on Beer's law, given by **Equation 8.7**, which states that absorbance (A) of UV radiation depends on the concentration of the analyte (c), the length the radiation travels through the solution (l), and the molar absorptivity (ϵ) of the compound.

$$A = \epsilon lc \quad \text{Equation 8.7}$$

The molar absorptivity is constant for a compound at a given wavelength. It expresses how much radiation the compound can absorb at the given wavelength [224 (p.436-437)]. There are three different factors that can affect the sensitivity of the detection. First, the wavelength of the UV radiation should be at or close to the ϵ_{max} of the compounds of interest, to maximize the absorptivity.

However, this must be balanced with not measuring at a wavelength where other compounds in the sample also absorb radiation. Second, the length the radiation travels through the solution should be maximized, for example done with a U-path flow cell, which is schematically illustrated in **Figure 8.19**. Third, the analytes should be as concentrated as possible when reaching the detector, *i.e.* band broadening should be minimized. Limiting band broadening requires an efficient column, tight couplings (no gaps that allow dwell volumes to collect), and narrow tubing [96 (p. 47-84), 224 (p. 686-687)].

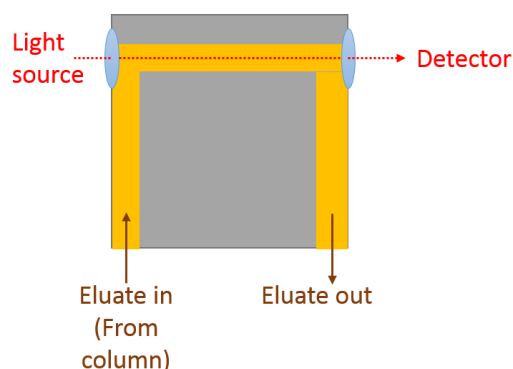


Figure 8.19: A schematic illustration of a U-path flow cell.

8.8 Nuclear magnetic resonance spectra

In the following section, various representative NMR spectra are presented. First, a representative spectrum for each experiment (one-dimensional ^1H NMR with excitation sculpting in **Figure 8.20**, JRES in **Figure 8.21**, and TOCSY in **Figure 8.22**) is given. A control sample from cell line T1548 was used in acquisition of the spectra (sample ID 76 in **Table 4.1**). Second, the one-dimensional spectrum of YM155-treated samples from a sensitive (**Figure 8.23**) and less sensitive (**Figure 8.24**) cell line is presented, to illustrate the difference in response to treatment. Finally, a representative spectrum of a TMZ-treated cell samples is given in **Figure 8.25**.

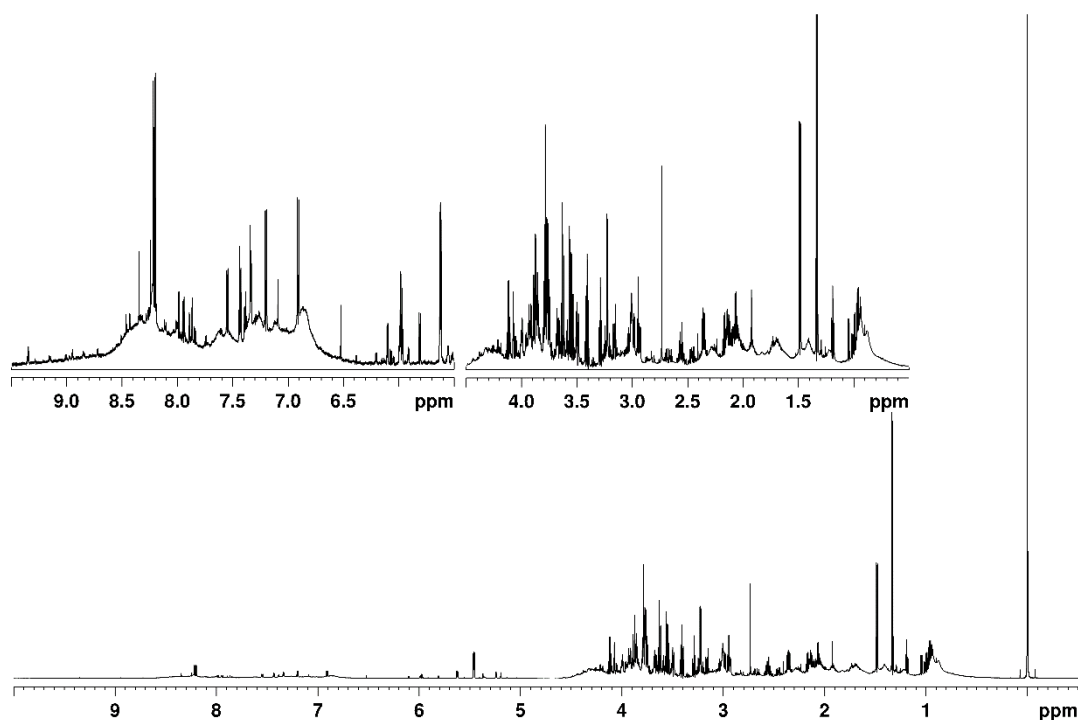


Figure 8.20: Representative one-dimensional spectrum with solvent suppression by excitation sculpting, NS = 4096. Spectrum obtained with a control sample (ID 76 in Table 4.1) from cell line T1548. The y-axis is adjusted to show maximum amount of detail.

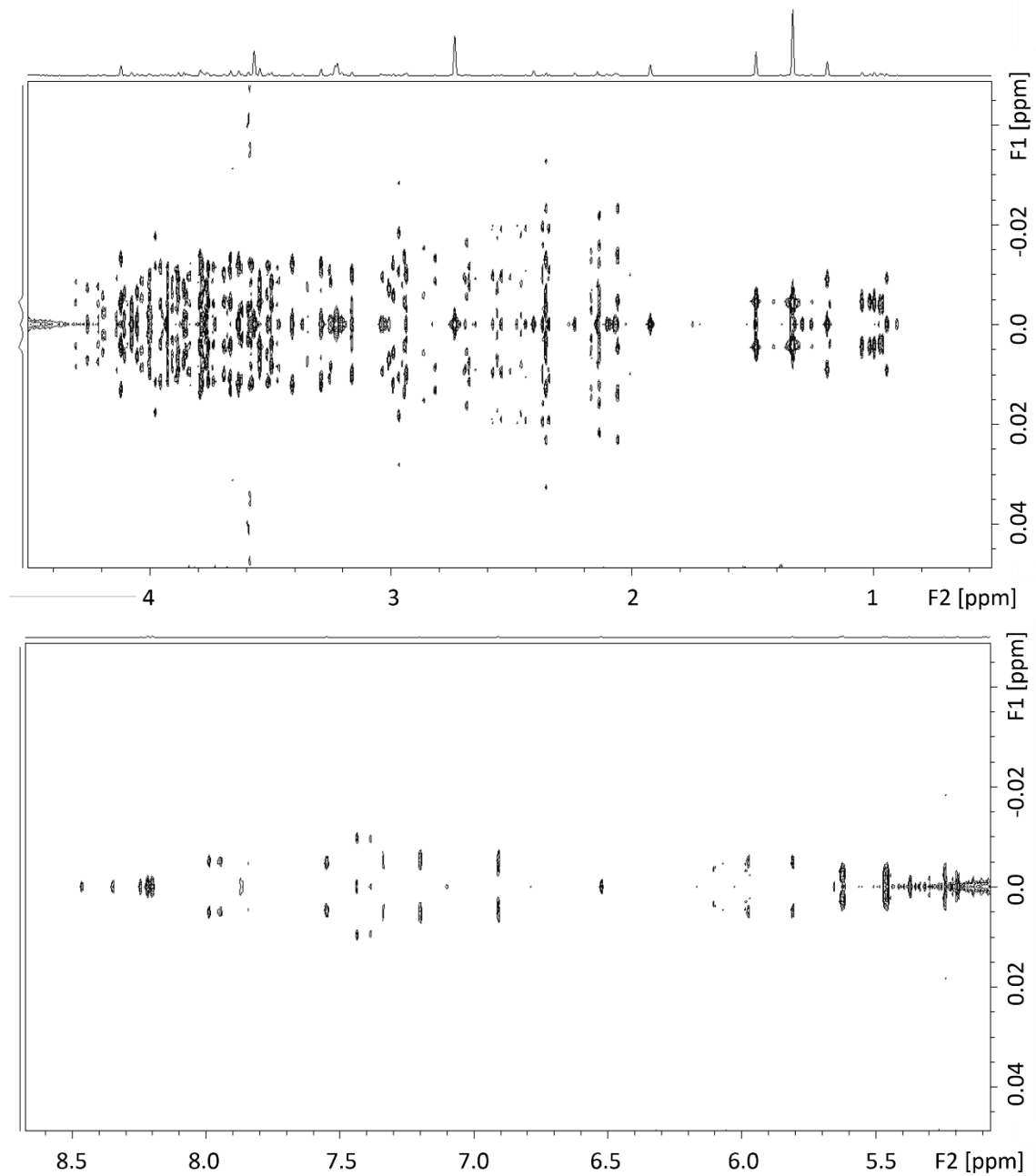


Figure 8.21: Representative JRES spectrum (f2: 0.5-4.5 ppm (top) and 5.0 to 8.6 ppm (bottom)), NS = 32 and number of t_1 increments was 160. Spectrum obtained with a control sample (ID 76 in Table 4.1) from cell line T1548.

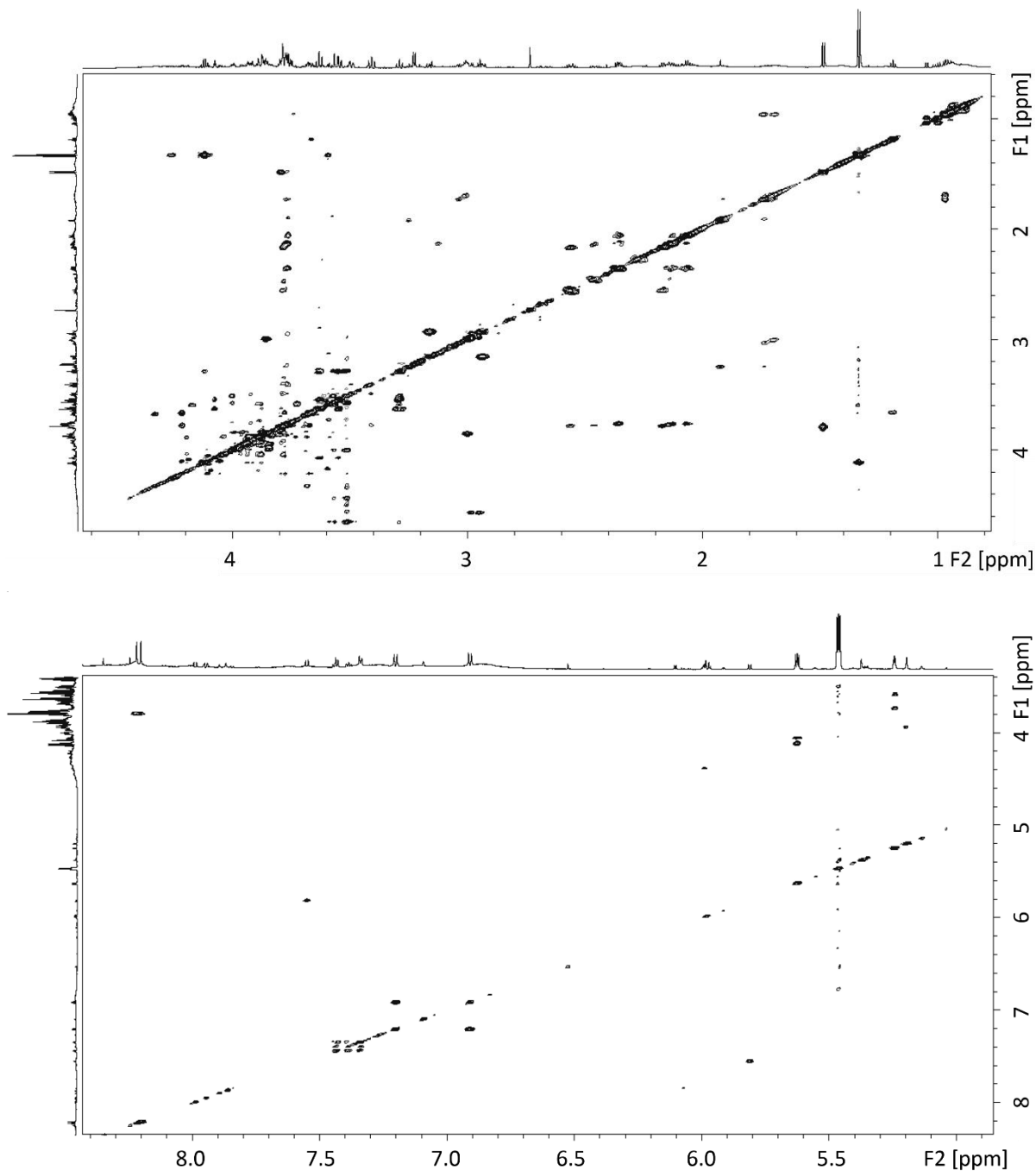


Figure 8.22: Representative TOCSY spectrum (f2: 0.5-4.5 ppm (top) and 5.0 to 8.5 ppm (bottom)), NS = 32 and number of t_1 increments was 1024. Spectrum obtained with a control sample (ID 76 in Table 4.1) from cell line T1548.

