

# **Characterisation of a novel bioactive complex polysaccharide from a marine invertebrate with potent anticancer and antimalarial activities**

**Chrow Ahmed Khurshid**

**A thesis submitted to The University of Salford Manchester in partial fulfilment of the requirements for the degree of Doctor of Philosophy**



**School of Environment & Life Sciences  
College of Science & Technology  
University of Salford**

**Salford, UK**

**September 2017**

## Contents

List of Tables .....	v
List of Figures .....	vi
Abstract.....	viii
Declaration.....	x
Acknowledgement .....	xi
List of Abbreviations .....	xii
<b>Chapter1: Introduction.....</b>	<b>1</b>
1.1 General research background .....	2
1.2 Glycosaminoglycans GAGs .....	3
1.2.1 GAG biosynthesis Pathway.....	4
1.2.2 HS and Heparin Chain Polymerization and Modifications .....	6
1.2.3 Heparin and Heparan Sulphate Domains.....	8
1.2.4 Patterns of 2- <i>O</i> - sulphation, 6- <i>O</i> - sulphation and 3- <i>O</i> - sulphation groups in HS and Heparin.....	9
1.2.5 CS/DS Chain Polymerisation and Modifications.....	11
1.3 De - Polymerisations Techniques .....	13
1.3.1 Enzymatic De-Polymerisation of HS/Hep .....	13
1.3.2 Chemical De-Polymerisation of HS/Hep.....	15
1.3.3 Enzymatic De-Polymerisation of CS/DS.....	16
1.4 Heparan sulphate/heparin-protein interactions.....	17
1.4.1 FGF and the FGFR .....	17
1.4.2 The FGF signalling pathways .....	18
1.4.3 HS/FGF/FGFR Complex interactions.....	19
1.7 Marine sulphated glycan .....	21
1.7.1 Structural composition of sulphated polysaccharides (SP) from marine sources .....	22
1.7.2 Potential drug development from marine glycan. ....	26
1.7.3 Therapeutic value of mammalian glycosaminoglycan in cancer.....	28
1.8 Breast cancer.....	29
1.8.1 Causes, Risk Factors, and Prevention of breast cancer .....	29
1.8.2 Epidemiology of breast cancer .....	30
1.8.3 Histology of normal breast vs histology of breast cancer .....	31
1.8.4 Molecular profiling of breast cancer .....	33
1.8.5 Role of metastasis in breast cancer.....	35
1.8.6 Treatment options for breast cancer .....	36
1.9 Malaria .....	38
1.10 Aims.....	41

<b>Chapter 2 : Materials and Methods</b> .....	43
2.1 Materials .....	44
2.2 Methods .....	44
2.2.1 Method of GAG isolation from soft body of <i>Whelk</i> . .....	44
2.3 Enzyme Preparations.....	45
2.3.1 Preparing of Hep lyase .....	45
2.3.2 Preparation of ABC lyase.....	46
2.4 Enzymatic digestion.....	46
2.4.1 Enzymatic degradation of Heparan Sulphate (GAG-HS) .....	46
2.4.2 Enzymatic degradation of Chondroitin Sulphate (GAG-CS) .....	47
2.5 Standard Preparation of HS and CS.....	47
2.6 Size exclusion analysis .....	47
2.7 Quantitative analysis of Monosaccharides by HPAEC-PAD.....	48
2.8 Ion exchange of intact GAG.....	49
2.9 PD-10 de-salting technique .....	50
2.10 Enzymatic degradation of ion-exchange fractions.....	51
2.10.1 Degradation of HS fractions .....	51
2.10.2 Degradation of CS fractions.....	52
2.11 LC/GC-MS Techniques .....	52
2.11.1 Isotopic Aniline tagging for GAG disaccharide analysis by mass spectrometry (GRIL-Glycan Reductive Isotope Labelling) .....	53
2.11.2 Trimethyl Silyl derivative for GC-MS (TMS).....	54
2.12 Periodic acid oxidation ( $H_3IO_4$ or $H_5IO_6$ ) .....	56
2.13 NMR spectrometry .....	57
2.14 Cell culture .....	58
2.14.1 Cell lines studied .....	58
2.14.2 Routine cell culture .....	58
2.14.3 MTT colorimetric Assay .....	60
2.14.4 Statistical approach.....	61
2.14.5 Protein extraction and quantification .....	62
2.15 Label-free protein quantitation proteomics .....	65
2.16 Mammosphere assay .....	66
2.17 General material and methods for <i>in vitro</i> culturing of malaria .....	67
2.17.1 Culture of <i>Plasmodium falciparum</i> .....	67
2.17.2 Drug susceptibility assays.....	69
<b>Chapter 3 : Biological activity of <i>whelk</i>-GAG crude extraction as an anticancer and antimalarial therapeutic agent</b> .....	71
3.1 Introduction.....	72

3.2	Results .....	74
3.2.1	Cytotoxicity assay on lung carcinoma .....	74
3.2.2	Cytotoxicity assay on two subtypes of positive breast cancer.....	76
3.2.3	Cytotoxicity assay of intact whelk GAG extracts on hepatoblastoma-derived cell line Hep G2.....	78
3.2.4	Cytotoxic assay of crude whelk GAG extracts on chronic myelogenous leukaemia.....	79
3.2.5	Crude Whelk GAG extracts as an antimalarial therapeutic agent .....	81
3.2.5.1	Blood smear of parasite .....	81
3.2.5.2	Cytotoxic activity of whelk GAG samples toward the malaria parasite using fluorescence-based drug susceptibility assay.....	82
3.3	Discussion.....	86
<b>Chapter 4 : Isolation and structural analysis of whelk GAGs.....</b>		<b>92</b>
4.1	Introduction.....	93
4.2	Results .....	94
4.2.1	Monosaccharides analysis of intact GAG by HPAEC-PAD .....	94
4.2.2	Enzymatic digestion and Gel filtration chromatography of Intact GAG.....	98
4.2.3	Ion exchange chromatography and PD-10 de-salting techniques .....	103
4.2.4	Quantitative analysis of monosaccharides on different –GAG fractions obtained from DEAE-fragmentation by HPAEC-PAD.....	104
4.2.5	GC-MS spectrum of fraction CK-1 GAG as trimethyl silyl derivative .....	108
4.2.6	Disaccharide analysis by enzymatic digestions of GAG fractions from ion-exchange chromatography .....	111
4.2.7	Glycan Reductive Isotope Labelling (GRIL) LC-MS analysis of Isotopic Aniline tagging for GAG disaccharide .....	112
4.2.8	Monosaccharides analysis by HPAEC-PAD technique of whelk-GAG sample oxidised by periodic acid.....	125
4.2.9	NMR spectra analysis .....	128
4.2.10	1D ( <sup>1</sup> H-NMR) and 2D (HSQC-NMR) experiments .....	129
4.3	Discussion.....	133
<b>Chapter 5 : The effect of whelk GAG on triple negative breast cancer and elucidation of their mechanisms of action .....</b>		<b>137</b>
5.1	Introduction.....	138
5.2	Results .....	141
5.2.1	Effect of intact GAG on triple negative breast cancer.....	141
5.2.2	Effect of HS/CS-disaccharide observed from digestion of whelk GAG and commercial porcine HS/CS on triple negative breast cancer .....	143
5.2.3	Cytotoxic effect of ion-exchange chromatography fractions of whelk GAG before and after oxidation with periodic acid on triple negative breast cancer cell lines.....	147
5.2.4	Cytotoxic activity of whelk GAG fractions CK1, CK2 and CK3 before and after enzymatic degradation on triple negative breast cancer cell lines.....	150

5.2.5	Cytotoxic effect of crude whelk GAG extracts on mammosphere formation in breast cancer.....	154
5.2.6	Label free protein quantification .....	157
5.2.6.1	Quantitative proteomics analysis.....	157
5.2.6.2	Mapping identified proteins to functional clusters in two triple negative breast cancer cell lines.....	159
5.2.6.3	Cell signalling pathway analysis .....	176
5.3	Discussion .....	193
<b>Chapter 6 : Conclusion and Future Works.....</b>		<b>200</b>
References .....		204
Appendices.....		217
Appendix I: Recipes for solutions and buffers.....		218
Appendix II: A: LC-MS spectra of aniline-tagged standards and whelk-GAG flow throw (FT) sample fractions. B: MS spectra of the unknown peak in CK1 fraction (CK-1 sample in negative mode direct injection MS), Standard CS in negative mode direct injection MS, Standard GlcA in negative mode direct injection MS, and Reagent blank run on MS .....		224
Appendix III: Monosaccharide quantification.....		241
Appendix IV: Monosaccharide analysis of the flow through fraction of the DEAE ion exchange column. ....		243
Appendix V: NMR spectra of whelk GAG. ....		244

## List of Tables

Table 2-1: The gradient settings used for monosaccharides. ....	49
Table 2-2: GCMS conditions setting. ....	56
Table 2-3: List of cell lines studied, classification, immune profile and complete growth media. ....	58
Table 3-1: IC50 Values represented in ( $\mu\text{g}/\text{ml}$ ) of different cancer cell lines over 5 day's incubation with crude whelk GAGs. ....	81
Table 4-1: Monosaccharide quantification in $\mu\text{g}/\text{ml}$ of whelk crude sample. ....	97
Table 4-2: Quantification of monosaccharide in three fractions of ionic exchange chromatography. ....	108
Table 4-3: HS and CS Disaccharides codes. ....	113
Table 4-4: Disaccharide quantities of fractions CK1, CK2 and CK3 in percentage mole. ....	117
Table 4-5: Type of sulphating in HS and CS derivatives of each whelk-GAG fractions. ....	119
Table 5-1: IC50 values of whelk GAG fragments libraries from Hep lyase and ABC lyase on TNBC cell lines. ....	144
Table 5-2: Adjusted functional annotation clustering of significant proteins from DAVID online analysis of label-free proteomic data from the MDA-MB-468 cell line. ....	169
Table 5-3: Adjusted functional annotation clustering of significant proteins from DAVID online analysis of label-free proteomic data from MDA-MB-231 cell line. ....	172
Table 5-4: Pathways enriched in MDA-MB-468 from PANTHER online analysis. ....	176
Table 5-5: Pathways enriched in MDA-MD-231 from PANTHER online analysis. ....	177
Table 5-6: Pathways enriched from REACTOME in MDA-MB-468 cell line. ....	180
Table 5-7: Pathway enriched from REACTOME in MDA-MB-231 cell line. ....	181

## List of Figures

Figure 1-1: The biosynthesis of GAG chain. ....	6
Figure 1-2: Typical HS and Heparin Chain Polymerization and Modifications.....	8
Figure 1-3: Typical organisation of heparin vs heparan sulphate domains. ....	9
Figure 1-4: Biosynthesis enzymes and 2- <i>O</i> - sulphation, 6- <i>O</i> - sulphation and 3- <i>O</i> - sulphation patterns found in HS/Hep.....	11
Figure 1-5: CS/DS Chain Polymerisations and Modifications.....	13
Figure 1-6: Heparinase (I, II, III) lyase action on each fractions substrate.....	15
Figure 1-7: Chondroitin (ABC) lyase action on GAG.....	16
Figure 1-8: HS/FGF/FGFR Complex interactions. ....	21
Figure 1-9: types of sulphated polysaccharide from marine organisms.....	23
Figure 1-10: Breast anatomy. A: Histology of normal breast. B: Histology of breast cancer (Howard M. Reisner 2015). ....	32
Figure 1-11: Metastasis in breast cancer.....	36
Figure 1-12: The life cycle of malaria. ....	40
Figure 2-1: Whelk GAG extraction process. ....	45
Figure 2-2: Size exclusion analysis of GAG samples and STDs. ....	48
Figure 2-3: Ion exchange chromatography and de-salting technique of intact GAG. ....	50
Figure 2-4: Sugars reduction with sodium borohydride following conversion of polyols to polyacetate esters, then formation of TMS ethers of sugars. ....	56
Figure 2-5: MTT assay, plate dose pattern. ....	61
Figure 3-1: Cytotoxic effects of different concentrations of whelk GAG crude extracts on pulmonary adenocarcinoma A549 cells. ....	75
Figure 3-2: Cytotoxic activity of whelk GAG extracts on two subtypes of breast cancer. Panel A: MCF-7. Panel B: SKBR3. ....	77
Figure 3-3: Cytotoxic activity of whelk GAG extracts on HepG2 cell line. ....	79
Figure 3-4: Cytotoxic activity of crude whelk GAG extracts on K562 cell line. ....	80
Figure 3-5: Thin blood smear for untreated and treated infected blood. ....	82
Figure 3-6: IC50 values of crude whelk GAG extracts on two mononuclear parasite stages. ....	84
Figure 3-7: Comparison of uninfected blood (negative control), infected blood (positive control) and sample treated with crude whelk GAG extracts. ....	85
Figure 4-1: Monosaccharide composition analysis.....	96
Figure 4-2: Gel filtration chromatography of intact GAG, digested GAG with HEP lyase and commercial HS. ....	101
Figure 4-3: Gel filtration chromatography of digested GAG with ABC lyase and commercial CS. ....	103
Figure 4-4: The monosaccharide analysis form HPAEC-PAD system for ionic exchange chromatography fractions. ....	106
Figure 4-5: Composition analysis of CK1- GAG fraction and standards by GC-MS. ....	111
Figure 4-6: Disaccharide yield from whelk-GAG sample using GRIL-LC-MS. ....	115
Figure 4-7: The interaction of amino group in the aniline tag with the reducing sugar form. ....	120
Figure 4-8: The separation technique of Nanosep Centrifugal Devices with Omega Membrane. ....	121
Figure 4-9: Monosaccharide analysis of retained fraction of the mixture in all sample fractions. ....	124
Figure 4-10: Monosaccharide composition analysis by HPAEC-PAD of whelk-GAG sample fractions oxidised with periodic acid. ....	127
Figure 4-11: Bruker NMR spectrometer at <sup>1</sup> D and <sup>2</sup> D, frequency of 800MHz at 70 °C and 600MHz at 37 °C. ....	132
Figure 5-1: Inhibition activity of intact whelk GAG sample on negative breast cancer cells.....	142
Figure 5-2: Anti-proliferation activity of crude whelk GAG extracts compared to the activity of commercial mammalian HS/CS before and after enzymatic depolymerisation.....	146

Figure 5-3: Inhibition activity of whelk GAG fractions (CK1, CK2, and CK3) from anion-exchange comparing with the biological activity of fractions oxidised with periodic acid on triple negative breast cancer cells.....	149
Figure 5-4: Anti-proliferation activity of whelk GAG fractions CK1, CK2 and CK3 before and after enzymatic degradation on triple negative breast cancer. ....	153
Figure 5-5: Mammosphere formation in three breast cancer cell lines. ....	156
Figure 5-6: Principal Component Analysis PCA of Triple Negative Breast Cancer TNBC cells proteome at protein level.....	159
Figure 5-7: Volcano plot images were produced using R, for each analysis, for corrected p-values. ....	162
Figure 5-8: Function classification from PANTHER online tool analysis of differential proteins expressed in the MDA-MB-468 cell line upon GAG treatment.....	164
Figure 5-9: Function classification from PANTHER online tool analysis of differential proteins expressed in the MDA-MB-231 cell line upon GAG treatment.....	166
Figure 5-10: The integrin cell surface interactions pathway. ....	192

## **Abstract**

Considerable excitement has been generated by the discovery of active compounds from marine organisms. These compounds have a remarkable role in cancer treatment, with fewer side effects on general health. Sulphated glycosaminoglycan (GAGs) are present in all animals, vertebrate and invertebrate. They are of proven economic importance, not only in the food industry but also in the pharmaceutical field.

The present study identifies the key structural differences between the GAGs isolated from whelk and mammalian GAG and investigates their biological activity in relation to cancer and malaria. Glycans from marine sources are unique in terms of their structure and function. They work as an alternative natural source to provide effective treatment for several types of cancer, especially triple negative breast cancer (TNBC), which is the most aggressive type of cancer with fewer available treatment options. In addition, marine GAGs have a proven influence on malaria, although their mechanisms of action are not fully understood.

Several methods have been employed to achieve the aims of this study. The structure of whelk GAG was determined using several analytical techniques such as gel filtration chromatography, ion exchange chromatography, mass spectrometry and nuclei magnetic resonance spectroscopy. Enzymatic depolymerisation was also used to generate a library of fragments; each was then evaluated for its biological activity on cancer growth. Anti-proliferation activity on several cancer cells, including pulmonary adenocarcinoma A459, the MCF-7 ER-positive cell line, overexpression of protein HER2- SKBR3 cell line, the hepatoblastoma-derived cell line (Hep G2), chronic myelogenous leukaemia (K562), TNBC and mammosphere formation in breast cancer subtypes, and on malaria has been investigated *in vitro* using drug sustainable assays. Focusing on TNBC, the label-free quantitative proteomic approach has been used to gain overall insight into the mechanisms of action of whelk GAG on two TNBC cell lines: MDA-MB-468 and MDA-MB-231.

The results from whelk GAG analysis suggest a complex fine structure with a high sulphation levels that is clearly distinct from mammalian GAG. Few impurities were detected within the whelk GAG structure, which exhibits enzymatic resistance; this generated structurally indeterminate fragments. These resistant fragments still have significant biological activity against cancer growth. *In vitro* assays demonstrated significant inhibition activity of whelk GAG toward all types of cancer cells, mammosphere formation from breast cancer cells and malaria.

The mechanisms of action by which whelk GAG inhibits the growth of two TNBC cell lines appear to involve influencing the cell integrin signalling cascade, extracellular organisation pathways including the regulation of fibroblast growth factor FGF signalling and fibroblast growth factor receptors FGFR, cell adhesion and reduced glycolysis metabolism. In addition, whelk GAG affected the cell mitosis pathways by downregulating DNA replication proteins.

To our knowledge, this is the first study to identify significant structural differences between whelk GAGs and mammalian GAGs, which helps to explain the structural-functional relationship of marine glycans in inhibiting cancer cell growth. It also examines the biological role of GAGs isolated from whelk as anti-proliferation agents toward a wide variety of cancer cell lines and mammosphere formation in breast cancer. Finally, this study is the first to highlight the unique activity of whelk GAG against malaria infection.

## **Declaration**

This thesis is submitted under the University of Salford requirements for the award of a PhD degree by research. I certify that this thesis has not been submitted for any other degree in any other university of higher learning, is not being submitted as part of candidature for any such degree, and does not contain any material, which has been accepted as part of the requirements for any such degree.

I clarify that the thesis does not contain any material previously published or written by another person except where due acknowledgment is made. Throughout the PhD study, some of the contents and findings have been published in conference posters only prior to the submission of the thesis.

Chrow Khurshid

## **Acknowledgement**

First and foremost, I would like to express my sincere gratitude to my supervisor Dr David Pye for all his help and continuous support for this research. Without his help, I would not be in the position that I am today. Special thanks for Dr. Biswa Choudhury, for his patience, encouragement and advice. We had few helpful discussions and he gave me very valuable points. I would like to express my gratitude to Dr. Niroshini and her PhD student Muna Abobaker for their collaboration at the parasitology labs.

I would like to thank all my colleagues at the biomedical and cancer research department for their kind help. Special thanks for the amazing lab teams at University of California San Diego, University of Manchester, NMR facility at MIB and University of Liverpool for their help throughout this project. Special thanks for University of Kirkuk and Ministry of Higher Education and Scientific Research for their funding throughout this project.

I would like to dedicate this thesis to my wonderful parents, sisters and my husband, who have supported me throughout my life in each, and every aspect.

**CHROW**

## List of Abbreviations

ΔUA	4, 5 unsaturated Uronic Acid
2OSTs	Uronosyl+2+O+sulphotransferase enzymes
3OSTs	Glucosaminyl 3+O+sulphotransferase enzymes
6OSTs	Glucosaminyl 6+O+sulphotransferase enzymes
ABC lyase	Chondroitin sulphate enzymes
AT III	Antithrombin III
BSA	Bovine serum albumin
CML	Chronic Myelogenous Leukaemia
CS	Chondroitin sulphate
CSC	Cancer stem cells
CTC	Circulating tumour cells
DAVID	Database for Annotation, Visualization and Integrated Discovery
DBA	Dibutylamine
DCIS	Ductal in situ carcinomas
DEAE	Diethylaminoethyl
DMEM	Dulbecco's Modified Eagle's medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dp	Degrees of polymerisation
DS	Dermatan sulphate
ECM	Extra cellular matrix
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Endothelial Growth Factor
EGFR	Epidermal growth factor receptor
EI	Electron impact
ER	Estrogen receptors
EXT	Exostoses
EXTL	Exostoses like
FASP	Filter-aided sample preparation
FBS	Fetal Bovine Serum
FDR	False discovery rate
FGFRs	Fibroblast Growth Factor receptors (IIIc isoform unless stated otherwise)
FGFs	Fibroblast Growth Factors
Fuc	Fucan
GAG	Glycosaminoglycan
Gal	Galactose
GalNAc	N+acetyl galactosamine
GalT+I	Galactosamine transferase I
GalT+II	Galactosamine transferase II
GalT+III	Galactosamine transferase III
GL	Gel filtration
Glc	Glucosamine
GlcNAc	N+acetyl glucosamine
GlcNS	Glucosamine N+sulphate
GlcNS (6S, 3S)	Glucosamine N+sulphate, 6+sulphate, 3+sulphate
GlcNS(6S)	Glucosamine N+sulphate, 6+sulphate
GlcUA	Glucuronic acid
GRIL	Glycan Reductive Isotope Labelling

GS	Gas chromatography
HA	Hyaluronic acid
HBGF	Heparan sulphate binding Growth Factor
Hep	Heparin
HEPES	4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid
HER2	Epidermal growth factor receptor
HexA	Hexose amine
HexUA	Hexuronic acid (either IdoUA or GlcUA).
HGF	Hepatocytes Growth Factor
HPAEC-PAD detection	High-Performance Anion-Exchange Chromatography- pulsed amperometric
HPLC	High performance liquid chromatography
HS	Heparan sulphate
HS	Heparan Sulphate
HSPG	Heparan Sulphate Proteoglycan
IDC	Invasive ductal carcinoma
IdoUA	Iduronic acid
IdoUA(2S)	2+O+sulphated Iduronic acid
ILC	Invasive lobular carcinoma
KS	Keratin sulphate
LCIS	lobular in situ carcinomas
LMWH	Low molecular weight heparin
MS	Mass spectrometry
NA+domain	N+acetylated domain
NDST	Glucosaminyl <i>N</i> -deacetylase/ <i>N</i> -sulfotransferase
NMR	Nuclear magnetic resonance spectroscopy
NS/NA+domain	N+sulphated/N+acetylated alternating domain
NSCLC	Non-small-cell lung carcinoma
OST	O+sulphotransferase
PANTHER	Protein ANalysis THrough Evolutionary Relationships
PAPS	3'+phosphoadenosine 5'+phosphosulphate
PARP	Poly (ADP-ribose) polymerase
PBS	Phosphate buffered saline
PCA	Principal component analysis
PG	Proteoglycan
PR	Progesterone receptor
PTM	Post-translational modifications
RBC	Red blood cells
RPMI	Roswell Park Memorial Institute Medium
SCLC	Small-cell lung carcinoma
SDS	Sodium dodecyl sulphate
Ser	Serine residue
ST	Serine/threonine rich domains
TFA	Trifluoroacetic acid
TKIs	Kinase inhibitors
TMS	Trimethylsilyl
TNBC	Triple negative breast cancer
Tyr	Tyrosine
UDP <sub>sugar</sub>	Uridine diphosphate
UV	Ultra violet
VEGF	Vascular endothelial growth factor
Xyl	Xylose
XylT	Xylose transferase enzyme

# **Chapter1: Introduction**

## 1.1 General research background

Marine biology is an exceptionally rich source of many biological compounds and exogenous sulphated glycans. They have important values in the pharmaceutical industry. Interest has significantly increased in this field with many reports of compounds that have diverse biochemical structure (Liu & Rein, 2010). Compositional diversity is predominant in marine organisms, almost each class revealing individually unique available compounds.

Marine species have been the source of a wide variety of metabolites with interesting biological effects, interest has increased significantly in this field which has led to considering as one of the natural and alternative sources of compounds with vital biological functions worthy of further in depth investigation (Nur Hanim Zainudin, 2014).

Some marine natural products have derived drugs which are registered and published either in the EU or in the US while more are under clinical and pre-clinically investigation. Over 14,000 natural compounds have been developed from marine organisms recently and hundreds of patents illustrating the activities of marine natural products have been submitted, most of their remarkable activity have been on cancer (MarinLit, 2003; Proksch *et al.*, 2002).

Marine-derived bio active molecules include polyunsaturated compounds with sulphated glycosaminoglycan as a part of their structure and they have been shown anticancer, anti-inflammatory, antioxidant and antimicrobial activities *in vitro* (Schmitz, 1994). Sulphated glycosaminoglycan is present in all animals, vertebrate and invertebrate (Yamada *et al.*, 2011). The polymeric form of sugars and amino sugars from marine organisms, known as glycosaminoglycans (GAGs), are of proven importance not only in the food industry but also in the pharmaceutical field. Interest in GAG has increased in response to their remarkable biological activity as co-receptors for a variety of growth factors, cytokines, chemokines, and their specific role in some enzyme activities and tumorigenesis (Rostand & Esko, 1997).

Many patent applications disclose GAG isolated from various marine species for anticoagulant and antithrombotic activity exclusively, while other patents present antitumor activity such as GAG extracted from cartilaginous fish. Anti-inflammatory properties reported from GAG isolated from *Perna canaliculis* due to sulphated hexosamine of GAG derived. Heparan sulphate glycosaminoglycans express the highest structural variability of their GAGs from other forms of sulphated glycosaminoglycan such as keratan sulphate (KS) (Newman *et al.*, 2000). Due to their pharmaceutical properties, various structures of heparan sulphate glycosaminoglycan have been extensively studied.

Cancer is one of the major causes of mortality worldwide; it is defined as abnormal cells dividing in an uncontrolled manner, which eventually could spread into other tissues. There are many types of cancer, this diversity in cancer types and the fact that they are initiated or generated from human body's own cells make it the most complicated disease to treat. There are several treatment options for cancer, mainly involve a combination of three approaches, surgery, radiotherapy and chemotherapy. All of these three options have severe side effects on the patient, especially chemotherapy. For example, doxorubicin is a useful chemotherapy agent, it used to treat several cancer types but after surviving cancer, the patient in serious danger of develop life- threatening conditions such as cardiomyopathy, nephropathy or congestive heart failure (Sharkey *et al.*, 2013).

## **1.2 Glycosaminoglycans GAGs**

GAGs are found covalently attached to proteins, giving conjugates known as proteoglycans. GAGs are a linear glycan chain composed of a repetitive structure of disaccharide units of uronic acid and glucosamine. Initially it thought that GAGs only exist in mammals; however, they found in wide variety of organisms such as fish, insects, fungi, yeast, and bacteria. It is a particularly rich component in the muscle of shellfish (Robyt, 1998).

GAGs found on the cell surface and the extracellular spaces in many tissues, especially connective tissues. Such as that found in cartilage and the walls of blood vessel. Connective tissues contain large amounts of GAGs, which also contains collagen and elastin fibres embedded in a gel-like matrix called the ground substance. There are many types of GAG families: hyaluronic acid (HA), chondroitin sulphate (CS), dermatan sulphate (DS), keratan sulphate (KS), heparin (Hep), and perhaps the most widely studied is heparan sulphate (HS). The main differences between the family members are the position and configuration of their glycosidic linkages, amount and location of their sulphate substituents and monosaccharide composition. Some of these species have very simple structure, for instance; hyaluronic acid, whereas the others such as CS, DS, Hep and HS contain extensive blocks of modified disaccharide units, with complex sulphation patterns which appear to be under tight biosynthetic control (Mathews, 1974).

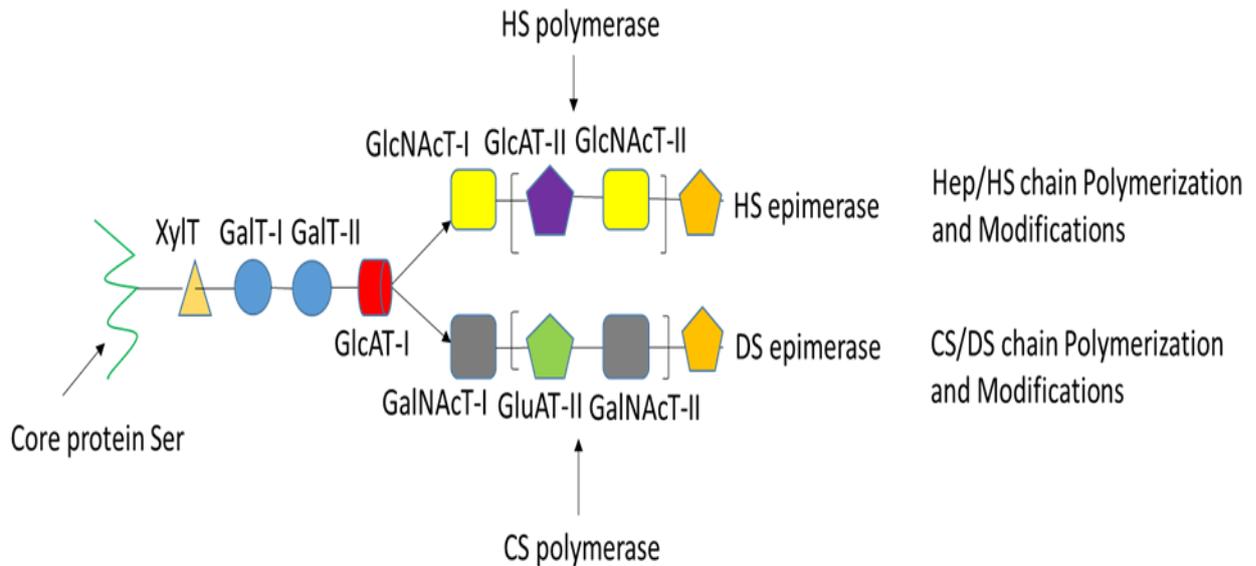
### **1.2.1 GAG biosynthesis Pathway**

GAG biosynthesis occurs through three steps; chain initiation, chain polymerization and polymer modification. It is generally believed that the polymerization takes place by stepwise addition of monosaccharide units to non-reducing polysaccharide chain ends. Those chains have various properties, which fixed at the initiation phase of polymerization and tightly controlled by glycosyltransferases, sulphotransferases and an epimerase enzyme in the lumen of the Golgi apparatus. The expression level of these enzymes during chain initiation is responsible for the display of high structural complexity within GAG chain later (Wei *et al.*, 1993). The position and arrangement of glycosides linkages that join the monomer sugar residues and determine, in part, the type of GAG synthesised are specified by the local configuration of four groups, sulphate, hydroxyl, carboxyl and acetyl within the monosaccharide building blocks (Roden & Schwartz, 1975).

Before the completion of GAG chain elongation, the polymerization products are extensively modified to produce sulphate sequences within the polysaccharide using 3-phosphoadenylylphosphorsulphate (PAPS) as a sulphate donor. The GAG chain contains uronic acid (either L-iduronic acid or D-glucouronic acid) and hexosamine (represented in D-glucosamine or D-galactosamine) residues respectively. The type of monosaccharide building block, the position of the glycosidic linkages and the location of sulphate groups on the monosaccharide rings will determine the subspecies of GAG synthesised.

The GAGs chain initiates with the protein core attached to tetrasaccharide contains GAG (n)-GlcA-Gal-Gal-Xyl-Ser sequence. The chain initiates by adding an *O*-serine residue using glycosyltransferases (Kim *et al.*, 2001). Glycosyltransferases such as XylT-1 responsible for adding next residue Xyl from UDP-Xyl to the hydroxyl group of the serine core proteins in the endoplasmic reticulum. Then two D-galactose monosaccharides are transferred to the Xyl residue, this occurs by the activity of two types of galactosyltransferases GalT-I and GalT-II (Carey, 1997). The GlcAT-I enzyme completes the biosynthesis of the linkage trisaccharide by addition of a D-glucuronic acid residue from the UDP - GlcA transferase I (Pedersen *et al.*, 2000; Perrimon & Bernfield, 2000; Prydz & Dalen, 2000).

Two extensive modifications are carrying on GAG chain after it has been synthesised. First, phosphorylation of the Xyl residue at position 2, which appears to happen in both CS and HS. The second modification takes place on Gal residues by sulfation on either one or both Gal residues at the 4 or 6 position. This step is found only in CS and DS, thus this might be important for the selective assembly of CS/DS rather than HS/heparin as the last two never shows any kind of modification in Gal residues onto the linkage region tetrasaccharide (Uyama *et al.*, 2007).



**Figure 1-1: The biosynthesis of GAG chain.**

Modified from (Moehler *et al.*, 2013)

### 1.2.2 HS and Heparin Chain Polymerization and Modifications

Chain polymerisation continues by adding hexosamine residues to the linkage-region tetrasaccharide core. The sugar backbone of HS/Heparin is initiated by adding GlcA and  $\alpha$ -GlcNAc by  $\alpha$ -GlcNAc transferase I (GlcNAcT-I) and  $\alpha$ -GlcNAc transferase II (GlcNAcT-II) enzymes respectively, these enzymes are also known as EXT1 and EXT2. In this step, the chain will be classified and growing as HS/Hep chain (Fritz *et al.*, 1994; Kim *et al.*, 2001; Sugahara & Kitagawa, 2000).

The addition of these residues results in the final length of the Hep/HS chain which could be typically between 50 and 200 disaccharide units (Duncan *et al.*, 2001; Turnbull *et al.*, 2001).

The next step is an extensive modification of the chain by enzymes, which yields specific sequences of sulphated and unsulphated regions within the Hep/HS chain. Glucosaminyl *N*-deacetylase/*N*-sulfotransferase (NDST) is a golgi enzyme that initiates modification steps by first replacing *N*-acetyl groups with *N*-sulphates on GlcNAc residues (Kjellen, 2003).

Phosphoadenosine 5-phosphosulphate PAPS is considered as the main donor for all sulphate groups (Kakuta *et al.*, 1999).

Recently, it was found that The NDST enzyme is active on both N-deacetyl and N-sulphate residues, *in vitro* studies have characterised four isoforms of NDST enzymes in which each enzyme's isoform carrying the same reaction, with regards to the chemical context. NDST isoforms 1 and 2 appear to be remarkably important in HS modification due to overlapping expression of two of these isoforms among others which might be produced by different genes (Pikas *et al.*, 2000).

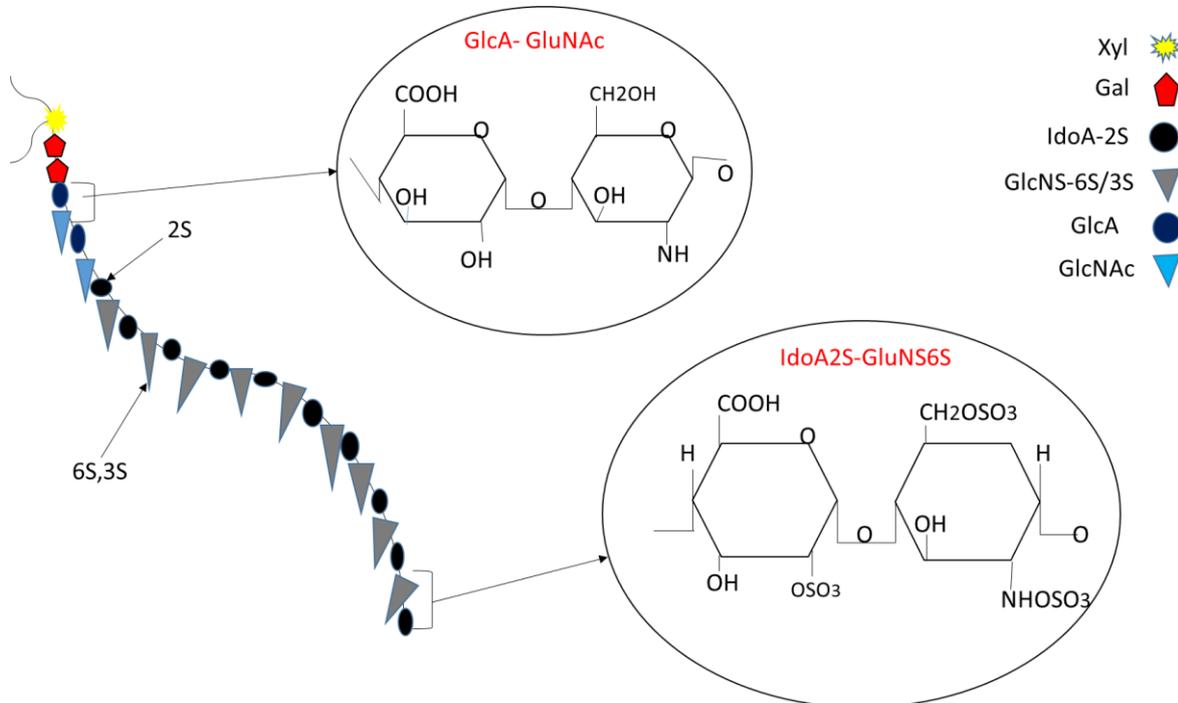
The next step in chain modification is epimerisation by GlcA C5 epimerase, this enzyme selectively epimerise GlcA to IdoA (Pinhal *et al.*, 2001). The reaction was found to be freely reversible in a solubilised enzyme system in the presence of  $^3\text{H}_2\text{O}$  trihydrate which is basically 3 molecules of water, and the *O*-sulphate group on the C2 position of the IdoUA residue which both seem to promote this reaction (Lindahl *et al.*, 1998; Pinhal *et al.*, 2001). The action of epimerase is to increase the flexibility of the GAG, thus the IdoUA units are considered as highly flexible units across the chain (Salmivirta *et al.*, 1996).

HS is further modified in the golgi, heparan sulphate *O*-sulfotransferases (HS-2/3/6OST) selectively add sulphate groups to IdoA or GlcA residues at the 2-*O* position, and the 3-*O* or 6-*O* positions of GlcNS or GlcNAc (Nakato & Kimata, 2002). Ultimately, the final structure of the polysaccharide chain is determined by the action of glucosaminyl *N*-deacetylase/*N*-sulphotransferase (the NDST's) golgi enzymes (Ringvall *et al.*, 2000).

In specific cells or tissues, further chain modification occurs by the action of 6-*O* sulfatases (sulfs) which selectively remove the 6-*O* sulfation (Ai *et al.*, 2003).

Despite the complicated modifications among HS chain, on average only 40-50% of HS disaccharides are sulphated, this giving HS great structural diversity leading to increase the ability to bind a variety of proteins. This converts crucial functional actions in many cellular

activities, such as anticoagulation when HS binding to antithrombin protein, also promote angiogenesis by HS binding to FGF (Lamanna *et al.*, 2006; Robinson *et al.*, 2006).



**Figure 1-2: Typical HS and Heparin Chain Polymerization and Modifications.**

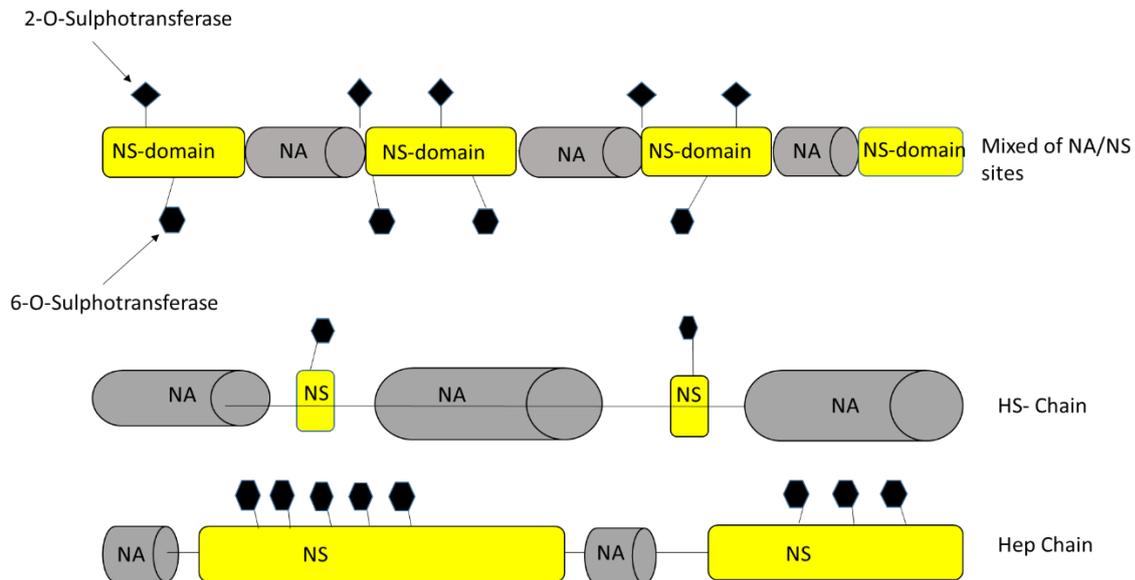
Modified from (Esko *et al.*, 2009)

### 1.2.3 Heparin and Heparan Sulphate Domains

The structural features between Hep/HS has been identified; the disaccharides units that configure the heparin chain are extensively modified unlike HS chain. In the HS biosynthesis and modification, selected GlcNAc residues likely undergo the *N*-deacetylation /*N*-sulfation modification with no further modification. This known as an incomplete biosynthetic modification pathway. Thus, HS chain appears as blocks of unmodified sequences (NAc-domains), separated by highly modified residues or (S- domains).