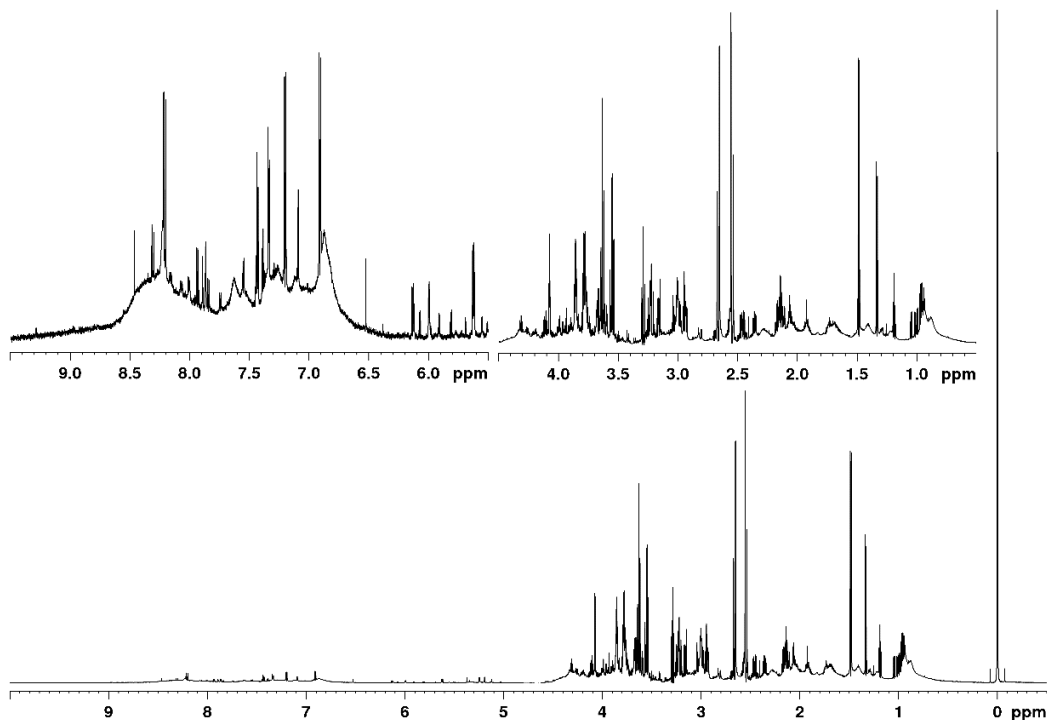


Figure 8.23: Representative one-dimensional spectrum of a sample from a less sensitive cell line treated with YM155. Solvent suppression was by excitation sculpting and NS = 4096. Spectrum obtained with an YM155-treated sample (ID 54 in Table 4.1) from cell line T1454. The y-axis is adjusted to show maximum amount of detail.

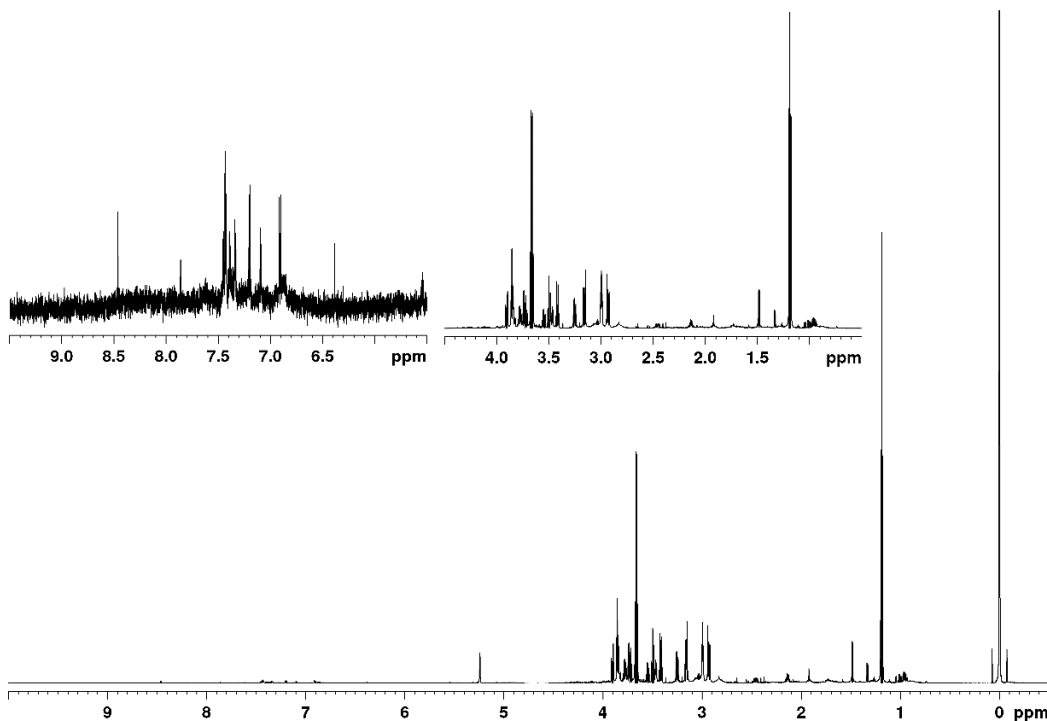


Figure 8.24: Representative one-dimensional spectrum of a sample from a sensitive cell line treated with YM155. Solvent suppression was by excitation sculpting and NS = 4096. Spectrum obtained with an YM155-treated sample (ID 102 in Table 4.1) from cell line T1459. The y-axis is adjusted to show maximum amount of detail.

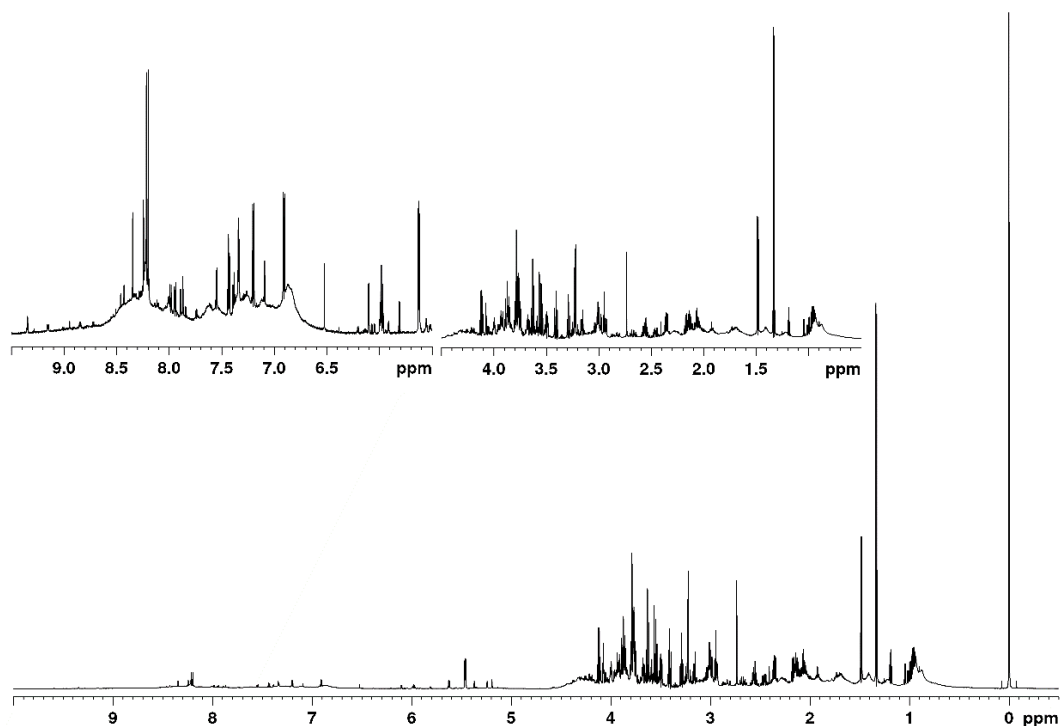


Figure 8.25: Representative one-dimensional spectrum of a sample treated with TMZ. Solvent suppression was by excitation sculpting and NS = 4096. Spectrum obtained with a TMZ-treated sample (ID 80 in Table 4.1) from cell line T1548. The y-axis is adjusted to show maximum amount of detail.

8.9 Chromatograms from liquid-chromatography-mass spectrometry analyses

Six chromatograms are presented below: two of control samples from cell line T1454 in **Figure 8.26**, two of YM155-treated samples from cell line T1459 in **Figure 8.27**, two of YM155-treated samples from cell line T1454 in **Figure 8.28**, and one of the standard citric acid solution (20 $\mu\text{g/mL}$) in **Figure 8.29**.

All the cell samples were dissolved in 80/20 (v/v) acetonitrile/buffer (30 mM ammonium formate, pH 4.5). The standard citric acid solution (20 $\mu\text{g/mL}$) was dissolved in the same solvent mixture. The injection loop had a volume of 500 nL. The MP flow rate was 0.60 $\mu\text{L/minute}$, and the MS was run in negative mode with parallel reaction monitoring (PRM) of m/z 191. All of the following chromatograms were for extracted fragment ions with m/z 111.005-111.009.

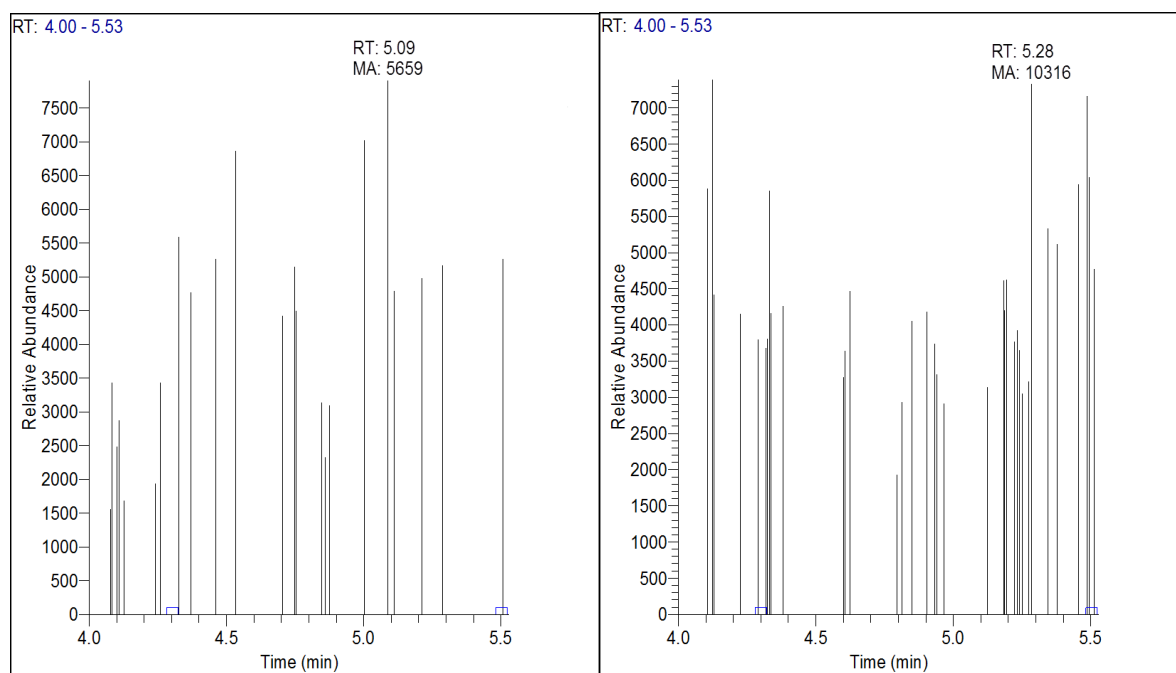


Figure 8.26: Chromatograms of two control samples from cell line T1454.

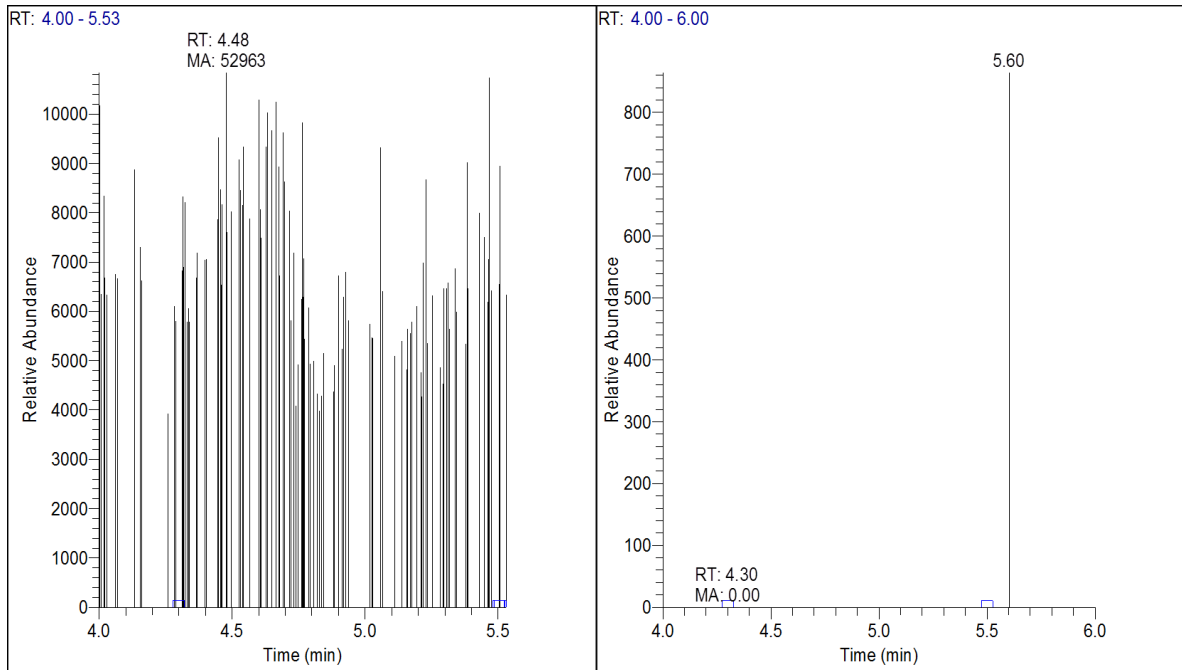


Figure 8.27: Chromatograms of two YM155-treated samples from cell line T1459.