However, in heparin the chain essentially contains S-domain interrupted by NAc domains. During the modification reactions, three types of domains are likely to be generated. First, unmodified *N*-acetylated regions (NA domains) with GlcA units, which are present in small

amounts in the predominant disaccharide unit in HS. The second type is adjacent *N*-sulphated regions with 2-*O*- and 6-*O*-sulphated in both IdoA and GlcA residues in (NS domains). Finally alternating regions of *N*-sulphated and *N*-acetylated disaccharides domains appear in HS chains (Mulloy *et al.*, 2011; Robinson *et al.*, 2006).



**Figure 1-3: Typical organisation of heparin vs heparan sulphate domains.**

Modified from (Chavaroche *et al.*, 2013)

#### 1.2.4 Patterns of 2-*O*- sulphation, 6-*O*- sulphation and 3-*O*- sulphation groups in HS and Heparin

Once the initiated synthesis and modification pathways in Hep/HS chain biosynthesis is completed, further modifications will occur to add additional complexity to its structure and function. Three types of HS sulphotransferases are responsible for the modification of HS/Hep domains. Heparan sulphate 2-*O*-Sulphotransferase (HS2ST) is one of these evaluative enzymes, and has a key role in forming 2-sulphated L-iduronic acid (IdoUA (2S)), which widely exists in HS chains. The mechanism of action is to catalyse the transfer of sulphate groups from PAPS to the C2 site on the L-iduronic acid (IdoUA) (Kobayashi *et al.*, 1997).

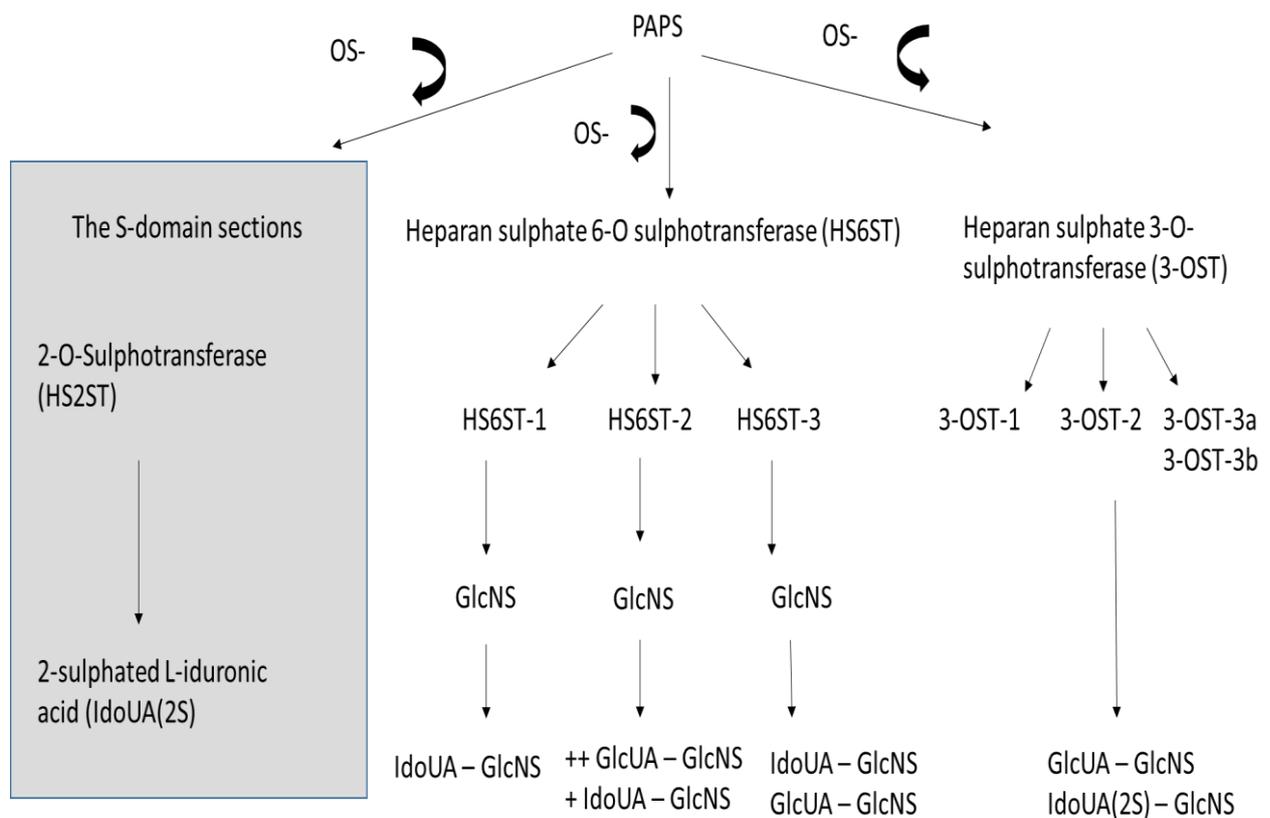
As this enzyme acts only on IdoUA units, it mainly modifies the S-domain sections of the GAG chain. It may also play a critical role in HS degradation especially within IdoUA (2S) – GlcNS (6S) domains (Habuchi *et al.*, 2000; Kobayashi *et al.*, 1997).

The second enzyme is heparan sulphate 6-*O*-sulphotransferase (HS6ST). This enzyme catalyses the process of adding a sulphate group to the C6 position of a glucosamine using PAPS as a sulphate donor. There are three isoforms of this enzyme, HS6ST-1, HS6ST-2 and HS6ST-3. The first enzyme can transfer sulphate groups to GlcNS sites of IdoUA – GlcNS disaccharide units (Sedita *et al.*, 2004).

The second isoform transfers sulphate groups to a wide variety of substrates. More sulphate groups may transfer to GlcUA – GlcNS domain, and less to IdoUA – GlcNS domain. Finally, the third isoform HS6ST-3 is thought to act on both domains, IdoUA – GlcNS and GlcUA – GlcNS (Habuchi *et al.*, 2000; Sedita *et al.*, 2004). In other words, HS6ST only act or transfer *O*-sulphate to GlcNS.

During the biosynthesis of the HS chain, the final and important modification occurs on carbon C3 of *N*- sulphated GlcN residues, this is the substrate that 3-*O*-sulphotransferase enzyme (3-OST-1) acts on. This enzyme is considered most important, as it is thought to be engaged in creating the antithrombin III (AT III) binding site within HS chains, thus converting non-anticoagulant HS into anticoagulant HS as the residue is critical for the formation of the antithrombin binding region the 3-*O*-sulphated GlcNS (3S) units (Liu *et al.*, 1999). 3-*O*-sulphation has been shown to be rare among other sulphations of glucosamine during the modification process. There are three different isoforms of this enzyme, 3-OST-1, 3-OST-2, 3-OST-3a and 3-OST-3b. Each of these isoforms is expressed in different tissues. 3-OST-1 is found in the brain, kidney and the heart, while 3-OST-2 is highly represented in brain tissue, the third isoform is found within a wide variety of different tissues (Habuchi *et al.*, 2000). The sulphation mechanism of 3-OST-2 is not limited to specific residues, it

transfers to the GlcNS residue regardless of the preceding UA unit hence the domain is GlcUA – GlcNS or IdoUA(2S) – GlcNS, whilst the 3-OST-3a and b enzymes act on specific sequences limited by IdoUA(2S) – GlcNS domains (Liu *et al.*, 1996).



**Figure 1-4: Biosynthesis enzymes and 2-O- sulphation, 6-O- sulphation and 3-O- sulphation patterns found in HS/Hep.**

### 1.2.5 CS/DS Chain Polymerisation and Modifications

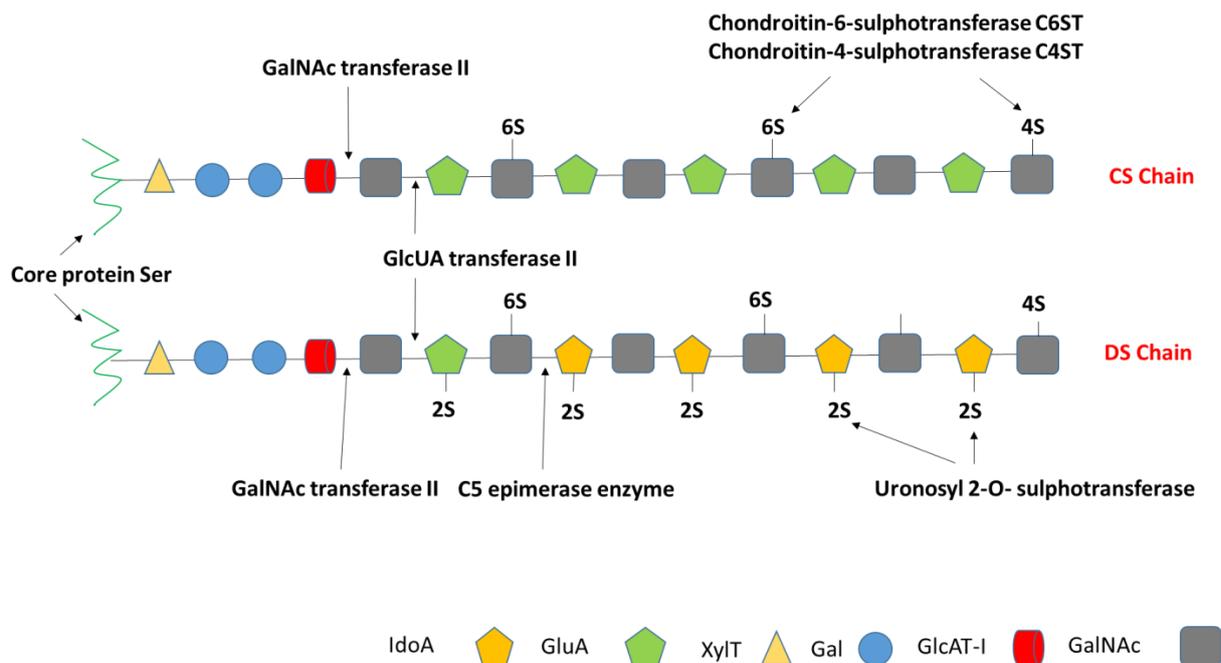
Biosynthesis of CS/DS seems to be similar to that in HS/Hep chain synthesis. The CS/DS biosynthesis enzymes are localised in trans-golgi complex, which is different from golgi network part in the lower end of the trans-golgi matrix where HS/hep biosynthesis enzymes are confirmed in golgi apparatus (Prydz & Dalen, 2000).

The diverging step in chain polymerisation begins after the formation of the linker tetrasaccharide section completed. The enzyme EXTL2, which also known as GalNAcT1 is different from the enzyme that adds Gal to the tetrasaccharide residues in chain initiation, it

acts by adding GalNAc to the chain, to produce CS/DS chains. EXTL2 is the enzyme that adds GlcNAc residue to HS chains. There are two distinct enzymes for the synthesis of the CS/DS backbone known as GlcUA transferase II and GalNAc transferase II, the action of these enzymes will lead produce chains with repeating 4- GlcUA  $\beta$ 1-3 GalNAc units (Sugahara *et al.*, 2003).

After the biosynthesis, the DS/CS chain undergoes a significant modification process, including the epimerization of GlcUA to IdoUA. The outcome of the C5 epimerase enzyme leads to the chain classified as DS. The main differences between CS and DS is that CS only has a glucuronic acid unit, while DS has both D-glucuronic acid and L-Iduronic acid units (Prydz & Dalen, 2000).

The final chain modifications are by sulphotransferase enzymes action, the chain eventually undergoes 2-*O*, 6-*O*-sulphation and 4-*O*-sulphation. The Chondroitin-6-sulphotransferase is known as C6ST, the action of this enzyme produces GlcUA – GalNAc (6S) domains from GlcUA – GalNAc by catalysing the conveyance of sulphate groups to the 6-*O* site of GalNAc. Unlike Chondroitin-4-sulphotransferase that shift the sulphate groups to the 4-*O* site of GalNAc residues, thus producing in GlcUA – GalNAc (4S) structure. Uronosyl 2-*O*-sulphotransferase is found to only be active on DS chains via the shifting of a sulphate group onto position 2 of either GlcUA/IdoUA (Sugahara *et al.*, 2003).



**Figure 1-5: CS/DS Chain Polymerisations and Modifications.**

Modified from (Maeda, 2015)

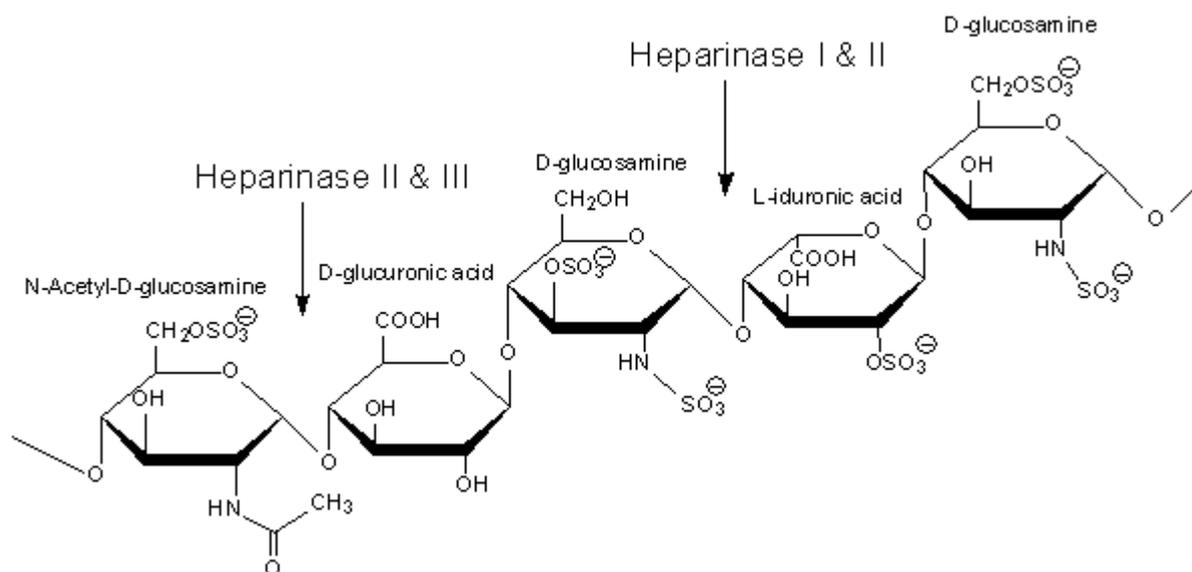
### 1.3 De - Polymerisations Techniques

There are two known methods for de-polymerisation of GAG chains, each one could be used separately or combined and yield different compounds. A method for determining the sequence type of GAG at the level of disaccharide provides key information for their structure-function relationships. Disaccharide is obtained by depolymerisation of the polysaccharide chains via enzymes and/or chemical degradation.

#### 1.3.1 Enzymatic De-Polymerisation of HS/Hep

The enzymes generally used for the enzymatic depolymerisation process have varying degrees of specificity for HS/Hep chains and are found in the soil bacterium, *Flavobacterium heparinum*. These enzymes act especially on the structure uronic acid-HexA releasing disaccharide or oligosaccharide compounds, with approximate fragment lengths of 10-20 disaccharides. The predictable structures produced containing a reducing end and  $\alpha/\beta$

unsaturated uronic acid at the non-reducing terminal (Galliher *et al.*, 1981). This will create an unsaturated double bond between C4 and C5 on the non-reducing end of the chain, which is known as unsaturated uronic acid ( $\Delta$ UA). This has UV sensitivity, which can be detected by wave length 232 nm, this allows enzyme digestion process to be detected (Linker, 1979). There are three different isoforms of heparanase known as lyases I, II and III, each of these lyases has a very high substrate specificity making them useful tools when comparing isolated GAGs from different cells or tissue types, also they play an important role in producing specific types of oligosaccharide fragments which vary in sulphate pattern, chain structure and length. Heparanase I (hep I) typically acts on the sequence IdoUA (2S) - GlcNS-(6S), thus it will release disulphate disaccharides and as a result releases the modified NS regions. The enzyme also cleaves the antithrombin III binding pentasaccharide domain in the heparin molecule. Heparanase II (hep II) acts on the sequences IdoUA-2S GlcNS/NA-3S/6S and GlcUA -GlcNS/NA-3S/6S. This is why Heparanase II is considered the least specific of the heparinase enzymes and able to cleave a wide range of substrates in HS/Hep chains. Heparanase III (hep III) acts on the sequence GlcUA -GlcNAc6S (Lindahl *et al.*, 1998; Pinhal *et al.*, 2001).



**Figure 1-6: Heparinase (I, II, III) lyase action on each fractions substrate.**

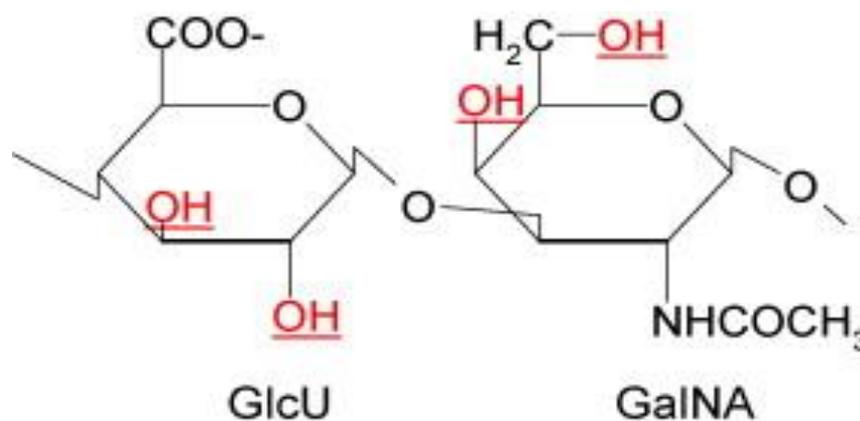
<http://www.sigmaaldrich.com>

### 1.3.2 Chemical De-Polymerisation of HS/Hep

Nitrous acid is involved in the chemical method used for generating disaccharide units from polysaccharide chains. Nitrous acid acts very selectively and quantitatively on the GlcNS sequence, and is also able to convert GlcNS, GlcNS (6S), GlcNS (3S) and GlcNS (3,6S) units into 2, and 5- anhydro D-mannose (Man). A specific range of pH is important to nitrous acid substrate cleavage sites, high pH (3-4) or a low pH (1.5) can be used (Rabenstein, 2002). Low pH nitrous acid-catalyzed deamination, which cleaves chains at *N*-sulphated GlcN residues while high pH nitrous acid-catalyzed deamination will not act on GluNS domains, instead the cleavage occurs at a small number of *N*-unsubstituted glucosamine residues GluN that exist in HS.

### 1.3.3 Enzymatic De-Polymerisation of CS/DS

Chondroitin sulphate ABC lyase is the enzyme used to de-polymerise chondroitin sulphate. It is purified first from *Proteus Vulgaris* with two isoforms, known as an endoeliminase, capable of depolymerizing chondroitin sulphate to produce, as products, a mixture of  $\Delta 4$ -unsaturated tetra- and disaccharides and an exoeliminase acting on chondroitin sulphate tetra- and hexasaccharides to yield the respective disaccharides. It acts on  $\beta 1, 4$ -galactosaminidic bonds between *N*-acetylgalactosamine and either D-glucuronic acid or L-iduronic acid in CS or DS and yield oligosaccharide mixtures of unsaturated tetra- and disaccharides (Yamagata *et al.*, 1968). The ABC lyase has a variety of applications, it has been used in the quantification of CS and DS from a crude mixture (Yamagata *et al.*, 1968), also it has an important role in structural analysis of the carbohydrate moiety of proteoglycans studies (Sugahara *et al.*, 1992). A complete depolymerisation of CS to unsaturated disaccharides yields from the combined activity of “endoeliminase (chondroitin sulphate ABC endolyase)” and exoeliminase (chondroitin sulphate ABC exolyase) (Hamai *et al.*, 1997).



**Figure 1-7: Chondroitin (ABC) lyase action on GAG.**

(Lauder, 2009)

## **1.4 Heparan sulphate/heparin-protein interactions**

Consensus sequences exist on the surface of some proteins with positively charged amino acids that are able to interact strongly with negative-charged groups on the HS chains. The interactions occurs within the binding regions in specific HS binding proteins (Goupille *et al.*, 1997).

The type of interaction is principally an ionic interaction involving negatively charged groups in HS chain specifically sulphate and the carboxylates, which exists in all GAG chains. So far, hundreds of proteins including in particular antithrombin III and the FGF family have been shown to have significant interactions with HS chains. In particular with NS domains, these domains provide multiple binding sites for protein ligands (Kreuger *et al.*, 2006). Their affinity for HS oligosaccharides basically correlates with the overall degree of sulfation (Jemth *et al.*, 2002). Heparin oligosaccharides were generally the most efficient complex promoters, whereas less sulphated HS species were less efficient (Jastrebova *et al.*, 2006). More selective interactions would require either sequences containing rare components or precise spacing of two or more sulphated domains *N*-sulphated/acetylated/sulphated (SAS) (Kreuger *et al.*, 2006).

### **1.4.1 FGF and the FGFR**

Fibroblast growth factors are known as signalling molecules, which have a capacity to promote angiogenesis by stimulating the proliferation of endothelial cells and can act directly on tumour cells by increasing their growth. They are considered as one of the principle angiogenic factors. Many members of the FGF family have the ability to bind to HS GAGs and to FGFR receptors. These receptors bind to members of the fibroblast growth factor family of proteins to form complexes on the cell surface.

Such interactions have been shown to be an absolute prerequisite for FGF mediated biological activity. There are a great number of fibroblast growth factor family members (FGF's) in humans and mice, all of them are similar in terms of structure, but FGF1 is the only one which is able to bind to all FGFR's (Rieckmann *et al.*, 2008; Wu *et al.*, 2003).

FGF1 and FGF2 are known as essential angiogenic factors more than vascular endothelial growth factor (VEGF) or platelet-derived growth factor (PDGF) (Cao *et al.*, 2003).

These FGFs bind the HSPGs chain with different affinities. There are two types of FGFR receptors on the surface of cells, each binding to FGF within different affinities, the low affinity cell surface receptors are known as HSPG's co-receptors, and the high affinity termed as tyrosine kinase receptors (FGFR's). Five distinct mammalian genes are encoding tyrosine kinase receptors. Each of these FGFR is expressed in variety of different tissues, for instance, FGFR1 is found mainly in connective tissue, FGFR2 mostly exists in bone, FGFR3 contributes to the structure of cartilage, and finally FGFR4 identified in muscle. There are two types of each aforementioned FGFR, unspliced receptors or binding receptors, which are divided into many variants known as A, B, and C. Their function secreted as receptor such as FGFR- type A or supply binding ligand requirements for interactions such as FGFR- 3c and FGFR -3b types. Some of the FGFRs binding to FGF1 and FGF2 equally such as FGFR- 2c, while FGFR-3 binds to FGF1 with approximately 1000 fold higher affinity than FGF2 (Goodger *et al.*, 2008; Rieckmann *et al.*, 2008).

#### **1.4.2 The FGF signalling pathways**

Signal transduction from FGF requires the interaction between the growth factors and their specific FGFR receptors including HSPG as a co-receptor, forming a complex that usually exists on most of cell surfaces. The interaction between FGF and HS chains occur with many residues but most of these interactions are likely to be inactive complexes.

Many factors can influence HS activation of FGFs including chain length and the presence of specific saccharide components, for e.g., HS oligosaccharides fractions of (dp6-dp12) have different capability to activate bFGF in a mitogenic assay and this was correlated with their length and disaccharide composition (Pye *et al.*, 1998).

Another factor correlates to the charge distribution, which is mainly depend on sulphated patterns. For e.g. the *N*- and 6-*O*-sulfation apparently sufficed to satisfy the requirement for HS in important VEGF and FGF signalling events (Kreuger *et al.*, 2006).

HSPGs are considered as co-receptors as they have secondary role in FGFR's signal transduction (Goodger *et al.*, 2008). Cellular response initially starts from cell surfaces where a particular type of FGF binding to their cell surface receptors, occurs the strength of the signalling from this interaction will mainly depend on the environment of the cell including the number of receptors available and the activated period of these receptors. All this will lead to receptor dimerization followed by tyrosine auto phosphorylation of the receptor and eventually activation of the appropriate signalling cascades (Pye *et al.*, 2000).

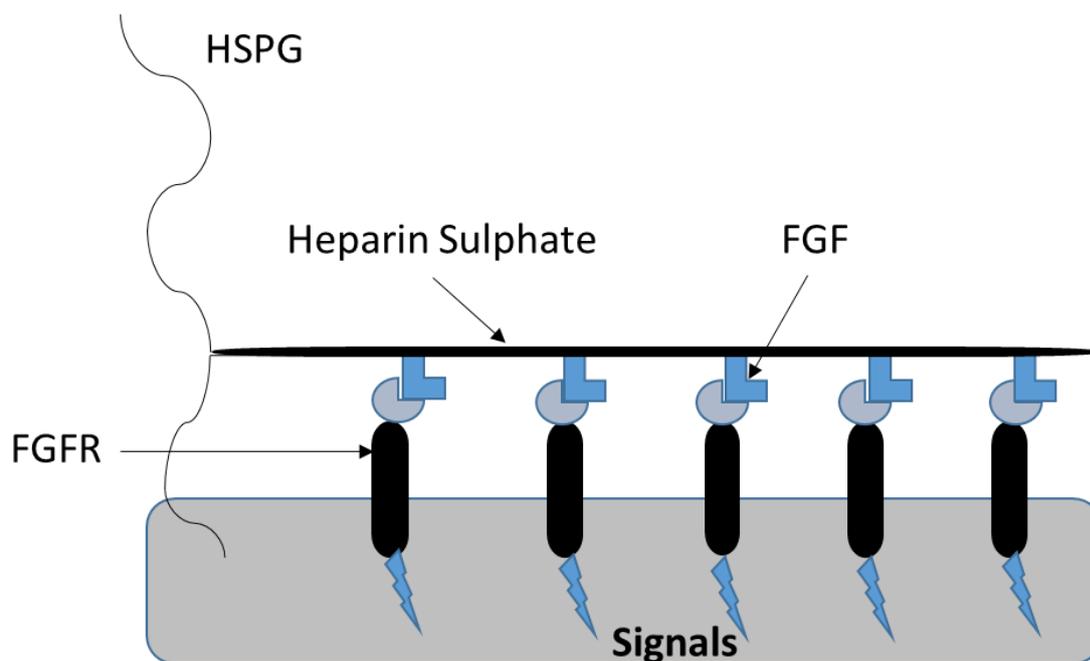
### **1.4.3 HS/FGF/FGFR Complex interactions**

Highly sulphated residues or NS domains within HS are important in FGF activation as they have the ability to bring together FGFs with their tyrosine kinase receptors, they also increase the stability of FGF-FGFRs interaction (Goodger *et al.*, 2008; Kreuger *et al.*, 2006). The ternary complex of FGF1: HS: FGFR1 is highly stable and loses its complex structure only when heated to 100 °C, while both of FGF1: HS and HS: FGFR1 binary complexes loss their binding structure when heated to 60 °C (Wu *et al.*, 2003).

The second crucial function of the complex interaction between HSPG and FGF-FGFRs is the ability to prevent the proteolysis of FGF proteins (Ornitz, 2000), and to protect both FGF1 and FGF2 from thermal denaturation (Zhang *et al.*, 2001). The FGF1: HS: FGFR1 is

known to occur in different ratio and ways. The FGF: HS: FGFR interactions happen in three modes known as Trans, cis and mix (contains both cis and Trans interactions). Multiple models generated from each of these modes (Wu *et al.*, 2003). In these complexes both ligand and receptor might interact with HSPGs, or only the ligand might interact with HS/Hep chain (Ornitz, 2000; Pellegrini *et al.*, 2000). For example, in Trans model, the HS derived hexasaccharide (dp6) is the smallest sequences that might interact with two FGF1's on each side of the isoform interaction. In the cis oriented model, HSPGs of (dp12) units is required and the minimum size required for the mix cis/trans to form 2:2 FGF1:FGFR2 complex is a HSPGs of dp16 (Ornitz, 2000).

There are two well-known structural models for this complex that have been confirmed by X-ray crystallography, which are the “symmetrical complex model” and the “asymmetrical complex model”. The symmetrical complex is more stable among the other complexes this due to extensive protein- protein contact (Goodger *et al.*, 2008). This model involves a 2:2:2 complex, two HS dp10 chains facilitate the complex and were confirmed by the X-ray crystal structure (Schlessinger *et al.*, 2000). The asymmetrical complex contains only a single chain of HS and is able to join the FGF and FGFR's together in a 2:2:1 ratio such as FGFR2 IIIc: FGF1: hep decasaccharide seen by Pellegrini *et al.*, (2000). The size of the HS chain and the amount of sulphation determines which HS chain will pick by the FGFR to create the complex.



**Figure 1-8: HS/FGF/FGFR Complex interactions.**

Modified from (Gallo *et al.*, 2015)

### 1.7 Marine sulphated glycan

Marine organisms have been shown to be one of the most extensive and diverse sources of many bioactive compounds. During the last 50 years, over 14,000 different natural products from marine organisms or species have been reported (MarinLit, 2003).

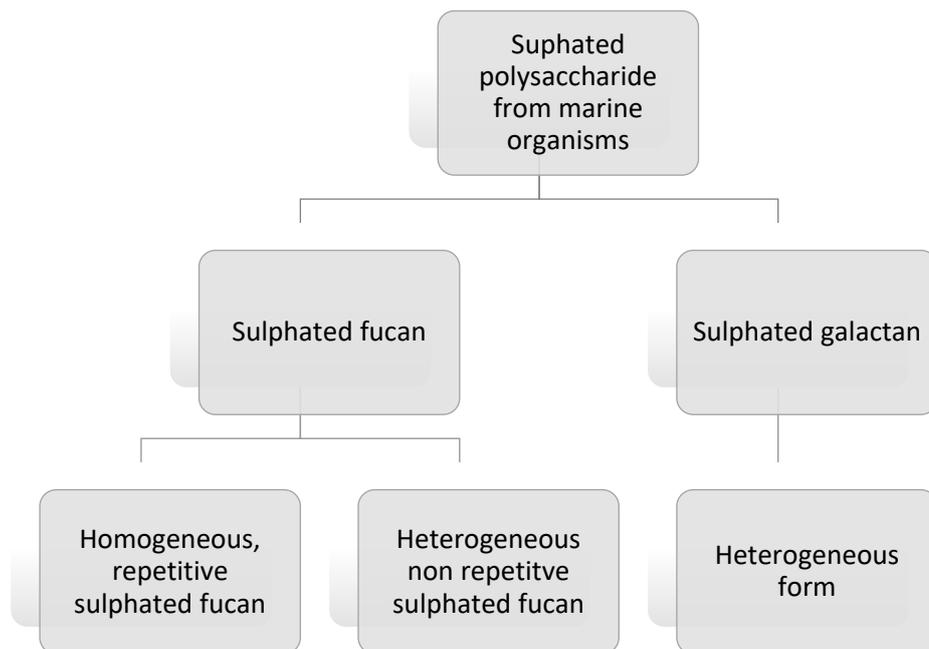
The first isolation of sulphated glycans from a marine organism was by Killing (1913), as he reported a structure of sulphated fucan from marine brown algae (Phaeophyta). Different types of sulphated GAGs have been purified from microalgae such as brown, green, and red algae, and other invertebrate species such as echinoderms (sea cucumber and sea urchins). Sulphated GAGs are found in other living phyla and have been confirmed in many other classes of invertebrates such as Porifera, Actinopterygii fish (Zierer & Mourão, 2000), also in commercially relevant species such as sharks, skate, codfish, and salmon. Most of these sulphated GAG's have shown biological activities in different aspects such as anti-pathogenic, antitumor and anticoagulant (Silva *et al.*, 2012).

Other species have been found with a rich content of sulphated GAG in their composition, such as *S. hermanni* and *S. vastus*. Other species of sea cucumbers known as 'GAMAT' are considered as a source of traditional medicines, due to containing high bioactive substances (Hawa *et al.*, 1999).

### **1.7.1 Structural composition of sulphated polysaccharides (SP) from marine sources**

Sulphated polysaccharide is a significant component in marine organisms' structures that have been extensively studied to date. Two types have been identified so far, sulphated fucan and sulphated galactans. Essentially, the first type occurs in two forms, heterogeneous non-repetitive sulphated fucans from algae and homogeneous, repetitive sulphated fucans of echinoderm. While heterogeneous sulphated galactans have been extracted from red algae, green algal species, some species of invertebrates and marine angiosperm polysaccharide (Pomin & Mourão, 2008). Many factors affect the composition of this polysaccharide; first, different species express different structures of sulphated polysaccharides. Second factor is the method of extraction

The third factor is the season or the environment of conditions, such as the crude fucan that is extracted monthly from makonbu, *Laminaria japonica* Areschoug, cultured from April to October at a southern site of the Hokkaido bay (Japan). The crude fucan tends to increase gradually from April to September, and rise markedly in October (Honya *et al.*, 1999; Mabeau *et al.*, 1990; Mian & Percival, 1973; Percival & Ross, 1950).



**Figure 1-9: types of sulphated polysaccharide from marine organisms.**

Heterogeneous sulphated fucans are more abundant in marine brown algae (Phaeophyta). Early study by Percival & Ross, (1950), reported the structure of these sulphated fucan, which mainly contain 1→2 glycosidic linkages and sulfation at the 4-position. Recently, another structure has been confirmed for brown algae's sulphated fucan which contains equal amounts of alternating units of 2,3-disulphated, 4-linked and 2-sulphated, 3-linked  $\alpha$ -L-fucopyranosyl units and has been confirmed by NMR analysis. The heterogeneous nature of these structures is thought to be a result of occurrence of branches of non-sulphated fucose residues. However, a unique structure is reported from other brown algae species with NMR studies suggesting the occurrence of *O*-acetylation in very high amounts (Chevolot *et al.*, 1999). The structure of a polysaccharide from brown algae is known as the most complex marine polysaccharide due to branching chains, diverse glycosidic linkages, acetylation, methylation and random sulphation pattern.

Homogeneous, repetitive sulphated fucans have been reported in marine invertebrates especially in sea urchins and sea cucumbers (Echinodermata, Holothuroidea) with simple repetitive structures composed of unbranched chain of  $\alpha$ -L-fucose either [ $\alpha$  (1 $\rightarrow$ 3) or  $\alpha$  (1 $\rightarrow$ 4)] type of glycosidic linkages and 4-*O* and/or 2-*O* sulfation positions (Pomin & Mourão, 2008).

Heterogeneous sulphated galactans are the other main type of sulphated glycan and are found principally in marine red algae. The main reason for heterogeneity of the glycans isolated from red algae is the sulphate distribution along the chain. There are two types of red algae termed as Carrageenans and Agarans both express an homogeneous galactan backbone unlike other red algae species (Bixler, 1994; Knutsen *et al.*, 1994; Lahaye, 2001; van de Velde *et al.*, 2004).

Both of these red algae have glycans with a non-branched (linear) chain comprised of alternating 3-linked  $\beta$ -D-galactopyranose and 4-linked  $\alpha$ -D/  $\alpha$ -L-galactopyranose residues. They have a variety of industrial applications due to their simple structure, they also have a unique ability to create strong aqueous gels. The physicochemical properties of these glycans are affected by variation in the backbone structure such as modified  $\alpha$ -D-galactopyranoses in carrageenans to  $\alpha$ -L-galactopyranoses in agaran (Lahaye, 2001). Sulphated galactans are also found in green algal species like *Codium* (Bilan *et al.*, 2007; Farias *et al.*, 2008). The chain here is without repetitive units of disaccharide and instead has a complicated sulphated pattern in positions 4 and/or 6 –sulfation unlike red algae. Chain complexity is higher in green algae than red algae but is considered as simpler structure in comparison to brown algae (Pomin, 2009). Sulphated galactans have also been reported in some species of invertebrates in ascidians (Urochordata, Ascidiacea), and in two species of sea urchins. The structures are typically homogeneous and contain linear chains as the backbone structure.

In general, the location of these sulphated glycans, either sulphated fucan or galactan, is determined within the body wall or extracellular matrix, this matches the amount of glycosaminoglycan found in the extracellular matrix of mammalian connective tissue. Marine sulphated fucans and galactans are more sulphated than CS and DS in vertebrate GAGs. Marine sulphated glycans are found with extremely high molecular weights that can be >1 million Da such as the mass of sulphated fucans that found in sea urchin egg jellies for the glycan fraction, as this sulphated glycan has high molecular weight itself without any core protein interaction, while the GAG molecular mass in mammals could reach ~15 and ~60 kDa, or approximately >100 kDa in case of covalent attach to core protein (Pomin & Mourão, 2008).

The presence of sulphated GAGs in some invertebrates is now well-documented. Different classes of invertebrates have shown that CS/DS, HS-like, and/or heparin-like compounds are present in many species (Medeiros *et al.*, 2000). studies have also shown that heparin is present in several species of molluscs. A compound from the clam *Mercenaria mercenaria* exhibits several structural similarities to heparin (Jordan & Marcum, 1986). Heparins with high anticoagulant activity have been isolated from the molluscs *Anomalocardia brasiliiana*, *Tivela mactroides* (Dietrich *et al.*, 1985; Pejler *et al.*, 1987) and *Tapes philippinarum* (Cesaretti *et al.*, 2004). CS, DS, and acharan sulphate have also been isolated and characterized from different families of molluscs (Cassaro & Dietrich, 1977; Kim *et al.*, 1996).

Whereas HSs and CSs are ubiquitous components of most tissues, heparin has shown a very peculiar distribution in mammalian and other vertebrate tissues, as well as invertebrates (Medeiros *et al.*, 2000). Furthermore, a large variation of the concentration of heparin among species is evident, with the non-mammalian vertebrate tissues showing considerably lower amounts. In invertebrates, heparin is found in few molluscs, crustaceans, annelida,

echinoderma, and cnidaria, with the anticoagulant activity varying according to the species analysed (Nader *et al.*, 2004).

### **1.7.2 Potential drug development from marine glycan.**

Marine sulphated polysaccharides can express their activity on mammal's biological system and provide pharmacological activity with different diseases, such as antithrombotic, anti-inflammatory, antiangiogenic, and antimetastatic. Sulphated fucan from brown and sulphated galactans from red and green algae are well known for their activity as anticoagulants as they interact with both antithrombin and heparin cofactor II (Berteau & Mulloy, 2003; Pereira *et al.*, 2005). However, the complex, heterogeneous structures of these polysaccharides has limited the possibility of identifying specific sequences required for their activity. Many researches, have worked in this field in order to confirm the structure of these polysaccharides and the relationships related to their activity. Chemical oversulphation or desulphation of the intact chain from these algae has been reported (Haroun-Bouhedja *et al.*, 2000; Soeda *et al.*, 1993). The results suggest that the activity increases with increasing sulphate regions and decreases when the crude or intact pattern of sulphation is reduced. The same role is applied to the molecular weight, the activity decreases with smaller molecular size of oligosaccharide or polysaccharide (Soeda *et al.*, 1993).