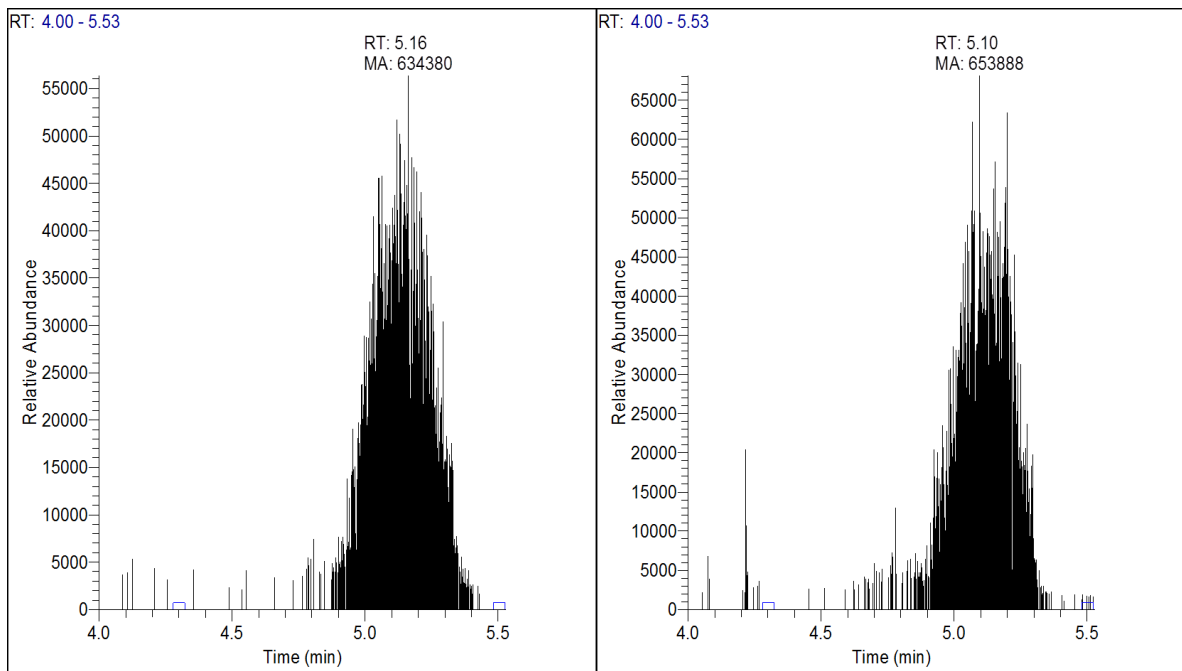


Figure 8.28: Chromatograms of two YM155-treated samples from cell line T1454.

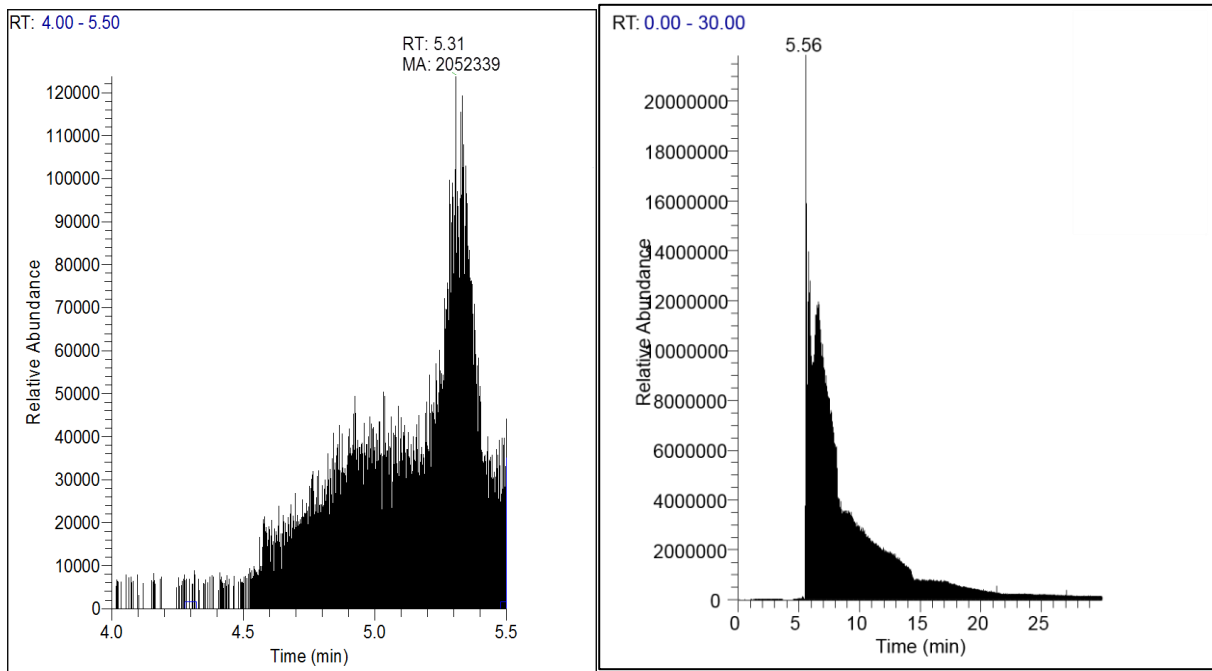


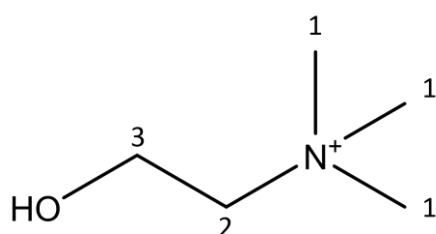
Figure 8.29: Chromatogram citric acid standard solution, from 4.00 to 5.50 minutes to the left and 0 to 30 minutes to the right. Due to the large amount of citric acid injected, the peak had a lot of tailing.

8.10 Molecule Encyclopedia

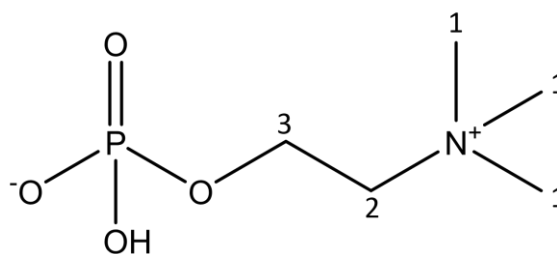
Peak ppm values written in **bold** are mainly used in identification of the metabolite, and only the protons with ppm values stated are numbered. Ppm values and coupling patterns are from the HMDB and an article by Govindaraju *et al.* [169], unless otherwise noted. The following abbreviations for splitting patterns are used: singlet = s, doublet = d, t = triplet, q = quartet, dd = double doublet, m = multiplet.

Total choline

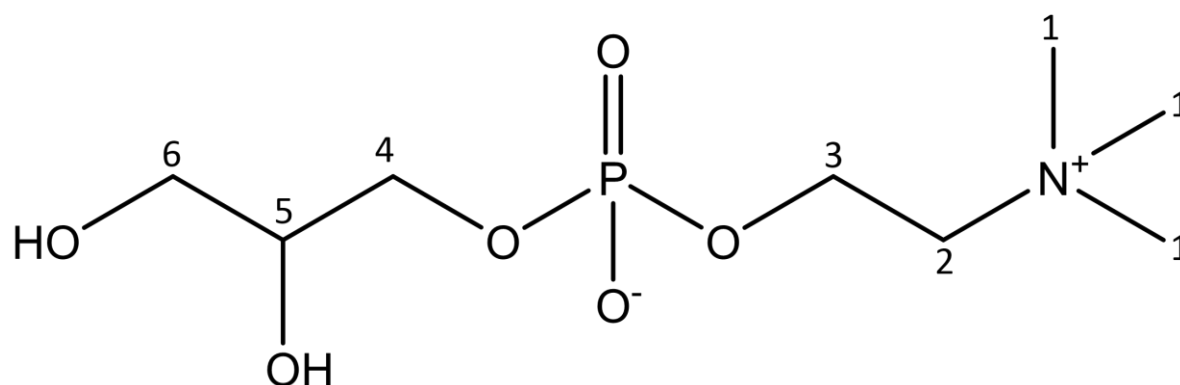
Choline (C)



Phosphorylcholine (PC)

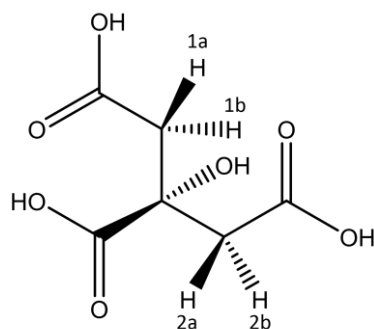


Glycerophosphorylcholine (GPC)



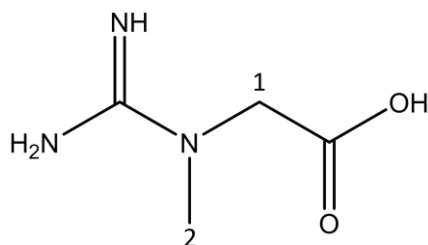
Monoisotopic mass	C: 104.10754 Da PC: 184.0739 Da GPC: 257.1028 Da
NMR peaks	C: 3.19 (#1, s), 3.51 (#2, dd), 4.06 (#3, m) PC: 3.208 (#1, s), 3.64 (#2, m), 4.28 (#3, m) GPC: 3.212 (#1, s), 3.66 (#2, m), 4.31 (#3, m), 3.90 (#4, 6, m), 3.87/3.95 (#5, m/m), 3.61/3.67 (#6, dd/dd)

Citric acid



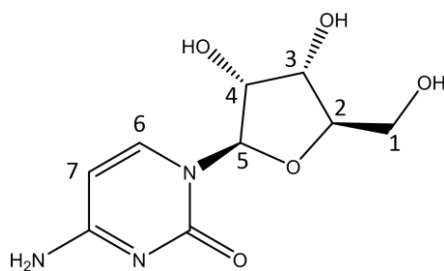
Monoisotopic mass	192.0270 Da
NMR peaks	2.53 (#1b and 2b, d), 2.66 (#1a and 2a, d)

Creatine



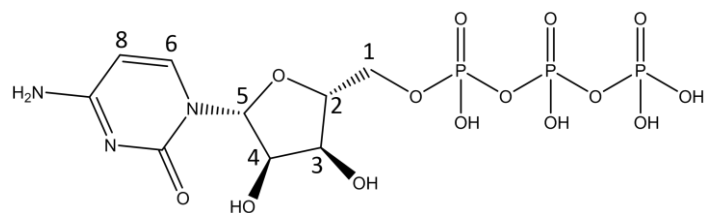
Monoisotopic mass	131.0695 Da
NMR peaks	3.02 (#2, s), 3.92 (#1, s)

Cytidine



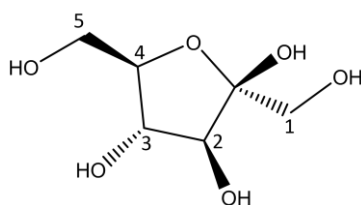
Monoisotopic mass	243.0855 Da
NMR peaks	3.80 (#1, dd), 3.92 (#1, m), 4.11 (#2, m), 4.20 (#3, t), 4.30 (#4, t), 5.89 (#5, d), 6.04 (#7, d), 7.83 (#6, d)

Cytidine triphosphate



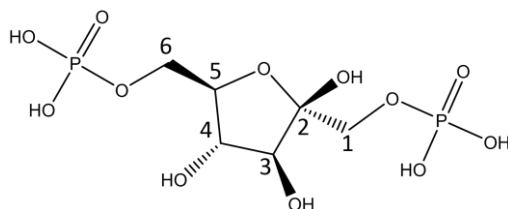
Monoisotopic mass	482.9845 Da
NMR peaks	4.25 (#1 and 2, m), 4.32 (#3, dd), 4.41 (#4, dd), 5.99 (#5, d), 6.13 (#8, d), 7.97 (#6, d)

D-Fructose



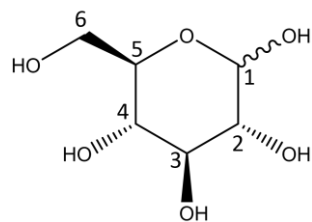
Monoisotopic mass	180.0634 Da
NMR peaks	3.58 (#1, m), 3.69 (#1 and 5, m), 3.82 (#2, 4, and 5, m), 3.90 (#3, dd), 4.00 (#4, m), 4.03 (#5, dd), 4.12 (#2 and 3, m)

D-Fructose 1,6-bisphosphate



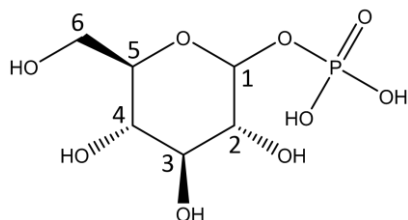
Monoisotopic mass	339.9960 Da
NMR peaks	3.77 (#1), 3.895 (#5), 3.925 (#4), 4.177 (#2), 4.19 (#3)
Note	NMR peaks found in Biological Magnetic Resonance Bank (BMRB). No splitting patterns were given by the database.