In order to identify the structure-function relationships of these polysaccharides, it was important to identify clearly oligosaccharides type, glycosidic bonds between the residues, and sulphate positions which are required for blood coagulation protein interaction to occur that exhibit the anticoagulant process. The effect of the type of saccharide residue, whether this was fucose or galactose on modifying anticoagulant activity has been reported by Pereira *et al.*, (2002). Comparisons were made between two different sulphated polysaccharides, the active 2-sulphated, 3-linked  $\alpha$ -L-galactan and inactive 2-sulphated, 3-linked  $\alpha$ -L-fucan, and

their ability to modify the activity from active to inactive sequence model. The position of the glycosidic linkage in this polysaccharide affected their activity; this was shown by comparing two models of sulphated galactans, the inactive 3-sulphated, 4-linked and the active 2-sulphated, 3-linked  $\alpha$ -L-galactans. The effect of the sulphation pattern is important in the structure-function relationship; increasing the sulphation content only  $\sim$ 1.8-fold caused extensive increases in their anticoagulant activity of nearly 38-fold. This was linked to the occurrence of two, 4-di-sulphated of 3-linked  $\alpha$ -L-fucans from 2-sulphated, 3-linked  $\alpha$ -L-fucan (Mourão, 2004).

Sulphated fucans have been shown to have angiogenesis inhibition activity. According to Koyanagi *et al.*, (2003), some fucoidans have been extracted from *Fucus vesiculosus* which interfere with the binding of VEGF with their receptors. The same interaction between these molecules and FGFs with their respective receptors is also reported which causes angiogenesis inhibition as a consequence (Soeda *et al.*, 2000).

Moreover, apoptosis and cell death due to autophagy in tumour cells was reported as a biological activity with some active compounds within sulphated fucans from brown algae. Sulphated heparin, fucoidan and Carrageenan lambda are reported to significantly inhibit activity toward lung adenocarcinoma in rat. This appears to be due to an interference with the passaging process in the tumour cells, which induces early apoptosis and leads to a remarkable loss of tumour cells from the lung after one hour, followed by cell death (Coombe *et al.*, 1987). Most of the bioactive compounds in algae are available commercially as an intact or crude compound containing sulphated polysaccharide but also contaminated with non-glycan compounds. Designing low molecular weight glycans with known sequences of saccharide remains unachieved as yet, due to the structural complexity of polysaccharide extracted from marine algae. In contrast, a correlation between structure and biological

activity is usually relatively easy to identify in marine invertebrates due to the repetitive nature of the backbone of disaccharide and their specific sulphation patterns.

### **1.7.3 Therapeutic value of mammalian glycosaminoglycan in cancer**

Glycosaminoglycan and proteoglycans play a diverse role in cancer growth where they control cancer cell proliferation and growth. As reported before, GAGs control cell growth via cell surface HSPGs that act as co-receptors for several growth factor tyrosine kinase receptors. In some cases, the ligands between HSPG and growth factors cleaves by the action of heparanase that is known as a  $\beta$ -endoglucuronidase enzyme. This releases the HSPG from FGF-FGFR's complex and leads to the loss of growth signalling (Bernfield *et al.*, 1999; Sanderson *et al.*, 2004).

Changes in the expression of ECM-degradative enzymes and the expression of cell-surface molecules is increasing the ability of cancer cell for invasion and metastasis. HS is degraded by heparanase that exists on the cell surface, this is reported to cause loss of syndecan-1 which in turn promotes association with different host cells and increases cancer cell dissemination to other tissues (Ludwig *et al.*, 2004; Ohkawa *et al.*, 2004).

In some studies, chondroitin sulphate proteoglycan supports cancer stem cell proliferation signals and provides a niche for preservation of stem cell properties in terms of DNA repairing capacity, apoptosis, and chemotherapy resistance. In some cases, like melanoma, chondroitin sulphate proteoglycans act as a specific marker for a class of epidermal stem cells (Legg *et al.*, 2003). Modulating glycosaminoglycan/ proteoglycan expression in cancer stem cells might cause apoptosis or reduce the malignant properties of the cells.

## **1.8 Breast cancer**

Among cancer types, breast cancer is well known as a heterogeneous and one that accounts for nearly 23% of all cancer diagnoses in women. George Gey was the first to establish culturing of cancer cells in the laboratory over 50 years ago. This paved the way to cell culture as we use it today, and remains a crucial and developmental experimental tool in cancer research. Long after this beginning, histopathologists successfully identified heterogeneity in the populations of cancer cells through morphological observations among breast cancer and the classification of breast cancer was based on several measures; including tumour grade, lymph node status, histological type and the presence of predictive markers such as oestrogen receptor ER and recently epidermal growth factor receptor 2(HER2). Breast cancer can be classified into at least five subtypes according to gene expression profiling and the immunohistochemically expression of ER $\alpha$ , progesterone receptor (PR) and HER2, which has been reported after molecular profiling development (Perou *et al.*, 2000). Each subtype has a different response for treatment depending on the therapeutic target. ER has become a more powerful therapeutic target for cancer research and drug discovery. Therefore, the luminal A and luminal B subtypes are susceptible to hormone therapy along with HER2 group. The more difficult type to treat are basal tumours, due to the lack of expression of ER $\alpha$ , PR and HER2 and they are collectively known as triple-negative (Badve *et al.*, 2011).

### **1.8.1 Causes, Risk Factors, and Prevention of breast cancer**

Many risk factors are associated with developing breast cancer. Gender, age and other biological factors like mutant genes, delay in pregnancy, early age at menarche, high level of hormones, and late menopause are typically linked with breast cancer. For instance, females are more likely to develop this disease especially above age 50 (Key *et al.*, 2001).

There is evidence of an increase in the possibility of developing breast cancer by 3% per one-year delay in pregnancy while childbearing at early ages of 20 years or younger might significantly decrease the chance of breast cancer in women (Key *et al.*, 2001). The risk of breast cancer is remarkably higher in women with high level of hormones such as oestrogen, also the risk might increase with exposure to exogenous hormones, especially for women on hormonal replacement therapy, this applies to oral contraceptives and hormonal replacement therapy (Chlebowski *et al.*, 2013). Higher risks are expected for patients of combined oestrogen and progesterone therapy compared to oestrogen therapy alone (Gierisch *et al.*, 2013).

Carrying the mutant genes BRCA1 and BRCA2 might increase the chance of developing breast cancer by 45-56% especially by the age of 70. Although 85% of women identified with breast cancer, do not have hereditary predisposition. By contrast 85% of women with a familial history unnecessarily develop this disease (Antoniou *et al.*, 2003). Environmental factors and life style are also considered as risk factors. High level of consumption of alcohol, obesity and less physical activity appear to have significant impact on developing breast cancer (Key *et al.*, 2001).

### **1.8.2 Epidemiology of breast cancer**

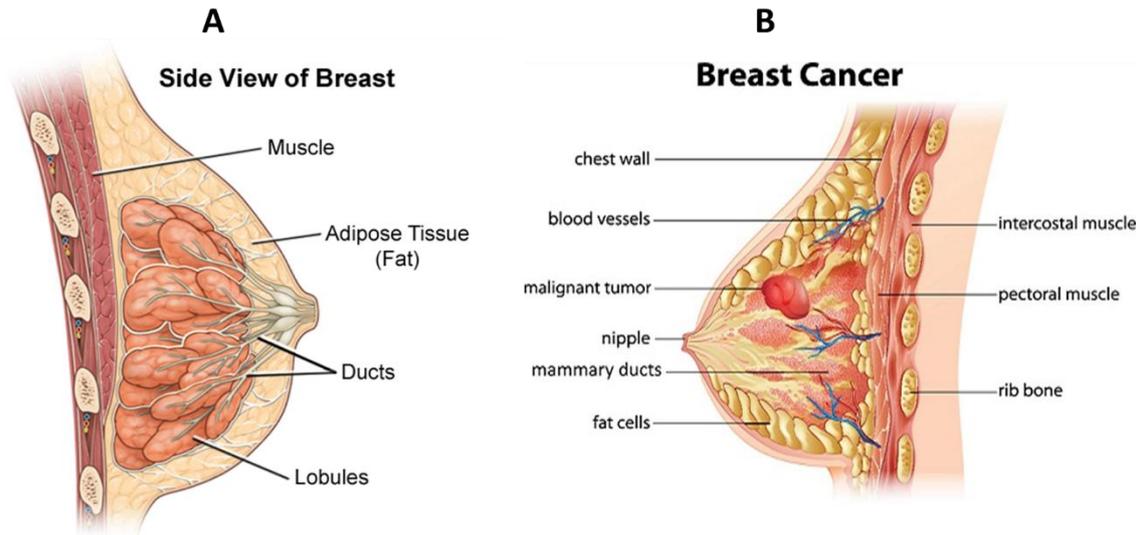
Breast cancer is one of the most widely spread cancers especially among women. Female mortality and breast cancer incidence rates are variable worldwide, with a high percentage of incidence is reported in developed countries unlike developing countries (Boyle & Ferlay, 2005). However, the percentage of mortality has dropped significantly in developed countries due to early diagnosis through implementation of screening programmes with mammography, improvement of adjuvant chemo and radiotherapies, decrease of hormonal replacement therapy and increase in awareness (Jemal *et al.*, 2010).

In the UK it is estimated 1 in 8 women develop breast cancer and the percentage of incidence has increased by 6% in last few years, but mortality is reported to have fallen by fifth (Arasta, 2015).

### **1.8.3 Histology of normal breast vs histology of breast cancer**

The structure of normal human breast is similar in male and female before puberty, there is one pair of mammary glands each situated over the pectoral muscles of the chest wall one the left and right side of the body. Both pendent from the anterior chest wall, and some fibrous strands are involved to assist this called cooper ligaments. There is a conical lifted part in each mammary gland called the nipple, which consists tiny size pores of 0.5 mm in diameter and the skin surrounding this part has more melanin than other areas of the breast, this called areola. In female, the breast structure changes after puberty due to an increase in the circulatory concentrations of ovarian hormones, which leads to an increase in the growth of the ductal-lobular system and adipose tissue (Russo & Russo, 2004).

There are 15 to 25 lobes that are surrounding nipple, connecting to each other by a huge amount of inter-lobar to produce a net of lobes each terminated in smaller lobes termed lobules that are open to the nipple and help with milk transport. The breast size becomes noticeably big during the premenstrual phase this due to size increases of ductal epithelial cells following the peak levels of oestrogen that happen in this phase in menstrual cycle. The morphological and functional changes are enormous during pregnancy and lactation due to the action of various hormones like oestrogen, progesterone, prolactin and human placental lactogen (Cotran *et al.*, 2010).



**Figure 1-10: Breast anatomy. A: Histology of normal breast. B: Histology of breast cancer (Howard M. Reisner 2015).**

In breast cancer, there is typical histological morphology that helps with the classification of breast cancer. The majority of breast tumours are adenocarcinomas; this could be an invasive or non-invasive carcinomas and could diversify as ductal in-situ (DCIS) or lobular in-situ (LCIS) carcinomas. In cases of non-invasive carcinomas, the malignant cells remain attached to the myoepithelial of the ductal or lobular system, with no invasive action to the surrounding area.

This kind of carcinoma might develop further if it is not treated (Stecklein *et al.*, 2013). Women with lobular in-situ (LCIS) carcinomas have a greater chance to develop invasive carcinoma and usually they are oestrogen and progesterone receptor positive. This type of carcinoma is not yet transformed into the myoepithelial cells but they appear as small uniform solid cells filling the central part of affected lobules. Ductal in-situ (DCIS) carcinoma is more heterogeneous than lobular carcinoma but the most diagnosed type of non-invasive malignancy. It could classify into many types such as comedo, cribriform, solid, and papillary and micropapillary. In this type, malignant cells appear in irregular shapes and contain a central necrosis area and large numbers of the mitotic figures. It has three different

morphologies, the solid sheet of cells form, called solid ductal carcinoma in situ, while it termed cribriform ductal carcinoma in-situ when there are spaces between cells, whereas if they form a finger-like structure, they are termed micropapillary ductal carcinoma in-situ (Viale, 2012).

Invasive carcinoma is the term used when epithelial tumours penetrate and affect the surrounding tissues. This type can again be classified into two groups invasive ductal carcinoma and invasive lobular carcinoma. Invasive ductal carcinoma (IDC) is the most widespread type, with no specific macroscopic shape; they have a good chance of metastasising to different parts of the body via the lymphatic system or bloodstream. They are reported as positive express HER-2 receptor within 30% (Stecklein *et al.*, 2013).

Invasive lobular carcinoma (ILC) is the second most popular type of breast tumour. Unlike invasive ductal carcinoma, this type has more regular shaped cells with irregular ovoid nuclei appearance, and are usually found as oestrogen and progesterone receptor positive (Viale, 2012).

#### **1.8.4 Molecular profiling of breast cancer**

Depending on the gene expression profile, breast cancer can be classified in to two essential groups, oestrogen receptor-positive and oestrogen receptor-negative tumours. These groups have further been classified into different subgroups depending on the cluster analysis of genome and transcriptome expression known as luminal A, luminal B, HER-2 positive, triple-negative or basal-like tumours, claudin-low and normal-like breast tumours (Malhotra *et al.*, 2010; Perou *et al.*, 2000).

Luminal A tumours are reported to show a variety of gene cluster expression. They are positive for an oestrogen receptor expression, negative for expression of HER-2 and have a low expression level of Ki-67 protein (Rakha *et al.*, 2008).

Luminal B tumours have positive expression for an oestrogen receptor-related gene cluster and HER-2. Unlike luminal A tumour, this subtype has a significant level of expression for Ki-67 and high proliferation rate. The third subtype of breast cancer is a HER-2 enriched subtype that reportedly overexpresses for HER-2 with poor prognosis and no expression for the oestrogen receptor-related gene cluster (Eroles *et al.*, 2012)

Basal-like tumours, classified as triple-negative tumours share many common characteristics, such as no expression profile for oestrogen receptor, progesterone receptors or HER-2. However, it has been reported that 15-45% of basal-like tumours might express either oestrogen or progesterone receptors or HER-2, unlike triple-negative tumours. This subtype expresses cytokeratin five and six (CK 5/6) and epidermal growth factor receptor (Nielsen *et al.*, 2004; Rakha *et al.*, 2008).

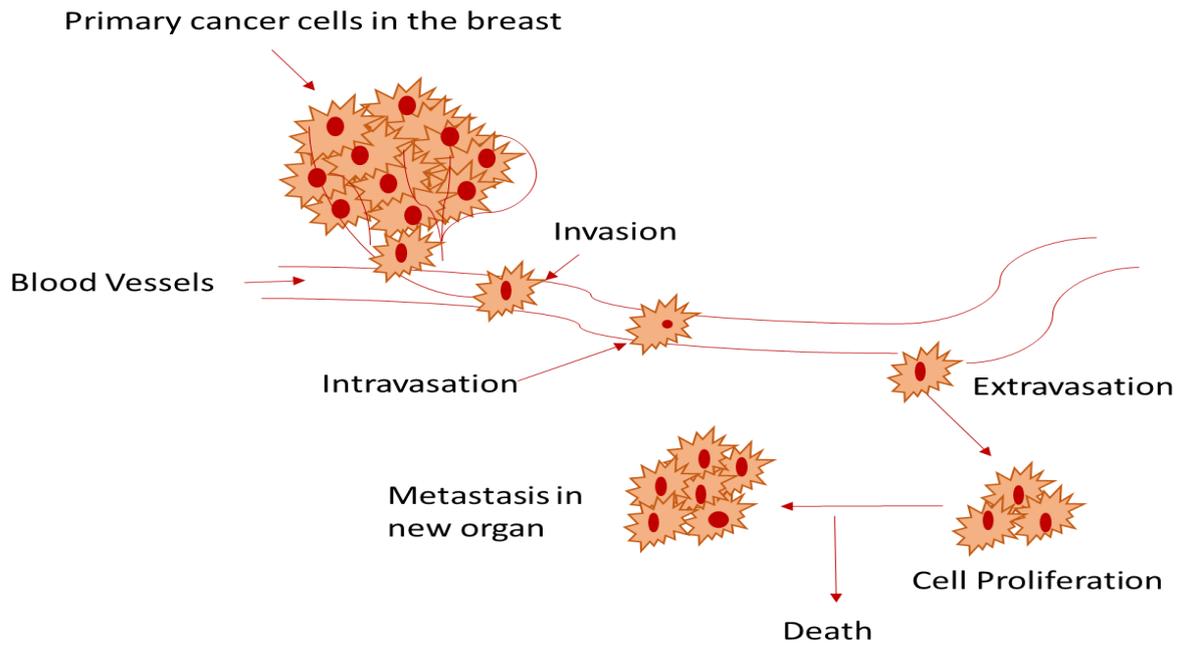
Triple-negative breast cancer is reported with the highest rate of mortality among other subtypes, particularly in black African American or British black younger women in age less than 50 years (Huo *et al.*, 2009). According to Turner *et al.*, (2010), there are several subtypes of triple-negative breast cancer confirmed based on gene their gene expression profile, which are mesenchymal, mesenchymal-stem like, immunomodulatory, a luminal androgen receptor-positive, and two basal-like BL1 and BL2 genes.

A new subtype of breast cancer was recently reported known as claudin-low subgroup (Herschkowitz *et al.*, 2007). This subtype shares many properties with the basal-like tumours regarding low gene expression profile for oestrogen, progesterone receptors or HER-2, luminal epithelial and also has a poor prognosis. However, it has a high level of expression for some markers like epithelial to mesenchymal transition genes, also cancer stem cell-like features and immune response genes (Prat *et al.*, 2010). Breast tumour gene expression profiles have been identified with significantly low expression of luminal epithelial cell genes (Parker *et al.*, 2009).

### **1.8.5 Role of metastasis in breast cancer**

Metastasis is the process by which the cancer spreads to distant parts of the body, and is considered as a final step in the cancer progression. The mechanism of metastasis starts from growing primary cancer cells in the regional tissue then detachment of the cancer cells from the primary tumour, invasion through the myoepithelium to the nearby lymph nodes, or travel into blood stream and subsequent spread to other organs. Both genetic alterations and environmental factors are involved, with monoclonal and polyclonal expansion of the tumour considered as one of the main factors for formation of the premalignant lesions that turn to invasive tumours later on (Valastyan & Weinberg, 2011). Only a small populations of cells within the primary tumour have the ability to induce metastasis, which are termed circulating tumour cells (CTC). Understanding the mechanism of invasion and transition is important in early detection of CTC in the blood stream, which is the key point to identify early spread of the tumour and prevent metastasis process in the patient (Maheswaran & Haber, 2010).

Breast cancer has a high capability to induce metastasis. The first potential recurrence site of breast cancer is the bone, followed by lung or liver and the central nervous system. Recurrence is reported between 2 to 3 years after the surgical treatment. This might be higher in oestrogen receptor-negative tumours (Schmidt-Kittler *et al.*, 2003).



**Figure 1-11: Metastasis in breast cancer.**

Modified from (Weigelt *et al.*, 2005)

### 1.8.6 Treatment options for breast cancer

There are many therapeutic options for breast cancer treatment. The early stage of identification, mainly starts with surgery which could involve either a lumpectomy or mastectomy operation, followed by a combination of adjuvant therapies such as radiotherapy, chemotherapy, endocrine therapy, molecular targeted therapies such immunological therapies and cytotoxic chemotherapy (Vilarino-Varela *et al.*, 2009).

Radiotherapy might assist a patient with oestrogen receptor-positive better than oestrogen receptor-negative and triple-negative breast cancer. In this therapy, high energy X-rays or other types of irradiation are employed and delivered to the lymph nodes or chest wall including the surgical scar. This increases the overall survival via reduced recurrence and breast cancer death among the breast cancer patients (Group, 2011).

Other treatment agents for breast cancer include systematic treatment, which is typically used for removing metastatic disease. It involves chemotherapy, hormonal therapy and other

targeted therapies. It might apply prior to surgery to shrink larger tumours or after lumpectomy or mastectomy. Conventional chemotherapy initially controls tumour growth by inducing cell death and preventing cell proliferation. However, small cell populations with tumorigenic potential may express resistance to chemotherapy, the number of these cells is expected to increase after the treatment. Consequently, the gene expression profile in tumour-resistant cells will be different before and after treatment, especially genes involved in cell cycle arrest and survival pathways (Li *et al.*, 2008). Chemotherapy is widely applied to Oestrogen receptor negative or HER-2 positive patients than other subtypes breast cancer patients.

In those patients with oestrogen and or the progesterone receptor positive, endocrine therapy is highly effective on ER+ metastatic breast cancer and improves patient survival. Like chemotherapy, this approach will lead to resistance over the time. The mechanism of action of endocrine therapy involves blocking the ER pathway, this induces deregulation within the ER pathway and alterations in cell survival signalling, such as increased signalling of growth factor receptor pathways, especially the EGFR/HER2 pathway. Therefore, a new combination therapy of endocrine therapy plus chemotherapy has been used in targeting both ER and growth factor receptor signalling which blocks the interaction between aforementioned pathways and reducing the possibility for creating escape-signalling routes (Clark *et al.*, 1984; Osborne & Schiff, 2011).

There are two anti-oestrogen drugs used in endocrine therapy known as Tamoxifen and aromatase inhibitors. Tamoxifen acts as anti-cellular proliferation agent via preventing oestrogen-regulated gene transcription, this could occur because of conformational changes in the receptor after Tamoxifen binding to oestrogen receptor. This modification in the receptor eventually inhibits access to one of the co-activator interaction surfaces that are important in cell proliferation process (Johnston & Yeo, 2014; Osborne, 1998).

Aromatase inhibitors act as an alternative or in a sequential order after initial tamoxifen therapy by reducing the level of circulating oestrogen by more than 80% in postmenopausal women, therefore, clinical studies have reported highly effective results on women with metastatic breast cancer (Brueggemeier, 1994).

Molecular targeting therapy is another example of a breast cancer therapeutic system. It involves targeting molecules that play exceptional roles in cell proliferation and promoting cancer. This type of treatment is known as biotherapy.

There are several examples of molecular that has been developed as biotherapy, such as tyrosine kinase inhibitors, which inhibit the additions of a phosphate group to the protein tyrosine kinase (phosphorylation) and lead to inactivation this protein, and PARP (Poly (ADP-ribose) polymerase) inhibitor enzyme activity, thus preventing the repair of DNA damage. The PARP is widely used in combined therapy for treatment of triple-negative breast cancer (Munagala *et al.*, 2011).

Some clinical data have shown sensitivity of triple-negative tumours to the above-mentioned therapeutic options. In addition, some resistance has reported with the new generation of targeted cancer therapeutics, which look similar in mechanisms of resistance to endocrine therapies, this due to reactivation of the signalling pathway that has not fully block at the first attempt.

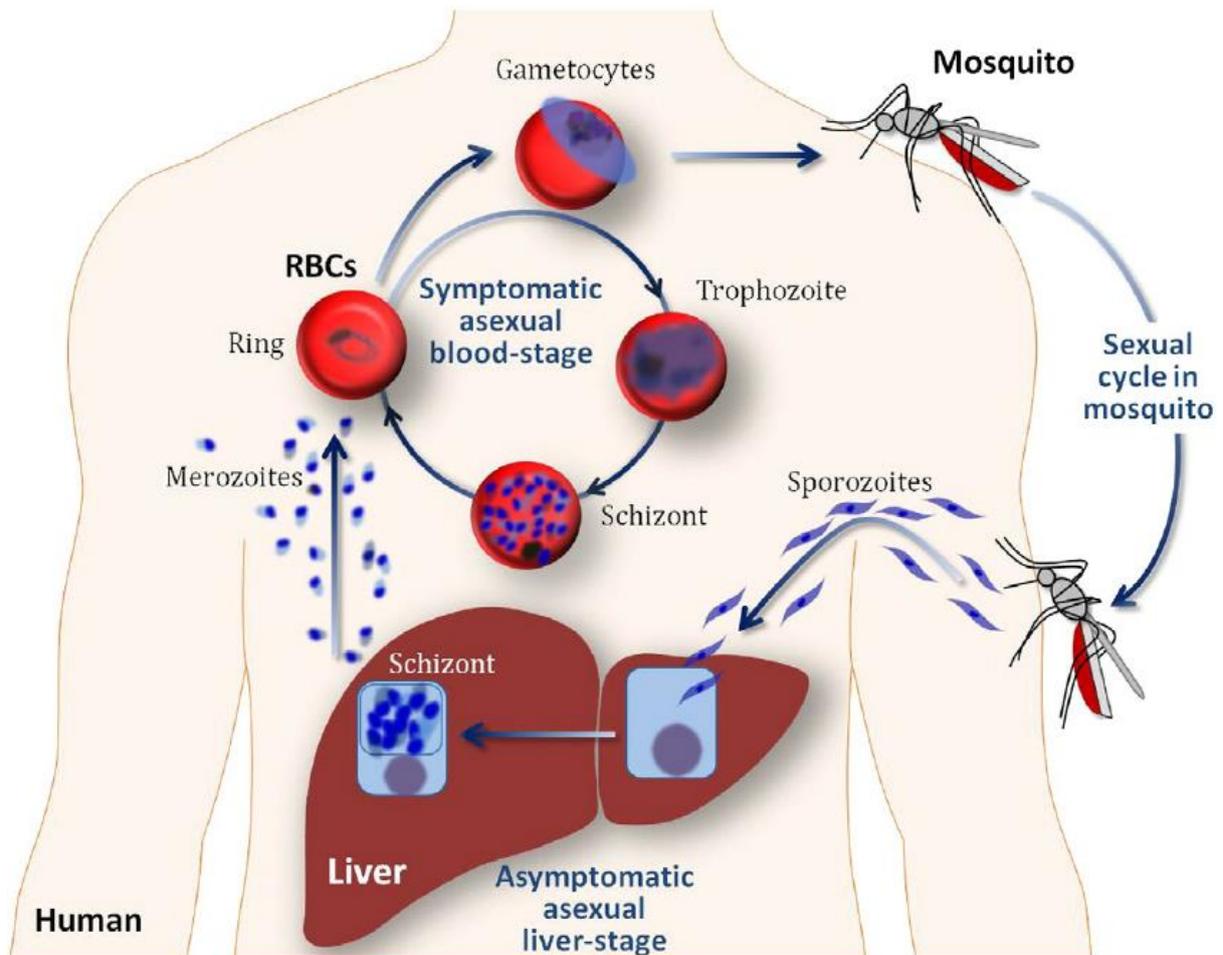
## **1.9 Malaria**

Malaria is the second disease that has been targeted by whelk GAG treatment in this study. Well known as a fatal tropical disease, the infection is caused by a single mosquito bite, through which the Plasmodium parasite that cannot survive outside the host, is transferred into the red blood cells. Only the female of certain species of Anopheles mosquito can transmit malaria when they take a blood meal for egg development. The parasite develops

inside the mosquito in stages, from gametocyte to sporozoite, and several factors, such as temperature and humidity, allow the parasite to complete its cycle in the mosquito host. The majority of malaria deaths are caused by a severe infection with *Plasmodium falciparum*, *Plasmodium vivax* being the second most common species to cause infection. Two other species, *P. malariae* and *P. ovale*, are responsible for a small percentage of infection across the world (Urbán *et al.*, 2011).

Symptoms appear between 7 to 18 days after infection and include fever, sweats and chills, headaches, vomiting, muscle pains and diarrhoea. In some cases, symptoms might not appear for up to a year or more after infection. According to recent estimates, approximately 300-500 million people suffer from this disease, the majority of them in sub-Saharan Africa and Colombia, and 2-3 million people die of malaria every year (Vangapandu *et al.*, 2007).

The life cycle of the Plasmodium parasite in the human body begins when an infected female Anopheles mosquito takes a blood meal and inoculates sporozoites into the bloodstream. Having entered the liver and infected hepatocytes, the parasite proliferates into thousands of merozoites, which then rupture from the hepatocytes and invade the red blood cells (RBCs), where they develop into early trophozoites known as ring forms and then into the late form schizonts. The blood cycle begins again when schizonts infect RBCs and release more merozoites that may develop into male and female gametes. At this point the parasite can be transmitted to another person through the bite of a female Anopheles mosquito once it completes its sexual reproductive cycle inside the mosquito (Figure 1.12).



**Figure 1-12: The life cycle of malaria.**

(Bautista *et al.*, 2014)

Most of the symptoms and pathologies appear at the blood cycle infection stage, which makes it the main target stage for most antimalarial therapeutic agents available currently (Griffith *et al.*, 2007).

Treatment options and length of treatment depend on several factors, such as the type of malaria, the severity of the symptoms, the country in which the infection took place, whether the patient took an antimalarial to prevent the disease and whether the patient is pregnant. There can be serious complications from malaria, including severe anaemia, in which the red blood cells are unable to carry enough oxygen around the body, and cerebral malaria in which small blood vessels in the brain become blocked, causing brain damage and coma. Currently,

antimalarial drugs have many limitations, such as varying degrees of toxicity and the existence of certain strains of parasite that have developed resistance to antimalarial treatments as a result of low doses that are required to avoid the toxicity; together these factors create an urgent need to discover and develop new strategies and novel approaches to treating malaria in addition to combined therapy approaches (Wells *et al.*, 2009).

### **1.10 Aims**

The bioactive compounds isolated from marine organisms play a key role in many biological processes according to many *in vivo* and *in vitro* models. However, many bioactive molecules and structures from different marine species remain to be discovered. The therapeutic values of these bioactive polymers is increased by identifying their fine structure and using this information to design small molecular weight fragments. Structure elucidation of GAG's from natural products is a big challenge with respect to their complexity.

In this study, the overall aims are to elucidate the structure of a complex GAG isolated from whelk (marine invertebrate), and monitor their bioactivity of on several types of cancer cell lines *in vitro*. Finally, the project focused on the mechanism of action of these compounds and their ability to control cell growth via inducing apoptosis or cell death among triple negative breast cancer. These types of breast cancer require urgent action to find alternative therapeutic agents.

This will be achieved in several stages; Firstly, identification of active compounds from the complex GAG chain mixtures isolated from whelk providing further understanding of the molecular composition of the whelk GAG derivative that may responsible for inhibition of cancer cell proliferation. Intact GAGs will then be isolated from the soft body of whelks and enzymatic degradation applied to produce fragments that will test for cancer cell growth inhibition. The fragments analysed by high performance liquid chromatography (HPLC-MS)

using isotopic aniline tagging, GC-MS, quantitative analysis of monosaccharides by high performance anionic exchange chromatography-pulse amperometric detection HPAEC-PAD and an extensive structure activity study and comparison with mammalian GAGs.

The second objective to examine the bioactivity of the intact and fragmented GAG chains on several cancer cells *in vitro*.

Final objective to characterise the mechanism of action of the bioactive compounds from whelk on triple negative breast cancer cell lines, by exploring the signalling or other pathway involved directly, or indirectly, in cell proliferation by quantifying the total protein differentially expressed with and without certain GAG concentrations using label free proteomics approaches and pathway analysis.

This study also aimed to determine the ability of these compounds to inhibit a small population within the breast cancer cells that have stem cell activity by using mammosphere assay on two subtypes of breast cancer.

Finally, this study aimed to examine the activity of whelk GAGs on another fatal diseases such as malaria, where there is an urgency to develop a new therapeutic agent against malaria, especially after significant increases in developing resistance toward several chemotherapeutic, this is further demonstrating in chapter 4.

# **Chapter 2 : Materials and Methods**

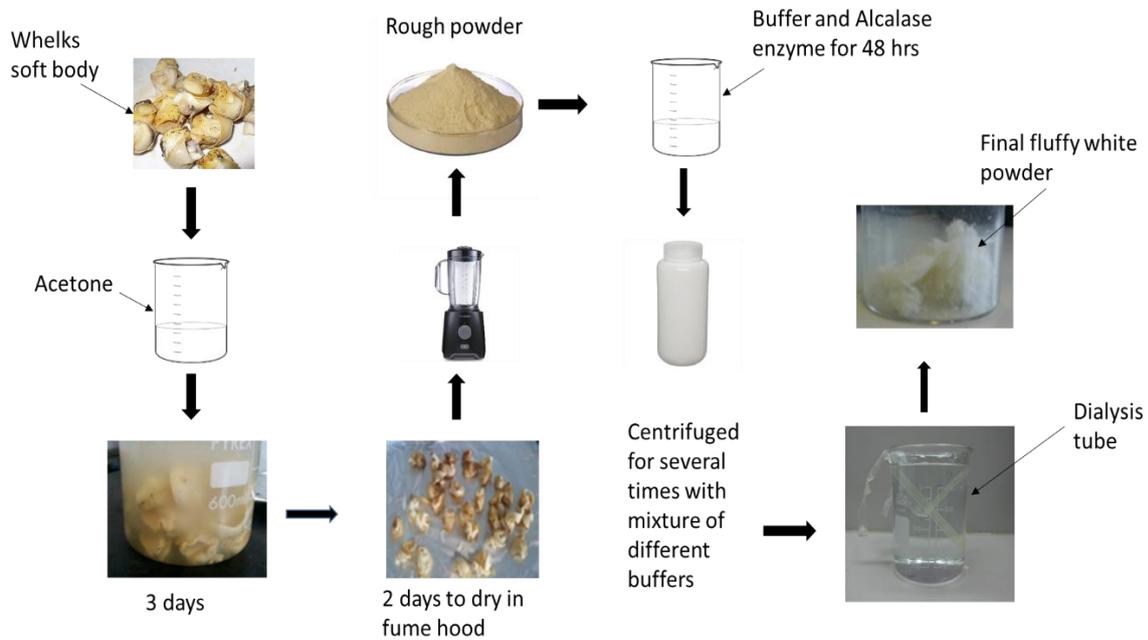
## **2.1 Materials**

All materials and buffer recipes listed in detail in appendix I.

## **2.2 Methods**

### **2.2.1 Method of GAG isolation from soft body of Whelk.**

Crude GAGs were extracted as a white powder from shellfish common whelks (*Buccinum undatum*), the shell was removed and the whole soft body was defatted by incubation with acetone for 72 hours. The acetone was changed every day and then the whelks left to dry for 24-48 hours. A blender was used to convert the fat-free dried whelks in to a powder. Following the extraction methods from (Kim *et al.*, 1996), 4 grams of the powder was suspended in 40 ml of 0.05 M of sodium carbonate buffer with a pH of 9.2, then 2 ml of Alcalase enzyme was added to degrade protein and release GAG chains, this step takes approximately 48 hours at 60°C with constant agitation of 200 rpm. The mixture was cooled at 4 °C then 5% (w/v) trichloroacetic acid was added to the sample and mixed to precipitate peptides. After 10 minutes the mixture was centrifuged at 8000 rpm for 20 minutes. Three volumes of 5% (w/v) potassium acetate in ethanol were added to one volume of supernatant and the solution was stored at 4 °C overnight then centrifuged for 30 minutes at 8000 rpm. The precipitate, approximately 1 g, was dissolved in 40 mL of 0.2 M of NaCl solution and centrifuged at 8000 rpm for 30 minutes and the supernatant was discarded. Cetylpyridinium chloride (0.5 mL of 5% (w/v) solution) was added to the supernatant followed by centrifugation to recover the precipitate, which was subsequently dissolved in 10 ml of 2.5 M of NaCl. Finally, 5 volumes of ethanol were added and the precipitate obtained by centrifugation for 30 minutes at 10000 rpm. The precipitate was then dialyzed against water for 72 hours. The water was changed twice every day and then the sample was lyophilized to obtain a white powder containing approximately 2.0 mg crude GAG.



**Figure 2-1: Whelk GAG extraction process.**

## 2.3 Enzyme Preparations

Two types of enzymes were used in de-polymerisation of intact GAG chain. First, heparinase fractions and second, chondroitinase ABC enzymes.

### 2.3.1 Preparing of Heparinase

Heparinases were supplied by Grampian Enzymes. Three different subtypes of this were used (Heparinase I (EC 4.2.2.7), 2 IU, Heparinase II, 0.25 IU and Heparinase III (EC 4.2.2.8), 0.5 IU). The enzyme buffer used was 50 mM sodium acetate, 0.5 mM calcium acetate pH 7.0. All enzymes were adjusted to 1.0 IU/ml by adding 1X heparinase buffer containing 10µg/ml of BSA as stabiliser, then split into 100 µl aliquots each containing enzymes at 1.0 IU/ml. These stocks were stored at -80 °C. The working stocks are prepared by taking one of the above aliquots and diluting to a concentration of 200 mIU/ml by adding 400 µl of fresh 1X heparinase buffer without BSA to make up to a volume of 500 µl of enzymes to a final concentration of 200 mIU/ml and stored at -20 °C.

### **2.3.2 Preparation of ABC lyase**

Chondroitinase supplied from IBEX Pharmaceuticals Inc. (Montreal, QC). Two fractions of chondroitin sulphate lyase from *Flavobacterium heparinum* used. AC (EC Number 4.2.2.5) enzyme cleaves sulphated and non-sulphated polysaccharide chains containing 1-4 linkages between hexosamines and glucuronic acid residues. The reaction yields oligosaccharide products mainly disaccharides containing unsaturated uronic acids, which can be detected by UV spectroscopy at 232 nm. The enzyme is active on chondroitin sulphates A and C and hyaluronic acid, but is not active on dermatan sulphate (chondroitin sulphate B), also active at pH range 3.5 – 9 with chondroitin sulphate A, and 4.5 - 9 with chondroitin sulphate C. The enzyme Chondroitin B (E.C. 4.2.2.19) cleaves polysaccharide chains containing 1-4 linkages between hexosamines and iduronic acid residues in dermatan sulphate (chondroitin sulphate B). The reaction yields oligosaccharide products mainly disaccharides containing unsaturated uronic acids, which can be detected by UV, spectroscopy at 232 nm, the pH range for activity is 5 – 10. The stock split into small Eppendorf tubes each containing 1.5 µl aliquots kept frozen at – 70 °C at 5IU/ µl concentration.

## **2.4 Enzymatic digestion**

### **2.4.1 Enzymatic degradation of Heparan Sulphate (GAG-HS)**

The starting material for this procedure was 100µg dried, desalted GAG's from whelk, which was dissolved in heparinase buffer prior to digestion of HS/Hep (50 mM sodium acetate, 0.5mM calcium acetate pH7.0). A mixture of Heparinase I, II and III (30mIU/ ml HepI, 16 mIU/ ml Hep II, 16 mIU/ ml Hep III) was added and the polysaccharide digested at 37 °C for 24 hours, the reaction was terminated by heating at 100 ° C for 5 minutes. The sample was run with gel filtration chromatography.

#### **2.4.2 Enzymatic degradation of Chondroitin Sulphate (GAG-CS)**

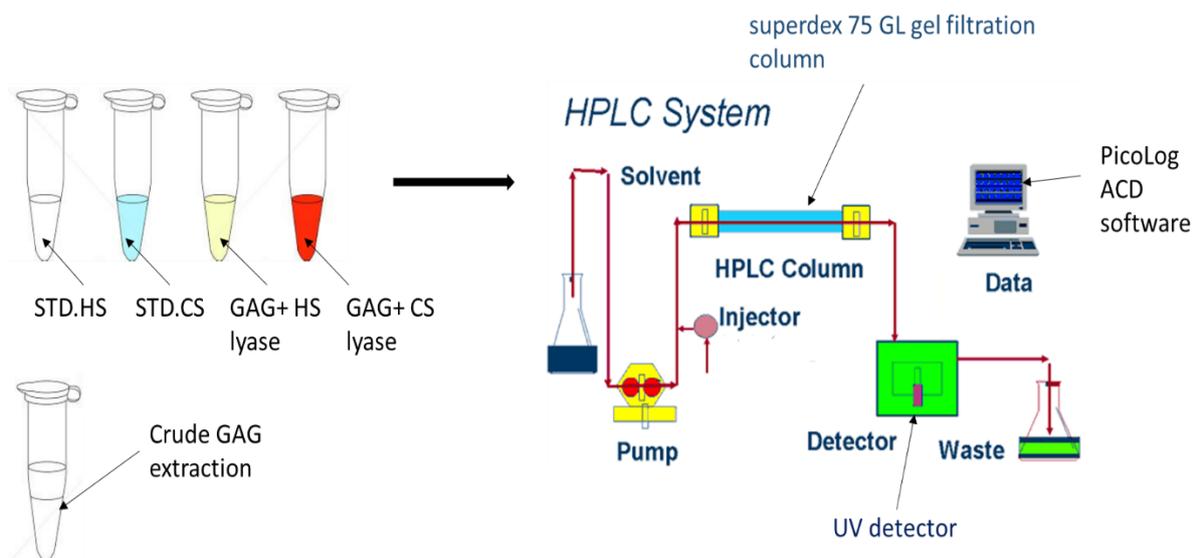
The starting material for this procedure was 100 µg dried, desalted GAG's from whelk, which were dissolved in 50 mM Tris-HCl/ 50 mM sodium acetate pH 8.0 buffer. A mixture of Chondroitinase ABC (1.5 µl) of 5 IU/µl AC + (1.5 µl) of 5 IU/µl B was added and the polysaccharide was digested at 37 °C for 24 hours, the reaction was terminated by heating at 100 °C for 5 minutes. Sample was run with gel filtration chromatography.