D-Glucose



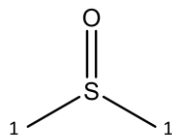
Monoisotopic mass	180.0634 Da
NMR peaks	3.233 (#3, dd), 3.398 (#4, m), 3.458 (#5, m), 3.524 (#2, dd), 3.728 (#3 and 6, m), 3.824 (#5 and 6, m), 3.889 (#6, dd), 4.634 (#1, d), 5.223 (#1, d)

D-Glucose 1-phosphate



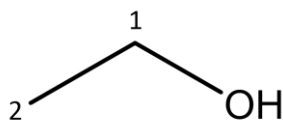
Monoisotopic mass	260.0297 Da
NMR peaks	3.39 (#4, t), 3.48 (#5, m), 3.75 (#3, m), 3.85 (#6, m), 3.90 (#2, m), 5.45 (#1, dd)

DMSO (Dimethyl sulfoxide)



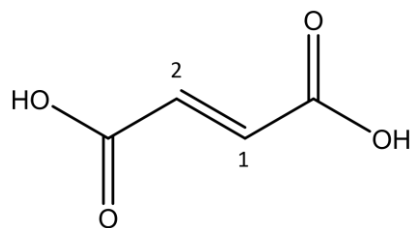
Monoisotopic mass	78.0139 Da
NMR peaks	2.62 (#1, s)

Ethanol



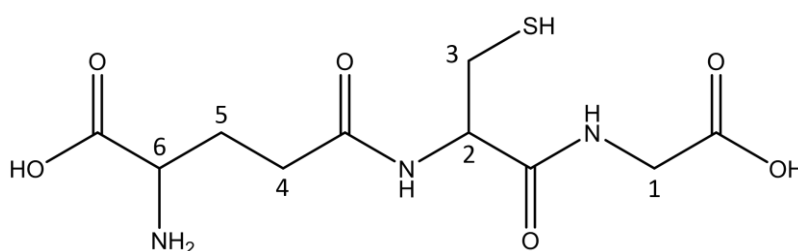
Monoisotopic mass	46.06844 Da
NMR peaks	1.17 (#2, t), 3.65 (#1, q)

Fumaric acid



Monoisotopic mass	116.0110 Da
NMR peaks	6.51 (#1 and 2, s)

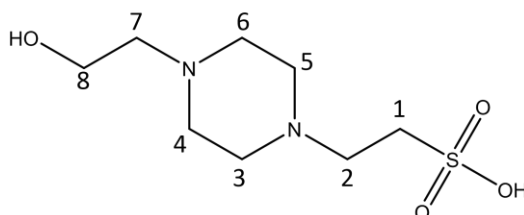
Glutathione



Monoisotopic mass	307.0838 Da
NMR peaks	2.15 (#5, m), 2.54 (#4, m), 2.97 (#3, dd), 3.78 (#6 and 1, m), 4.20 (#2, q)

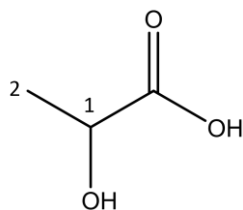
HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)

NMR peaks found in Biological Magnetic Resonance Bank (BMRB). No splitting patterns were given by the database.



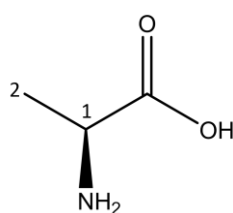
Monoisotopic mass	238.3045 Da
NMR peaks	2.851 (#3 and 5), 2.936 (#2), 3.151 (#1), 3.151 (#4 and 6), 3.056 (#8), 3.851 (#7)

Lactic acid



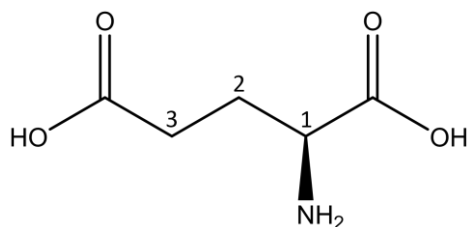
Monoisotopic mass	89.0244 Da
NMR peaks	1.32 (#2, d), 4.10 (#1, q)

L-Alanine



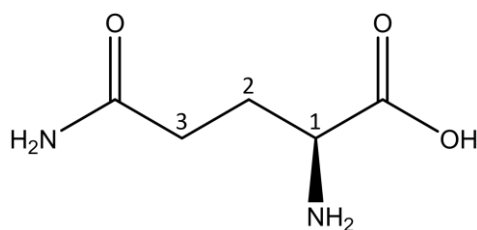
Monoisotopic mass	89.0477 Da
NMR peaks	1.47 (#2, d), 3.77 (#1, q)

L-Glutamic acid



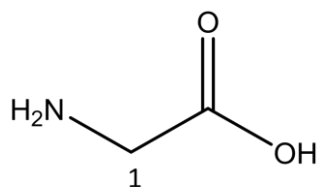
Monoisotopic mass	147.1293 Da
NMR peaks	~2.04 (#2, m), 2.119 (#2, m), 2.341 (#3, m), 3.748 (#1, dd)

L-Glutamine



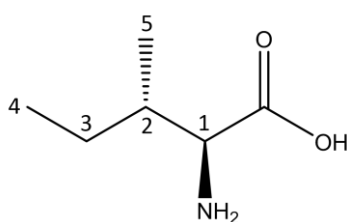
Monoisotopic mass	146.0691 Da
NMR peaks	2.12 (#4, m), 2.45 (#5, m), 3.77 (#3, t)

L-Glycine



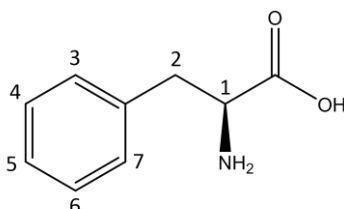
Monoisotopic mass	75.032028409 Da
NMR peaks	3.54 (#1, s)

L-Isoleucine



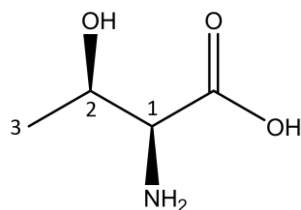
Monoisotopic mass	131.0946 Da
NMR peaks	0.93 (#4, t), 1.00 (#5, d), 1.25 (#3, m), 1.46 (#3, m), 1.97 (#2, m), 3.66 (#1, d)

L-Phenylalanine



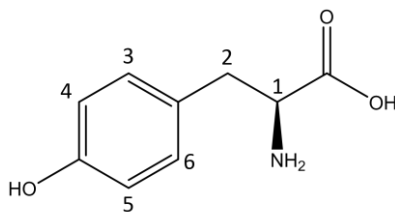
Monoisotopic mass	165.0790 Da
NMR peaks	3.11 (#2, m), 3.27 (#2, m), 3.98 (#1, m), 7.32 (#3 and 7, d), 7.36 8#5, m), 7.42 8#4 and 6, m)

L-Threonine



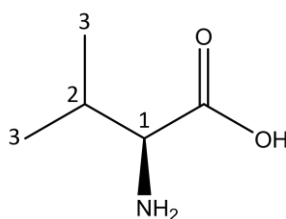
Monoisotopic mass	119.0582 Da
NMR peaks	1.32 (#3, d), 3.58 (#1, d), 4.24 (#2, m)

L-Tyrosine



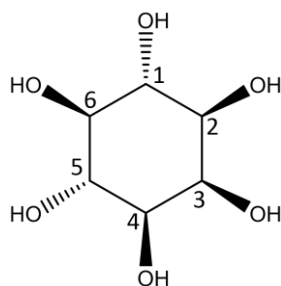
Monoisotopic mass	181.0739 Da
NMR peaks	3.024 (#2, dd), 3.17 (#2, dd), 3.921 (#1, dd), 6.877 (#4 and 5, d), 7.17 (#3 and 6, d)

L-Valine



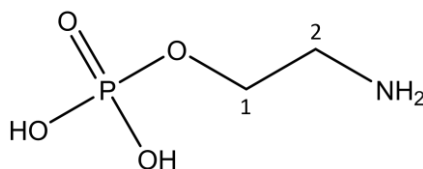
Monoisotopic mass	117.1463 Da
NMR peaks	0.98 (#3, d), 1.03 (#3, d) , 2.26 (#2, m), 3.60 (#1, d)

Myo-inositol



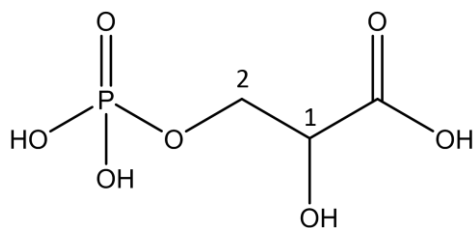
Monoisotopic mass	180.0634 Da
NMR peaks	3.28 (#6, t), 3.52 (#2 and 4, dd), 3.61 (#1 and 5, t), 4.05 (#3, t)

O-Phosphoethanolamine



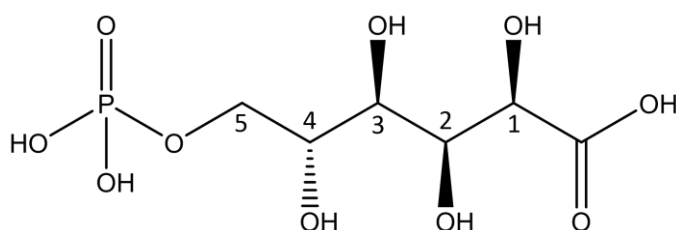
Monoisotopic mass	141.0191 Da
NMR peaks	3.24 (#2, t), 4.01 (#1, m)

3-phosphoglyceric acid



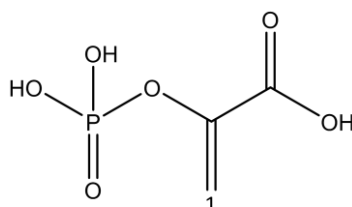
Monoisotopic mass	185.9929 Da
NMR peaks	3.88 (#1, m), 4.00 (#1, m), 4.19 (#1, dd)

6-Phosphogluconic acid



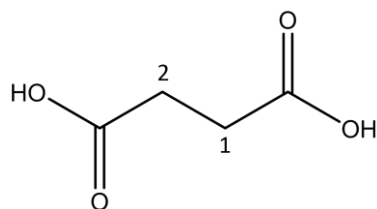
Monoisotopic mass	276.0246 Da
NMR peaks	3.84 (#2 and 3, m), 3.96 (#4, m), 4.10 (#1, m), 4.19 (#5, d)

Phosphoenolpyruvic acid



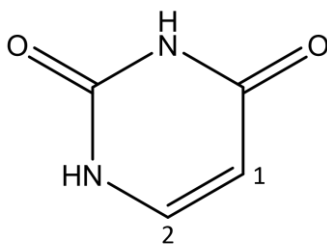
Monoisotopic mass	167.9824 Da
NMR peaks	5.18 (#1, t), 5.26 (#1, t)

Succinic acid



Monoisotopic mass	118.0266 Da
NMR peaks	2.39 (#1 and 2, s)

Uracil



Monoisotopic mass	112.0272 Da
NMR peaks	5.79 (#2, d), 7.52 (#1, d)

8.11 R scripts

Daniel Sachse has written the scripts presented here. The author has made minor changes, but Sachse did the main bulk of coding. The code in sections **8.11.1** and **8.11.2** were colored by the author. The script presented in section **8.11.3** contains functions made by Sachse, which are used in the two preceding scripts.

8.11.1 R script: pre-processing and PCA of preliminary data

```
require(pcaMethods);
require(FTICRMS); #for the baseline correction
require(pls);
require(plotrix);
source("functions.R");# Loading and processing data

#
#### Remember to set session working directory to source file location! ####
#

# # where are the spectra files, relative to this R script?
folder = "./Old 1D NMR txt files/";

# Get the file names
vnames <- list.files(folder)

# We convert the txt files to appropriate, two-column matrices with a for loop
# First, always use the length command so the code is robust to data changes!

numObs <- length(vnames)
for(i in 1:numObs){

  # Create a temporary variable tmp. It has four columns. (index, Re, Im, ppm)
  tmp <- read.csv(paste0(folder, vnames[i]), header=FALSE)
  cat(vnames[i], ":", nrow(tmp),"lines read. Ppm: ", range(tmp[,4]), "\n")

  # Only take the "safe" window from -1 til +11 ppm to combat Bruker's varying ppm scales
  right = which(tmp[,4] < (-1))[1]
  left = which(tmp[,4] < 11)[1]
  window = left:right
  cat("\tleft", left, tmp[left,4], "right", right, tmp[right,4],
      "\t-> window", length(window), tmp[left,4] - tmp[right,4], "\n")
  if (i == 1) {
    windowLenFixed = length(window) # remember for later, so all spec have same length
```

```

}
right = left + windowLenFixed; # -1, ah w/e
window = left:right

# Put together a big matrix:
# On first iteration, initialize matrix with correct size, but all nulls
if (i == 1) {
  rownames = paste("m",gsub(".txt","",vnames[1:numObs]),sep="")
  colnames = tmp>window,4] # ppm
  spec0 <- matrix(0, nrow = numObs, ncol = length(window),
                 dimnames = list(rownames, colnames))
}
# Then, for each iteration, copy the intensities from file to matrix row:
spec0[i,] = tmp>window,2] # intensities Re
}

# Clip water and ethanol, remove empty space (by defining regions we want to keep)
spec = spec0[ , c(sclr(spec0, 10, 5.0), sclr(spec0, 4.5, 3.7), sclr(spec0, 3.62, 1.22),
sclr(spec0, 1.15, -0.5))]

# Save this preliminary data matrix
save(spec, file = "Glioblastoma-Spec-Matrix.RData")
load("Glioblastoma-Spec-Matrix.RData")

# Make a data frame
data = data.frame(fnames = I(rownames(spec)))
data$spec = spec
data$cellLine = substring(data$fnames, 2,2)
data$treatment = substring(data$fnames, 3,3)
data$id = substring(data$fnames, 4)

cbind(data$fnames, data$treatment)
with(data, cbind(fnames, cellLine, treatment, id))

# More cleaning: Measure TSP, normalize
# Look at TSP
plotspectra(spec[,sclr(spec, -0.02, 0.02)], ylim=c(1,1e9))
# integrate TSP:
data$tspsint = rowSums(spec[,sclr(spec, -0.02, 0.02)])
# integrate entire spectrum (except TSP)
data$totalint = rowSums(spec[,sclr(spec, 0.2, 9.9)])

# Remove TSP as well:
data$spec = data$spec[,sclr(data$spec, 0.2, 10)]

#Normalize to total integral (different number of cells in the various samples...)

```

```

data$spec.totint = sweep(data$spec,1,data$totalint,"/")
plotspectra(data$spec.totint[1:10,], ylim=c(0,0.001), col=data$treatment)

# Save this.
save(data, file="Glioblastoma-Spec-Dataframe.RData")
load(file="Glioblastoma-Spec-Dataframe.RData")

symbTr = function(x) { (x=="1")*1 + (x=="2")*15 + (x=="3")*17 + (x=="4")*8 + (x=="5")*5 +
(x=="6")*11}

tmp = data$spec.totint

# Run a PCA with pareto scaling and no baseline:
# Run with one cell line of choice (flt = data$cellLine=="X") or all cell lines (flt = TRUE)
#flt=TRUE
flt = data$cellLine=="1"
cat("Doing PCA, this will take a while.\n");
spec.pca <- pca(tmp[flt,],method="nipals",nPcs=4,scale="pareto")
summary(spec.pca)
pc <- sprintf("%.1f",100*R2cum(spec.pca)[1])
pc <- c( pc , sprintf("%.1f",100*diff(R2cum(spec.pca))) )
pcs <- c(1,2) #which to show in plot
xlim <- NULL #c(-104,120);
ylim <- NULL #c(-700,100);

plot(pcaMethods::scores(spec.pca)[,pcs[1]],pcaMethods::scores(spec.pca)[,pcs[2]],type="p",cex.
axis=1.75,main="PCA of all cell lines",cex.lab=1.5,xlab=paste("PC",pcs[1],":",
pc[pcs[1]],"%",sep=""),ylab=paste("PC",pcs[2],":",pc[pcs[2]],"%",sep=""),xlim=xlim,
ylim=ylim, col=data$cellLine[flt],pch=symbTr(data$treatment[flt]),cex=3);

# Strict baseline:
mxx=nrow(data);
max.iter=30;
plot.this = TRUE;
water = c(4.5,4.9) #leave this region out of the baseline correction; careful not to exceed
the clipping region below
data$basl <- data$spec; #initialize #matrix(0, mxx, length(ppm), dimnames=list(NULL,ppm));
data$spec.bas <- data$spec; # initialize #matrix(0, mxx, length(ppm),
dimnames=list(NULL,ppm));
for (i in 1:mxx) {
  cat(i, "/",mxx," \n");
  mindex = min(sclv(data$basl,water))
  maxdex = max(sclv(data$basl,water))
  #from 1 to mindex:
  region=data$spec.totint[i,1:mindex]
  tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
                max.iter=max.iter);
  if (tmp$iter == max.iter) {

```

```

cat("2nd try\n")
tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
               max.iter=max.iter, init.bd=rep(0,mindex));
}
# if (tmp$iter == max.iter) {
#   cat("3rd try\n")
#   tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
#                 max.iter=max.iter, halve.search=TRUE);
# }
data$basl[i,1:mindex] = tmp$baseline;
cat("  Baseline (1/2) :",tmp$iter,"iterations.\n");
#in the water region, copy the original spec, makes the result zero:
data$basl[i,(mindex+1):(maxdex-1)] = data$spec.totint[i,(mindex+1):(maxdex-1)];
#from maxdex to end:
region=data$spec.totint[i,maxdex:ncol(data$spec.totint)]
tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
               max.iter=max.iter);
if (tmp$iter == max.iter) {
  cat("2nd try\n")
  tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
                 max.iter=max.iter, init.bd=rep(0,ncol(data$spec.totint)-maxdex+1));
}
# if (tmp$iter == max.iter) {
#   cat("3rd try\n")
#   tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
#                 max.iter=max.iter, halve.search=TRUE);
# }
data$basl[i,maxdex:ncol(data$spec.totint)] = tmp$baseline;
cat("  Baseline (2/2) :",tmp$iter,"iterations.\n");

#subtract baseline from spec:
data$spec.bas[i,] = data$spec.totint[i,] - data$basl[i,]

if (plot.this)
  plotspectra(data$spec.bas[i,], main=i)
}

badones = c("m21608", "m23603", "m24606")
good = !(data$fnames %in% badones)

flt = !good

with(data[flt,], plotspectra(spec.tsp[,sclr(spec.tsp, 6, 7.5)], col=treatment))
with(data[flt,], plotspectra(spec.tsp[,], col=treatment))

# One colum is now all 0 after baseline. Remove it to allow UV scaling.

```

```

allnull = which.min(apply(data$spec.bas,2,"sd"))
data$spec.bas = data$spec.bas[,-allnull]
# Run a PCA after baseline with pareto scaling:
tmp = data$spec.bas

# Run with one cell line of choice (flt = data$cellLine=="X") or all cell lines (flt = good)
#flt=good
flt = good & data$cellLine=="1"
cat("Doing PCA, this will take a while.\n");
spec.pca <- pca(tmp[flt,],method="nipals",nPcs=4,scale="pareto")
summary(spec.pca)
pc <- sprintf("%.1f",100*R2cum(spec.pca)[1])
pc <- c( pc , sprintf("%.1f",100*diff(R2cum(spec.pca))) )
pcs <- c(1,2) #which to show in plot
xlim <- NULL #c(-104,120);
ylim <- NULL #c(-700,100);

plot(pcaMethods::scores(spec.pca)[,pcs[1]],pcaMethods::scores(spec.pca)[,pcs[2]],type="p",cex.
axis=1.75,main="PCA of all cell
lines",cex.lab=1.5,xlab=paste("PC",pcs[1],":",pc[pcs[1]],"%",sep=""),
ylab=paste("PC",pcs[2],":",pc[pcs[2]],"%",sep=""),xlim=xlim,ylim=ylim,col=data$cellLine[flt],p
ch=symbTr(data$treatment[flt]),cex=3);

```

8.11.2 R script: pre-processing, PCA and PLS

```

require(pcaMethods);
require(FTICRMS); #for the baseline correction
require(pls);
require(plotrix);
source("functions.R"); # Loading and processing data
#
#### Remember to set session working directory to source file location! #####
#
# # where are the spectra files, relative to this R script?
folder = "./1D NMR txt files/";
# Get the file names
vnames <- list.files(folder)
# We convert the txt files to appropriate, two-column matrices with a for loop
# First, always use the length command so the code is robust to data changes!
numObs <- length(vnames)
for(i in 1:numObs){

  # Create a temporary variable tmp. It has four columns. (index, Re, Im, ppm)

```

```

tmp <- read.csv(paste0(folder, vnames[i]), header=FALSE)
cat(vnames[i], ":", nrow(tmp), "lines read. Ppm: ", range(tmp[,4]), "\n")

# Only take the "safe" window from -1 til +11 ppm to combat Bruker's varying ppm scales
right = which(tmp[,4] < (-1))[1]
left = which(tmp[,4] < 11)[1]
window = left:right
cat("\tleft", left, tmp[left,4], "right", right, tmp[right,4],
    "-> window", length(window), tmp[left,4] - tmp[right,4], "\n")
if (i == 1) {
  windowLenFixed = length(window) # remember for later, so all spec have same length
}
right = left + windowLenFixed; # -1, ah w/e
window = left:right

# Put together a big matrix:
# On first iteration, initialize matrix with correct size, but all nulls
if (i == 1) {
  rownames = paste("m", gsub(".txt", "", vnames[1:numObs]), sep="")
  colnames = tmp[window,4] # ppm
  spec0 <- matrix(0, nrow = numObs, ncol = length(window),
    dimnames = list(rownames, colnames))
}
# Then, for each iteration, copy the intensities from file to matrix row:
spec0[i,] = tmp[left:right,2] # intensities Re
}

# Clip water and ethanol, remove empty space (by defining regions we want to keep)
spec = spec0[ , c(sclr(spec0, 10, 5.0), sclr(spec0, 4.5, 3.7), sclr(spec0, 3.62, 1.22),
sclr(spec0, 1.15, -0.5))]

# Save this preliminary data matrix
save(spec, file = "Glioblastoma-Spec-Matrix.RData")
load("Glioblastoma-Spec-Matrix.RData")

# Make a data frame
data = data.frame(fnames = I(rownames(spec)))
data$spec = spec
data$cellLine = substring(data$fnames, 2,2)
data$treatment = substring(data$fnames, 3,3)
data$id = substring(data$fnames, 4)

cbind(data$fnames, data$treatment)
with(data, cbind(fnames, cellLine, treatment, id))

```

```

# Look at TSP
plotspectra(spec[,sclr(spec, -0.02, 0.02)])

# integrate TSP:
data$tspsint = rowSums(spec[,sclr(spec, -0.02, 0.02)])

# integrate entire spectrum (except TSP)
data$totalint = rowSums(spec[,sclr(spec, 0.2, 9.9)])

# Remove TSP as well:
data$spec = data$spec[,sclr(data$spec, 0.2, 10)]
plotspectra(data$spec[1:10], ylim=c(0,1e7))

# Normalize to TSP:
data$spec.tsp = sweep(data$spec,1,data$tspsint,"/")
plotspectra(data$spec.tsp[1:10,], ylim=c(0,0.001), col=data$treatment)

# Save this.
save(data, file="Glioblastoma-Spec-Dataframe.RData")
load(file="Glioblastoma-Spec-Dataframe.RData")

# Symbols for plotting
symbTr = function(x) { (x=="1")*1 + (x=="2")*15 + (x=="3")*17 }

tmp = data$spec.tsp

# Strict baseline:

mxx=nrow(data);
max.iter=30;

plot.this = TRUE;

water = c(4.5,4.9) #leave this region out of the baseline correction; careful not to exceed
the clipping region below

data$basl <- data$spec; #initialize #matrix(0, mxx, length(ppm), dimnames=list(NULL,ppm));
data$spec.bas <- data$spec; # initialize #matrix(0, mxx, length(ppm),
dimnames=list(NULL,ppm));

for (i in 1:mxx) {

  cat(i, "/", mxx, "\n");

  mindex = min(sclv(data$basl,water))

  maxdex = max(sclv(data$basl,water))

  #from 1 to mindex:

  region=data$spec.tsp[i,1:mindex]

  tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
                 max.iter=max.iter);

  if (tmp$iter == max.iter) {

    cat("2nd try\n")

    tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",
                  max.iter=max.iter, init.bd=rep(0,mindex));

  }

  # if (tmp$iter == max.iter) {

  #   cat("3rd try\n")

  #   tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",

```

```

#             max.iter=max.iter, halve.search=TRUE);
#   }

data$basl[i,1:mindex] = tmp$baseline;

cat("   Baseline (1/2) :",tmp$iter,"iterations.\n");

#in the water region, copy the original spec, makes the result zero:
data$basl[i,(mindex+1):(maxdex-1)] = data$spec.tsp[i,(mindex+1):(maxdex-1)];

#from maxdex to end:

region=data$spec.tsp[i,maxdex:ncol(data$spec.tsp)]

tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",

               max.iter=max.iter);

if (tmp$iter == max.iter) {

  cat("2nd try\n")

  tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",

                max.iter=max.iter, init.bd=rep(0,ncol(data$spec.tsp)-maxdex+1));

}

#   if (tmp$iter == max.iter) {

#     cat("3rd try\n")

#     tmp = baseline(region,sm.norm.by="constant",neg.norm="constant",

#                   max.iter=max.iter, halve.search=TRUE);

#   }

data$basl[i,maxdex:ncol(data$spec.tsp)] = tmp$baseline;

cat("   Baseline (2/2) :",tmp$iter,"iterations.\n");

#subtract baseline from spec:

data$spec.bas[i,] = data$spec.tsp[i,] - data$basl[i,]

if (plot.this)

  plotspectra(data$spec.bas[i,], main=i)

}

badones = c("m41067", "m52088", "m53082", "m53090")
good = !(data$fnames %in% badones)

flt = !good

with(data[flt,], plotspectra(spec.tsp[,sclr(spec.tsp, 6, 7.5)], col=treatment))
with(data[flt,], plotspectra(spec.tsp[,], col=treatment))

# One colum is now all 0 after baseline. Remove it to allow UV scaling.

allnull = which.min(apply(data$spec.bas,2,"sd"))

data$spec.bas = data$spec.bas[,-allnull]