#### **2.5 Standard Preparation of HS and CS**

A standard commercial porcine intestinal mucosal heparan sulphate and chondroitin sulphate were provided from Celsus (Cincinnati, OH, USA). All standards were digested with hepinases and chondroitinase lyase at 37 °C for 24 hours, the reaction was then terminated by heating at 100 °C for 5 minutes, the mixture was fractionated using a 10,000 MWCO filter to remove enzyme and undigested GAG chains, which were then used as standards in all analytical techniques and bioactivity screening assays.

#### **2.6 Size exclusion analysis**

Aliquots of 100 µl containing 100 µg of crude mixture of whelk GAG and a corresponding concentration of each intact HS/CS standards and digested GAG/HS, GAG/CS were applied to a superdex 75 10/300 GL gel filtration column from GE healthcare (Bed Volume 24 ml), The column was pre-equilibrated with PBS (phosphate buffered saline, 0.15 M NaCl, 20 mM Na<sub>2</sub>HPO<sub>4</sub>) pH 7.2, then the sample was applied and eluted from the column at a flow rate of 1.0 ml/min, this was monitored for absorbance at 232 nm. The data was validated by using a PicoLog ACD software (Pico technology Ltd, Cambridge, UK).



**Figure 2-2: Size exclusion analysis of GAG samples and STDs.**

## 2.7 Quantitative analysis of Monosaccharides by HPAEC-PAD

Samples containing 1 mg of crude GAG were dissolved in 1 ml deionized water, and a 100  $\mu$ l (10% GAG) aliquot was treated with 100  $\mu$ l of 4N trifluoroacetic acid, 4N TFA (prepared by adding 2.375 ml deionised water to 1.00 ml of 13.5 M TFA, this provides a stock with 4N TFA) at 100°C for 4-6 hours to cleave all glycosidic linkages including uronic acids. Following hydrolysis sample tubes were removed from the heating block and allowed to cool to room temperature. The samples were centrifuged at 27000 rpm for 3 minutes to bring down any condensation on the sides and cap of the tube. Excess acid was evaporated by dry nitrogen flush for 10 minutes, and repeated co-evaporation with 70  $\mu$ l of 50% isopropanol to get rid of the acids completely. Finally, the samples were dissolved in 100  $\mu$ l deionised water and spun down, then analysed by HPAEC-PAD using a CarboPac PA-1 column 4 mm x 250 mm, 4 $\mu$ m, with a 4 mm x 50 mm Guard. Solvents were used as below:

A: Water.

B: 100 mM sodium hydroxide (NaOH) with 5 mM (sodium acetate trihydrate) NaOAc (Appendix I).

Initial conditions of 16% B at 1.0 mL/min were applied, a pulsed amperometric detector PAD was used in this system. In PAD applications, the electrode is automatically cleaned and prepared for detection by each cycle of the programmed series of potentials (i.e., the waveform), thereby minimizing electrode fouling by oxidation products of the analyte and other sample compounds, and thus maintaining consistent response.

Monosaccharide standards were treated in parallel and used for calibration of HPAEC-PAD response. Table 2-1 shows the gradient settings used for monosaccharides analysis.

**Table 2-1: The gradient settings used for monosaccharides.**

<b>Time/mins</b>	<b>%A</b>	<b>%B</b>
<b>0</b>	84	16
<b>20</b>	84	16
<b>21</b>	0	100
<b>31</b>	0	100
<b>32</b>	84	16
<b>50</b>	84	16

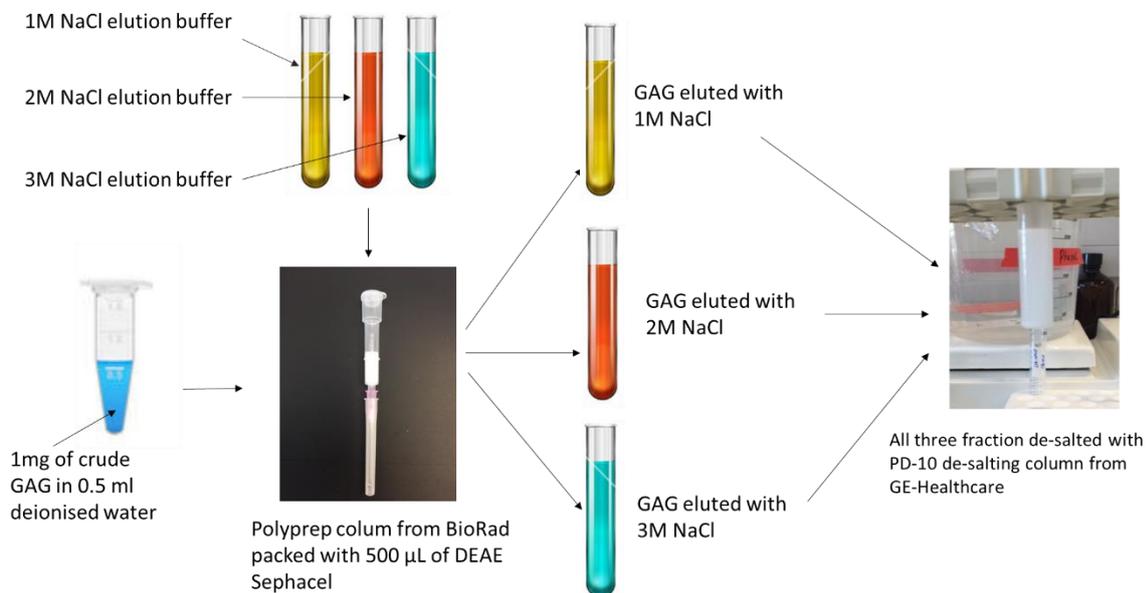
## **2.8 Ion exchange of intact GAG**

Small Polyprep column from BioRad was packed with 500 µL of DEAE Sephacel (GE-Healthcare), the bed was calibrated with 5 mL of Diethylaminoethyl DEAE Pre-wash buffer (see Appendix I). 1 mg crude GAG was dissolved in 500 µl deionised water, 250 µl was

loaded onto the column, then the column was washed with 5 mL DEAE washing buffer (see Appendix I), this is called flow through buffer (more information provided in appendix III). The GAGs were eluted by adding 2.5 mL of DEAE elution buffers (1,2 and 3) (see Appendix I), each fraction was collected separately. All samples were de-salted using PD-10 column.

## 2.9 PD-10 de-salting technique

PD10 columns from GE-Healthcare were used for desalting and buffer exchange. samples preparation in the range of 1.0 to 2.5 ml by gravity flow was carried out. Desalting capacity was greater than 90%, the recovery was in the range of 70 to >95%. Bed dimensions were 14.5× 50 mm and the columns were prepacked with Sephadex G-25 medium. The column was prepared by washing 6X with 10% ethanol, then 1 ml of sample was loaded and washed with 2.5 mL of 10% ethanol, the first 2.5 mL of elutant discarded as it is the void volume of the column. The next 3.5 mL of 10% ethanol were collected and lyophilised.



**Figure 2-3: Ion exchange chromatography and de-salting technique of intact GAG.**

## **2.10 Enzymatic degradation of ion-exchange fractions**

The crude GAG was purified depending on the charge groups using ion exchange chromatography as mentioned above. Three fractions were collected, each eluted with different gradient of NaCl salt. Samples were coded as:

- CK1 for 1M NaCl.
- CK2 for 2M NaCl.
- CK3 for 3M NaCl.

100µl (10%) of each fraction was prepared for monosaccharide quantification analysis, and 70µl (7%) was injected in the system (more details about the calculation provided in appendix III). According to the amount of substrate provided in each fraction, 10 µl from CK1 and 62.5µl of both CK2 and CK3 were used as starting material for enzymatic depolymerisation.

### **2.10.1 Degradation of HS fractions**

Each ion exchange fraction CK1, CK2 and CK3 was incubated with a mixture of Heparinases (I, II and III) in the amounts set out below for 24 hours. The same protocol was applied to standard porcine HS.

CK1- 10µl Substrate + 30mU HepI + 16mU Hep II+ 16mU Hep III.

CK2- 62.5 µl Substrate + 30mU HepI + 16mU Hep II+ 16mU Hep III.

CK3- 62.5 µl Substrate + 30mU HepI + 16mU Hep II+ 16mU Hep III.

HS-Std.- 10 µl Substrate + 30mU HepI + 16mU Hep II+ 16mU Hep III.

All samples were fractionated using a 10,000 MWCO filter to remove enzyme and undigested GAG chains, then centrifuged at 14000Xg for 20 minutes. The flow through (the material in the bottom of the Eppendorf tubes) was collected as digested oligosaccharide, mainly disaccharide, and lyophilised for further analysis. The retained fraction of the filtrate

(the material on the top of the filter in the Eppendorf tubes) which is mainly undigested material or enzyme was also dried down for further investigations.

### **2.10.2 Degradation of CS fractions**

Each sample fraction CK1, CK2 and CK3 was incubated with CS-ABC lyase for 24 hours. The same protocol was applied to standard CS. The amounts of enzyme used are set out below:

CK1- 10 $\mu$ l substrate + (1.5  $\mu$ l) of 5 IU/ $\mu$ l AC + (1.5  $\mu$ l) of 5 IU/ $\mu$ l B.

CK2- 62.5  $\mu$ l substrate + (1.5  $\mu$ l) of 5 IU/ $\mu$ l AC + (1.5  $\mu$ l) of 5 IU/ $\mu$ l B.

CK3- 62.5  $\mu$ l substrate + (1.5  $\mu$ l) of 5 IU/ $\mu$ l AC + (1.5  $\mu$ l) of 5 IU/ $\mu$ l B.

CS-Std.- 10  $\mu$ l substrate + (1.5  $\mu$ l) of 5 IU/ $\mu$ l AC + (1.5  $\mu$ l) of 5 IU/ $\mu$ l B.

All samples were fractionated by 10,000 MWCO filter to remove enzyme and undigested glycosaminoglycan chains, then centrifuged 14000Xg for 20 minutes. The flow through in the bottom of the Eppendorf tubes was collected as digested oligosaccharide mainly disaccharide, and lyophilised for further purification methods. The retained fraction of the filtrate which is mainly undigested material or enzyme was also dried down for further analysis.

### **2.11 LC/GC-MS Techniques**

TMS derivatisation analysis of polysaccharide by GC-MS and isotopic aniline tagging for GAG disaccharide analysis by mass spectrometry (GRIL-Glycan Reductive Isotope Labelling) were used for composition analysis and mass detection of disaccharide yield from both HS lyase and CS lyase digestions and filtration (Lawrence, Olson, *et al.*, 2008).

### **2.11.1 Isotopic Aniline tagging for GAG disaccharide analysis by mass spectrometry (GRIL-Glycan Reductive Isotope Labelling)**

15  $\mu\text{l}$  of  $^{12}\text{C}_6$  aniline and 15  $\mu\text{l}$  of 1 M of 95% Sodium cyanoborohydride/Sodium cyanotrihydridoborate ( $\text{NaCNBH}_3$ ) (Sigma-Aldrich) freshly prepared in dimethyl sulfoxide: acetic acid (65:35 v/v), and added to each GAG fractions from ion exchange chromatography (CK1, CK2 and CK3), each at concentration mentioned in sections (2.11.1 and 2.11.2) 8 pmol of HS and CS standard disaccharides were used and mixed with the tagging buffer above.

Reactions were carried out at 37 °C for 16 hours, and then dried in a centrifugal evaporator. After complete sample tagging with  $^{12}\text{C}_6$  aniline, 17 $\mu\text{l}$  of LC-MS grade water (water, 8 mM AcetAcid, 5 mM DBA) was added to each fraction (CK1, CK2 and CK3), mixed gently to produce homogenate mixture, sample spun down at 14000Xg for 7 minutes, then 5  $\mu\text{l}$  was taken to LC-MS vials and spiked with 2  $\mu\text{l}$  of 8pmol internal  $^{13}\text{C}_6$  aniline tag, 1  $\mu\text{l}$  of 10X GRIL buffer and 2  $\mu\text{l}$  of LC-MS water were added to bring up the mixture to the total volume of 10  $\mu\text{l}$ , only 4  $\mu\text{l}$  was injected to the system. A quaternary high-performance liquid chromatography pump (Thermo-Finnigan, San Jose, CA) was used for disaccharide analyses. Aniline isotopic disaccharide and non-isotopic disaccharides were separated on a C18 reversed-phase column (0.46 cm x 25 cm, Vydac) with the ion pairing agent dibutylamine (DBA, Sigma-Aldrich). The gradient system used is shown below:

- 100% buffer A contained (8 mM acetic acid, 5 mM DBA) - 10 minutes
- 17% buffer B contained (70% methanol, 8 mM acetic acid, 5 mM DBA)- 15 minutes.
- 32% buffer B for 15 minutes.
- 40% buffer B for 15 minutes.
- 60% buffer B for 15 minutes.
- 100% buffer B for 10 minutes.
- 100% buffer A for 10 minutes.

Most disaccharides eluted in 60% buffer B (42% methanol), classic quadrupole ion trap mass spectrometer equipped with an electrospray ionization source was used, negative ion mode was used to observe ions of interest. The capillary temperature and spray voltage were kept at 140 °C and 4.75 kV to decrease in-source fragmentation of sulphated disaccharides.

### **2.11.2 Trimethyl Silyl derivative for GC-MS (TMS)**

GC-MS (TMS) is a powerful tool that successfully analyses methyl glycosides of neutral and *N*-acetylated amino sugars and methyl esters of uronic acids using either DB-1 or DB-5 capillary column. Trimethylsilyl derivatives of these sugars can be detected from their retention times by electron impact (EI) ionization mass spectrum comparing to standards.

For the material used in this experiment see appendix I. The standard and sample preparation were as below:

- **1 M Methanolic-HCl:** 1 mL of 3 M Methanolic-HCl slowly dissolved with 2 mL of cold Methanol on ice in a glass screw-top tube.
- **Standard Preparation:** To prepare the TMS standard 10 µg of each sugar placed into a glass sample tube along with 1 µg of myo-inositol, as an internal standard and lyophilize.
- **Sample Preparation:** 100 µg of sample was taken in a glass screw cap tube and 1 µg of myo-inositol added and lyophilize.

### **PROCEDURE**

1. 200 µL of 1 M Methanolic-HCl added to the sample and tap gently to mix. The screw cap tightens, and incubated overnight (16-18 hr) at 80 °C.

2. The sample cooled to room temperature, and then the Methanolic-HCl removed by dry nitrogen flush, and then 100 µL of methanol added to remove residual Methanolic-HCl and the sample brought to dryness using nitrogen flush.

3. Next re-*N*-acetylate the samples by adding 4:1:1 Methanol:Pyridine:Acetic Anhydride to a final volume of at least 200  $\mu$ L. the tubes tightly capped and incubate at 100 °C for 1 hr.
4. The sample cooled again to room temperature and brought to dryness using dry nitrogen flush.
5. Samples derivatized with the addition of 200  $\mu$ L of Tri-Sil HTP Reagent, re-capped and heated at 80 °C for 30 min. Samples cooled down and removed excess Tri-Sil HTP by dry the samples by dry N<sub>2</sub> flush.
6. 1 mL of hexane added to each sample and then vortexed and sonicated to break up the dried sample. Next, centrifuged for 2 min at 2000 rpm.
7. The sample filtered through a Pasteur pipette packed with glass wool. The glass wool prepared first by washing with a small amount of hexane.
8. The filtrate collected and the hexane removed by dry nitrogen flush.
9. The samples reconstitute in 100-150  $\mu$ L of dichloromethane and transfer into a sample vial for analysis by GC-MS.

**GC-MS SETTINGS:**

Carrier Gas: Helium

Inlet Conditions :

Temperature: 220 °C

Pressure : 12.897 psi

GC-MS Transfer Line Temp: 280 °C

Column: DB-5 or Equivalent, 30 m x 0.25 mm x 0.25  $\mu$ m

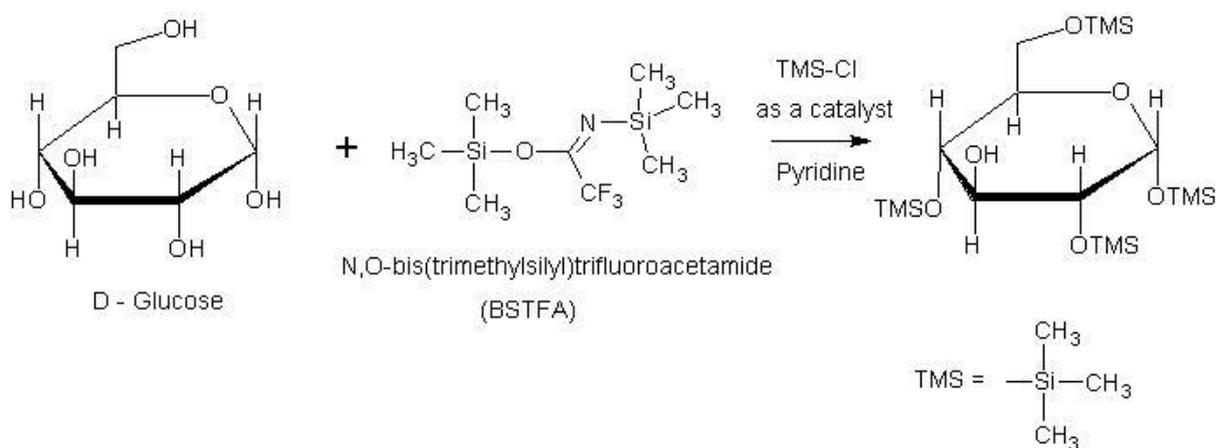
Column Flow: 1.1971 mL/min

Injection Volume: 1  $\mu$ L

Run Time: 45.667 min

**Table 2-2: GCMS conditions setting.**

Conditions	°C/min	°C	Hold (min)	Time	Run Time (min)
Initial		100	0		0
Ramp1	5	120	1		5
Ramp2	3	230	4		45.667



**Figure 2-4: Sugars reduction with sodium borohydride following conversion of polyols to polyacetate esters, then formation of TMS ethers of sugars.**

### 2.12 Periodic acid oxidation (H<sub>3</sub>IO<sub>4</sub> or H<sub>5</sub>IO<sub>6</sub>)

Periodic acid oxidation is used for polysaccharide structure elucidation and to determine the size of the ring by oxidation of carbon-carbon bonds between the alcohol and carbonyl group or adjacent diol (vicinal), each broken bond will be replaced with an OH group, in the case of two OH group at one carbon it will transform to a carbonyl group by losing water and no effects reported on ethers or acetals. Carbonyl compounds in formaldehyde might oxidize to formic acid or aldehydes, while ketones oxidize to carboxylic acids then to CO<sub>2</sub>.

975 µl from CK1 was mixed with an equal volume of 2M buffer containing 0.1 M sodium acetate, pH 5.8, 40 µg of sodium metaperiodate, dissolved at 4 °C for 56 hours in the dark. Ethylene glycol (150 µl) was added to stop the reaction, the mixture incubated for 30 mins.

100 µg of sodium borohydride was then added and the mixture neutralised with 25% acetic acid on ice. Dialysis was carried out for two days and the material dried down for further monosaccharide analysis and bioactivity screening.

### **2.13 NMR spectrometry**

NMR spectra were recorded at Manchester Institution of Biotechnology/ NMR service using Bruker NMR spectrometer operating at a 1D and 2D frequency of 800 MHz at room temperature and 70°C. <sup>1</sup>H-NMR for whelk-GAG sample fractions from ion exchange chromatography were recorded at pharmaceutical research centre in UCSD using Bruker NMR for 1D frequency of 600 MHz at room temperature. For NMR experiments, the total amount of 5mg of each fractions CK1, CK2, CK3 and crude GAG was dissolved in 500 µl deuterium oxide (Sigma Aldrich 151882-10X1ML). The chemical shifts from 1H-NMR spectra, total correlated spectroscopy TOCSY, and hetero nuclear single quantum coherence HSQC were recorded in ppm. The NMR experiments were analysed using “TopSpin” software.

## 2.14 Cell culture

### 2.14.1 Cell lines studied

Several human cancer cells were used in this study and are illustrated in detail in table 2-3.

**Table 2-3: List of cell lines studied, classification, immune profile and complete growth media.**

	<b>Cell line</b>	<b>Classificatio n</b>	<b>Immuno- profile</b>		<b>Growth medium</b>
<b>1</b>	MDA-MB-468 Breast adenocarcinoma	Basal	ER-, HER2-	PR-,	DMEM +10%fetal bovine serum+2 mM Glutamine
<b>2</b>	MDA-MB-231 Breast adenocarcinoma	Claudin-low	ER-, HER2-	PR-,	DMEM +10%fetal bovine serum+2 mM Glutamine
<b>3</b>	SK-BR3 Breast adenocarcinoma	HER2	ER-, HER2+	PR-,	DMEM +10%fetal bovine serum+2 mM Glutamine
<b>4</b>	MCF-7 Breast adenocarcinoma	Luminal A	ER+, HER2-	PR+/-,	DMEM +10%fetal bovine serum+2 mM Glutamine
<b>5</b>	A549 Lung Carcinoma	-	-	-	DMEM +10%fetal bovine serum+2 mM Glutamine
<b>6</b>	K562 chronic myelogenous leukemia	-	-	-	RPMI1640+10%fetal bovine serum+2 mM Glutamine

### 2.14.2 Routine cell culture

All adherent cell lines (MDA-MB-468, MDA-MB-231, SK-BR3, MCF-7, and A549-Lung Carcinoma) were cultured as monolayers while K562 were cultured as suspension cells. All cells were maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. The complete growth media used to culture adherent cells contain Dulbecco's Modified Eagle's Medium (DMEM-w/1,0 g/L glucose w/o L-Glutamine/ BE12-707F) from Lonza then 10% FBS (500 mL- FB-

1090/500) and 2 mM L-Glutamine (100 mL- XC-T1715/100) from Lab Tech (UK) were added. K562 also were cultured in Lonza RPMI1640 w/ 25 mM HEPES w/o Gln, 500 ml- BE04-558F without L-glutamine, liquid, sterile-filtered, suitable for cell culture from SLS, then 10% FBS (500 mL- FB-1090/500) and 2 mM L-Glutamine (100 mL- XC-T1715/100) from Lab Tech (UK) were added. Cells were grown in tissue culture flasks T-25 (Vented Filter Cap, TC-Treated-Green-30 x 10- CC7682-4825) from Star lab to approximately 70-80% confluence then media were removed and cells washed with warm PBS pH 7.4 (10010023- Thermo fisher scientific) prior to trypsinization. Cells were detached by incubation in 1ml of 1x trypsin-EDTA solution at 37 °C for 5-10 minutes (Trypsin 0.25% - EDTA in HBSS w/o Calcium w/o Magnesium w/phenol red 100ml- CC-5012) from Lonza (UK). Cells were collected by centrifugation at 1500 rpm for 5 minutes and resuspended into fresh media. Medium was renewed 3 times per week from 24 hours after each seeding. Each cell line reseeded at specific dilution ratio as below:

- MDA-MB-468: 1:4 seeding at 1-3 x10,000 cells/ml.
- MDA-MB-231: 1:3 seeding at 1-3 x10,000 cells/ml.
- SK-BR-3: 1:3 seeding at 1-3 x10,000 cells/ml.
- MCF-7: 1:4 seeding at 2-4 x10,000 cells/ml.
- A549: 1:6 seeding at 2-4x10,000 cells/ml.
- K562: 1:6 seeding at 100,000 to 1,000,000 cells/ml.

Cryopreservation of all cells performed with 90% FBS and 10% of dimethyl sulphoxide (DMSO) in 1.0 ml aliquots either in liquid nitrogen for long term storage, or -80 for short term storage.

### 2.14.3 MTT colorimetric Assay

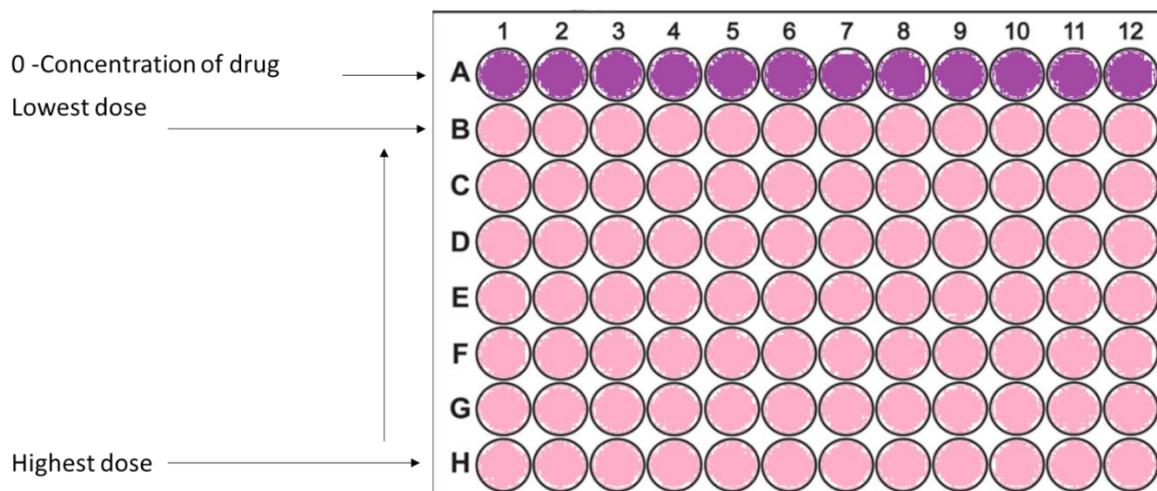
MTT assay was first described by Mosmann, (1983), who introduced a colorimetric assay for measuring cell viability, which therefore is used as a method for cell proliferation and a cytotoxicity assay. The basic principle of this assay is based on the ability of living and metabolically-active cells to reduce MTT (3, (4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide) salt to formazan product. MTT is soluble in water and, after being cleaved by the activity of mitochondrial enzymes, is converted to an insoluble crystal-form purple formazan by the cleavage of the tetrazolium ring.

Healthy cells, which had been sub-cultured at least once and subsequently reached 70-80% confluence were used for the MTT assay.

The first step was to count cells from culture using a haemocytometer. For adherent cell lines, cells have detached as per subculture (i.e. adding trypsin) then 4.5 ml media were added, then 10  $\mu$ l of cells were taken, placed on a haemocytometer for cell counting. A concentration of 3000 cells per 100  $\mu$ l per well was used. The template for the anticancer assay of the extracts is shown in Figure 2-5. The cells were then incubated at a temperature of 37 °C with 5% CO<sub>2</sub>. After 24 hours incubation, cells were examined microscopically to ensure that there were no signs of contamination and the cells were healthy. Subsequently, 100  $\mu$ l extracts of whelk GAG, filtered by a 0.2  $\mu$ m sterile filter, was added into the corresponding wells. The concentration of the extract was prepared in a serial dilution according to the microplate template, which started with a negative control in the first well (row A) until it reached the highest concentration in row H (Figure 2-5). The assay was performed in triplicate. Cis-platinum, a known anticancer drug was used as a positive control.

Afterward, the plates were incubated to allow the cell proliferation with or without the presence of the whelk GAG. After five days of incubation, the plates were taken out from the incubator and 50  $\mu$ l of MTT solution at a concentration of 5 mg/ml was added into each well.

The plates were then kept in the incubator for the next 3-4 hours until the purple crystals of formazan salt were produced, followed by removing the media through gentle aspiration of the liquid, leaving the crystals at the bottom of the wells. Subsequently, 200 µl DMSO was added into each well to dissolve the formazan crystal, followed by the reading of the absorbance of the formazan solution in microplate reader using Sigma Plot 2000 software, at the test wavelength of 450 nanometre (nm) (OD1) and reference wavelength at 690nm (OD2).



**Figure 2-5: MTT assay, plate dose pattern.**

#### **2.14.4 Statistical approach**

##### **2.14.4.1 Calculation of IC50 values**

The results were expressed as the value of optical density (OD) generated from OD1 subtracted from OD2, which represented the amount of the purple formazan crystal formation after elimination of the background reading generated from the DMSO, remaining media or the microplate plastic base. The calculation of cell viability on each well of the corresponding column on the plate is performed by the following equation:

$$\text{Cell viability (\%)} = ( \text{OD}_0 / \text{OD}_n \times 100 ).$$

Where OD<sub>0</sub> is the OD from the negative control (row A) and OD<sub>n</sub> is the OD of the corresponding well. The mean value of viable cells from each extracts then plotted on a graph and used for calculation of the 50% inhibitory concentration (IC<sub>50</sub>) of each tested extract.

The data was further analysed using Graphpad prism 5.0 (GraphPad Software, San Diego California USA). IC<sub>50</sub> values were determined using non-linear regression to fit a dose response curve to the data after transform by the drug concentrations values to log values on X and the response normalised and represented as Y values.

#### **2.14.4.2 Statistical analysis**

All data set was analysed using Graphpad prism 5.0. Data were presented as mean ± standard deviation (SD). Statistical analysis was performed by two-way ANOVA test with Bonferroni's multiple comparison post-test. A p-value of < 0.05 was considered statistically significant and R with Benjamini-Hochberg.

#### **2.14.5 Protein extraction and quantification**

Total protein extraction was performed on two triple negative breast cancer cell lines for further investigate the mechanisms of action of GAG on these cell lines by using advanced proteomic approaches at the Biological Mass Spectrometry Core Facility, Faculty of Life Sciences, University of Manchester.

##### **2.14.5.1 Cell lysates and protein concentration measurement**

MDA-MB-231 and MDA-MB-468 were cultured in T-25 culture flask to 80% confluence, then each cell line sub-cultured at an intensity of  $5 \times 10^4$  cells per well in two strains of 6 well plate to around 90% confluence. Cells were treated with whelk GAGs at a concentration equivalent to their IC<sub>50</sub> values prior to further analysis for 24 hours. The 6-well dishes were

placed on ice, and approximately 400  $\mu\text{L}$  of 0.05% SDS with 50 mM Tris HCl pH 8 added to each dish. Using a scraper, cells were removed from dish into a pre-labelled Eppendorf tube then heated at 90°C for 10 minutes. The protein quantification was measured using Direct Detect® spectrometer (DDHW00010-WW) from Millipore. The protein was measured using the average protein concentration from sample replications to get an estimated of how much to digest, in this experiment 25  $\mu\text{g}$  for each replicate was digested and processed by label-free protein quantification proteomic analysis.

#### **2.14.5.2 Proteins extraction in solution digestion technique FASP**

The proteins extracted underwent further sample preparation steps using filtered-aided sample preparation FASP. The procedure used as below:

1. 25  $\mu\text{g}$  of protein added to a spin filter tube with 200  $\mu\text{L}$  of UA2 buffer (see appendix I). Centrifuged at 14000Xg at 20°C for 15 minutes.
2. Next, 100  $\mu\text{L}$  of UA2 buffer added and centrifuged again at 14000Xg at 20°C for 15 minutes.
3. To alkylate the samples which is required to ensure maximal coverage in LC/MS-MS, 50  $\mu\text{L}$  of UA1 (see appendix I) buffer with 0.05 M iodoacetamide was added to the filters and the samples were incubated in darkness at room temperature for 30 minutes.
4. The solution centrifuged through and then wash the filters twice with 100  $\mu\text{L}$  of UA2 buffer followed by a further two washes of UA3 buffer.
5. 50  $\mu\text{L}$  of UA3 buffer was added to the filter and the protein was digested using endoproteinase LysC at protein: enzyme ratio of 1:20 at 37°C for 2 hours and a fresh collection tube was used for subsequent spins. Lyophilized LysC was dissolved in ultrapure water. Once made, the solution is stable at least until the expiration date printed on the label at -80 °C.

6. Following this the solution was diluted to 300  $\mu\text{L}$  with the addition of 250  $\mu\text{L}$  of 50 mM Tris-HCl (pH 8.5). This brings the urea concentration down from 6M to 1M. The protein was further digested with trypsin at a protein: enzyme ratio of 1:20 overnight at 37°C.

7. After digestion peptides were collected by centrifugation at 4000Xg at 20°C for 15 minutes and the filtration units were washed once with 50  $\mu\text{L}$  of UA1 buffer and subsequently with two 50  $\mu\text{L}$  washes of 50 mM ammonium bicarbonate.

8. Peptides were cleaned up with R3 beads and lyophilized and stored dry at -20°C until analysis.

Samples were de-salted prior to LC-MS/MS analysis, the de-salting procedure used here in 96 well format is as below:

- 1 mg (100  $\mu\text{L}$  of 10 mg/mL stock) of POROS R3 beads added to each well in a Corning 96 well plate.
- The plate centrifuged at 200Xg (1400 rpm) for 1 minute.
- 50  $\mu\text{L}$  of wet solution (50% acetonitrile) were added, the pellet re-suspended gently, and centrifuge at 200Xg (1400 rpm) for 1 minute. This repeated once.
- 50  $\mu\text{L}$  of wash solution (0.1% formic acid) added, resuspending gently, and centrifuge at 200Xg (1400 rpm) for 1 minute. This repeated once.
- The flow through was discarded.
- The filters removed from the FASP tubes and 100  $\mu\text{L}$  of the protein sample added to the corresponding well, resuspended gently, and centrifuge at 200Xg (1400 rpm) for 1 minute.
- another 100  $\mu\text{L}$  of sample was added and centrifuge at 200Xg (1400 rpm) for 1 minute. This has repeated until the entire sample was added.
- Washed and centrifuged with wash solution (0.1% formic acid) twice.
- The old flow through plate discarded and replaced with a fresh one.

- 50  $\mu$ L of elution solution (50% acetonitrile, 0.1% formic acid) added and centrifuged at 200Xg (1400 rpm) for 1 minute. This repeated once.
- The eluted sample transfer into chromatography sample vials and dried in the Heto SpeedVac for 2 hours.
- 10  $\mu$ L of 5% acetonitrile with 0.1% formic acid added to re-suspend the dried peptides. Vortex mix and ensured that the solution was at the bottom of the tube with no bubbles present.
- Samples were ready for LC/MS.

All materials, buffers and reagents used in this experiment are listed in details in Appendix I.

### **2.15 Label-free protein quantitation proteomics**

Digested samples were analysed by LC-MS/MS using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a QE HF (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. Peptide mixtures were separated using a gradient from 93% A (0.1% Formic Acid (v/v) FA in water) and 7% B (0.1% Formic Acid (v/v) FA in acetonitrile) to 18% B over 57 min followed by second gradient to 27% B over 14 min both at 300 nL per minute, using a 75 mm x 250  $\mu$ m i.d. 1.7  $\mu$ M CSH C18, analytical column (Waters). Peptides were selected for fragmentation automatically by data dependant analysis on a basis of the top 12 peptides with m/z between 300 to 1600Th and a charge state of 2, 3 or 4. Raw data was searched using Mascot (Matrix Science UK), against the [database selected] database with taxonomy of [species] selected. Data was validated using PROGENESIS QI. The biological role, molecular function, and cellular localization were identified using PANTHER and DAVID classification system. Further pathway analysis was performed using PANTHER and Reactome analysis software.

## 2.16 Mammosphere assay

Identification of CSC renewal activity in three types of breast cancer cells, MCF-7, MDA-MB-468 and MDA-MB-231 carried out. All buffers recipes are provided in details in appendix I. Non adherent plates were prepared as below:

1. 10X poly-HEMA stock solution was prepared by dissolving 2.4 g of Poly (2-hydroxyethyl methacrylate) in 20 ml of 95% ethanol.
2. The mixture was steering constantly on a heated plate at 55°C overnight.
3. To prepare 1 x poly-HEMA working solution, 1 part of 10x poly-HEMA stock solution was added into 9 part of 95% (v/v) ethanol and mixed well, then 200 µl added per well in 12 well plate.
4. The plate was left in the sterile hood overnight without lid to dry.
5. When the plate dried each well was washed with sterile PBS. Mammosphere culture media was prepared as below:

- DMEM/F12 (1:1).
- 10 ng/mL b-FGF.
- 20 ng/ml EGF.
- 5 µg/ml ITS.
- 1x B-27 supplement (optional) (Wang *et al.*, 2014) .

Cells were cultured to 70–80% confluence, according to standard protocols then detached using trypsin and centrifuged at 580 g for 2 minutes, media was then removed and cell pellets re-suspended in 1–5 ml of ice cold PBS. Cells were passed three times through a 25G needle to obtain single cell suspension. A Haemocytometer was used to confirm that a single cell suspension was present and to determine number of cells, then 1000 cells/ml were plated in culture plates coated with poly-HEMA to prevent cell attachment. Cells were then incubated in a humidified atmosphere at 37 °C and 5% CO<sub>2</sub> for 5 days without moving or replenishing

the media. After 5 days, media containing mammosphere was pipetted into a centrifuge tube and centrifuged at 115 g for 5 min. The supernatant was discarded and the pellet re-suspended in 300µl of 0.5% trypsin/0.2% EDTA then incubated at 37 °C for 2–3 minutes and the trypsin neutralised with double the volume of serum containing media. Again a 25 G needle was used to obtain single cell suspension and cells counted with haemocytometer. Cells were next plated in culture plates coated with poly-HEMA in triplicate and dosed with GAG in four different concentrations, first the IC50 of each corresponding cell, 20 µg/ml, 50 µg/ml, and 100 µg/ml. Plates were incubated for 5 days. Finally, images of the wells were taken using Zeiss Axio Microscope (Axio Imager 2) equipped with camera for advanced imaging.

## **2.17 General material and methods for *in vitro* culturing of malaria**

### **2.17.1 Culture of *Plasmodium falciparum***

Parasite culture procedures were carried out in the collaboration with pathogen laboratory researchers (University of Salford) using a sterile hood (ESCO class II Biological safety cabinet) and pre-sterilised equipment. Routine culture methods were optimised from (Read & Hyde, 1993). Disinfection of waste material was carried out using Virkon (Antec International, UK) before autoclaving and disposal. Media used comprised of RPMI 1640 1x (+) L-Glutamine (+) 25 mM HEPES (4-(2-Hydroxyethyl) piperazine-1-ethanesulfonic acid) (Gibco, Life Technologies, UK). Then 2.5 g Albumin bovine serum fraction V (Sigma, UK), 2.5 ml 1mg/ml hypoxanthine (Sigma, UK) in phosphate buffered saline (PBS) (Fisher Chemical,UK), 2.5 ml 40% glucose (Dextrose Anhydrous, Fisher Scientific, UK) in sterile water and 0.5 ml of 50 mg/ml gentamycin (Sigma, UK) in PBS were transferred to a 50 ml falcon tube. Then approximately 20 ml RPMI 1640 from a newly opened bottle was added and the mixture passed through a 0.22 µm filter directly into the 500 ml bottle of RPMI 1640 medium using a 20 ml syringe, the mixture gently mixed and the complete medium was stored at 2-8 ° C. The washing media used was contain RPMI 1640 1x (+) L-Glutamine (+)

25mM Hepes (Gibco,Life Technologies, UK) without additives and was stored at 2-8 °C for up to 2 weeks.

The human blood (O+ whole blood obtained from the human blood bank) was prepared for culture of Plasmodium falciparum by washing the blood immediately to remove leukocytes before use. The whole blood was centrifuged for 5 minutes at 3400 rpm using 50ml falcon tubes then blood plasma and the pale layer (the buffy coat-containing white blood cells) that forms on top of the red blood cells was removed then re-suspended in an equal volume of washing medium and following centrifugation (as described previously) the supernatant along with the buffy coat, was discarded. The washing process was carried out 3X then the haematocrit was re-suspended in complete medium and stored at 2-8 °C until further use.

*In vitro* culture of plasmodium falciparum was started by mixing 10 ml of complete medium and 0.5 ml of washed blood in 50 ml culture flask, warmed to 37 °C prior to the addition of the parasites, then 0.5 ml of parasitised blood (predominantly ring stage culture retrieved from liquid nitrogen) was then added and the parasite culture was gassed with a 5 % CO<sub>2</sub>, 5 % O<sub>2</sub> and 90 % N<sub>2</sub> gas mixture (BOC Limited, UK) and placed in the incubator (Leec culture safe touch 190 CO<sub>2</sub>, Leec Limited, Uk) at 37°C. In this experiment, smaller flasks (12.5 cm<sup>3</sup>) with a final culture volume of 5 ml were used, for every 10 ml of complete medium 1 ml of 50 % haematocrit blood was added and the media was changed every 48 or 72 hours. Routine maintenance of a Plasmodium falciparum culture included removing and discarding the old medium from the flask without disturbing the parasitized blood layer at the bottom of the flask, then the estimation of parasitaemia calculated by transferring a single drop of concentrated parasitised blood on to a microscope slide, thin blood smear used to obtain monolayer cells, then the slide dried in open air and fixed by rinsing in 100% methanol. The slide was immersed in Giemsa stain at room temperature for ~20 minutes (Gurr's Giemsa

stain solution (BDH/VWR international limited, UK) diluted at ratio 1:10 with Giemsa buffer), the buffer recipe is shown in appendix I.

The slide was gently washed with tap water, dried and viewed under oil immersion (X100) using a Leica DM 500 compound microscope. Estimation of parasitaemia was calculated as an average percentage of at least 3 fields of view and the estimated parasitaemia determined by counting the total number of red blood cells per field of view (approx. 100-200) and noting those containing parasites. Next the culture was diluted to 1 % parasitaemia with washed blood (50 % haematocrit) and new fresh complete medium was added to bring it to final haematocrit concentration of 5 %.

In order to obtain a predominantly ring stage parasite culture, the sorbitol synchronisation technique was used by adding sorbitol (5 % w/v) solution at 9 ml to 1 ml of culture pellet and incubated for 5 mins then centrifuged at 3400 rpm for 5 minutes and the supernatant was discarded. The parasite pellet then washed 3X with complete medium, eventually re-suspension in complete medium at 50 % haematocrit.

### **2.17.2 Drug susceptibility assays**

To evaluate antimalarial activity of the whelk-GAG samples, a stock solution at concentration of 25 µg/ml was prepared by dissolving 25 µg of the GAG in 1ml complete medium, the solutions were passed through a 0.22 µm filter then the infected blood treated by serial dilution of GAG stock solution from highest 25 µg/ml to lowest of 3.13 µg/ml for 48h and 72 hours respectively in black bottom 96 well plates (Nunc, Denmark). Untreated infected blood (positive control) and uninfected blood (negative control) samples were suspended at final volume of 100 µl of complete medium and a haematocrit of 5 %.

At time point of 48h and 72 hours of GAG treatment, a 100 µl of each replicate from serial concentration was transferred into separate Eppendorf tubes and washed once with PBS and

centrifuged for 90 seconds at 140000 rpm, the pellet was then re-suspended in 1 ml 5 x SYBR Green 1 solution (in PBS) and incubated in the dark for 20 minutes at room temperature followed by centrifugation for 90 seconds at 14,000 rpm. The samples were fixed by re-suspending the pellet in 250  $\mu$ l of 0.37 % formaldehyde solution in PBS (formaldehyde solution for molecular biology 36.5 %, Sigma, UK), followed by incubation for 10-15 minutes at 4 °C. The fixed sample was washed 3X in PBS and re-suspended in 1 ml of PBS, flow cytometry method was used, the FITC channel of the BD FACs Verse flow cytometer system used to record fifty thousand events per sample and the IC50 values calculated using nonlinear regression Gaphpad prism 5.0. Data was normalised and the log-transformed drug concentrations were then plotted against the dose response, the data represented as log (inhibitor) vs. normalised response-variable slope.

# **Chapter 3 : Biological activity of whelk-GAG crude extraction as an anticancer and antimalarial therapeutic agent**

### 3.1 Introduction

Natural products have been extensively studied over a period of many years. A diverse class of compounds that are associated with many desirable therapeutic properties, including anti-inflammatory, anti-oxidant and anticancer activity.

While there are many examples of therapeutic natural products that have been identified, most are not effective as drugs when introduced to biological systems. Maintaining activity can be problematic for a variety of natural compounds due to poor pharmacology linked to their being excreted, being destroyed by enzymes, or being too large to cross the cell membrane independently (Jamieson *et al.*, 2013).

The nature of GAGs and their role in cells may ultimately allow their development as effective therapeutic agents once the molecular structure that is responsible for their biological and pharmacological interaction is determined.

Researchers have identified bioactive molecules from marine organisms that are considered to be drugs and have been approved for human use in different parts of the world. These include chemical compounds isolated from sponge-derived microorganisms known as *Sorbicillactone A* which have been found to be effective against leukaemia cells without showing notable cytotoxicity (Bringmann *et al.*, 2007). Another example is Didemnin B, a compound isolated from the Caribbean tunicate *Ecteinascidia turbinata*, which showed antitumour activity against a variety of cancer cell lines including breast, ovarian, cervical, myeloma, glioblastoma/astrocytoma and lung cancers (Rinehart Jr *et al.*, 1981). This was the first marine extraction to be investigated and developed specifically for its anticancer activity (Nuijen *et al.*, 2000).

Although several studies have been conducted on marine algae, many bioactive extractions from invertebrates such as whelk (a type of mollusc) have received little attention. A few compounds have been characterised in this phylum although heterogenic, heavy sulphated

GAGs and extremely modified HS/CS are described in various species of molluscs (Dietrich *et al.*, 1989).

Despite several limitations in the discovery of bioactive lead compounds, many studies have been undertaken to find a new, interesting molecular species with potential to be developed as a therapeutic agent for certain diseases such as cancer. Some of these are cytotoxic agents against cancerous cells, target cell division and proliferation, damaging DNA, leading to cell death and evading apoptosis, depending on their overexpression and dysfunction or, for instance, deregulated signal transduction in cell cycle pathways (Fabbro & Garcia-Echeverria, 2002).

Natural products are lead compounds in drug discovery, especially for the development of new therapeutic agents. Some plant-derived products such as terpenes, phenolics and alkaloids have revealed anti-parasitic properties with high efficacy and selectivity. Quinine is a subtype of alkaloid, which has been reported to have antiplasmodial and anti-parasitic activity. Terpenes also has been reported to show some activity against protozoan parasites and the potential of tetrahydrofuran lignans such as grandisin, a phenolic subtype, to prevent the transmission of *T. cruzi* by blood transfusion has been reported (Christensen & Kharazmi, 2001; Kayser *et al.*, 2003).

This chapter focuses initially on screening the bioactivity of crude whelk GAG extracts on cell proliferation of several types of cancer, including lung carcinoma, two subtypes of positive breast cancer, liver cancer and leukaemia. The aim is to determine any selectivity of whelk GAG crude extracts toward specific types of cancerous cells. The second part of this chapter demonstrates the potential antimalarial activity of crude whelk GAG extraction, using SYBR Green dye and flow cytometry in drug susceptibility assays' as described in detail in Chapter 2 (section 2.16).

## 3.2 Results

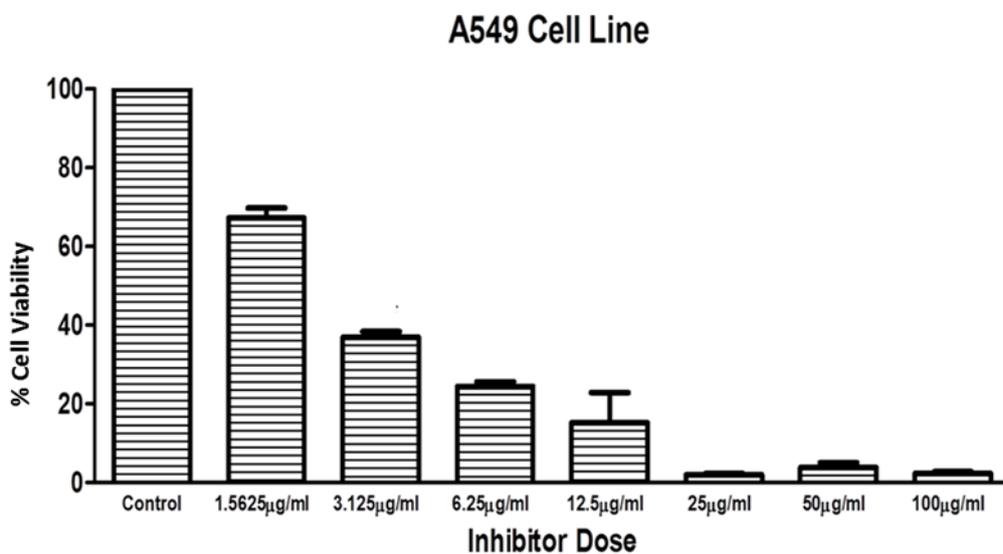
### 3.2.1 Cytotoxicity assay on lung carcinoma

Uncontrolled cell growth in the lung tissue is known as lung carcinoma. Most primary lung carcinomas cells are classified histologically into two main types: small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma (NSCLC). The latter includes several subtypes, including squamous cell carcinoma, large cell carcinoma and adenocarcinoma. Small-cell lung carcinomas grow more quickly and thus are more likely to be discovered when they have spread outside of the lung. The majority of lung cancer cases are correlated with a patient history of smoking cigarettes, although non-smokers can develop lung cancer due to other factors such as genetic predisposition or environmental hazards. Adenocarcinoma is the most common lung cancer, especially among non-smokers. Although there are several treatment options for lung cancer, including surgical removal combined with adjuvant radiation and chemotherapy, this particular cancer is associated with poor outcomes and high mortality rates (Bender, 2014; Dong *et al.*, 2016).

However, the treatment of lung cancer has improved following investigation of the molecular mechanisms involved in tumorigenesis that enable multiple oncogenic alterations to be determined. Epidermal growth factor receptor (EGFR) plays a remarkable role in treating refractory lung cancer in the patients carrying EGFR mutations. Such patients respond to the EGFR tyrosine kinase inhibitors (TKIs) gefitinib and erlotinib, achieving outcomes that could not be observed from traditional chemotherapy (Lynch *et al.*, 2004).

Identifying alternative treatment options for lung cancer that appear to be more effective and less toxic is crucial. Thus, the effect of whelk GAG crude extracts on pulmonary adenocarcinoma A549 cell line has been examined. The results shown in Figure 3.1 demonstrate the percentage of estimated cell viability over a five-day incubation period with whelk GAG intact extraction. A range of concentrations from 100  $\mu\text{g/ml}$  to 1.5625  $\mu\text{g/ml}$  was used, and the cells were cultured in monolayer per well and treated with serial dilution of

whelk-GAG concentrations and treated with serial dilution of whelk-GAG concentrations ranging from 100 $\mu$ g/ml to 1.5625 $\mu$ g/ml. Cell viability was determined by MTT colorimetric assay (purple formazan crystals absorbance) and the data was validated using GraphPad software to calculate the inhibition concentration value (IC<sub>50</sub>), as shown in Table 3.1. The results show significant inhibition activity of intact crude whelk-GAG on the A549 cell line with IC<sub>50</sub> value of 2.44 $\pm$ 0.02  $\mu$ g/ml.



**Figure 3-1: Cytotoxic effects of different concentrations of whelk GAG crude extracts on pulmonary adenocarcinoma A549 cells.**

The data is represented as the percentage of viable cells for sample size (n=3) as mean  $\pm$  SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay was performed using MTT colorimetric assay. The IC<sub>50</sub> values were calculated using nonlinear regression analysis (GraphPad Prism 5.0).

### 3.2.2 Cytotoxicity assay on two subtypes of positive breast cancer

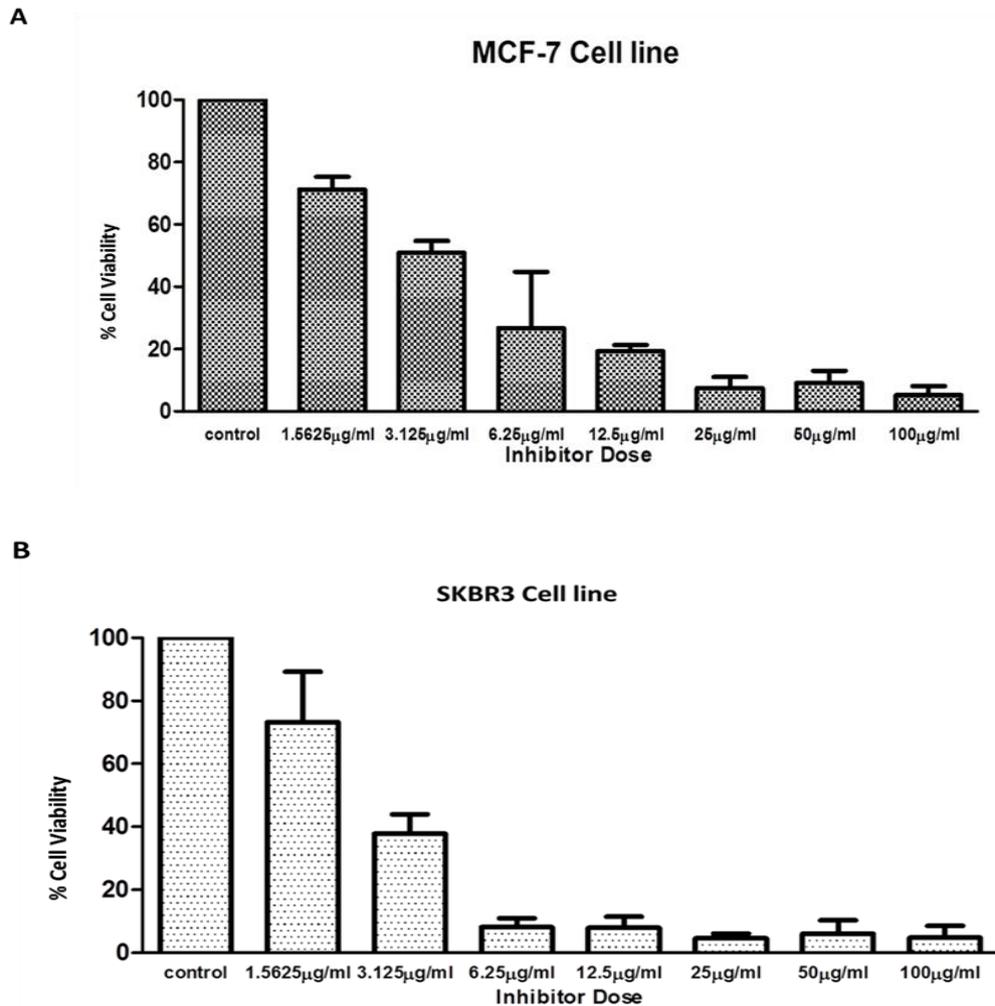
The heterogeneity of breast cancer was demonstrated in Chapter 1. Each subtype of breast cancer has a specific immune profile and other characteristics, such as low or high expression of the proliferation marker Ki67, lack or over expression of ER $\alpha$ , PR and HER2 or enrichment for markers associated with mammary cancer stem cells (CSCs) (for example, CD44<sup>+</sup> CD24<sup>-</sup>/low phenotype). Each subtype also has a different treatment response. Some are amenable to hormone therapy due to ER expression, others subtypes could subject for trastuzumab therapy, which involves the use of a monoclonal antibody to target cancer cells that are overexpressing the protein HER2 (human epidermal growth factor receptor 2). Another subtype, known as triple negative breast cancer, nonresponsive for endocrine therapy but often responsive to chemotherapy (Holliday & Speirs, 2011).

Here we aimed to investigate the anti-proliferation activity of intact crude whelk-GAG on different subtypes of positive breast cancer. The activity on triple negative breast cancer is extensively studied in Chapter 5.

Intact Whelk-GAG extracts were incubated with MCF-7 and SKBR3 cell lines. Both cell lines often are responsive to chemotherapy. However, MCF-7 is known to be endocrine responsive and low expression of the proliferation marker Ki67, unlike SKBR3 cell line. The other treatment option for this cell line is trastuzumab therapy, as SKBR3 is not responsive to endocrine therapy and is known for high expression of the proliferation marker Ki67.

The results shown in Figure 3.2 demonstrate the percentage of viable cells over a five-day incubation period treated with serial dilution of whelk GAG concentrations ranging from 100  $\mu\text{g/ml}$  to 1.5625  $\mu\text{g/ml}$ . Cell viability was determined by MTT colorimetric assay (purple formazan crystals absorbance) and the data validated using GraphPad software to calculate the inhibition concentration (IC<sub>50</sub>) value, as shown in Table 3.1.

The results illustrate significant inhibition activity of intact whelk-GAG extracts on both cell lines with IC<sub>50</sub> values of 3.24±0.03 µg/ml and 2.54±0.02 µg/ml for MCF-7 and SKBR3 respectively.



**Figure 3-2: Cytotoxic activity of whelk GAG extracts on two subtypes of breast cancer. Panel A: MCF-7. Panel B: SKBR3.**

The data is represented as the percentage of viable cells for sample size (n=3) as mean ± SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay was performed using MTT colorimetric assay. The IC<sub>50</sub> values were calculated using nonlinear regression analysis (GraphPad Prism 5.0).

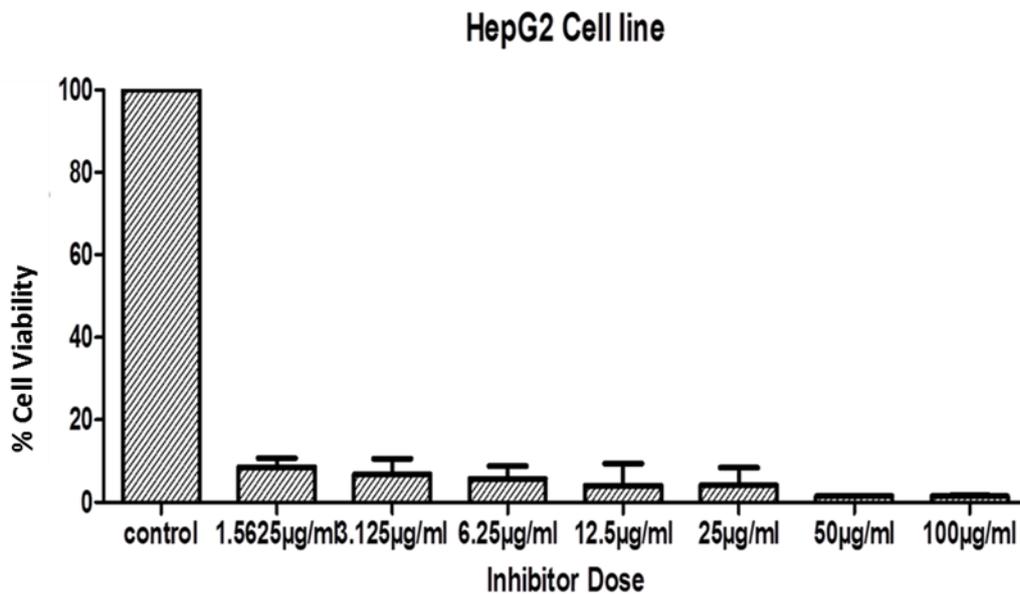
### **3.2.3 Cytotoxicity assay of intact whelk GAG extracts on hepatoblastoma-derived cell line Hep G2**

Hepatoblastoma is a primary human liver carcinoma and widespread childhood tumour among other type of cancer. It is considered a rare tumour and is less responsive to chemotherapy. This cell line has been used to study a wide variety of biological processes, including liver metabolism, hepatotoxicity and oncogenesis. Another type of liver cancer, known as secondary liver cancer, develops when the original cancer metastasises and spreads throughout the body, as is especially common in the case of lung, breast, pancreatic and stomach cancers. Metastasis begins when the cancer cells manage to enter the bloodstream and are allocated in other organs. Because the liver works as a filter for the blood, there is a high risk that cancer nodes will form in it, causing secondary liver carcinoma (López-Terrada *et al.*, 2009).

Treatment options for this type of cancer are very complicated. Some of these options are available for only primary human liver carcinoma. The options available to treat secondary liver carcinoma depend on the origin or source of the primary tumour, the stage and location of the cancer cells within the liver and the general health of the patient. Treatment may involve surgery, liver transplantation in rare cases, chemotherapy, radiotherapy and other biological therapies. Because all these options have serious consequences in terms of the extensive damage they do to the liver, there is an urgent need to discover an alternative therapeutic agent with less toxic properties.

The aim of this experiment is to examine the inhibition activity of crude whelk GAG extracts on this cell line and thereby gaining some idea of the hepatotoxicity of the crude whelk GAGs. The result obtained from MTT experiments are illustrated in Figure 3.3 and show that the percentage of viable cells drops off dramatically after the cancer cells are treated with serial dilution of crude whelk GAG extracts at a range of concentrations varying in dosage from 100 µg/ml to 1.5625 µg/ml over a five-day incubation period. The IC<sub>50</sub> value obtained

was  $0.85 \pm 0.01 \mu\text{g/ml}$ , which is very small compared to the activity of crude whelk GAG samples on other cell lines. This suggests a kind of specific toxicity of the crude extracts toward liver cells.



**Figure 3-3: Cytotoxic activity of whelk GAG extracts on HepG2 cell line.**

The data is represented as the percentage of viable cells for sample size ( $n=3$ ) as mean  $\pm$  SD. Cells were cultured in monolayers and maintained at  $37^\circ\text{C}$  in a humidified 5%  $\text{CO}_2$  atmosphere. Cytotoxicity assay was performed using MTT colorimetric assay. The  $\text{IC}_{50}$  values were calculated using nonlinear regression analysis (GraphPad Prism 5.0).

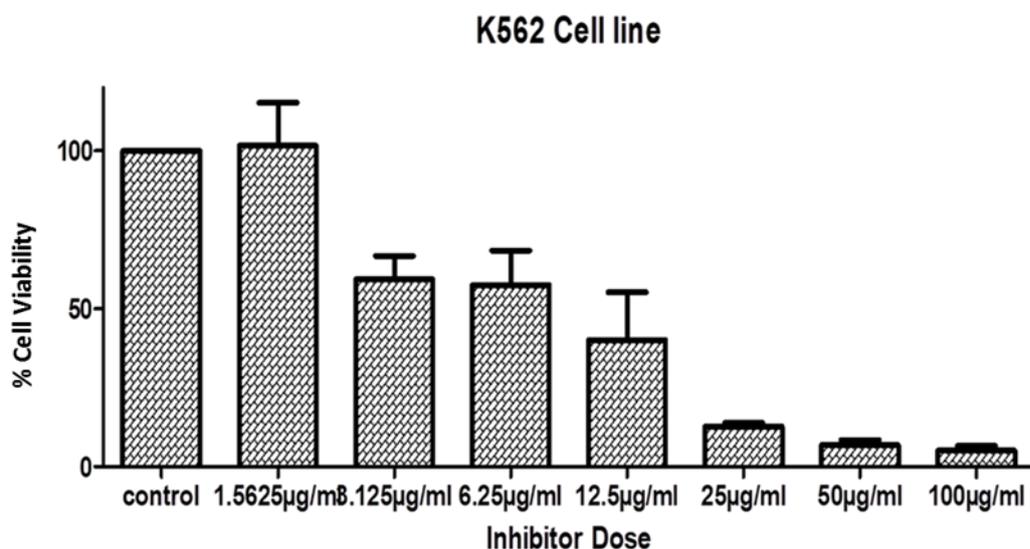
### **3.2.4 Cytotoxic assay of crude whelk GAG extracts on chronic myelogenous leukaemia**

The K562 cell line is a type of chronic myelogenous leukaemia (CML), which is diagnosed as a clonal myeloproliferative disorder in the stem cells that form the primitive hematopoietic resulting cancer of the marrow and blood. This type of cancer is heterogenic among patients, has two or three phases and is characterised by the presence of a chromosomal marker, the Philadelphia chromosome (Ph) in the leukaemia. The three phases of chronic myelogenous leukaemia (CML) comprise an initial phase or chronic phase, which is more likely to be controlled by therapy than the other phases, which is followed by an accelerated phase and

finally the blast crisis phase, resulting in the patient's death within 3 to 6 months. Some patients die from complications arising during the accelerated phase; others might develop blast phase CML without transferring from initial to accelerated phase (Kantarjian *et al.*, 1993).

Chronic myelogenous leukaemia is less severe than acute leukaemia as the CML does not overlap with the mature red cells, white cells and platelets. The most reliable treatment option is tyrosine-kinase inhibitors (TKI). Chemotherapy drugs, including protein synthesis inhibitor therapy, is another treatment option that is effective for chronic and accelerated phase CML patients in addition to many other therapies such as stem cell transplantation.

The aim of this experiment is to identify the inhibitory activity of crude whelk GAG extracts on this particular cancer type that is known for its multidrug resistance. Promising results with a half inhibition concentration value IC<sub>50</sub> was achieved (see Table 3.1), as was a decrease in the cell survival. The percentage of viable cells shown in Figure 3.4.



**Figure 3-4: Cytotoxic activity of crude whelk GAG extracts on K562 cell line.**

The data is represented as the percentage of viable cells for sample size (n=3) as mean  $\pm$  SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay has performed using MTT colorimetric assay. The IC<sub>50</sub> values were calculated using nonlinear regression analysis (GraphPad Prism 5.0).

**Table 3-1: IC50 Values represented in ( $\mu\text{g/ml}$ ) of different cancer cell lines over 5 day's incubation with crude whelk GAGs.**

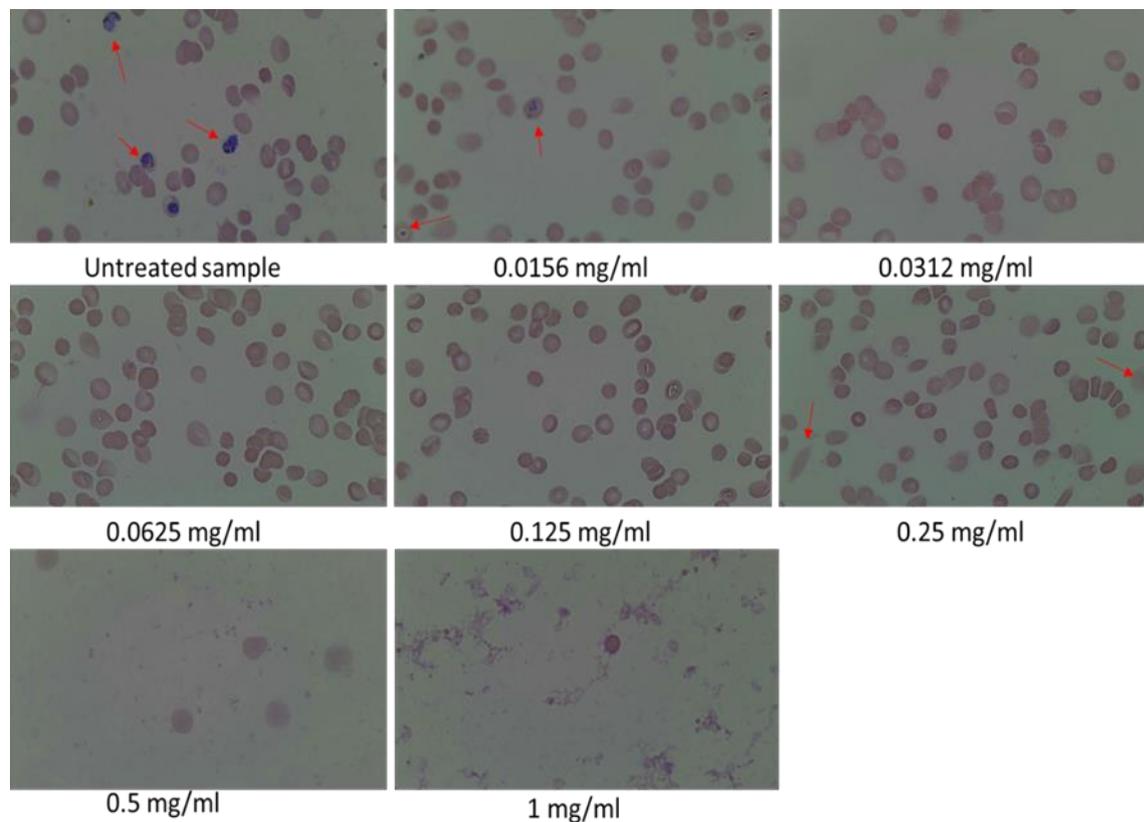
	<b>Cell Line</b>	<b>IC50 Values <math>\mu\text{g/ml}</math></b>
1	A549	2.44 $\pm$ 0.02
2	MCF-7	3.24 $\pm$ 0.03
3	SKBR-3	2.54 $\pm$ 0.02
4	HepG2	0.85 $\pm$ 0.01
5	K562	5.35 $\pm$ 0.03

### **3.2.5 Crude Whelk GAG extracts as an antimalarial therapeutic agent**

#### **3.2.5.1 Blood smear of parasite**

A thin blood smear was used to obtain monolayer cells the estimation parasitaemia according to the procedure described in Chapter 2. The same method was used to microscopically monitor the effect of different concentrations of crude whelk GAG extracts on the infected blood. This experiment was done using serial dilution of whelk GAG samples over a range of concentrations from 1 mg/ml to 0.015 mg/ml. The results shown in Figure 3.5 demonstrate the blood films of single slides at single concentration. Interestingly, the red blood cells begin to undergo lysis at 0.25 mg/ml; complete blood lysis is observed at 0.5 and 1 mg/ml, while the level of parasite appears to be reduced at a concentration of 0.0156 mg/ml. By increasing the concentration of GAG to 0.0312, 0.0615 and 0.125 mg/ml, the infection appears to be eliminated from the RBC. This preliminary data shows a potential toxicity effect of whelk GAG crude extraction on infected RBC at particular concentrations. All the experiments

performed subsequently were designed according to the apparent toxicity, and the whelk GAG concentrations used were below the doses that cause RBC lysis.



**Figure 3-5: Thin blood smear for untreated and treated infected blood.**

The microscopic monitoring of blood smear shows clean RBC at doses of 0.0312 mg/ml and so on. RBC lysis is obvious at high doses of crude whelk GAG extracts (0.5 mg/ml and 1 mg/ml).

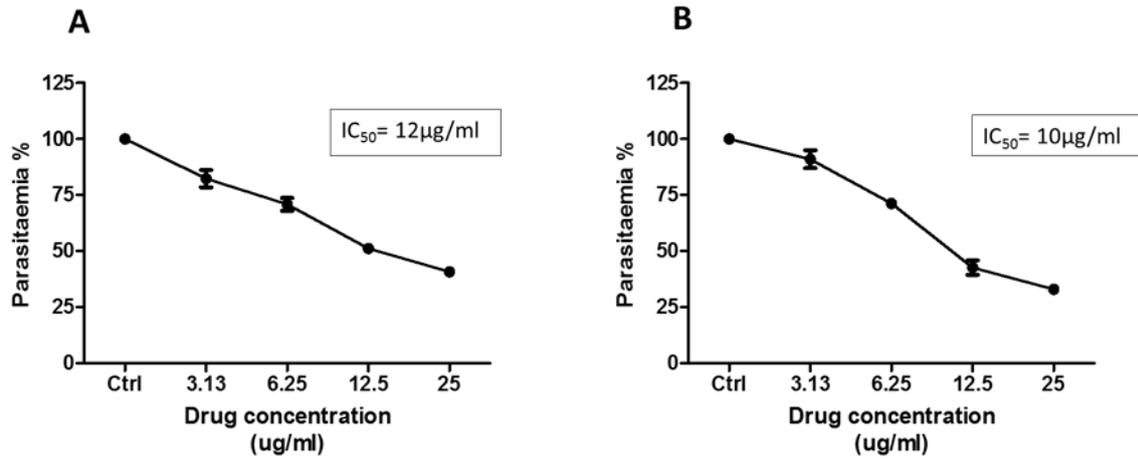
### **3.2.5.2 Cytotoxic activity of whelk GAG samples toward the malaria parasite using fluorescence-based drug susceptibility assay**

Over the years, a variety of drug screening techniques and assays have become available. However, because of the importance of the erythrocyte stages of the parasite life cycle as noted previously, most of these assays target screening drug activity on this particular stage. This includes the fluorescence-based drug susceptibility assays that utilise relatively simple, quick and high sensitive DNA-binding fluorescent dyes such as SYBR Green which depend

on the fact that the parasite's host cells (human RBC) are free from DNA. This means that the fluorescence observed in the sample can be attributed to parasite double stranded DNA only. Here, the uninfected blood, infected blood and the sample treated with different concentrations of whelk GAG extracts were stained in accordance with the methods described in Chapter 2. Using synchronised culture at ring stage revealed only one population of fluorescence intensity on the cytometer output, showing good labelling of the parasites at this stage. The BD-FACSVerse software was used to accurately determine parasitaemia by recording the percentage of infected cells relative to the total number of events recorded (red blood cells). The gating strategy was adapted from (Karl *et al.*, 2009).

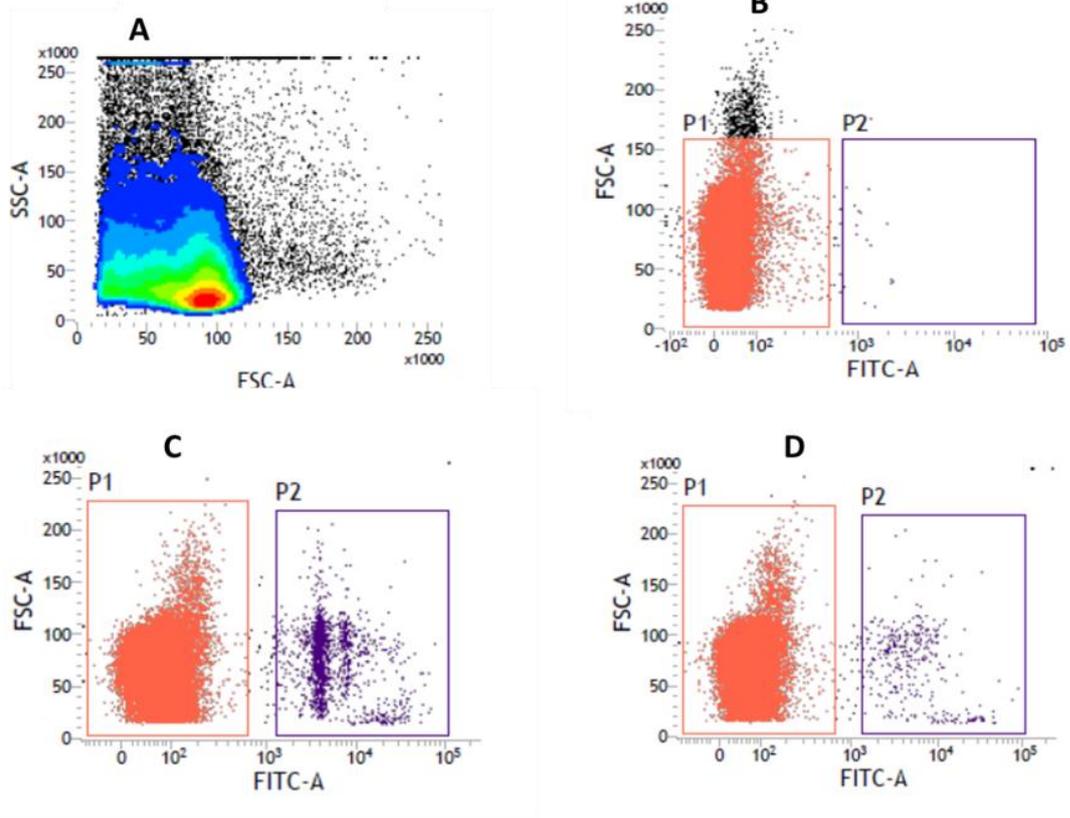
The IC<sub>50</sub> values were calculated using nonlinear regression analysis (GraphPad Prism 5.0). The percentage of parasitaemia plotted against drug concentration in  $\mu\text{g/ml}$  and the graphs of the two parasite stages (ring and schizonts) are presented in Figure 3-6. The serial dilution of whelk-GAG used in this experiment ranged in dosage from 25  $\mu\text{g/ml}$  to 3.13  $\mu\text{g/ml}$ . Output data was produced from technical triplicate tests on each dose. Interestingly, the results showed significant inhibition activity of whelk-GAG intact compound toward *Plasmodium falciparum* proliferation at both aforementioned stages. The IC<sub>50</sub> values generated from the ring and schizonts phases of the parasite life cycle after 48 and 72 hours of incubation with whelk-GAG intact extraction were 12  $\mu\text{g/ml}$  and 10  $\mu\text{g/ml}$  respectively.

Comparison of negative control (uninfected blood), positive control (infected blood without whelk GAG treatment) and infected blood that treated with whelk GAG crude extracts shown in Figure 3-7. Dot plot of forward light scatter (FSC) versus side scatter in BD flowcytometry has been used to determine the overall population of RBC. Scatterplots based on forward scatter (FSC-A) and SYBR Green dye fluorescence (FITC-A) at 494/521 nm was used to monitor negative control, positive control and the effect of whelk GAG on parasite viability in the infected RBC.



**Figure 3-6: IC<sub>50</sub> values of crude whelk GAG extracts on two mononuclear parasite stages.**

**Panel A:** The percentage of viable parasitaemia (trophozoites stage) against drug concentration after 48 hours of incubation. **Panel B:** The percentage of viable parasitaemia (ring stage) against drug concentration after 72 hours of incubation. The IC<sub>50</sub> values were calculated using nonlinear regression analysis (GraphPad Prism 5.0).



**Figure 3-7: Comparison of uninfected blood (negative control), infected blood (positive control) and sample treated with crude whelk GAG extracts.**

**Panel A:** Dot plot of forward light scatter (FSC) versus side scatter. **Panel B:** Scatterplots based on forward scatter (FSC-A) and SYBR Green dye fluorescence (FITC-A) (494/521) representing uninfected blood (negative control). **Panel C:** Example of infected blood sample in P2. **Panel D:** Reduced number of mononuclear parasite level due to whelk GAG treatment. The gating strategy was employed to distinguish between the RBC population (P1) and the mononuclear rings of the Plasmodium life cycle (P2).

### **3.3 Discussion**

The aim of this chapter has been to evaluate the biological activity of whelk GAG crude extraction on a variety of cancer cells and on malaria. The main finding of this chapter is that whelk GAG crude extracts expressed significant inhibition activity toward all cancer cells and toward malaria.

Cancer is defined as uncontrolled growth of mutant cells among normal tissue. Multiple oncogenic alterations may occur in the cells, including mutations in some genes. These mutations are responsible for clear modifications in cell function, which offer potential targets in cancer therapy.

Over the last few decades, treating and understanding cancer has become the main aim of many studies, especially in the developed countries where cancer is one of the major causes of mortality. Different cancer types have different immune profiles and their response to the treatment options available vary depending on the tissue type and other histological properties, cancer stage, degree of metastasis and multiple oncogenic alterations. Regardless of the cancer type, the treatment options available for cancer patients generally begin with surgery and is often followed by chemotherapy, radiotherapy and in some cases hormone therapy. Conventional cancer treatments do not always have a desirable result. In fact, they can make cancer more malignant. For instance, chemotherapy and radiotherapy have been used as powerful tools to destroy the DNA of the cancer cells in the hope of causing more harm to them than to the normal cells. However, this is not the case most the time. Very significant side effects have been observed from these treatment options; moreover, even after treatment, there is a very high risk that cancer will recur and even develop drug resistance. In some cases, treating particular type of cancer with chemo agents ends with more deadly type of cancer like liver cancer.