```



```

# Run a PCA after baseline:

tmp = data$spec.bas

#All cell lines

flt=good

cat("Doing PCA, this will take a while.\n");

spec.pca <- pca(tmp[flt,],method="nipals",nPcs=4,scale="uv")

summary(spec.pca)

pc <- sprintf("%.1f",100*R2cum(spec.pca)[1])

pc <- c( pc , sprintf("%.1f",100*diff(R2cum(spec.pca))) )

pcs <- c(1,2) #which to show in plot

xlim <- NULL #c(-104,120);

ylim <- NULL #c(-700,100);

plot(pcaMethods::scores(spec.pca)[,pcs[1]],pcaMethods::scores(spec.pca)[,pcs[2]],type="p",cex.
axis=1.75,main="PCA of all cell lines",cex.lab=1.5,xlab=paste("PC",pcs[1],":
",pc[pcs[1]],"%",sep=""),ylab=paste("PC",pcs[2],":
",pc[pcs[2]],"%",sep=""),xlim=xlim,ylim=ylim,col=data$cellLine[flt],pch=symbTr(data$treatment[
flt]),cex=3);

#Loadings:

lwplot(pcaMethods::loadings(spec.pca)[,1],pcaMethods::scl(spec.pca),main="PC1", cex.main=1.75,
cex.axis=1.75)

lwplot(pcaMethods::loadings(spec.pca)[,2],pcaMethods::scl(spec.pca),main="PC2", cex.main=1.75,
cex.axis=1.75)

# Run with one cell line

#Adjust number in cellLine prior to all. Also remember to change the title in plot.

flt = good & data$cellLine=="1"

cat("Doing PCA, this will take a while.\n");

spec.pca <- pca(tmp[flt,],method="nipals",nPcs=4,scale="uv")

summary(spec.pca)

pc <- sprintf("%.1f",100*R2cum(spec.pca)[1])

pc <- c( pc , sprintf("%.1f",100*diff(R2cum(spec.pca))) )

pcs <- c(1,2) #which to show in plot

xlim <- NULL #c(-104,120);

ylim <- NULL #c(-700,100);

plot(pcaMethods::scores(spec.pca)[,pcs[1]],pcaMethods::scores(spec.pca)[,pcs[2]],type="p",cex.
axis=2,main="PCA of cell line T1454",cex.lab=1.75,xlab=paste("PC",pcs[1],":
",pc[pcs[1]],"%",sep=""),ylab=paste("PC",pcs[2],":
",pc[pcs[2]],"%",sep=""),xlim=xlim,ylim=ylim,col=data$cellLine[flt],pch=symbTr(data$treatment[
flt]),cex=3);

```

```

#Loadings (1 and 2):

lwplot(pcaMethods::loadings(spec.pca)[,1],pcaMethods::scl(spec.pca),main="PC1", cex.main=1.75,
cex.axis=1.75)

lwplot(pcaMethods::loadings(spec.pca)[,2],pcaMethods::scl(spec.pca),main="PC2", cex.main=1.75,
cex.axis=1.75)

#Plot spectral areas of interest, color code by treatment (black=control, red=TMZ,
green=YM155)

with(data[flt,], plotspectra(spec.bas[,sclr(spec.bas, 2.35, 3.0)], cex.axis=1.7, cex.lab=1.5,
col=treatment))

with(data[flt,], plotspectra(spec.bas[,sclr(spec.bas, 1.2, 1.6)], cex.axis=1.7, cex.lab=1.5,
col=treatment))

with(data[flt,], plotspectra(spec.bas[,sclr(spec.bas, 2.8, 4.0)], cex.axis=1.7, cex.lab=1.5,
col=treatment))

##### PLS #####

data$control = as.numeric(data$treatment == 1)
data$tmz = as.numeric(data$treatment == 2)
data$ym = as.numeric(data$treatment == 3)
with(data, cbind(id, control, tmz, ym))

### YM med valideringssett ###

validering = data$cellLine=="2" | data$cellLine=="5"
flt = good & data$tmz==0 & !validering #data$time_followup %in% smp;
spec2.pls = pls( ym~spec.bas , data=data[flt,], ncomp=10 , validation="LOO", scale=TRUE )
#hent ut R2 og Q2 - husk å tilpasse variabelnavn i for-løkken

rsq=c();
rsq=as.vector(as.table(R2(spec2.pls,intercept=FALSE,estimate = "train")$val)[,,])
qsq=as.vector(as.table(R2(spec2.pls,intercept=FALSE)$val)[,,])
#x11();
plot(qsq/rsq,type="b",ylim=c(-1,1),main=paste("Q2/R2"));
abline(0.8,0,col="red")
abline(0,0,col="black")
lines(qsq)
lines(rsq, col="blue")
print(rsq);
print(qsq);
print(qsq/rsq);

### validation set: Calculate Q2 (?) ###

fltv = good & data$tmz==0 & validering #same as above, but YES validering
vsq=as.vector(as.table(R2(spec2.pls,intercept=FALSE,newdata = data[fltv,])$val)[,,])

```

```

# Add them to same plot above
lines(vsqr, col="green")

### PLS, other treatment TMZ ###

#Without validation, results were bogus/bad

flt = good & data$ym==0; #data$time_followup %in% smp;

spec2.pls = plsr( tmz~spec.bas , data=data[flt,], ncomp=10,
validation="CV", segments=4, segment.type="random", scale=TRUE )

#hent ut R2 og Q2 - husk å tilpasse variabelnavn i for-løkken

rsq=c();

rsq=as.vector(as.table(R2(spec2.pls,intercept=FALSE,estimate = "train")$val)[, ,])

qsq=as.vector(as.table(R2(spec2.pls,intercept=FALSE)$val)[, ,])

#x11();

plot(qsq/rsq,type="b",ylim=c(-1,1),main=paste("Q2/R2"));

abline(0.8,0,col="red")

abline(0,0,col="black")

lines(qsq)

lines(rsq, col="blue")

print(rsq);

print(qsq);

print(qsq/rsq);

#### End of document ###

```

8.11.3 R code for file 'functions.R'

```

danielclip <- function(x) { x[x<0]=0; x } #alternatively: { x-abs(x) } #which is faster? or
"better"?

#returns the column names of a matrix, converted to numeric (e.g. the ppm scale of a set of
spectra)

scl <- function(m) as.numeric(dimnames(m)[[2]])

#returns the closest matrix index to a given value among its column names (e.g. the closest
index to a given ppm value)

sclv <- function(mat,v) { # the matrix m is not duplicated in memory as long as it's not
changed, see tracemem() or duplicate()

  sapply(v,function(x) which.min(abs(scl(mat)-x)) ); #automatically resolves ties by
taking the first that fits - just like me.

}

#returns a range of matrix indices between the two given x-axis (e.g. ppm) values

sclr <- function(mat,v1,v2) {

  sclv(mat,v1) : sclv(mat,v2) ;

}

```

```

# Binning to (quite arbitrary) new bins:

# takes an x-axis (the old bin midpoints), a y-matrix (i.e. sets of spectra) and "breaks" (the
new bin edges).

# Note 1: The resulting spectrum is scaled such that its y values are on the same order of
magnitude as the original ones, facilitating comparing them.

# Note 2: "breaks" can be a vector or a list of vectors, each vector specifying the edges of
"target bins".

#       The latter case is useful when excluding e.g. the clipped water signal from the
procedure.

# Note 3: Within the scope of each "target bin" vector, the original x-axis is ASSUMED TO BE
EQUIDISTANT

#
danielbin <- function(x, y, breaks) {

  #test and process input

  x <- as.numeric(x); #if the x scale came from e.g. some dimnames

  if (!is.matrix(y)) { y <- t(as.matrix(y)); } #if single spectrum, make it a one-row
matrix

  if (!is.list(breaks)) { breaks <- list(breaks) }

  #prepare the output list (this has -unfortunately- to be done beforehand in order to
save memory and avoid cbind() operations)

  newx=c(); #empty vector, to contain the new bin midpoints

  for(k in 1:length(breaks)) { newx = append(newx,(breaks[[k]][-1]+breaks[[k]][-
length(breaks[[k]])])/2); } #iterate through the "breaks" input and calculate the midpoints
for each segment; append to x

  widths=c(); #same for the target bin widths

  for(k in 1:length(breaks)) { widths = append(widths,breaks[[k]][-1]-breaks[[k]][-
length(breaks[[k]])]); }

  outlist <- list( x=newx, widths=widths,
y=matrix(0,nrow(y),length(newx),dimnames=list(dimnames(y)[[1]],newx) ); #a matrix of zeros,
with as many rows as the input, as many columns as target bins, and proper dimnames

  c = 0; #the current index offset, will help us while filling the output "y", i.e.
overwriting the zeros

  #now for each segment of "breaks", do a regular binning:

  for(k in 1:length(breaks)) {

    cat("k =",k,"\n");

    #find the indices of x that definitely can contribute to the new bins,
calculate "old" resolution in that region...

    possib = which( (x>=min(breaks[[k]])) & (x<=max(breaks[[k]])) ); #so far: old
bins *within* the new breaks - might be missing some at the edges!

    oldres <- (x[possib][length(x[possib])]-x[possib][1])/(length(x[possib])-1);
#calculate resolution

    #...then, just to be safe and include *all possible* contributing old bins,
expand the list of indices by one on each side,

```

```

mi = min(possib); if (mi>1) { possib=c(mi-1,possib); }

ma = max(possib); if (ma<length(x)) { possib=c(possib,ma+1); }

#and construct the old bin edges (assumes equidistant x in each segment)

oldleft  <- x[possib]-oldres/2;

oldright <- x[possib]+oldres/2;

if ( (breaks[[k]][1])<oldleft[1] ||
(breaks[[k]][length(breaks[[k]])]>oldright[length(oldright)] ) { warning("New bin edges
exceed old bin edges."); }

#construct new bin edges from "breaks" (trivial)

newleft  <- breaks[[k]][-length(breaks[[k]])]; #chop last one

newright <- breaks[[k]][-1];

#now iterate over the new bins:

for(i in 1:length(newleft)) {

cat(".");

#for each new bin, look at the overlaps between the new bin and the old
bins

#and sum up the values found in the old bins accordingly. So, iterate
over the old bins:

for(j in 1:length(oldleft)) {

fact = ( min(newright[i],oldright[j]) -
max(newleft[i],oldleft[j]) ) / (oldright[j]-oldleft[j]) #find overlap "ratio" of the old with
the new bin

scalfac = oldres/(newright[i]-newleft[i]); #introduce scaling
factor to average the new bin (two effects: 1) scales spectrum back to original size after
"summing the contributions", 2) in the case of unequally wide new bins, corrects for the size
of the bins)

#cat("i=",i,"/",length(newleft)," j=",j,"/",length(oldleft),"
overlap=",fact,"\tc+i=",c+i,"/",ncol(outlist$y),"\\n",sep="");

if (fact>0) { outlist$y[,c+i] <- outlist$y[,c+i] +
scalfac*fact*y[,possib[j]]; } #sum up

}

}

cat("\\n");

#finally, increase current column index by number of new bins filled:

c = c+length(breaks[[k]])-1;

}

return (outlist);

}

#color map for heatmaps: red-black-green

```

```

danielheatcols <- function(n) {
  sapply(seq(-1,1,length.out=n), function(x) rgb( (-x+abs(-x))/2 , (x+abs(x))/2 , 0 ))
#from red to black to green
}

danieloplscols <- function(n=5) {
  rgb( approx(c(0,0.5,0,1,0.7),n=n)$y , approx(c(0,0.5,1,1,0),n=n)$y ,
approx(c(0.7,1,0,0,0),n=n)$y )
}

danieloplscols.mapped <- function(y,scale=c(min(y),max(y)) ) {
  y = pmax(y[-1],y[-length(y)]);
  rgb( approx(seq(scale[1],scale[2],length.out=5),c(0,0.5,0,1,0.7),xout=y)$y ,
approx(seq(scale[1],scale[2],length.out=5),c(0,0.5,1,1,0),xout=y)$y ,
approx(seq(scale[1],scale[2],length.out=5),c(0.7,1,0,0,0),xout=y)$y )
}

#color map for correlation matrix maps: blue-white-red
danielcm.cols <- function(n) {
  sapply(seq(-1,1,length.out=n), function(x) rgb( 1+(x-abs(x))/2 , 1-abs(x) , 1+(-x-abs(-
x))/2 )) #from blue to white to red
}

danielcm.cols2 <- function(n=16) {
  x = c( -1, -0.5, 0, 0.5, 1 )
  yR = c( 0, 0, 0, 1, 1 )
  yG = c( 1, 0, 0, 0, 1 )
  yB = c( 1, 1, 0, 0, 0 )
  # cyan, blue, black, red, yellow (read column-wise)
  rgb( approx(x,yR,n=n)$y, approx(x,yG,n=n)$y, approx(x,yB,n=n)$y )
}

#Plot one or more spectra. X-axis is taken from colnames() and inverted. Color-coded cyclically
plotspectra <- function(inspec, col=c("black","red","green","blue","magenta"),
ylim=c(min(inspec),max(inspec)), xlab="Chemical Shift (ppm)", ylab="Intensity (a.u.)",
xaxs="i", ... ) {
  if (!is.matrix(inspec)) {
    inspec <- t(as.matrix(inspec)); #if input is only one spectrum: coerce into a
one-row matrix
  }
  ppm <- as.numeric(colnames(inspec));
  maxminppm <- c(max(ppm),min(ppm));
  for (i in 1:nrow(inspec)) {
    if (i>1) {par(new=TRUE);}
    plot(ppm, inspec[i,], type="l", xlim=maxminppm, ylim=ylim, col=col[((i-
1)%length(col))+1], xlab=xlab, ylab=ylab, yaxs=xaxs, ... );
  }
}