As with chemotherapy, the risk of inducing secondary cancers with self-renewing capacity is high, especially for some cancers such as breast cancer. With these concerns in mind, specific biological approaches to treating cancer have been developed, including target therapies that employ purified molecules such as monoclonal antibodies (MABs), cancer vaccines, cell growth factors, cancer growth blockers, drugs that block cancer blood vessel growth for i.e. angiogenesis inhibitors such as interferon and interleukin 2 (immunotherapy) and gene therapy (Oldham, 1984).

This study demonstrates that whelk GAG crude extracts inhibits the growth of HepG2 cells and A549 cells. The results are broadly consistent with those of a study that investigated the biological activity of brown seaweed extraction, known as *Sargassum pallidum*, which reported significant cytotoxic activity on HepG2 cells and A549 cells. The cytotoxicity effect was explained by the correlation between the activity and their molecular weight and sulphation pattern (Ye *et al.*, 2008). Another study has reported antitumor activity from brown algae (*U. pinnatifida*), which inhibited proliferation of prostate cancer PC-3, cervical cancer HeLa, alveolar carcinoma A549, and hepatocellular carcinoma HepG2 cells in a similar pattern to the commercial fucoidan from *F. vesiculosus* as reported by Synytsya *et al.*, (2010). However, the molecular structure of these active compounds seems to be different from that of crude whelk GAG extracts, which contains marine non- GAG material. The chemical structures of marine sulphated GAGs are very unique and distinct from GAG structures. Although they bear sulphation and are sometimes composed of disaccharide repeating units like GAGs, they are much more homogeneous and regular in terms of backbone composition and sulphation patterns than mammalian derived GAGs (Pomin, 2012a, 2012b, 2015; Pomin & Mourão, 2008). Over-sulphated fucoidans is another type of sulphated glycan from marine organisms that inhibit tumour-induced angiogenesis by

preventing the binding of VEGF<sub>165</sub> to its cell surface receptor and inhibiting the VEGF-mediated signalling transduction (Narazaki *et al.*, 2008).

The results of this study describe for the first time a strong cytotoxic activity of whelk GAG crude extracts against two subtypes of breast cancer cell, MCF-7 and SKBR-3. The most likely explanation for this finding is that whelk GAG extract acts as an anti-proliferation agent. This process is an important part of cancer development, as many proteins and signal transduction pathways resulting from mutations that favour proliferation is involved in stimulating cell growth. Administration of exogenous GAGs from whelk with active structural features may lead to a competitive process, with the functional endothelial GAGs binding to the growth factors that are commonly seen during the cell differentiation event, thereby giving rise to an anti-proliferation outcome.

Dose-response bioassays usually involve one assessment, which takes place at the end of a defined exposure period. More information is needed on the effect of whelk GAGs on each cell line generated by adapting cell cycle analysis and apoptosis assay, which could carry on using flow cytometry. Although it was not possible to cover this aspect in the present study, an earlier study by our group using a similar compound isolated from whelk toward breast cancer (MDA-MB468-NQ01 and MDA-MB-468), leukaemia (MOLT-4 and K562) and ovarian (HeLa) cell lines demonstrated significant perturbations in the cancer cell cycle, showing cell cycle arrests at different stages. Similarly, significant apoptosis inductions were induced by whelk GAGs on the cancer cells (Ogundipe, 2015).

The anticancer activity of sulphated polysaccharides isolated from marine animals such as sea cucumbers is well established. The secondary metabolites characterised from this species suppress the proliferation of various human tumour cell lines *in vitro*. The molecular mechanisms of these glycosides include cell apoptosis through the activation of intracellular caspase cell death pathways, arrest of the cell cycle at S or G2/M phases and dysregulation of

certain cellular receptors and enzymes participating in cancerogenesis, such as epidermal growth factor receptor (EGFR), extracellular signal-regulated kinases (ERK) and others. Administration of some glycosides leads to many cellular changes including reduction of cancer cell adhesion, suppression of cell migration and tube formation in those cells, suppression of angiogenesis, inhibition of cell proliferation and tumour invasion. As a result, growth inhibition of tumours occurs *in vitro* and *in vivo* (Aminin *et al.*, 2015).

Targeting molecules such as receptors, signalling molecules etc., based on alterations in their expression and/or function that promote proliferation in many different human cancers may block tumour maintenance. Targeting cell cycle proteins in sustained proliferative signalling especially at the G1 phase of the cell cycle when the progression depends on the balance of proliferative and anti-proliferative signals is important. It is believed that cancer cells ignore the signals to stop proceeding into the S phase and arrests in G1, here resulting in excessive DNA replication and induced mutations (Malumbres & Barbacid, 2009).

Moreover, this is the first study, to our knowledge, to examine the activity of whelk GAG crude extracts against plasmodia growth. From the output data presented in this chapter, intact whelk GAGs can be considered as a highly antiplasmodial agent with respective IC<sub>50</sub> values of 10 and 12 µg/ml at the ring (trophozoites) and schizont stages. This finding is in line with those reported from marine sea cucumbers *Ludwigothurea grisea* and *Isostichopus badionotus*, from the red alga *Botryocladia occidentalis* and from the marine sponge *Desmapsamma anchorata*, which exhibited significant antimalarial activity *in vitro* model that inhibit *Plasmodium falciparum* growth at low-anticoagulant concentrations. The mechanism of action of the active compounds from these marine organisms has been indicated by the inhibition of erythrocyte invasion by *Plasmodium* through coating of the parasite. Heparin-related polysaccharides of low anticoagulating activity could play a crucial role as potentiates of immune responses, such as invasion of human RBC mediated by sulphated

polysaccharides is followed by prolonged exposure of Plasmodium to the immune system (Marques *et al.*, 2016).

Interestingly, the important role of high negatively charged sulphated glycol-conjugates such as dextran sulphate 500K, dextran sulphate 5K, sulphatides, fucoidan and heparin in inhibiting the invasion of erythrocytes by merozoites and the cytoadherence of parasitized erythrocytes PRBC to endothelial cells, especially by highly negatively charged proteoglycan, has been reported. The mode of action of these compounds in inhibiting the invasion of erythrocytes by Plasmodium falciparum merozoites is unclear, but it has been confirmed that it was not mediated by killing the parasites.

Furthermore, the correlation between the inhibition activity and the highly negatively charged glycol-conjugate is critical. Both invasion and cytoadherence are inhibited by highly sulphated dextran sulphate 500K, while un-sulphated dextran 500K and hyaluronic acid have no significant inhibition effects. In fact, the positively charged protamine sulphate could promote cytoadherence.

Sulphated proteoglycans are involved in the binding of sporozoites of malaria parasites to hepatocytes (Xiao *et al.*, 1996). Adherence is important for parasite development as parasites use specific protein for adherence to specific receptors on host cells surfaces. This increases the potential of utilising such properties to identify malaria vaccine candidates.

The possibility of targeting plasmodial invasion proteins for antimalarial drug development has been demonstrated by numerous studies. Among the major surface proteins of the sporozoites of the various Plasmodium species such as *P. yoelii* is circumsporozoite (CS), which selectively binds to heparin, fucoidan and dextran sulphate Sepharose, but binds with lower affinity to chondroitin sulphate A and C-Sepharose. This illustrates the capability of GAGs to block merozoite and sporozoite invasion (Pancake *et al.*, 1993). An indispensable antigen MSP-1, which plays an important role in invasion and is localized in abundance on

the merozoite surface, is another example of targeting invasion proteins for the development of novel antimalarial drugs. This protein was targeted by heparin, which binds to the specific domain of MSP-1 (termed MSP-133) and prevents secondary proteolytic processing (Boyle *et al.*, 2010).

The above result does not enable us to determine the exact mechanism of action of whelk GAGs toward malaria. However, the findings suggest a strong anti-proliferation effect from whelk GAG, which would also be beneficial in the treatment of certain other diseases such as malaria.

These findings can contribute to the development of anticancer and antimalarial agents. Further studies are required to investigate their mode of action, given the urgent need to develop new therapeutic agents following a significant increase in the development of resistance toward several of the available chemotherapeutic approaches (Mita *et al.*, 2009; O'Brien *et al.*, 2011).

# **Chapter 4 : Isolation and structural analysis of whelk GAGs**

## 4.1 Introduction

The structure elucidation is the most important process for understanding the biological function of the GAGs isolated from whelk. Whelk GAGs have potentially powerful therapeutic effects as proven by their activity as anticancer agents (chapter 3 and 5).

The extraction methods of crude GAGs from whelk, as mentioned in previous chapter, included de-fatting the soft body with acetone wash, followed by pronase treatment to break down the protein part and precipitate by addition of TCA, ammonium sulphate or ammonium bicarbonate. These methods are often used in carbohydrate extraction because conjugates have been found with protein or lipids in the majority of biological samples. GAGs were then further purified using step elution anion exchange chromatography and their biological activities assessed. The ion exchange chromatography helps to remove non GAG contamination and enables the desirable sugar to be eluted at void volume of the column. GAG chains are then desalted using either PD-10 column or extensive dialysis against deionized water.

The complex sulphate-rich polysaccharides produced in this procedure contain repetitive units of certain monosaccharides. The monosaccharide composition analysis performed in this study utilised HPAEC-PAD (High pH anion exchange chromatography coupled with pulsed amperometric detector), which characterised the types and amounts of monosaccharide in the crude whelk GAGs and the purified fractions. Monosaccharide identification was achieved in comparison with authentic standards (Willför *et al.*, 2009). In addition, GC-MS was used as an additional technique to further verify the results obtained from monosaccharide composition HPAEC-PAD analysis.

A variety of additional analytical techniques and methods can also be used to help further understand the structure- function relationship of the GAGs (Pomin & Mourão, 2008), such as chemical and enzymatical degradations of complex polysaccharides. Glycan chains maybe

further processed by enzymatic degradation to remove impurities such as proteins or lipids. This is important in two aspects, firstly, to improve their detection and secondly, it facilitates their separation.

Gel filtration chromatography is a useful technique for carbohydrate purification, especially after enzymatical degradation of conjugated glycan chains. Stable isotopes labelling through reductive amination of the derivative sugar has been used in GAGs analysis, followed by reverse phase HPLC coupled with MS (Lawrence, Olson, *et al.*, 2008).

The aim of this chapter is to analyse the structure of crude biologically active GAGs isolated from whelk. To address this a variety of analytical techniques, such as HPAEC-PAD, gel filtration chromatography, enzymatic digestion and GRIL-LC-MS methods have been used. Further investigation was carried out by 1D/2D-NMR experiments, and finally all ion exchange GAG fractions were transferred to bioactivity screening assay, to investigate their biological activity toward cancer cells, which is demonstrated in chapter 5.

## **4.2 Results**

### **4.2.1 Monosaccharides analysis of intact GAG by HPAEC-PAD**

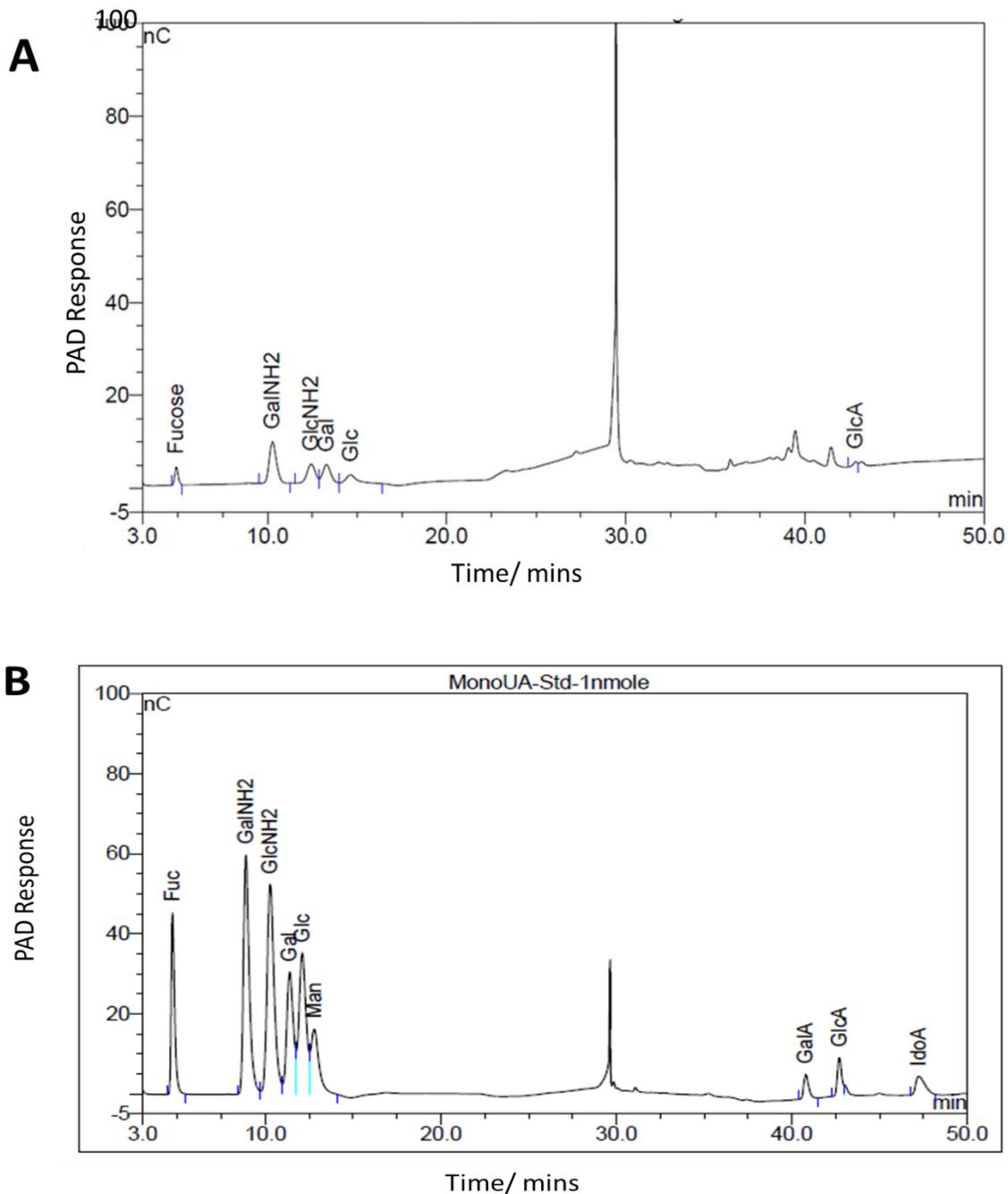
To understand the fine structure and backbone chain composition of unknown polysaccharides from whelk-GAG sample, it is required to breakdown the complex polysaccharide into constituent monosaccharide units. A simple and sensitive method has been used to quantify and identify the monosaccharide composition in whelk GAG using high-performance anion exchange chromatography (HPAEC) with electrochemical detection by pulsed amperometric detection (PAD). This method has previously been applied to sulphated glycans from different human tissues at low-picomole level (Adamo *et al.*, 2009).

This method is a powerful and sensitive procedure to identify monosaccharide composition and total carbohydrate present in the sample. Based on the weakly acidic properties of carbohydrates that allows selective separation of carbohydrate using strong ion-exchange at

high pH. This provide simple and direct quantification of non-derivatized monosaccharides at low-picomole concentrations. The HPAEC is coupled with pulsed amperometry detection that uniquely detects the oxidizable function groups at specific voltages used within the detection.

The column used for monosaccharide analysis of whelk-GAG sample was CarboPac PA1, which is designated specially to analyse mono-, di- or oligosaccharides. These columns are classified under type one (nonporous pellicular resin) and have excellent mechanical stability that can tolerate back pressures of more than 4000 psi, and wide stable range of pH (0-14).

The theory behind pulsed amperometry detection is that the pulse sequences are typically recorded in less than one second from the measured potentials on the working gold electrode surface. The first potential measure on the working electrode surface is recorded as a pulse electrical current generated from oxidation of hydroxyl groups in the carbohydrates, followed by reduction that effectively cleans it of unwanted products from the carbohydrate oxidation. The final potential step completely reduces the gold oxide on the electrodes surface back to gold, which renews the electrode surface after removal of the bound analyte. Monosaccharides from neutral and amino sugar are routinely separated by HPAEC-PAD and this is a standard protocol for glycan chain analysis (Weitzhandler *et al.*, 1993).



**Figure 4-1: Monosaccharide composition analysis.**

**Panel (A)**- Monosaccharide profile of GAG isolated from whelk species. The hydrolyzed sample was dissolved in water and injected on Dionex-ICS3000 HPAEC-PAD, the column used for monosaccharide profiling was Carbopac PA-1 (4mm x 250mm). The solvents used for HPAEC was 100mM NaOH and 100mM NaOH + 250mM NaOAc. The amount injected for monosaccharide analysis of intact GAG was 5 $\mu$ g/ml. **Panel (B)**- Standard monosaccharides were dissolved in water and injected on Dionex-ICS3000 HPAEC-PAD, the column used for monosaccharide profiling was Carbopac PA-1 (4mm x 250mm). The solvents used for HPAEC was 100mM NaOH and 100mM NaOH + 250mM NaOAc. The amount injected for each monosaccharide was 1 nmole of MonoUA-Std.

The monosaccharide profile in whelk GAG, as shown in figure (4-1 panel A), was compared to the MonoUA-Standards figure (4-1 panel B), the results show the crude GAG mixtures contain a mixture of glycoproteins in relatively high amounts, most of these sugars are generally present in different glycoconjugates found in eukaryotic system such as various glycoproteins and proteoglycans. GalNH<sub>2</sub> can be a component of *O*-glycans, chondroitin sulphate and dermatan sulphate; likewise, GlcNH<sub>2</sub> and Glu are components of *N*-linked glycan, heparin/heparan sulphate, hyaluronan and keratan sulphate (Wiederschain, 2009). The presence of GlcA along with the amino-sugars GlcNH<sub>2</sub> and GalNH<sub>2</sub> indicates the presence of a significant quantity of GAG in whelk extracts. A noticeable absence of IdoA might indicate the absence of any dermatan sulphate in this sample. Or an unusual heparan sulphate structure. The amount of each monosaccharide units in the whelk sample are represented in µg/ml in the table below.

**Table 4-1: Monosaccharide quantification in µg/ml of whelk crude sample.**

Sample	GalNH <sub>2</sub>	GlcNH <sub>2</sub>	Gal	Glc	Man	GlcA
GAG-Sample	0.048	0.023	0.046	0.017	0.0	0.018

## **4.2.2 Enzymatic digestion and Gel filtration chromatography of Intact GAG**

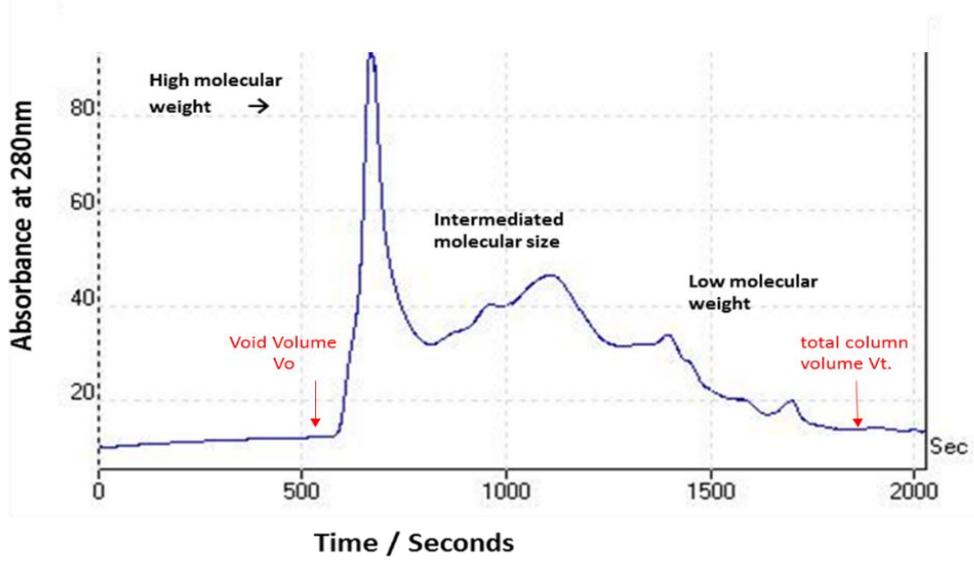
### **4.2.2.1 Enzymatic digestion and gel filtration chromatography of digested GAG with Heparinase Lyase**

The enzymatic de-polymerisation is a powerful tool for fragmentation of long chain of complex polysaccharide into smaller oligosaccharides as part of a structure study. The enzymes act on specific substrates sequences, within the GAG chains. The complete digestion provides a range of simple known sequences of repetitive disaccharides or oligosaccharides. The whole GAG and standard commercial porcine intestinal mucosal heparan sulphate was subjected to enzymatic digestion.

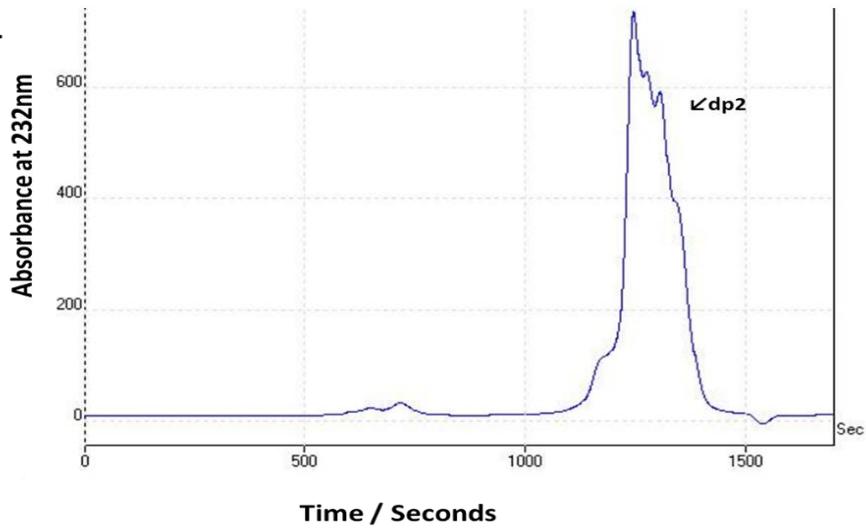
Size exclusion chromatography has been used to purify carbohydrates following enzymatic digestion. The chromatogram of digested GAG was compared to undigested GAG sample (figure 4-2). However, the results obtained demonstrated a low-resolution profile. In undigested GAG, a broad peak was observed, which indicates a complex mixture of glycan, that may still be conjugated to some contamination peptides.

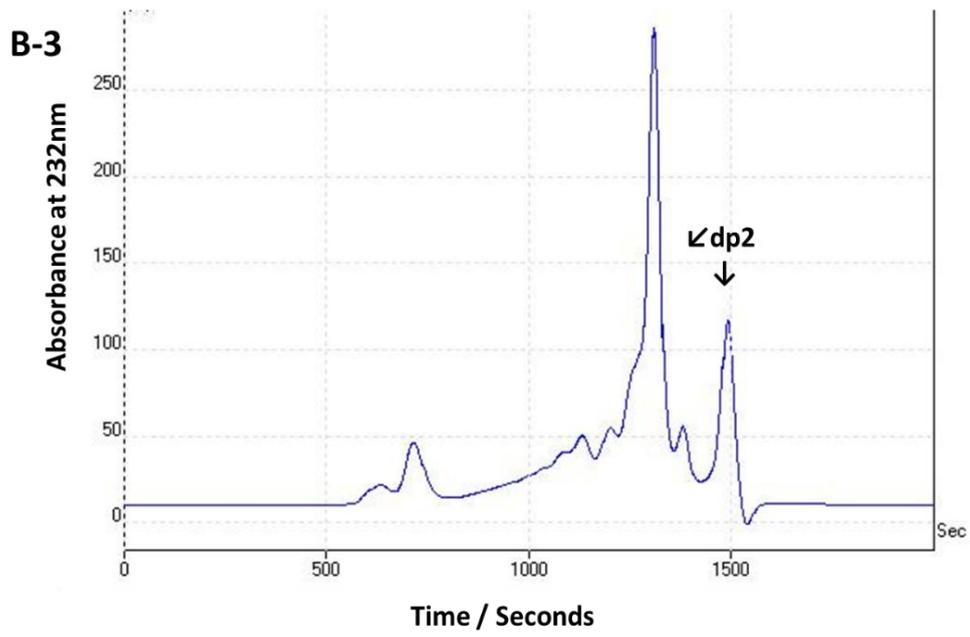
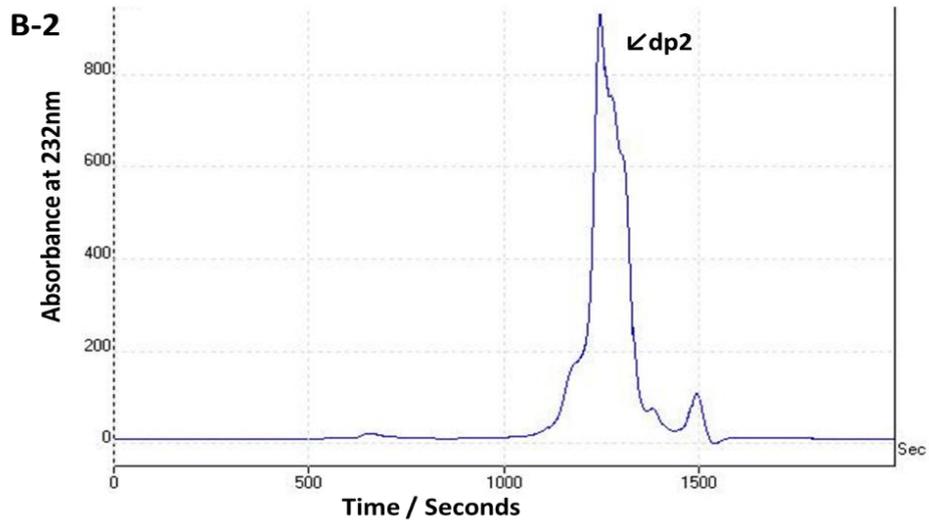
Heparinase enzyme has been used to release the glycan from the conjugated glycan chains and the disaccharide released could eluted in column total volume. The elution profiles observed with all hep lyase fractions showed more asymmetrical peaks, that eluted at total column volume as expected. This compared favourably to the elution profile of disaccharide units from commercial HS. Complete de-polymerisation was expected after using mixture of three heparin lyase enzymes. However, the possibility of a small proportion incompletely digested chains cannot be ruled out. Longer incubations with increasing amount of enzymes did not affect the elution profile.

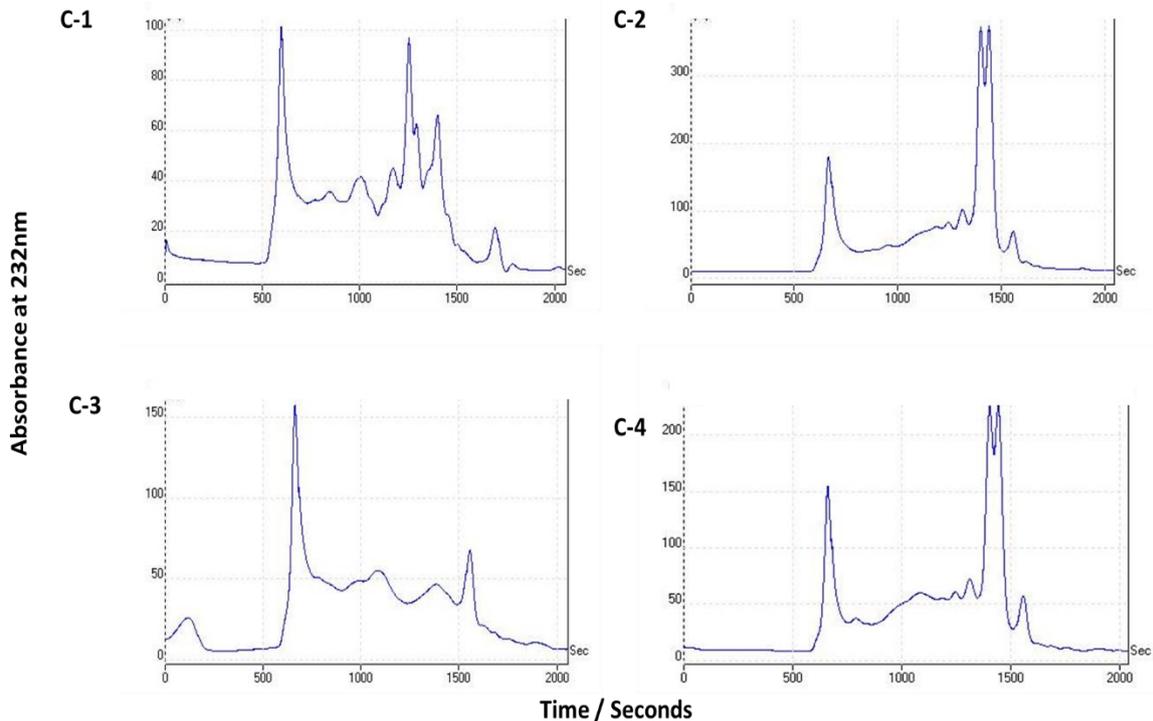
**A**



**B-1**





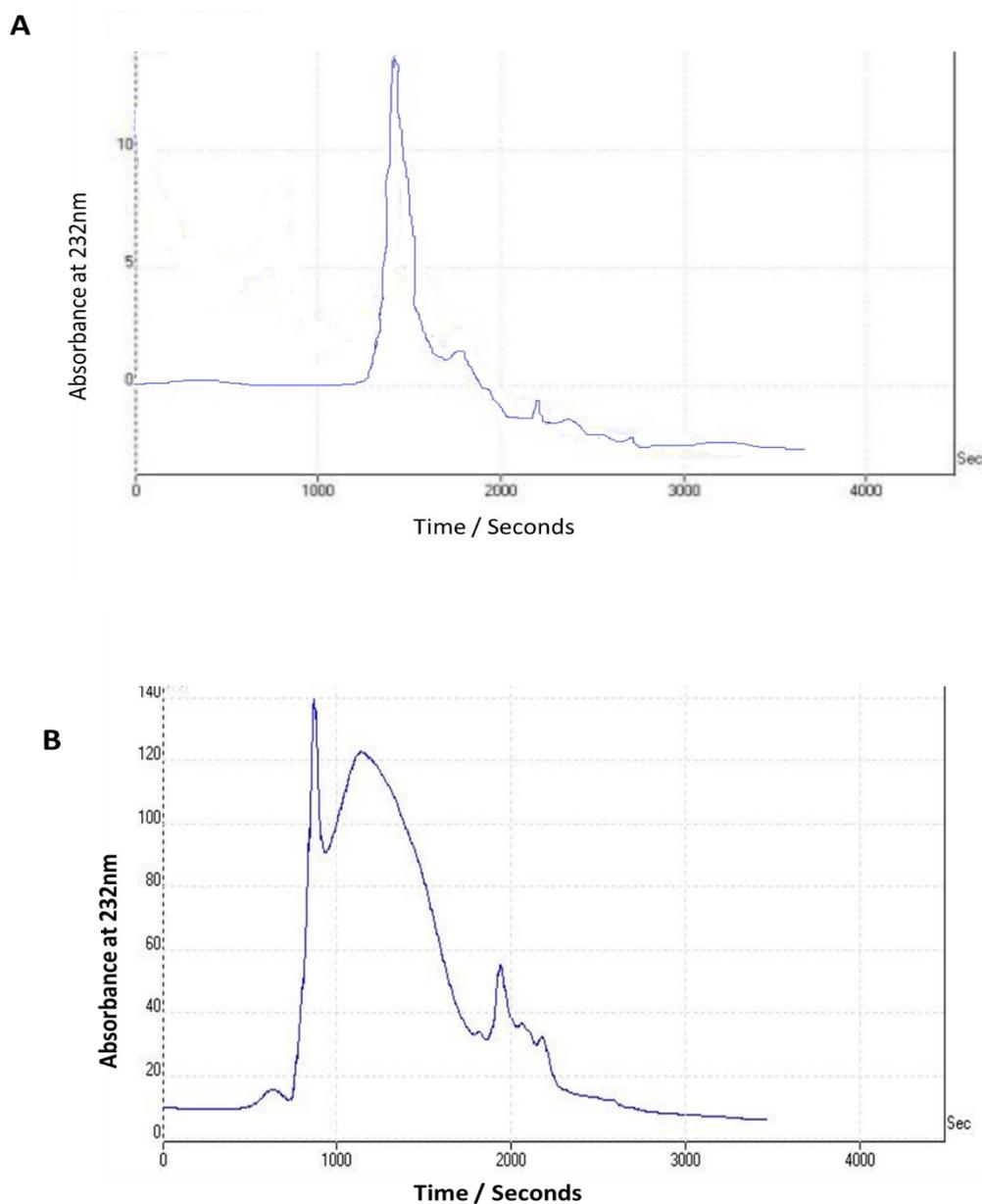


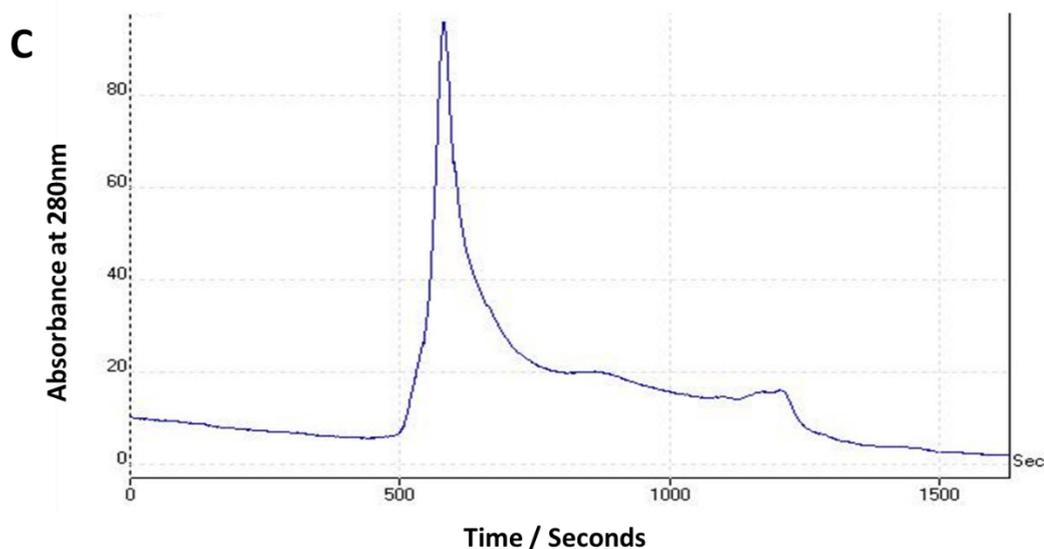
**Figure 4-2: Gel filtration chromatography of intact GAG, digested GAG with HEP lyase and commercial HS.**

**Panel A-** size exclusion profile of intact GAG from whelk. 100 $\mu$ l aliquot containing 100 $\mu$ g of crude mixture of whelk GAG is prepared and 50 $\mu$ l injected in HPLC system, sample separated using superdex 75 10/300 GL gel filtration column. Column pre-equilibrated with PBS-pH7.0, UV detection is used at 280nm and result recorded per second for maximum 2000 seconds. **Panel B-** size exclusion profile of standard commercial Porcine intestinal mucosal heparan sulphate digested with HEP lyase fractions for 24hours at 37 °C, 50 $\mu$ l applied to superdex 75 10/300 GL gel filtration column, UV detection is used at 232nm and result recorded per second for maximum 2000 seconds. B-1, is commercial HS digested with HepI. B-2, is commercial HS digested with HepII. B-3, is commercial HS digested with HepIII. **Panel C-** size exclusion profile of whelk-GAG sample. 100 $\mu$ l aliquot containing 100 $\mu$ g of crude mixture of whelk GAG is digested by Hep lyase fractions for 24hours at 37 °C, 50 $\mu$ l applied to superdex 75 10/300 GL gel filtration column, UV detection is used at 232nm and results recorded per second for maximum 2000 seconds. C-1, is crude sample digested with Hep I. C-2, is crude sample digested with Hep II. C-3, is crude sample digested with Hep III. C-4, is crude sample digested with Hep I, II and III.

#### 4.2.2.2 Enzymatic digestion and gel filtration chromatography of digested GAG with ABC Lyase

To investigate the presence of chondroitin sulphate in whelk-GAG, the crude GAG was digested with ABC lyase to release CS disaccharides. The results in figure (4-3) show the elution profile of whelk GAG along with the elution profile from standard porcine intestinal mucosal chondroitin sulphate digested with ABC-chondroitin lyases.





**Figure 4-3: Gel filtration chromatography of digested GAG with ABC lyase and commercial CS.**

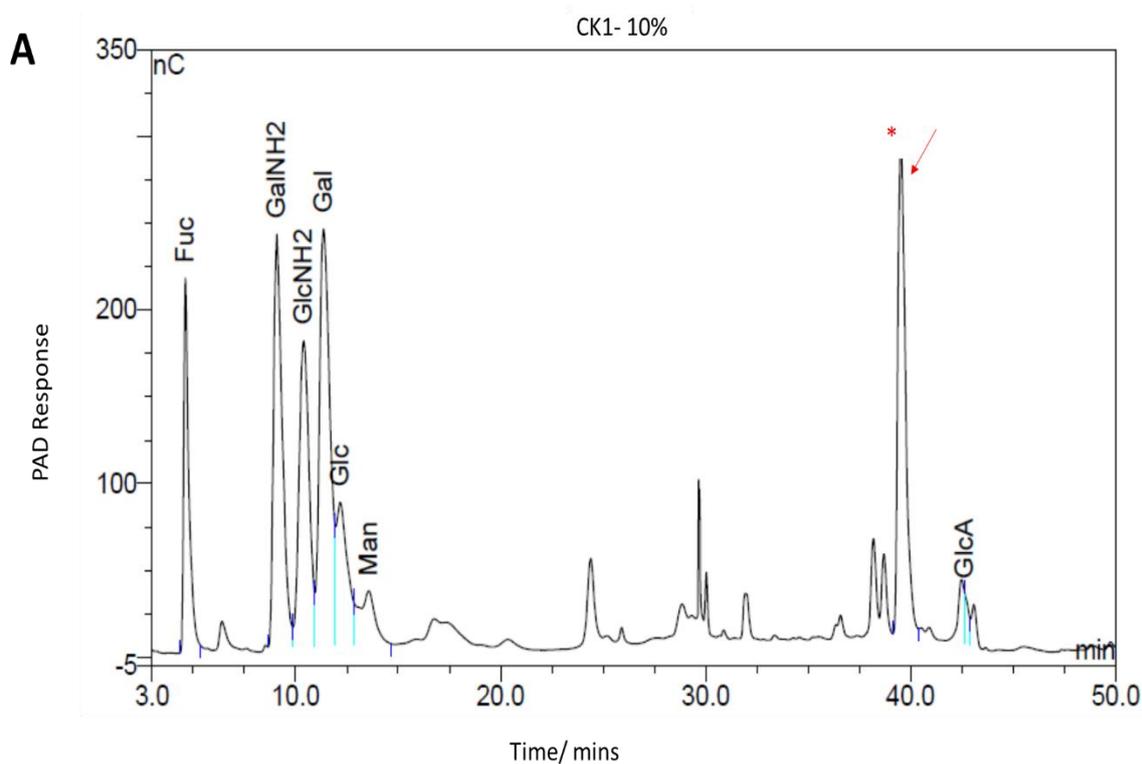
**Panel A-** size exclusion profile of standard commercial Porcine intestinal mucosal chondroitin sulphate digested with ABC lyase fractions for 24hours at 37 °C, 50µl applied to superdex 75 10/300 GL gel filtration column, UV detection is used at 232nm and results recorded per second for maximum 4000 seconds. **Panel B-** size exclusion profile of digested whelk-GAG sample. 100µl aliquot containing 100µg of crude mixture of whelk GAG was digested by ABC lyase fractions for 24hours at 37 °C, 50µl applied to superdex 75 10/300 GL gel filtration column, UV detection is used at 232nm and result recorded per second for maximum 4000 seconds. Panel C: Elution profile of whelk-GAG intact sample.