```

```

    }
}

#Plot density profiles by factor. X-axis is taken from colnames() and inverted. Color-coded
cyclically

#Removes NA values automatically

facdensiplot <- function(var, fac, xlim=c(min(var,na.rm=TRUE),max(var,na.rm=TRUE)),
col=c("red","green","blue","magenta","black"), bw=(xlim[2]-xlim[1])/100, ... ) {

  #ufac=sort(unique(fac));

  ufac=sort(unique(fac),na.last=NA); #this removes "NA" as a factor

  for (i in 1:length(ufac)) {

    if (i>1) {par(new=TRUE);}

    plot(density(var[fac==ufac[i]&!is.na(var)],bw=bw,na.rm=TRUE), xlim=xlim,
col=col[((i-1)%length(col))+1], xlab=deparse(substitute(var)), ylab="Density (not
normalized)" , main=paste("Density Plot factored by",deparse(substitute(fac))) , ... )

  }

  #make a legend:

  coor = par("usr");

  legend( x=(coor[1]+0.1*(coor[2]-coor[1])),y=(coor[4]-0.1*(coor[4]-coor[3]))
,legend=ufac,col=col,lwd=1,bty="n")

}

#Function that plots correlation matrices with proper axes

#Parameters: the correlation matrix (corm), the correlation cutoff and the number of colors
(ccut and ncol).

#           Further: The colormap function (must have same ncol as stated before)

plotcorm <- function(corm, projX=NULL, projY=NULL, ccut=0.7, ncol=17, col=danielcm.cols(ncol),
widths=c(6,1), bounce=TRUE, Xrange=NULL, Yrange=NULL, ...) {

  #general settings:

  Xppm=as.numeric(rownames(corm))

  Yppm=as.numeric(colnames(corm)) #for some reason, "image" plots matrices sideways

  mamix=c(max(Xppm),min(Xppm))

  mamiy=c(max(Yppm),min(Yppm))

  if (bounce) {

    bounce=0.5*((mamix[1]-mamix[2])/length(Xppm)); # bounce out plot limits to
avoid "half pixels" in "image". assumes evenly-

    mamix=c(mamix[1]+bounce,mamix[2]-bounce); # spaced axis (not necessarily true,
see clipped water), may therefore overestimate the necessary bounce... which is OK.

    bounce=0.5*((mamiy[1]-mamiy[2])/length(Yppm)); # bounce out plot limits to
avoid "half pixels" in "image". assumes evenly-

    mamiy=c(mamiy[1]+bounce,mamiy[2]-bounce); # spaced axis (not necessarily true,
see clipped water), may therefore overestimate the necessary bounce... which is OK.

  }

}

```

```

colb=c( seq(from=-1,to=-ccut,length.out=(ncol+1)/2) ,
seq(from=ccut,to=1,length.out=(ncol+1)/2) )

#Case A: if no projections specified, just plot the correlation matrix:
if (is.null(projX) && is.null(projY)) {

    image(Xppm,Yppm,corm,col=col,breaks=colb,xlim=mamix, ylim=mamiy, ... )

    box()

#Case B: if *either* of the projections is set, make a 2-by-2 layout (even if maybe
only one projection is defined)

} else {

    #some settings and calculations:
    pr.ax.col = "gray";
    projCol = "black";

    if (is.null(projX)) {projX=rep(0,length(Xppm))}
    if (is.null(projY)) {projY=rep(0,length(Yppm))}

    #save figure parameters to reset later
    def.par <- par(no.readonly = TRUE) # save default, for resetting...

    # make a layout matrix like this:  1 -
                                     3 2   which means that the top will be
processed first, then the right, then the left bottom (the central correlation plot)

    nf <- layout( matrix(c(1,0,3,2),2,2,byrow=TRUE), widths=widths,
heights=rev(widths), respect=FALSE)

    layout.show(nf)

    # subfigure 1: top projection. set margins (bottom, left, top, right) to
(0,5,0,.1)

    par(mar=c(0,5,0,.1))

    plot(Xppm,projX,type="l",xlim=mamix,ylim=Yrange,xaxs="i",yaxs="i",xaxt="n",col.axis=pr.
ax.col,fg=pr.ax.col,col=projCol)

    # subfigure 2: right projection. set margins to (5,0,.1,0)

    par(mar=c(5,0,.1,0))

    plot(projY,Yppm,type="l",ylim=mamiy,xlim=Yrange,xaxs="i",yaxs="i",yaxt="n",col.axis=pr.
ax.col,fg=pr.ax.col,col=projCol)

    # subfigure 3: correlation matrix. set margins to (5,53,.1,.1)

    par(mar=c(5,5,.1,.1))

    image(Xppm,Yppm,corm,col=col,breaks=colb,xlim=mamix, ylim=mamiy, ... )

    abline(0,1,col="gray")

    box(); #redraw the box around the matrix.

    #reset figure parameters.

    #par(def.par)

```



```

    }
}

# Function that measures the area under the curve in a specified region of a plot.

# - Assumes that a spectrum (or any other graph) has already been drawn, expects the very same
spectrum as its input argument

# - Lets the user pick two points (the end points of the region)

# - Calculates the area in that region in three different ways:

#   - naive: altitude over zero line times width

#   - baseline from spectrum (auto.baseline): assumes x/y coordinates of spectrum closest to
user-selected points to define a baseline, subtracts that from area

#   - baseline from user (user.baseline): assumes the points chosen by the user to define the
baseline and subtracts it (no, even better: only takes what lies ABOVE it)

measurepeak <- function(spec) {

  spec=t(as.matrix(spec));

  if (nrow(spec)!=1) { warning("Expecting a one-row matrix as input.\n"); } else {

    cat("measurepeak(): Please locate a peak or group to be measured: Click two positions
at the \"foot\" of the signal.\n");

    #get 2 points from user:

    co=locator(n=2,type="o",col="red")

    #find nearest integer indices of spectrum

    index.x=sort(sclv(spec,co$x));

    #find autobaseline

    base.y=spec[index.x]

    scal=as.numeric(colnames(spec))

    # calculate areas:

    an <- sum( (spec[index.x[1]:(index.x[2]-1)]+spec[(index.x[1]+1):index.x[2]]) *
abs(scal[index.x[1]:(index.x[2]-1)]-scal[(index.x[1]+1):index.x[2]]) / 2 )

    m.a <- (spec[index.x[2]]-spec[index.x[1]]) / (scal[index.x[2]]-scal[index.x[1]]); n.a
<- spec[index.x[1]]-m.a*scal[index.x[1]];

    bl.a <- m.a*(scal[index.x[1]:index.x[2]])+n.a;

    aa <- sum( ( (spec[index.x[1]:(index.x[2]-1)]+spec[(index.x[1]+1):index.x[2]]) -
(bl.a[-1]+bl.a[-length(bl.a)]) ) * abs(scal[index.x[1]:(index.x[2]-1)]-
scal[(index.x[1]+1):index.x[2]]) / 2 )

    m.u <- (co$y[2]-co$y[1]) / (co$x[2]-co$x[1]); n.u <- co$y[1]-m.u*co$x[1];

    bl.u <- m.u*(scal[index.x[1]:index.x[2]])+n.u;

    au <- sum( danielclip( (spec[index.x[1]:(index.x[2]-
1)]+spec[(index.x[1]+1):index.x[2]]) - (bl.u[-1]+bl.u[-length(bl.u)]) ) *
abs(scal[index.x[1]:(index.x[2]-1)]-scal[(index.x[1]+1):index.x[2]]) / 2 )

```

```

#return output list

list( x=co$x, y=co$y, index.x=index.x, base.y=base.y, area.naive=an,
area.auto.baseline=aa, area.user.baseline=au )

} #end if(nrow(spec)!=1)
}

#function that plots color-coded, rescaled loading weights with a color legend.
#Parameteres: lw is the K vector from (a uv-scaled) O2-PLS or the loading weights /
coefficients from a (uv-scaled) PLS; scalfac is the scaling factor from the uv scaling

#Note: color scale will be adjusted between the minimal and maximal absolute correlation (i.e.
value in variable "lw"). Otherwise you can set colscal=c(0,1) and get a "true" correlation
scale

lwplot <- function (lw, scalfac=1, colscal=c(min(abs(lw)),max(abs(lw))), ncol=32,
widths=c(10,1), ... ) {

  #prepare splitscreen layout for plot and legend

  nf <- layout(matrix(c(2,1),1,2,byrow=TRUE), widths=widths, heights=c(1), respect=FALSE)
  layout.show(nf)

  #plot legend:

  par(mar=c(3,0,2,2.5)) #margins in subplot: down, left, up, right

  mima=matrix(seq(colscal[1],colscal[2],length.out=ncol),nrow=1)

  #plot legend w/o axes

  image(1,seq(colscal[1],colscal[2],length.out=ncol),mima,col=danieloplscols(ncol),xaxt="
n",yaxt="n",xlab="",ylab="Correlation")

  axis(4); #4=on right

  box();

  #plot rescaled loading weights, color coded by their original values:

  par(mar=c(3,2.5,2,0.5))

  x=as.numeric(names(lw));

  mamix=c(max(x),min(x));

  plot(x,lw*scalfac, type="n", xlim=mamix,xaxs="i",yaxs="i",...); #type="n" produces an
empty plot, which plotrix uses to draw lines into

  color.scale.lines(x,lw*scalfac,xlim=mamix,
col=danieloplscols.mapped(abs(lw),scale=colscal),colvar=lw,lwd=2)

  box();

}

```

8.12 Nuclear magnetic resonance spectroscopy pulse program scripts

Pulse program code obtained from the software TopSpin. Code denoted with .nn was written for this specific study by Nils Nyberg from Bruker.

Pulse program code for *noesygppr1d.comp*

```
;noesygppr1d.comp
;avance-version (06/11/09)
;1D version of noesyprtp
;with presaturation during relaxation delay and mixing time
;  and spoil gradient
;
;$CLASS=HighRes
;$DIM=1D
;$TYPE=
;$SUBTYPE=
;$COMMENT=
;
;$OWNER=stsc
#include <Avance.incl>
#include <Grad.incl>
;d12=20u"
;
;acqt0=-p0*2/3.1416"
;
1 ze
2 30m
;
4u BLKGRAD
;
d12 p19:f1
;
d1 cw:f1 ph29
;
4u do:f1
```

```

50u UNBLKGRAD

p16:gp1

d16 p11:f1

p1 ph1

4u

p1 ph2

d12 p19:f1

d8 cw:f1

4u do:f1

p16:gp2

d16 p11:f1

;4u BLKGRAD

p0 ph3

go=2 ph31

30m mc #0 to 2 F0(zd)

exit

ph1=0 2

ph2=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2

ph3=0 0 2 2 1 1 3 3

ph29=0

ph31=0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3

;p11 : f1 channel - power level for pulse (default)

;p19 : f1 channel - power level for presaturation

;p0 : for any flip angle

;p1 : f1 channel - 90 degree high power pulse

;p16: homospoil/gradient pulse

;d1 : relaxation delay; 1-5 * T1

;d8 : mixing time

;d12: delay for power switching [20 usec]

;d16: delay for homospoil/gradient recovery

;NS: 8 * n, total number of scans: NS * TD0

```

```

;DS: 4

;use gradient ratio:   gp 1 : gp 2
;
;                   50 : -10

;for z-only gradients:
;gpz1: 50%
;gpz2: -10%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100

;$Id: noesygppld,v 1.3.2.1 2006/11/10 11:02:26 ber Exp $

```

Pulse program code for *zgesgp.dp*

```

;zgesgp
;avance-version (07/10/04)
;1D sequence
;water suppression using excitation sculpting with gradients
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
; Series A 112 275-279 (1995)
;
;$CLASS=HighRes
;$DIM=1D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple>

#include <Avance.incl>

```

```
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"d12=20u"

1 ze
2 30m

d12 p11:f1 BLKGRAD
d1
p1 ph1

50u UNBLKGRAD
p16:gp1
d16 p10:f1
(p12:sp1 ph2:r):f1
4u
d12 p11:f1

p2 ph3

4u
p16:gp1
d16
50u
p16:gp2
d16 p10:f1
(p12:sp1 ph4:r):f1
4u
d12 p11:f1

p2 ph5
```

```

4u

p16:gp2

d16

go=2 ph31

30m mc #0 to 2 F0(zd)

4u BLKGRAD

exit

ph1=0

ph2=0 1

ph3=2 3

ph4=0 0 1 1

ph5=2 2 3 3

ph31=0 2 2 0

;p10 : 120dB

;p11 : f1 channel - power level for pulse (default)

;sp1 : f1 channel - shaped pulse 180 degree

;p1 : f1 channel - 90 degree high power pulse

;p2 : f1 channel - 180 degree high power pulse

;p12: f1 channel - 180 degree shaped pulse (Squa100.1000) [2 msec]

;p16: homospoil/gradient pulse

;d1 : relaxation delay; 1-5 * T1

;d12: delay for power switching [20 usec]

;d16: delay for homospoil/gradient recovery

;NS: 8 * n, total number of scans: NS * TD0

;DS: 4

;use gradient ratio: gp 1 : gp 2

; 31 : 11

;for z-only gradients:

```

```
;gpz1: 31%
;gpz2: 11%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100

;$Id: zgesgp,v 1.5.6.1 2007/10/04 16:52:07 ber Exp $
```

Pulse program code for *zgesgp.nn*

```
;zgesgp
;avance-version (12/01/11)
;1D sequence
;water suppression using excitation sculpting with gradients
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
; Series A 112 275-279 (1995)
;
;$CLASS=HighRes
;$DIM=1D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple>

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"di2=20u"

"TAU=de+p1*2/3.1416+50u"
```