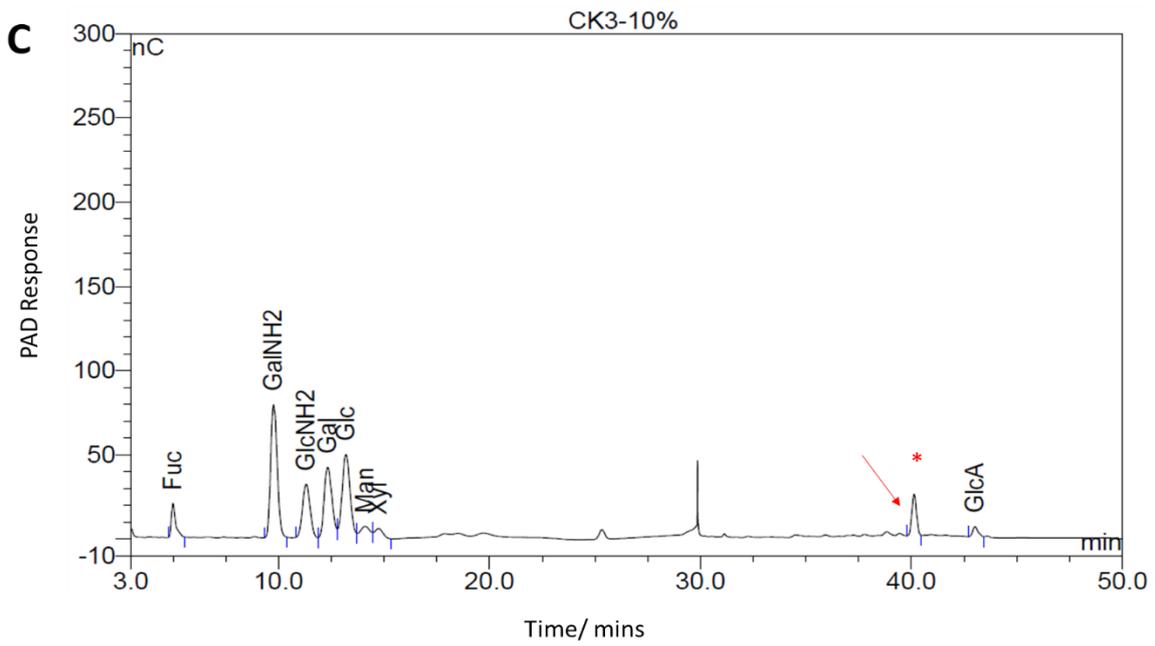
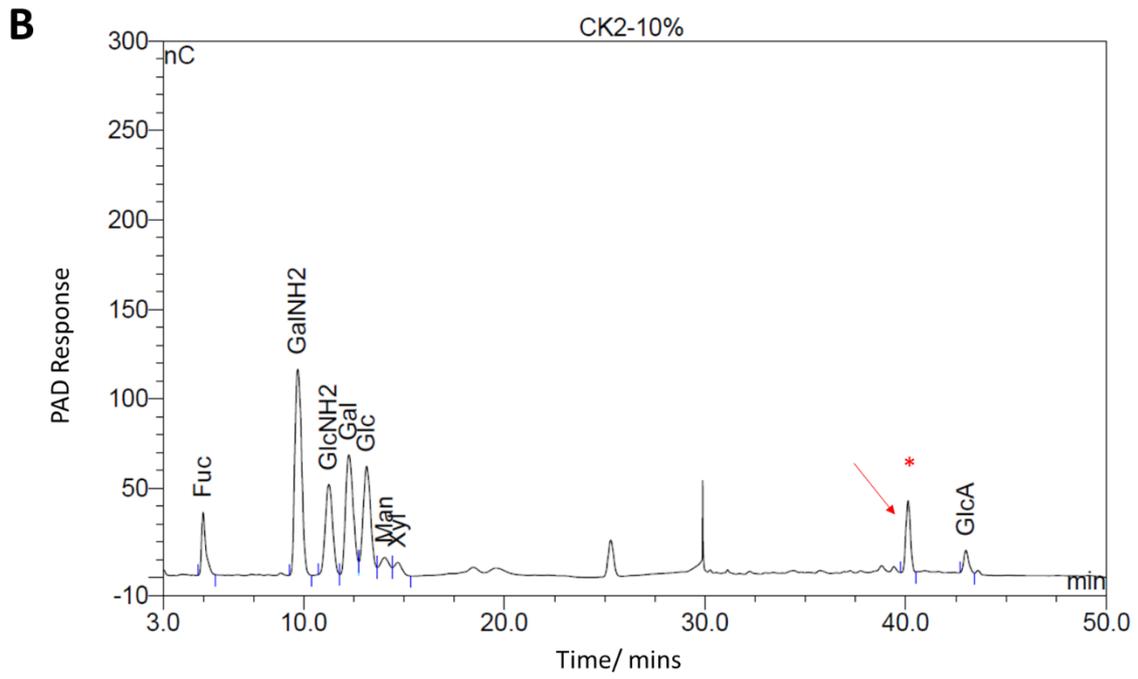
#### 4.2.3 Ion exchange chromatography and PD-10 de-salting techniques

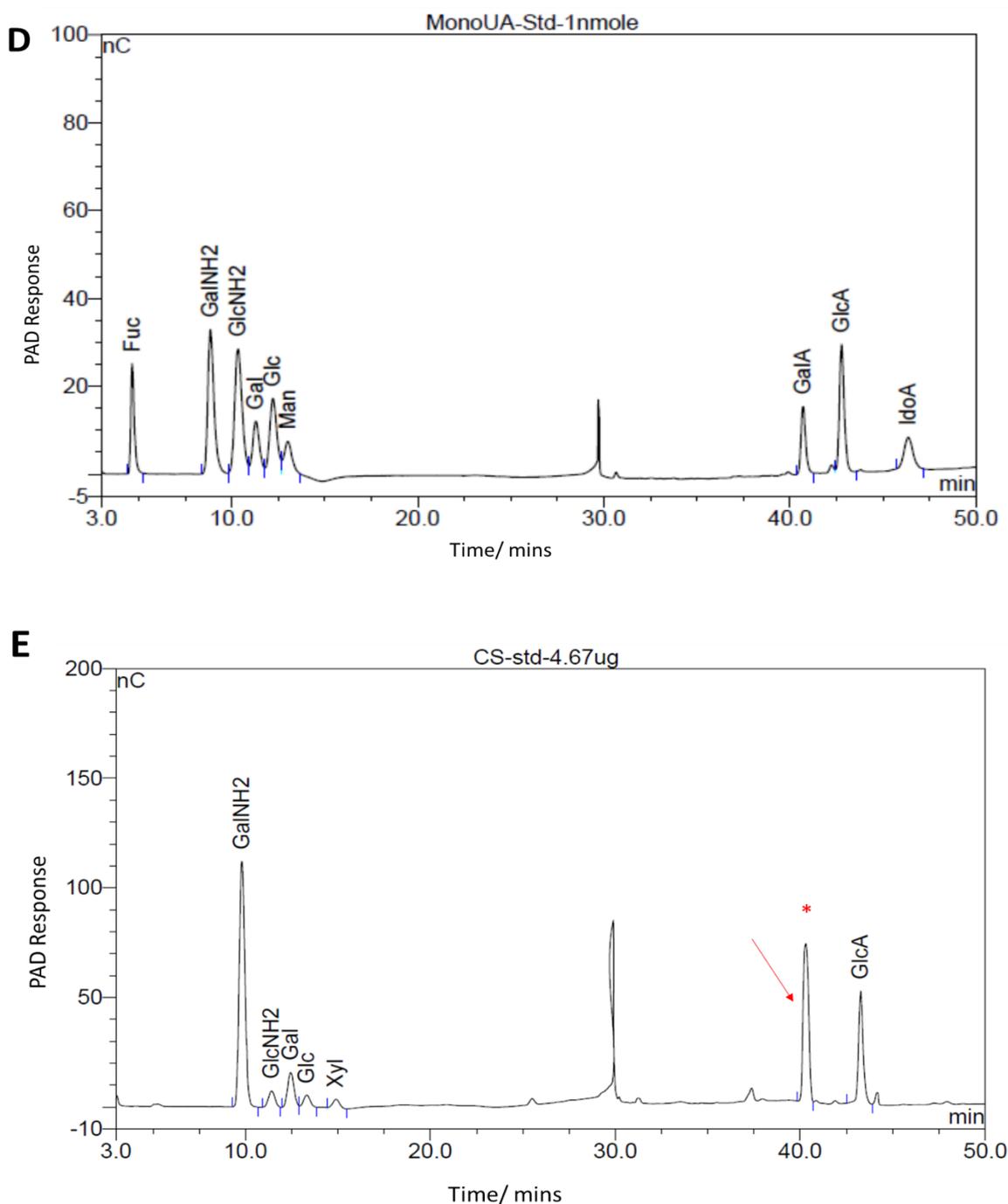
Whelk GAG was separated according to the GAG charged group using anion exchange chromatography prior to structural analysis and assessment of biological activities. The DEAE-Sepharose was used to fractionate the GAG mixture according to net negative charge. The whelk-GAG sample was fractionated into three fractions by step-wise addition of buffer with increasing chloride ion concentrations from 1.0 to 3.0 M. Each fractions were then de-salted using a PD-10 column and GAG containing fractions were pooled, lyophilized and stored for further structural investigation.

#### 4.2.4 Quantitative analysis of monosaccharides on different –GAG fractions obtained from DEAE-fragmentation by HPAEC-PAD

The three fractions of dry de-salted GAG eluted from ionic exchange chromatography were labelled as follows: CK1 (sample eluted at 1M NaCl), CK2 (sample eluted at 2M NaCl), and CK3 (sample eluted at 3M NaCl). These fractions were then subjected to, a monosaccharide analysis and quantification. Results in figure 4-4 panel A, B, C illustrate the types of monosaccharide found in each sample with the quantities being listed in the table (4-2), while panel D and E represent spectra of MonoUA-Std and CS- Std.







**Figure 4-4: The monosaccharide analysis form HPAEC-PAD system for ionic exchange chromatography fractions.**

**Panel A-** is spectra of monosaccharide analysis of 7% injected material of fraction CK1- eluted from ionic exchange chromatography in gradient of 1M NaCl. **Panel B-** is spectra of monosaccharide analysis of 7% injected material of fraction CK2- eluted from ionic exchange chromatography in gradient of 2M NaCl. **Panel C-** is spectra of monosaccharide analysis of 7% injected material of fraction CK3- eluted from ionic exchange chromatography in gradient of 3M NaCl. **Panel D-** is spectra of MonoUA-Std at 1nmole injected value. **Panel E-** is monosaccharide analysis of CS- Std from HPAEC-PAD.

Briefly, monosaccharides were produced by subjecting 10% of dry GAG material from each fraction to acid hydrolysis for 4 hours with TFA, and then each fraction was injected in HPAEC-PAD system. Typically, common monosaccharides were found in all three fractions. The results demonstrated the types of monosaccharide that exists in fractions CK1, CK2, CK3 and CS-Std and their relative amounts. A peak indicated with (\*) is unknown compound and is potentially of huge interest in relating the relative biological activity of the whelk GAG chains. The unknown peak was found in all samples, in differing amounts, as well as in the standard CS. In order to identify this unknown peak, Fraction CK1 was further analysed using GC-MS after conversion of sugars into their volatile derivatives by TMS. The mass of the unknown peak in fraction CK1 was identified and it has a mass of 226.97 (M-H) (see Appendix II). The samples were hydrolysed using 2N TFA at 100°C for 4h and acid removed, followed by LCNH<sub>2</sub> cartridge purification to remove all neutral components. The bound fractions were eluted by 0.1M ammonium formate dried and re-suspended in 50% methanol and injected directly on MS. The spectra are acquired in negative mode. We can see uronic acid as an ion of 193 (M-H) and the 226.97 (M-H) ion which is not present in the blank or GlcA std, and only present in CS-std and CK1(Appendix II).

The GalNH<sub>2</sub> monosaccharide was most abundant in all samples, which could be a result of *O*-glycans. TFA cleavage conditions with *O*-glycan units increase the chance of producing sugar moiety, as using TFA broadly not recommended especially with acid instable glycosidic linkage of some certain *O*-glycan molecules (Miertus & Fassina, 2005).

**Table 4-2: Quantification of monosaccharide in three fractions of ionic exchange chromatography.**

Sample	Fuc	GalNAc	GlcNAc	Gal	Glc	Man	GalA	GlcA	Total Carb	Amount taken for GRIL-LCMS( $\mu$ l)
CK-1 (1M NaCl fr)	26.334	29.299	42.564	126.237	29.758	-	0.000	1.903	256.095	10
CK-2 (2M NaCl fr)	1.619	5.409	5.814	15.110	4.511	-	0.000	0.000	32.464	62.5
CK-3 (3M NaCl fr)	3.012	11.660	11.056	20.320	12.079				58.125	62.5

#### 4.2.5 GC-MS spectrum of fraction CK-1 GAG as trimethyl silyl derivative

GC analysis of carbohydrate was an alternate method to verify monosaccharide composition of the crude and fractionated whelk GAG. Application of GC requires volatile stable derivatives, this what introduce us to trimethyl silyl derivatives methods which is first described by Sweeley *et al.*, (1963).

Nowadays, GC-MS is a well established techniques for use in carbohydrate analysis. Three main steps are involved in this methods, first each species is converted into a derivative, by converting the hydroxyl groups of free monosaccharides generated by glycan hydrolysis to trimethylsilyl ethers, then the sample is run with GC-MS, which give qualitative and some limited quantitative data on the whelk GAGs structural composition.

Poor solubility of some sugar can sometimes be seen with the TMS derivatives, along with a variations in the sensitivities of different linkages toward hydrolysis especially *O*-glycan linkage. Multiple peaks of same monosaccharide can also be presents for some TMS derivative as the ring forms of the sugar can adopt different anomeric configurations. The

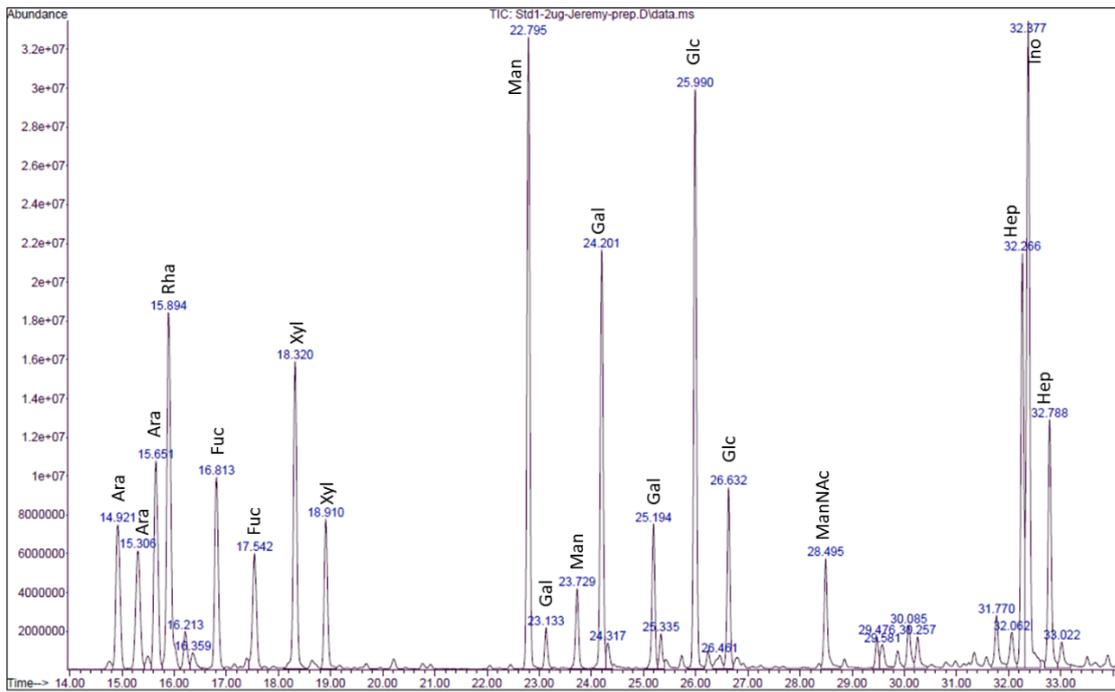
combination with trimethylsilylation (TMS) allows the GC-MS separation of the D and L pair of isomers, thus it helps to identify the absolute configuration of single monosaccharide that might difficult to confirm their formation by TFA hydrolysis of the glycan (Varki *et al.*, 2009).

The result from GC-MS of fraction CK-1 shown in figure (4-5), multiple peaks of the same monosaccharide are present in TMS derivative produced from this fraction. High abundance of GlcNAc and GalNAc was also reconfirmed. The peaks were matched with the fragmentation pattern in two prepared TMS standards, in order to confirm the peak assignments. The TMS standards used contained 1 mg/mL solution of standard sugars: rhamnose, fucose, ribose, arabinose, xylose, mannose, galactose, glucose, *N*-acetyl glucosamine, *N*-acetyl galactosamine, *N*-acetyl mannosamine, glucuronic acid and galacturonic acid along with 1 µg of myo-inositol, as an internal standard were prepared, lyophilize and used as standard.

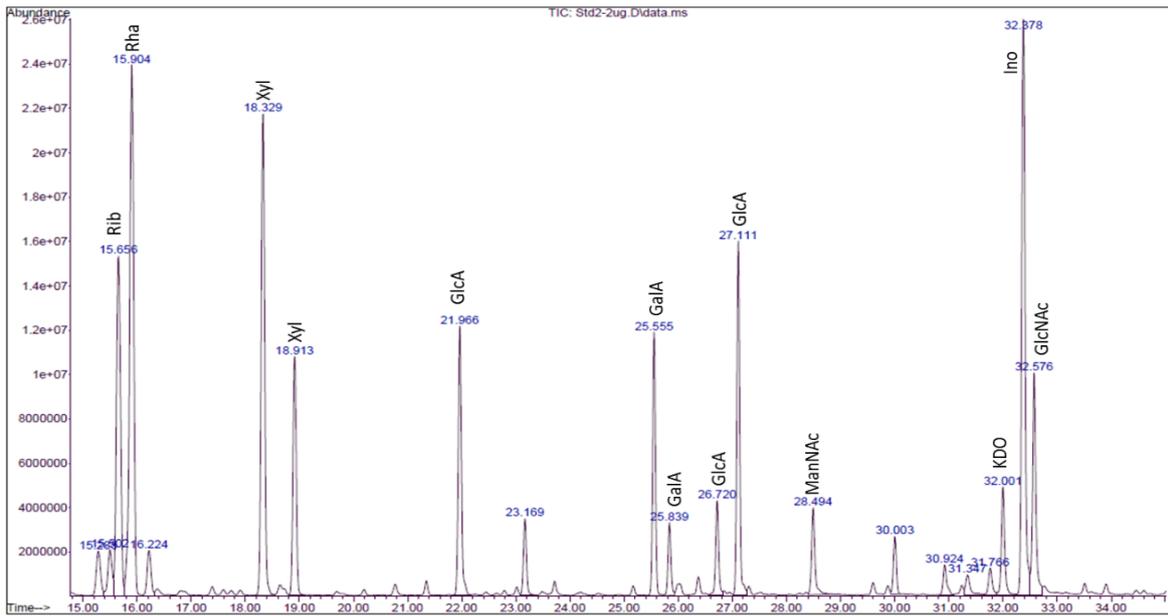
This result confirms that the unknown peak present in all active fractions may be due to a modification in one of the composition, which turned the single monosaccharide to be more acidic and it is very high in most CS containing samples hence in this sample is highly enriched in CS.

**A**

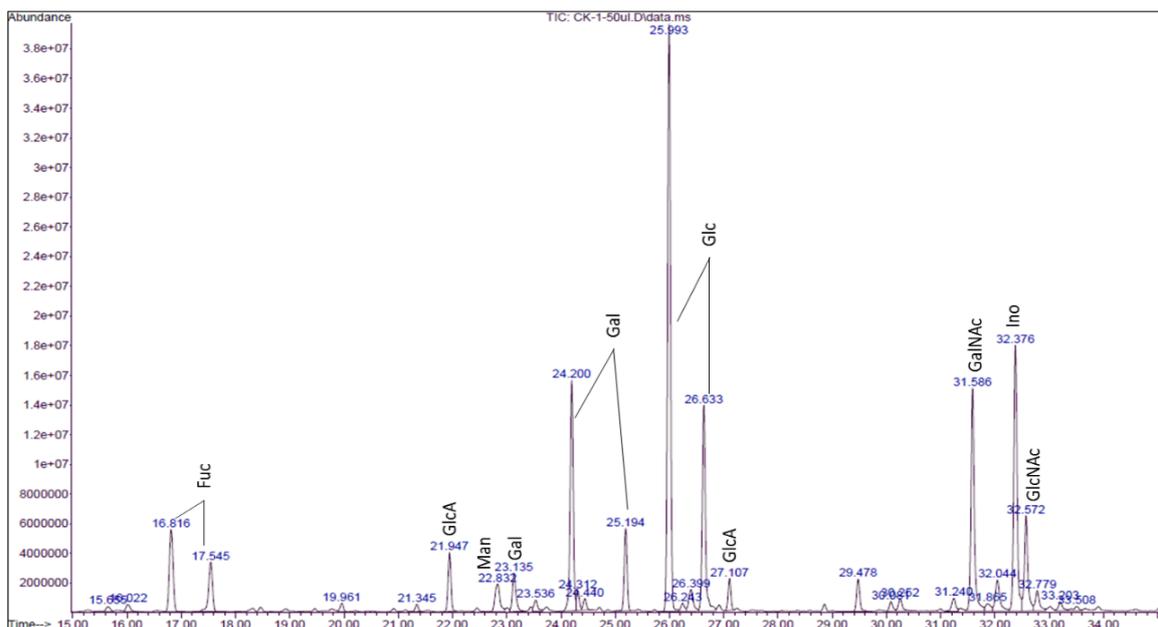
## Standard-1 TMS

**B**

## Standard-2 TMS



### C Composition analysis of GAG fraction (CK-1) by GCMS as TMS derivative



**Figure 4-5: Composition analysis of CK1- GAG fraction and standards by GC-MS.**

**Panel A** and **B**- are standard 1 and 2- TMS. **Panel C**- composition analysis of fraction CK1 by GC-MS as TMS derivative. Multiple peaks of same monosaccharide are presents for TMS derivative as the ring forms of the sugar can adopt different anomeric configurations of D and L forms.

#### 4.2.6 Disaccharide analysis by enzymatic digestions of GAG fractions from ion-exchange chromatography

The amount of each fractions (CK1, CK2 and CK3) required for enzymatic digestion was determined as in in table (4-2). 10 $\mu$ l of CK1, 62.5 $\mu$ l of both CK2 and CK3 were digested with a mixture of Heparinase enzymes and CS-ABC lyase for 24 hours at 37 °C, samples were next fractionated with 10K MWCO filter to remove enzyme and undigested GAG chains, then the flow throw fractions, containing the disaccharide produced by the digestion, were tagged with Aniline isotopic label for future quantitative analysis by Glycan Reductive Isotope Labelling (GRIL) LC-MS.

#### 4.2.7 Glycan Reductive Isotope Labelling (GRIL) LC-MS analysis of Isotopic Aniline tagging for GAG disaccharide

The fundamental idea of isotopic labelling is to replace a single atom by one of its isotopes, thus the single atom can be detected and localised in the molecule structure in the products by detecting the isotopes. In biochemistry there are several types of isotopes such as  $^2\text{H}$  (deuterium),  $^3\text{H}$  (tritium),  $^{13}\text{C}$ ,  $^{14}\text{C}$ ,  $^{15}\text{N}$ , and  $^{18}\text{O}$ , but not all of them yield structural information. Different isotopes can be classified into different groups, some of them are known as radioactive isotopes such as  $^3\text{H}$  and  $^{14}\text{C}$ , others are not radioactive but stable isotopes like  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^{18}\text{O}$  and easy to detect by analytical methods such as NMR especially  $^{13}\text{C}$  and  $^{15}\text{N}$ . Here we use aniline-containing stable isotopes ( $^{12}\text{C}_6$  and  $^{13}\text{C}_6$ ) in Glycan Reductive Isotope Labelling (GRIL) LC-MS as very sensitive and reliable methods to identify the disaccharide even rare type such as D2A0. The (GRIL) LC-MS method can be applied for any type of GAG sourced from human or invertebrate tissues. This technique is much sensitive than UV or fluorescent derivatisation to analyse the disaccharide following enzymatic degradation (Kinoshita *et al.*, 1997).

It provides accurate peak retention time, absolute mass and fragment ion analysis of each disaccharide. The only limitation of analysing carbohydrate by LC-MS methods is the solvent effects on disaccharide ionization that produces variation in the ionization efficiency, which in turn leads to incomplete quantification of disaccharide. However, to overcome this problem the concentration of disaccharide correlated with the detection of a molecular standard. This addressed by applying linear best fit equations (Saad & Leary, 2003), or by tagging techniques using different isotopic tags (Yuan *et al.*, 2005).

The procedure of sample preparation and tagging methods used in this study have described in details in sections (2.9.1, 2.9.2 and 2.10.1).

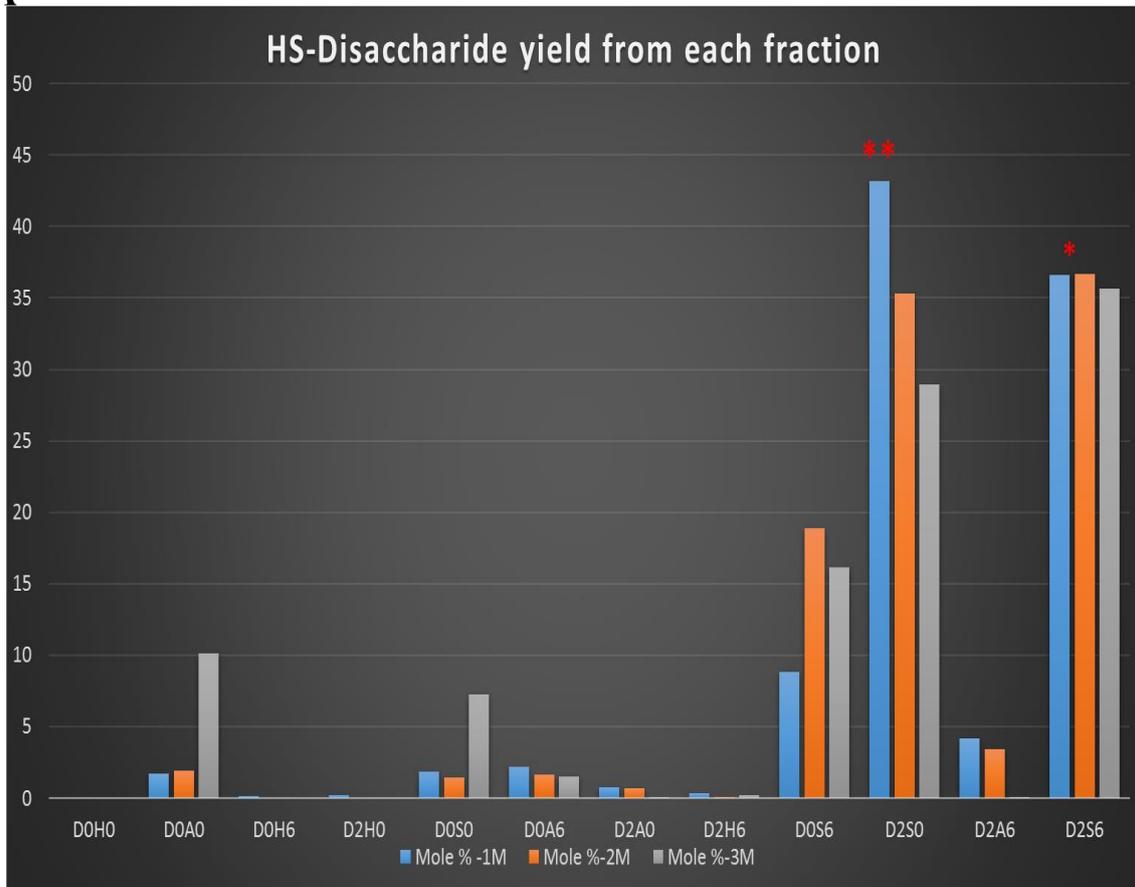
The results shown in Figure 4-6 demonstrate the types and % molar quantities of disaccharide composition identified in flow through FT fraction (the part of material that digested with

Hep lyase and ABC lyase and passed through the 10K filter) of whelk GAG from GRIL-LC-MS. The disaccharide coded according to Lawrence, Lu, *et al.*, (2008) as below:

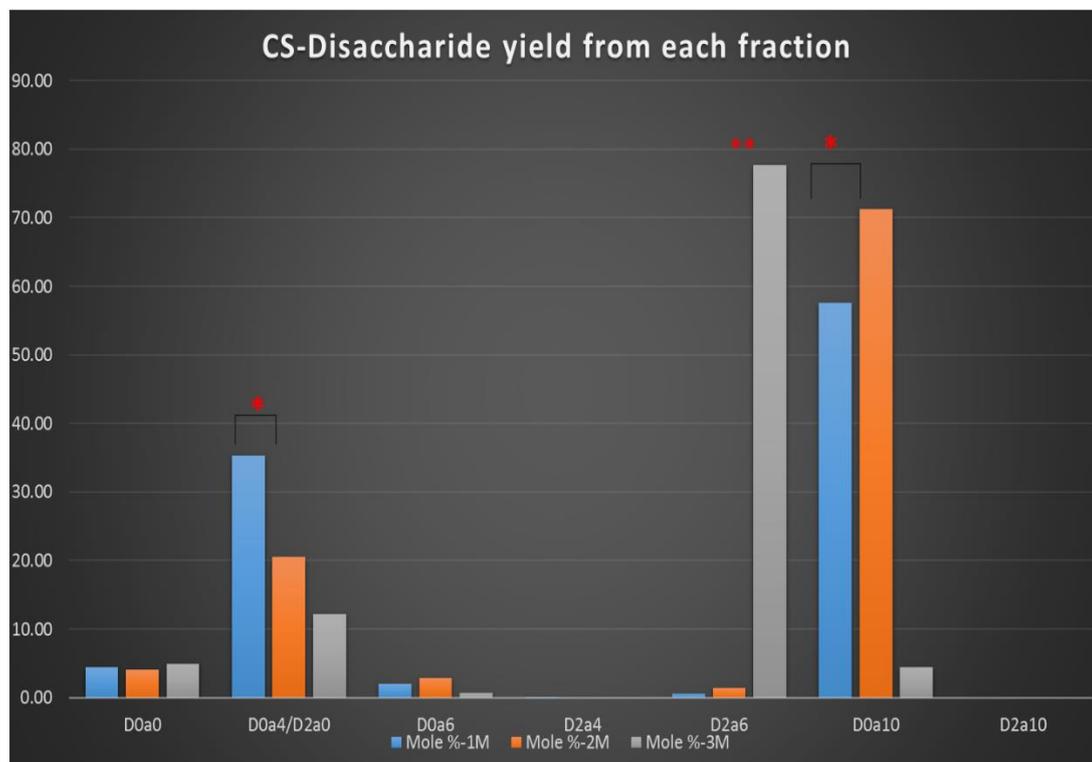
**Table 4-3: HS and CS Disaccharides codes.**

<b>HS</b>	<b>Codes</b>	<b>CS</b>	<b>Codes</b>
<b>Disaccharides</b>		<b>Disaccharides</b>	
D0H0	▲ UA-GlcNH <sub>2</sub>	D0a0	▲ UA-GalNAc
D0H6	▲ UA-GlcNH <sub>2</sub> -6S	D0a4	▲ UA-GalNAc4s
D2H0	▲ UA-2S-GlcNH <sub>2</sub>	D2a0	▲ UA2S-GalNAc
D2H6	▲ UA2S-GlcNH <sub>2</sub> 6S	D0a6	▲ UA-GalNAc6S
D0A0	▲ UA-GlcNAc	D2a4	▲ UA2S-GalNAc4S
D0A6	▲ UA-GlcNAc6S	D2a6	▲ UA2s-GalNAc6s
D2A0	▲ UA2S-GlcNAc	D0a10	▲ UA-GalNAc4S6S
D2A6	▲ UA2S-GlcNAc6S	D2a10	▲ UA2S-GalNAc4S6S
D0S0	▲ UA-GlcNS		
D0S6	▲ UA-GlcNS6S		
D0S3	▲ UA-GlcNS3S		
D2S0	▲ UA2S-GlcNS		
D2S6	▲ UA2S-GlcNS6S		
D2S3	▲ UA2S-GlcNS3S		
D0S9	▲ UA-GlcNS3S6S		
D2S9	▲ UA2S-GlcNS3S6S		

A



**B**



**Figure 4-6: Disaccharide yield from whelk-GAG sample using GRIL-LC-MS.**

**Panel A-** bar graph shows 12 HS disaccharides, (\*\*) indicate most abundant disaccharide in all fractions, (\*) indicated second highest amount of disaccharide type in all flow through FT part in all fractions. **Panel B-** bar graph of 7 CS disaccharides yield from enzymatic degradation of intact whelk-GAG sample identified and quantified by GRIL-LC-MS after tagging with  $^{12}\text{C}_6$  and  $^{13}\text{C}_6$ . (\*\*) indicate most abundant disaccharide in flow through FT fraction in fraction CK3, (\*) indicate highest amount of disaccharide found in all flow through FT fraction in fraction CK1 and CK2.

The results show that there is a significantly high amount of DS20 ( $\Delta\text{HexA}2\text{S-GlcNS}$ ) in whelk GAG fractions (43.15%) relative to commercial porcine HS (10.4%) as identified by Gill *et al.*, (2013) and porcine heparin (0.9%) as detected by Saad *et al.*, (2005). The second noticeable disaccharide detected in all whelk GAG fractions CK1, CK2 and CK3 was D2S6 ( $\Delta\text{HexA}2\text{S-GlcNS}6\text{S}$ ), in contrast, this disaccharide is representing most abundant disaccharide in commercial porcine heparin and HS. Types of CS-disaccharide found within whelk GAG fractions are also significantly different from commercial porcine CS, the

amount of specific disaccharide between each fraction are clearly various. In fraction CK1 and CK2, D0a10 ( $\Delta$ HexA-GalN4S6S) disaccharide was the highest at 57.60% and 71.22% respectively, while in fraction CK3 it was 4.42%, comparing to the commercial porcine which is represent the least amount 1.39%. The remarkable amount of D2a6 ( $\Delta$ HexA2s-GalN6S) is found only in fraction CK3 at 77.74%, while it was 0.59%, 1.43% and 0.04% in CK1, CK2 and CS-Std respectively. The percentage of molarity of each HS disaccharide characterised in fractions CK1, CK2, and CK3 illustrated in details in table (4-4). The percentage of *N*, 2-*O*, 6-*O*-sulfation patterns in HS disaccharide has been calculated in each fraction and the average number of *SO*<sub>3</sub> per disaccharide quantified. The results show high amounts of *N-SO*<sub>3</sub> at about 90.23% average for all GAG sample fractions comparing to other type of sulphating pattern. In addition, the percentage 2-*O*, 4-*O*, 6-*O*-sulfation pattern and average *SO*<sub>3</sub> per disaccharide has clearly identified in CS disaccharide yield in all fractions in table (4-5). A high amount of 6-*O-SO*<sub>3</sub> sulphating pattern (72.82%) found in all CS disaccharide fractions, almost twice as much as the amount found in CS-Std (36.24%). This has followed by 2-*SO*<sub>3</sub> sulphating pattern (an average of 26.59% of all fractions) which is represent 0.30% of the total compound sulphating pattern in CS-Std. The 2-*SO*<sub>3</sub> pattern, representing 77.74% of all sulphating types, found mostly in fraction CK3, which was eluted in the highest salt strength of 3M NaCl. The amount of 4-*O*-sulfation in CS disaccharide in whelk GAG fractions found very similar to the amount in the CS-standard, as represent at an average of 67.11% in the sample and 61.42% in the CS-Std.

**Table 4-4: Disaccharide quantities of fractions CK1, CK2 and CK3 in percentage mole.**

A: percentage mole of HS disaccharide for all whelk-GAG sample fractions. B: Percentage mole of CS-Std and CS disaccharide for all whelk-GAG sample fractions. (\*) indicates most abundant HS/CS disaccharide in whelk GAG fractions.

<b>A</b>	Disaccharide from Heparin lyase of intact whelk-GAG sample	% Mole in CK1-1M NaCl derivative	% Mole in CK2-2M NaCl derivative	% Mole in CK3-3M NaCl derivative	5% Mole in commercial porcine HS
1	D0H0	0.00	0.00	0.00	0.00
2	D0A0	1.70	1.91	10.10	28.35
3	D0H6	0.12	0.00	0.00	0.13
4	D2H0	0.23	0.00	0.00	0.03
5	D0S0	1.85	1.42	7.28	12.39
6	D0A6	2.19	1.63	1.48	10.55
7	D2A0	0.79	0.69	0.09	1.33
8	D2H6	0.36	0.06	0.24	0.41
9	D0S6	8.83	18.90	16.17	11.21
10	D2S0	43.15*	35.28	28.93	8.81
11	D2A6	4.18	3.45	0.08	0.98
12	D2S6	36.60*	36.66	35.62	25.81
	total	100.00	100.00	100.00	100.00

**B**

	Disaccharide from CS-ABC lyase of intact whelk-GAG sample	% Mole in CK1-1M NaCl derivative	% Mole in CK2-2M NaCl derivative	% Mole in CK3- 3M NaCl derivative	5% Mole in commercial porcine CS
1	D0a0	4.43	4.02	4.95	3.83
2	D0a4/D2a0	35.36	20.53	12.21	59.77
3	D0a6	2.00	2.81	0.67	34.71
4	D2a4	0.02	0.00	0.00	0.16
5	D2a6	0.59	1.43	77.74*	0.04
6	D0a10	57.60*	71.22*	4.42	1.39
7	D2a10	0.00	0.00	0.00	0.10
	total	100.00	100.00	100.00	100.00

**Table 4-5: Type of sulphating in HS and CS derivatives of each whelk-GAG fractions.**

A: represent the type of sulphating in each HS fractions and the average concentration of sulphates per HS disaccharide. B: represent the type of sulphating in each CS fractions and average concentration of sulphates per CS disaccharide. (\*) indicates most abundant HS/CS disaccharide sulphating patterns in whelk GAG fractions.

**A**

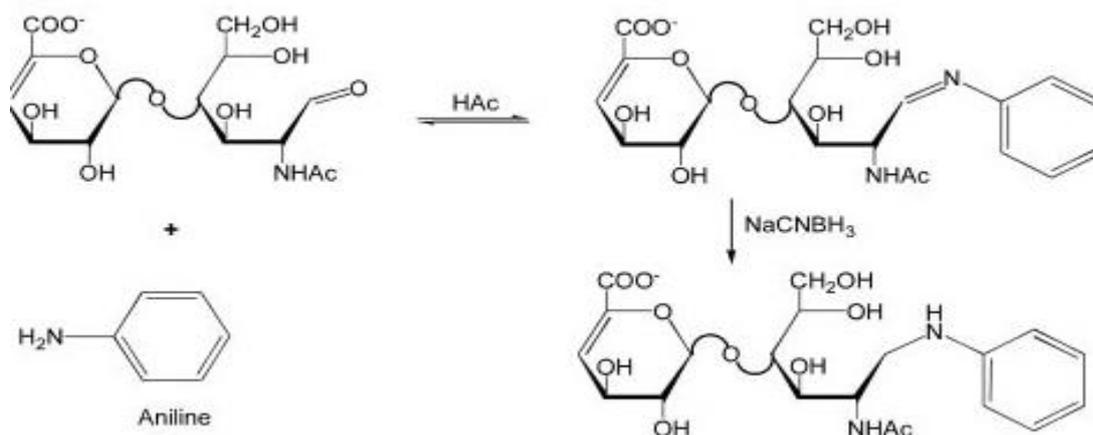
Type of sulphating	of HS-CK1	HS-CK2	HS-CK3	HS-Std	Average of all fractions
% <b>Unsulphated</b>	1.70	1.91	10.10	28.35	4.57
% <b>N-SO<sub>3</sub></b>	90.44*	92.26*	88.01*	58.21	90.23
% <b>2-O-SO<sub>3</sub></b>	85.31	76.15	64.96	37.37	75.47
% <b>6-O-SO<sub>3</sub></b>	52.28	60.71	53.59	49.09	55.52
<b>Average SO<sub>3</sub> per disaccharide</b>	2.28	2.29	2.07	1.45	2.21

**B**

Type of sulphating	CS-CK1	CS-CK2	CS-CK3	CS-Std	Average of all fractions
% <b>Unsulphated</b>	4.43	4.02	4.95	3.83	4.46
% <b>2-SO<sub>3</sub></b>	0.61	1.43	77.74*	0.30	26.59
% <b>4-O-SO<sub>3</sub></b>	92.98*	91.74*	16.63	61.42	67.11
% <b>6-O-SO<sub>3</sub></b>	60.18	75.45	82.84	36.24	72.82
<b>Average SO<sub>3</sub> per disaccharide</b>	1.54	1.69	1.77	0.98	1.66

A LCQ classic quadrupole ion trap mass spectrometer equipped with an electrospray ionization source, and a quaternary high-performance liquid chromatography pump (Thermo-Finnigan, San Jose, CA) was used for disaccharide analyses. Derivative and non-derivative disaccharide residues were separated on a C18 reversed-phase column (0.46 x 25 cm, Vydac) with the ion pairing agent dibutylamine (DBA, Sigma-Aldrich). Ions of interest were monitored in negative ion mode. To minimize in-source fragmentation of sulphated disaccharides, the capillary temperature and spray voltage were kept at 140 °C and 4.75 kV, respectively (Katajamaa & Orešič, 2005).