"acqt0=0"

baseopt_echo

1 ze

2 30m

d12 p11:f1 BLKGRAD

d1

p1 ph1

50u UNBLKGRAD

p16:gp1

d16 p10:f1

(p12:sp1 ph2:r):f1

4u

d12 p11:f1

p2 ph3

4u

p16:gp1

d16

TAU

p16:gp2

d16 p10:f1

(p12:sp1 ph4:r):f1

4u

d12 p11:f1

p2 ph5

4u

p16:gp2

d16

go=2 ph31

```

30m mc #0 to 2 F0(zd)

4u BLKGRAD

exit

ph1=0
ph2=0 1
ph3=2 3
ph4=0 0 1 1
ph5=2 2 3 3
ph31=0 2 2 0

;p10 : 0W
;p11 : f1 channel - power level for pulse (default)
;sp1 : f1 channel - shaped pulse 180 degree
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p12: f1 channel - 180 degree shaped pulse (Squal00.1000) [2 msec]
;p16: homospoil/gradient pulse
;d1 : relaxation delay; 1-5 * T1
;d12: delay for power switching [20 usec]
;d16: delay for homospoil/gradient recovery
;ns: 8 * n, total number of scans: NS * TD0
;ds: 4

;use gradient ratio: gp 1 : gp 2
; 31 : 11

;for z-only gradients:
;gpz1: 31%
;gpz2: 11%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100

;$Id: zgesgp,v 1.9 2012/01/31 17:49:32 ber Exp $

```

```

;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;use gradient ratio:  gp 1 : gp 2 : gp 3 : gp 4
;
;           80 : 20.1 : 11 : -5   for C-13
;           80 :  8.1 : 11 : -5   for N-15

;for z-only gradients:
;gpz1: 80%
;gpz2: 20.1% for C-13, 8.1% for N-15
;gpz3: 11%
;gpz4: -5%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
;gpnam3: SMSQ10.100
;gpnam4: SMSQ10.100

;cnst17: Factor to compensate for coupling evolution during a pulse
;
;      (usually +1). A positive factor indicates that coupling
;
;      evolution continues during the pulse, whereas a negative
;
;      factor is necessary if the coupling is (partially) refocussed.

;$Id: hsqcedetgpsisp2.2,v 1.10 2012/01/31 17:49:26 ber Exp $

```

Pulse program code for *jresgppraf*

```

;jresgpprqf

;avance-version (12/01/11)
;homonuclear J-resolved 2D correlation
;with presaturation during relaxation delay
;using gradients
;
;$CLASS=HighRes

```

```

; $DIM=2D
; $TYPE=
; $SUBTYPE=
; $COMMENT=

#include <Avance.incl>
#include <Grad.incl>

"p2=p1*2"
"d11=30m"
"d12=20u"

"in0=in0/2"

"d0=3u"

1 ze
2 d11
3 d12 p19:f1
  d1 cw:f1 ph29
  4u do:f1
  d12 p11:f1
  50u UNBLKGRAD
  p1 ph1
  4u
  d0
  p16:gp1
  d16
  p2 ph2
  4u
  p16:gp2
  d16
  d0 BLKGRAD
  go=2 ph31
  d11 mc #0 to 2 F1QF(caldel(d0, +in0))
exit

```

```

ph1=0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph2=0 2 1 3 1 3 2 0 1 3 2 0 2 0 3 1
ph29=0
ph31=0 0 2 2 1 1 3 3

;p11 : f1 channel - power level for pulse (default)
;p19 : f1 channel - power level for presaturation
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p16: homospoil/gradient pulse [1 msec]
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O [30msec]
;d12: delay for power switching [20 usec]
;d16: delay for homospoil/gradient recovery
;inf1: 1/w, w = max. width of multiplet
;in0: 1/(2 * w), w = max. width of multiplet
;nd0: 2
;ns: 4 * n
;ds: 16
;td1: number of experiments
;FnMODE: QF

;use gradient ratio: gp 1 : gp 2
; 10 : 10
;for z-only gradients:
;gpz1: 10%
;gpz2: 10%
;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
;$Id: jresgpprqf,v 1.4 2012/01/31 17:49:26 ber Exp $

```

Pulse program code for *dipsi2esgpph*

```
;dipsi2esgpph
;avance-version (12/01/11)
;homonuclear Hartman-Hahn transfer using DIPSI2 sequence
;   for mixing
;phase sensitive
;water suppression using excitation sculpting with gradients
;
;A.J. Shaka, C.J. Lee & A. Pines, J. Magn. Reson. 77, 274 (1988)
;T.-L. Hwang & A.J. Shaka, J. Magn. Reson.,
;   Series A 112 275-279 (1995)
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=

prosol relations=<triple>

#include <Avance.incl>
#include <Delay.incl>
#include <Grad.incl>

"p2=p1*2"
"d11=30m"
"d12=20u"
"d13=4u"

"in0=infl"

"d0=in0*0.5-p1*4/3.1416"

"TAU=de+p1*2/3.1416+4u"

"FACTOR1=(d9/(p6*115.112))/2"
```

"l1=FACTOR1*2"

"acqt0=0"

baseopt_echo

1 ze

2 d11

3 d12 p132:f1

d1 cw:f1 ph29

d13 do:f1

d12 p11:f1

p1 ph1

d0

p1 ph2

50u UNBLKGRAD

p16:gp1

d16 p110:f1

;begin DIPSI2

4 p6*3.556 ph23

p6*4.556 ph25

p6*3.222 ph23

p6*3.167 ph25

p6*0.333 ph23

p6*2.722 ph25

p6*4.167 ph23

p6*2.944 ph25

p6*4.111 ph23

p6*3.556 ph25

p6*4.556 ph23

p6*3.222 ph25

p6*3.167 ph23

p6*0.333 ph25

p6*2.722 ph23

p6*4.167 ph25

p6*2.944 ph23

p6*4.111 ph25

p6*3.556 ph25

p6*4.556 ph23

p6*3.222 ph25

p6*3.167 ph23

p6*0.333 ph25

p6*2.722 ph23

p6*4.167 ph25

p6*2.944 ph23

p6*4.111 ph25

p6*3.556 ph23

p6*4.556 ph25

p6*3.222 ph23

p6*3.167 ph25

p6*0.333 ph23

p6*2.722 ph25

p6*4.167 ph23

p6*2.944 ph25

p6*4.111 ph23

lo to 4 times l1

;end DIPSI2

4u

p16:gp2

d16 p11:f1

p1 ph3

p16:gp3

d16 p10:f1

(p12:sp1 ph4:r):f1

4u

d12 p11:f1


```

p2 ph5

4u
p16:gp3
d16
TAU
p16:gp4
d16 p10:f1
(p12:sp1 ph6:r):f1
4u
d12 p11:f1

p2 ph7

4u
p16:gp4
d16
4u BLKGRAD

go=2 ph31
d11 mc #0 to 2 F1PH(calph(ph1, +90) & calph(ph29, +90), caldel(d0, +in0))
exit

ph1=0 2
ph2=0 0 0 0 2 2 2 2
ph3=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2
ph4=0 0 1 1
ph5=2 2 3 3
ph6=0 0 0 0 1 1 1 1
ph7=2 2 2 2 3 3 3 3
ph23=3
ph25=1
ph29=0
ph31=0 2 2 0 0 2 2 0 2 0 0 2 2 0 0 2

```

```

;p10 : 0W
;p11 : f1 channel - power level for pulse (default)
;p110: f1 channel - power level for TOCSY-spinlock
;p132: f1 channel - power level for low power presaturation
;sp1 : f1 channel - shaped pulse 180 degree
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p6 : f1 channel - 90 degree low power pulse
;p12: f1 channel - 180 degree shaped pulse (Squa100.1000) [2 msec]
;p16: homospoil/gradient pulse
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d9 : TOCSY mixing time
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d13: short delay [4 usec]
;d16: delay for homospoil/gradient recovery
;l1: loop for DIPSI cycle: ((p6*115.112) * l1) = mixing time
;inf1: 1/SW = 2 * DW
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;ns: 8 * n
;ds: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEQ

;use gradient ratio: gp 1 : gp 2 : gp 3 : gp 4
;
; 1 : 3 : 31 : 11

;for z-only gradients:
;gpz1: 1%
;gpz2: 3%
;gpz3: 31%
;gpz4: 11%

;use gradient files:

```

```

;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
;gpnam3: SMSQ10.100
;gpnam4: SMSQ10.100

;set pl32 to 0W when presaturation is not required
; use pl1 + 75 to 80dB to reduce radiation damping

;Processing
;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1

;$Id: dipsi2esgpph,v 1.14 2012/01/31 17:49:22 ber Exp $

```

Pulse program code for *hsqcdietgpsisp.2*

```

;hsqcedetgpsisp2.2
;avance-version (12/01/11)
;HSQC
;2D H-1/X correlation via double inept transfer
; using sensitivity improvement
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition
;using trim pulses in inept transfer
;with multiplicity editing during selection step
;using shaped pulses for all 180degree pulses on f2 - channel
;with gradients in back-inept
;
;A.G. Palmer III, J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
; Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
; 10663-5 (1992)
;J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky,
; S.J. Glaser, O.W. Sorensen & C. Griesinger, J. Biomol. NMR 4,
; 301-306 (1994)
;W. Willker, D. Leibfritz, R. Kerssebaum & W. Bermel, Magn. Reson.

```

```

; Chem. 31, 287-292 (1993)
;
; $CLASS=HighRes
; $DIM=2D
; $TYPE=
; $SUBTYPE=
; $COMMENT=

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"d4=1s/(cnst2*4)"
"d11=30m"

"d0=3u"

"in0=inf1/2"

"DELTA=d21-cnst17*p24/2-p16-d16-p2-d0*2"
"DELTA1=p16+d16-p1*0.78+de+8u"
"DELTA2=d4-larger(p2,p14)/2"
"DELTA3=d24-cnst17*p24/2-p19-d16"
"DELTA4=d4-larger(p2,p14)/2-p16-d16"
"DELTA5=d21-cnst17*p24/2"

"acqt0=0"
baseopt_echo

1 ze
    d11 p112:f2
2 d1 do:f2
3 (p1 ph1)
    DELTA2 p10:f2
    4u

```

(center (p2 ph1) (p14:sp3 ph6):f2)
4u
DELTA2 p12:f2 UNBLKGRAD
p28 ph1
4u
(p1 ph2) (p3 ph3):f2
d0
(p2 ph7)
d0
p16:gp1*EA
d16
DELTA
4u
(center (p2 ph1) (p24:sp7 ph4):f2)
4u
DELTA5 p12:f2
(center (p1 ph1) (p3 ph4):f2)
p19:gp3
d16
DELTA3
(center (p2 ph1) (p24:sp7 ph1):f2)
DELTA3
p19:gp3
d16 p12:f2
(center (p1 ph2) (p3 ph5):f2)
p16:gp4
d16
DELTA4 p10:f2
(center (p2 ph1) (p14:sp3 ph1):f2)
DELTA4
p16:gp4
d16
(p1 ph1)
DELTA1
(p2 ph1)
4u

```

p16:gp2

d16 p112:f2

4u BLKGRAD

go=2 ph31 cpd2:f2

d1 do:f2 mc #0 to 2

      FlEA(calgrad(EA) & calph(ph5, +180), caldel(d0, +in0) & calph(ph3, +180) & calph(ph6, +180)
& calph(ph31, +180))

exit

ph1=0

ph2=1

ph3=0 2

ph4=0 0 2 2

ph5=1 1 3 3

ph6=0

ph7=0 0 2 2

ph31=2 0 0 2

;p10 : 0W

;p11 : f1 channel - power level for pulse (default)

;p12 : f2 channel - power level for pulse (default)

;p112: f2 channel - power level for CPD/BB decoupling

;sp3: f2 channel - shaped pulse (180degree inversion)

;spnam3: Crp60,0.5,20.1

;sp7: f2 channel - shaped pulse (180degree refocussing)

;spnam7: Crp60comp.4

;p1 : f1 channel - 90 degree high power pulse

;p2 : f1 channel - 180 degree high power pulse

;p3 : f2 channel - 90 degree high power pulse

;p14: f2 channel - 180 degree shaped pulse for inversion

;      = 500usec for Crp60,0.5,20.1

;p16: homospoil/gradient pulse

;p19: gradient pulse 2 [500 usec]

;p24: f2 channel - 180 degree shaped pulse for refocussing

;      = 2msec for Crp60comp.4

;p28: f1 channel - trim pulse

;d0 : incremented delay (2D) [3 usec]

```

```

;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J)XH
;d11: delay for disk I/O [30 msec]
;d16: delay for homospoil/gradient recovery
;d21: set d21 according to multiplicity selection
;      1/(2J(XH)) XH, XH3 positive, XH2 negative
;d24: 1/(8J)XH for all multiplicities
;      1/(4J)XH for XH
;cnst2: = J(XH)
;cnst17: = -0.5 for Crp60comp.4
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;ns: 1 * n
;ds: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;use gradient ratio: gp 1 : gp 2 : gp 3 : gp 4
;      80 : 20.1 : 11 : -5 for C-13
;      80 : 8.1 : 11 : -5 for N-15

;for z-only gradients:
;gpz1: 80%
;gpz2: 20.1% for C-13, 8.1% for N-15
;gpz3: 11%
;gpz4: -5%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
;gpnam3: SMSQ10.100
;gpnam4: SMSQ10.100

```

;cnst17: Factor to compensate for coupling evolution during a pulse
;
; (usually +1). A positive factor indicates that coupling
;
; evolution continues during the pulse, whereas a negative
;
; factor is necessary if the coupling is (partially) refocussed.

;\$Id: hsqcedetgpsisp2.2,v 1.10 2012/01/31 17:49:26 ber Exp \$