Samples were treated with Heparinases I, II, and III, similar amounts digested separately with ABC lyase to yield derivatives, the reducing form of each GAG disaccharide can form a Schiff base with the amine group of aniline. The resulting imine is then reduced with sodium cyanoborohydride resulting in a stable derivative (Lawrence, Olson, *et al.*, 2008).



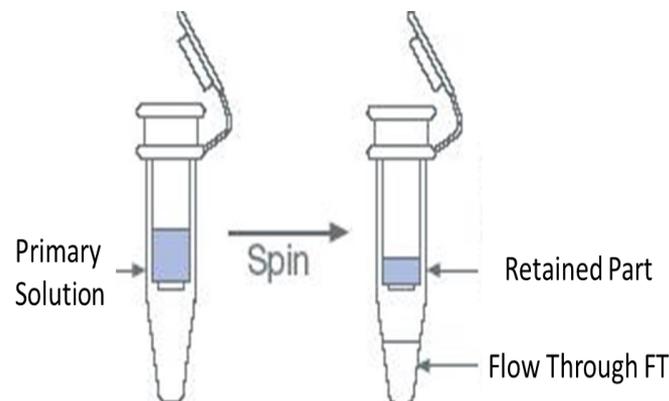
**Figure 4-7: The interaction of amino group in the aniline tag with the reducing sugar form.**

The LC-MS spectra of the flow through (FT) fraction from nanosep centrifugal devices with omega membrane of all fractions CK1, CK2, CK3, HS-Std and CS-Std shown in appendix II. Disaccharides and standards produced by treatment with CS-ABC/ Hep lyase, tagged with <sup>12</sup>C-aniline and spiked with known amount of <sup>13</sup>C-aniline tag as an internal standard. Ions of interest were monitored in negative ion mode. The <sup>13</sup>C<sub>6</sub> aniline-derivative forms of commercially provide HS and CS were used for quantitative analysis by GRIL-LC/MS.

The m/z values of the <sup>13</sup>C<sub>6</sub> aniline-derivative standards differ from the <sup>12</sup>C<sub>6</sub> aniline derivatives by + 6 atomic mass units, more illustration found in appendix II.

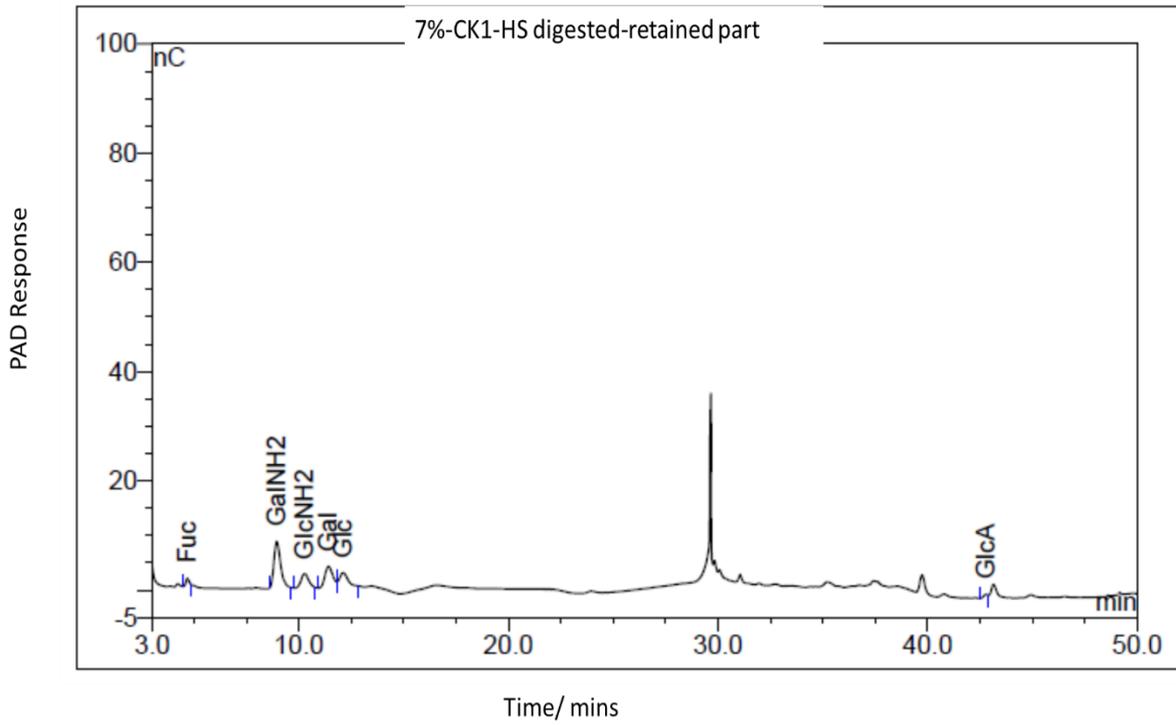
The GRIL-LC-MS is applied to all flow through (FT) of the whelk-GAG fractions CK1, CK2 and CK3, the FT is the material that has been completely digested by Hep lyase or ABC lyase into di or oligosaccharide of (2-6 disaccharide) that passed through the 10K filter (Nanosep

centrifugal devices with omega membrane) (Figure 4-8), the filter media contain omega membrane that has low protein-binding, modified polyethersulfone on polyethylene substrate, the filtrate receiver is polypropylene, for rapid purification, separation of mixture and sample fractionation, the membrane pore size is less than 5nm and able to fractionate biomolecule with rang of 30K-90K molecular weight. Any material longer in size than the membrane pore size will remain on the top of the filter as retained fraction, this could include buffers, enzymes, and incomplete digested substrates. The retained fraction of all sample fractions were subjected for monosaccharide analysis to determine the monosaccharide in the mixture, using HPAEC-PAD. The results in figure 4-9 confirmed the monosaccharide composition of the retained fraction from all fractions, which is clearly suggest incomplete digestion of the whelk GAG with both Hep lyases and chondroitinases ABC.

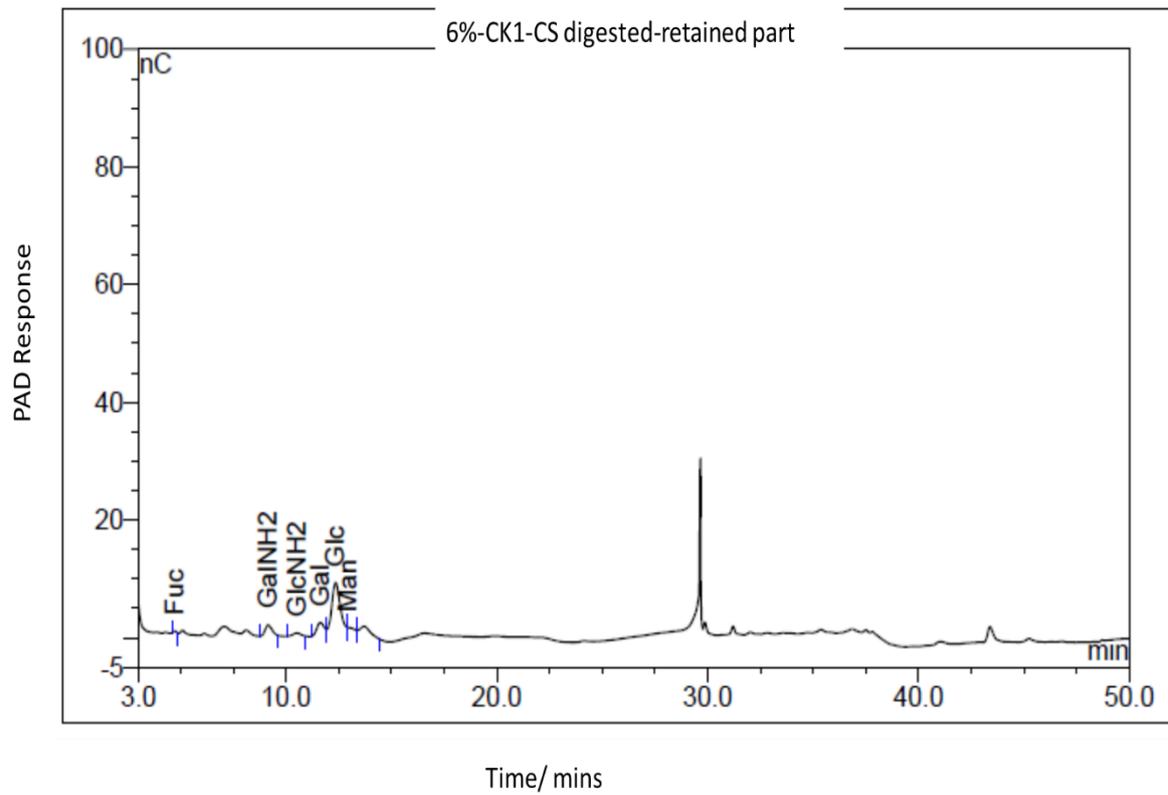


**Figure 4-8: The separation technique of Nanosep Centrifugal Devices with Omega Membrane.**

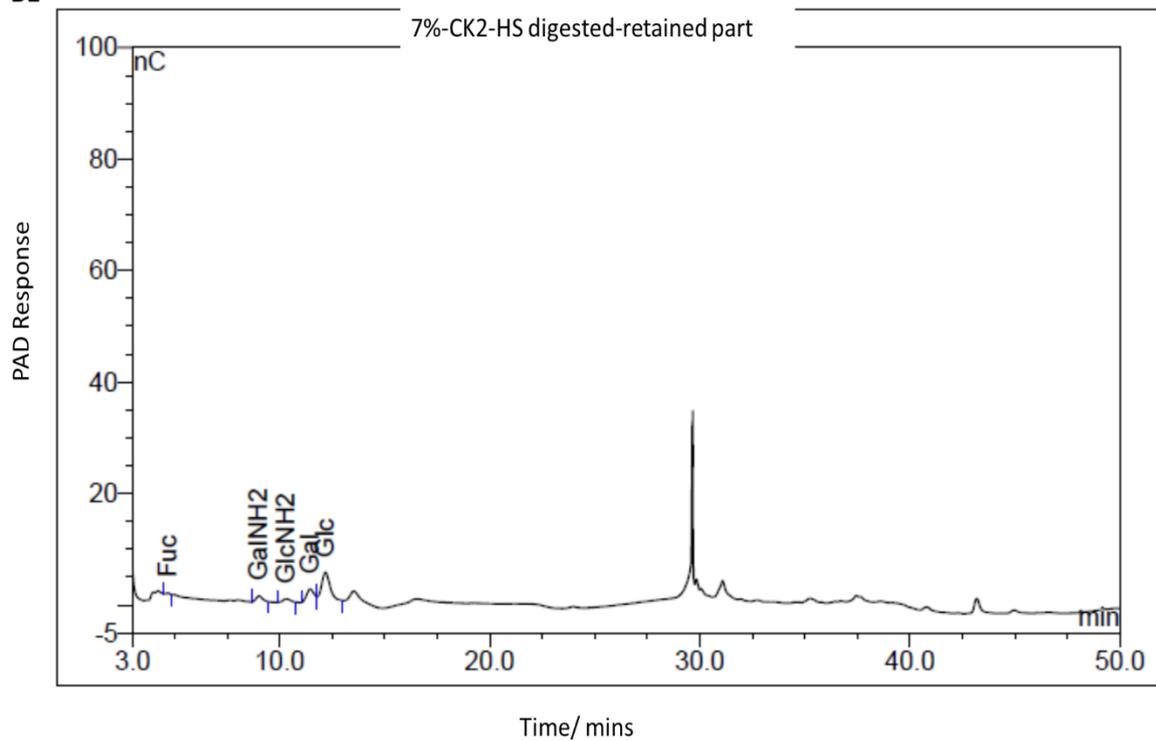
A1



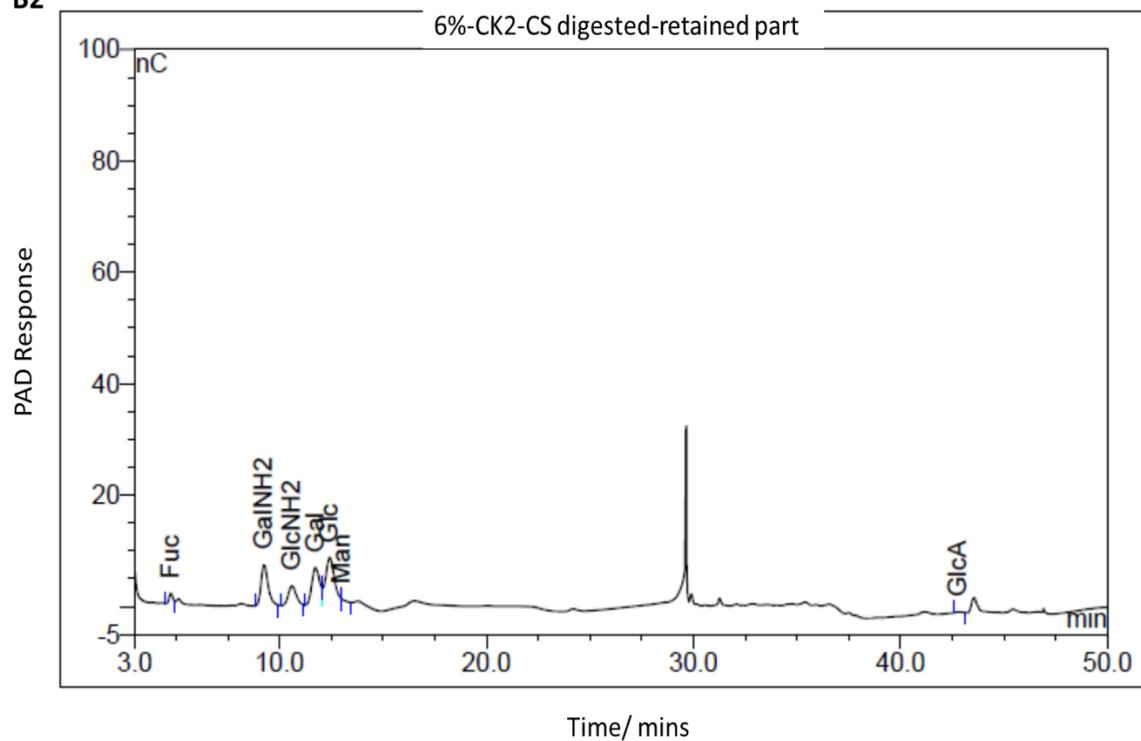
A2

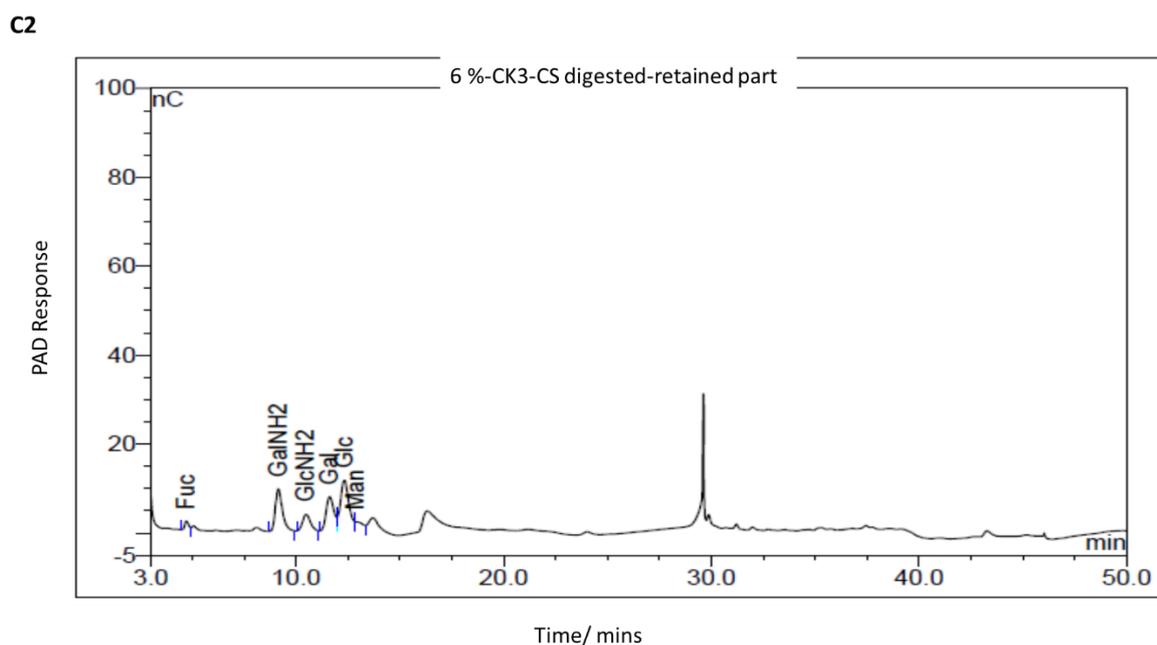
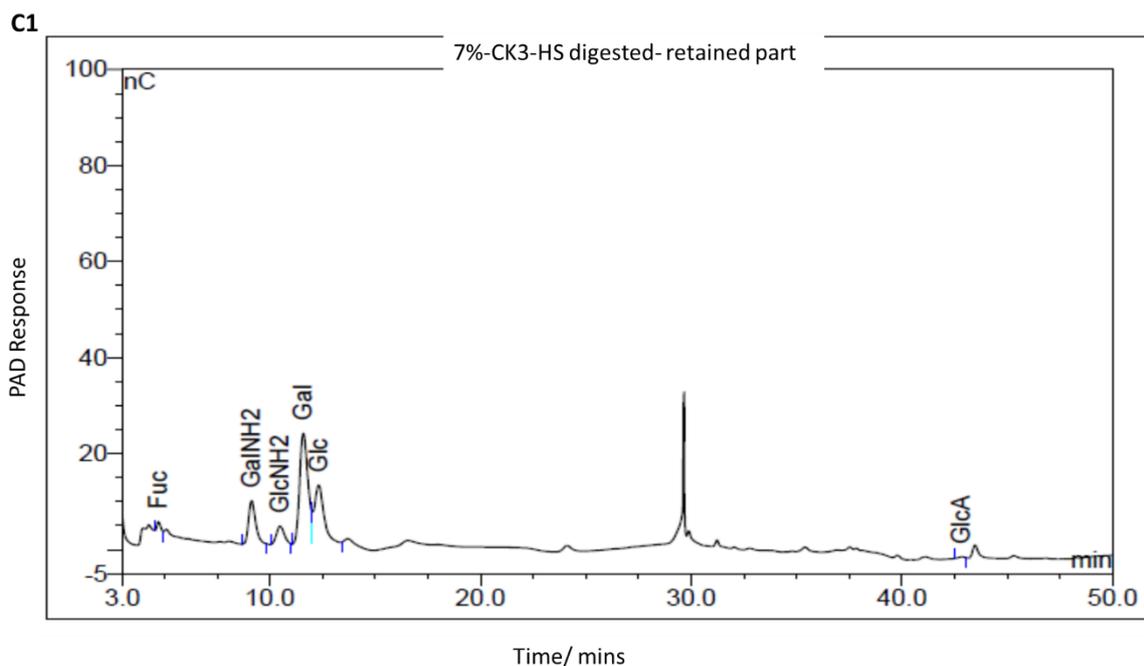


**B1**



**B2**





**Figure 4-9: Monosaccharide analysis of retained fraction of the mixture in all sample fractions.**

**Panels A1, B1 and C1:** Represent the monosaccharide that exist in the retained fraction of 10% CK1, CK2 and CK3 fractions respectively that digested by HS lyase. **Panels A2, B2 and C2:** Represent the monosaccharide that exist in the retained fraction of 10% CK1, CK2 and CK3 fractions respectively that were digested by CS-ABC lyase.

#### **4.2.8 Monosaccharides analysis by HPAEC-PAD technique of whelk-GAG sample oxidised by periodic acid.**

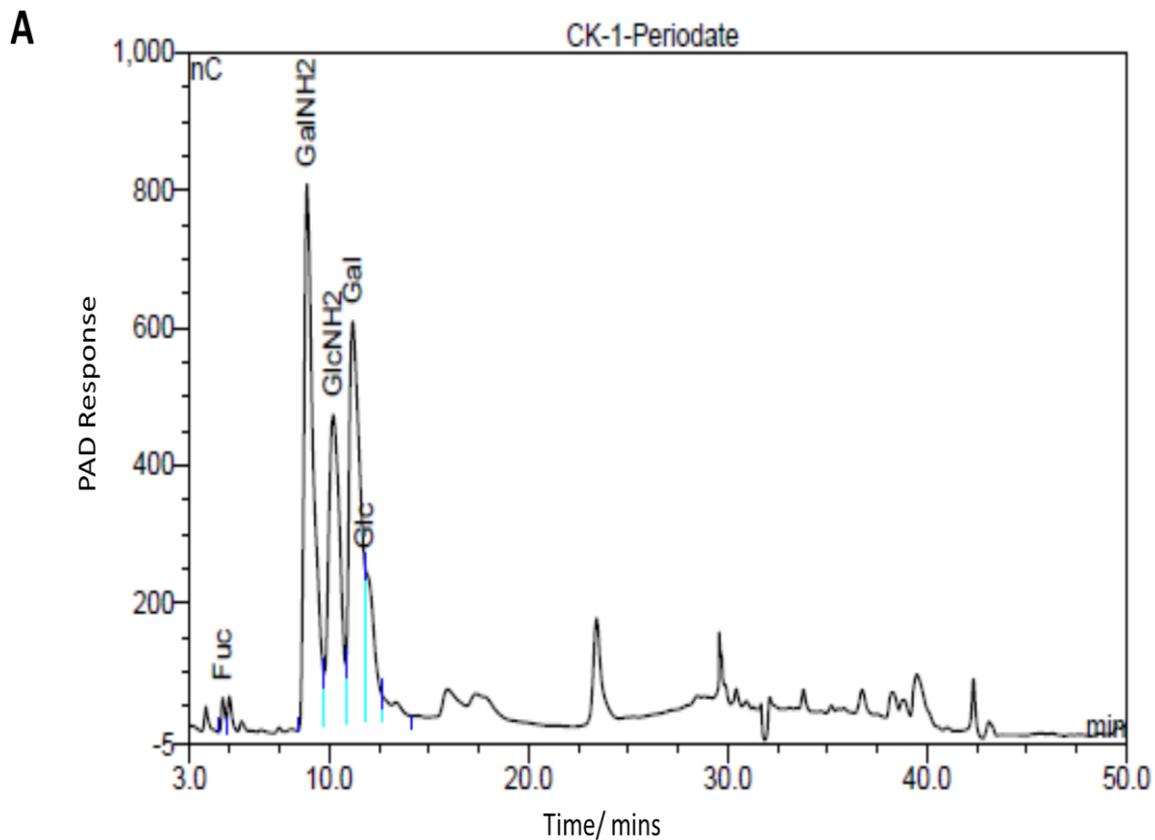
Periodic acid (HIO<sub>4</sub>) oxidation cleaves the vicinal diol group on a carbohydrate molecule to generate an aldehyde groups. This oxidation is often applied to polysaccharides that have vicinal hydroxide groups to generate oligosaccharide for structural elucidation. The monosaccharide residues of a polymer containing hydroxyl groups which are involved in an ester or glycosidic linkage are found to be resistant to periodic acid oxidation.

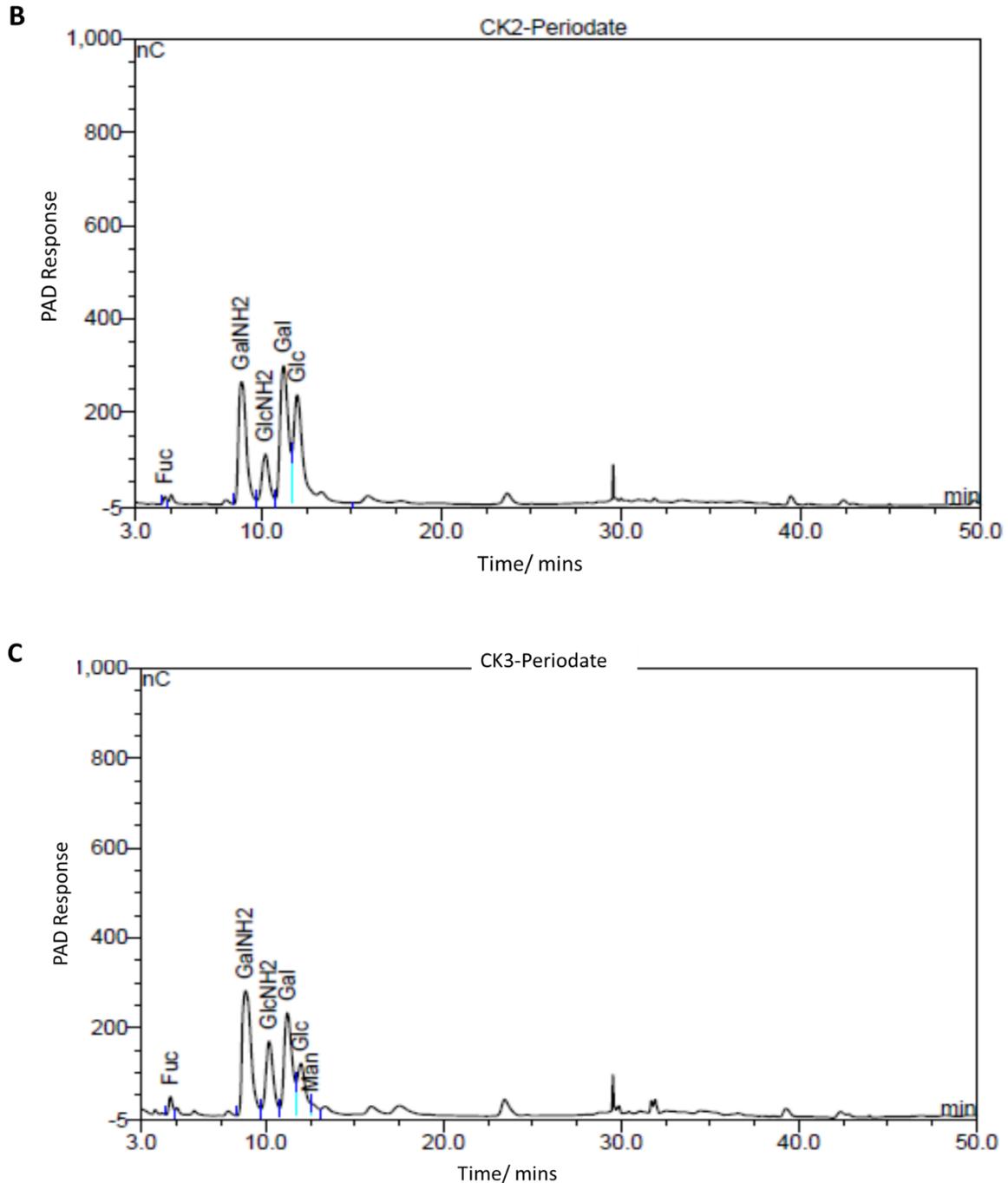
Most of neutral sugars are considered to be reactive monosaccharides, such as mannose, fucose, glucose and galactose. It has been reported that periodic acid might not produce aldehydes in uronic acid and highly sulphated hexosamines in the glycol-conjugated or proteoglycan due to electrostatic repulsion condition (Leblond *et al.*, 1957).

The aim of this experiment is to determine the sugar ring structure and in addition, to examine the possibility glucose are oxidised to aldehydes in the whelk-GAG samples, which might then in turn lose their bioactivity on cancer cells. To validate the results from the periodic acid oxidation treatment, all whelk-GAG sample underwent monosaccharide compositional analysis to provide a clear evidence of whether the mixture of intact whelk-GAG is completely oxidised or that the ring form of the monosaccharide remained intact. The results shown in figure 4-10 demonstrated the monosaccharide composition analysis of whelk-GAG fractions (CK1, CK2 and CK3) after the oxidation process by periodic acid treatments. It is obvious that the respective amounts of monosaccharide remained non-oxidised and the chains are still highly polymerised. This was expected, as such compounds show no sign of aldehyded formation by oxidation process with periodic acid due to the strong negative charge pass on by ionised carboxylate groups of the glucouronic acid and the sulphated groups of hexosamine that cause electrostatic repulsion with the negatively charged periodate ions, which inhibit the oxidation of acidic polysaccharide (Bancroft & Gamble, 2008). However, no uronic acid was detected after the oxidation reaction in all fractions,

whilst high abundance of neutral sugars is observed. Prolonged periodic acid treatment might affect uronic acid in the GAG sample fractions and hence make it positive for periodic oxidation treatment (Kiviranta *et al.*, 1985).

According to Scott & Harbinson, (1969), targeting the uronic acid moiety in chondroitin sulphates A, B and C, and seaweed with oxidation reaction by medium containing 0.2 M sodium perchlorate to avoid electrostatic field effect, showed that polymer-bound uronic acids whether this was iduronic or guluronic are more quickly oxidised than their glucuronic and mannuronic D-analogues.





**Figure 4-10: Monosaccharide composition analysis by HPAEC-PAD of whelk-GAG sample fractions oxidised with periodic acid.**

**Panel A:** Monosaccharide composition of fraction CK1 after treating with periodic acid to oxidise the sugar rings. **Panel B:** Monosaccharide composition of fraction CK2 after treating with periodic acid to oxidise the sugar rings. **Panel C:** Monosaccharide composition of fraction CK3 after treating with periodic acid to oxidise the sugar rings

#### 4.2.9 NMR spectra analysis

Nuclear Magnetic Resonance Spectroscopy (NMR) is considered as a powerful analytical method for a variety of analytical applications. First observation is back to 1940s. All elements can be usefully studied using NMR. However, some elements are difficult to observe due to their nuclear properties such as oxygen and aluminium.

This technique could be defined as “the absorption and emission of electromagnetic radiation by the nuclei of certain atoms when they are placed in magnetic field”. During the last decade NMR has become increasingly attractive for use in biology and medicine applications including studies of bio-macromolecular structure and function.

It believes that both standards (the proton  $^1\text{H}$  and carbon  $^{13}\text{C}$ ) along with other spectroscopic can provide confirmation of the structure of drug substances. In case of  $^1\text{H}$  NMR could find a wide range of protons in different chemical groups can be found. Chemical shift measurements are formally based on the resonance position of the bare proton nucleus as the primary standard; therefore, we could define the chemical shift in terms of the difference in resonance frequencies between the nucleus of interest (i.e.  $^1\text{H}$ ) and a reference nucleus (i.e.  $^1\text{H}$  of TMS) by means of a dimensionless parameter (Earl & VanderHart, 1982).

Some atoms have no overall spin as they spin are paired and cancel each other such as  $^{12}\text{C}$ ,  $^{16}\text{O}$ ,  $^{32}\text{S}$  atoms, while others like  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{31}\text{P}$ ,  $^{15}\text{N}$ ,  $^{19}\text{F}$  have nucleus with an overall spin. There is a correlation between overall spin and the number of neutrons and protons, if they both balanced then the nucleus has no spin. But if the number of neutrons and the number of protons are disproportionate then the nucleus has a half-integer spin, or it has an integer spin if there is not appropriate number of neutrons or protons (Edwards, 2009).

In drug discovery, NMR has an important role in structure elucidation of unknown bioactive macromolecular compounds from first identification to optimisation steps. It is an efficient and accurate tool to successfully screen the compounds with powerful action as drugs for

curing diseases. All the structure information of the compounds can be monitored and extracted based on the chemical shift.

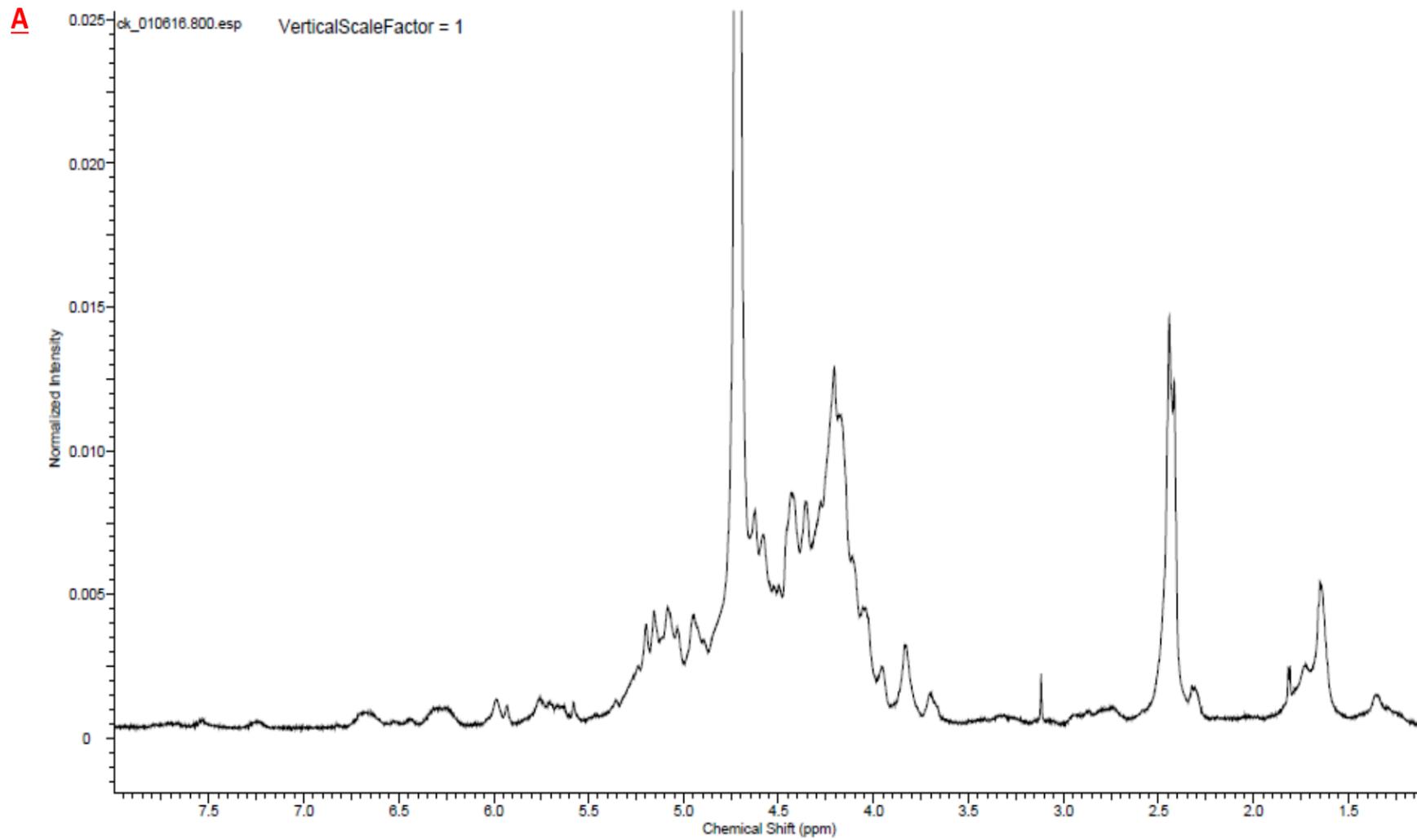
#### **4.2.10 1D (<sup>1</sup>H-NMR) and 2D (HSQC-NMR) experiments**

Two types of NMR experiments (<sup>1</sup>D and <sup>2</sup>D) were run in order to identify the structure of the whelk-GAG samples, using Bruker NMR spectrometer operating at a <sup>1</sup>D and <sup>2</sup>D frequency of 800MHz. AVIII NMR spectrometer with TCI cryoprobe equipped with Z gradients at 70°C. Also 600MHz to operate <sup>1</sup>H-NMR at room temperature was used to monitor all whelk-GAG fractions from ion exchange chromatography CK1, CK2 and CK3 using Deuterium Oxide (D<sub>2</sub>O) as the solvent.

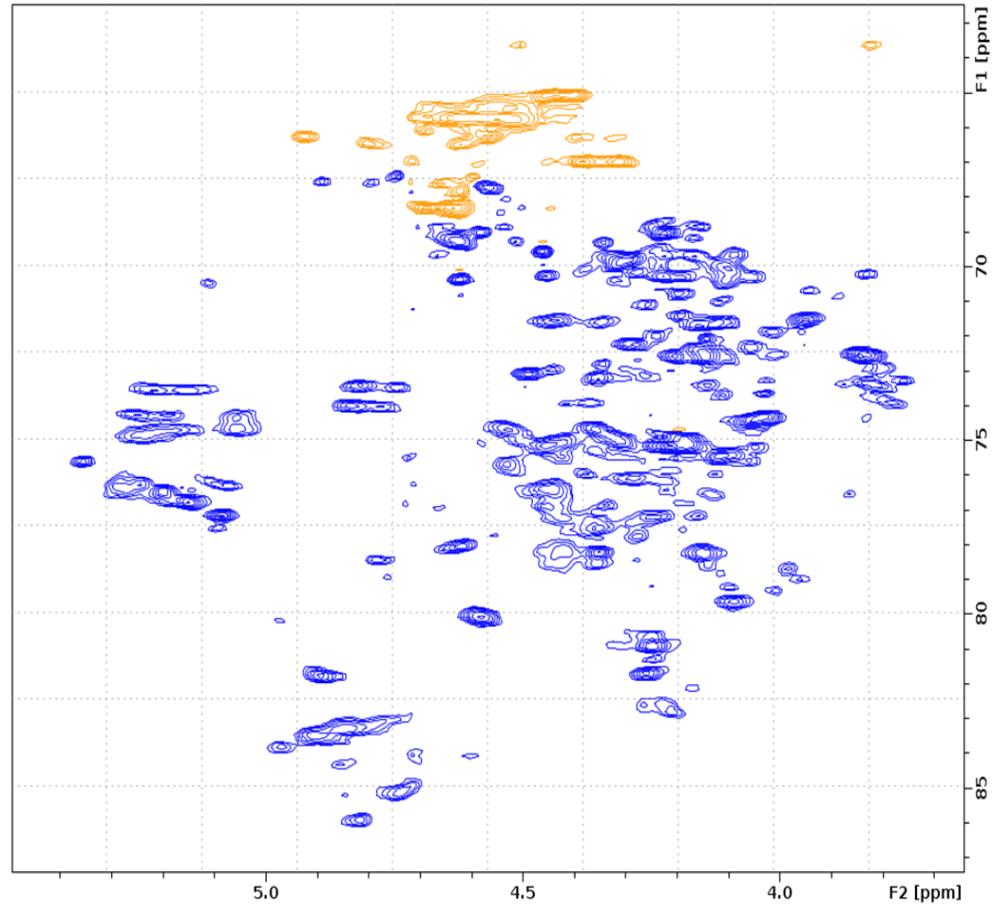
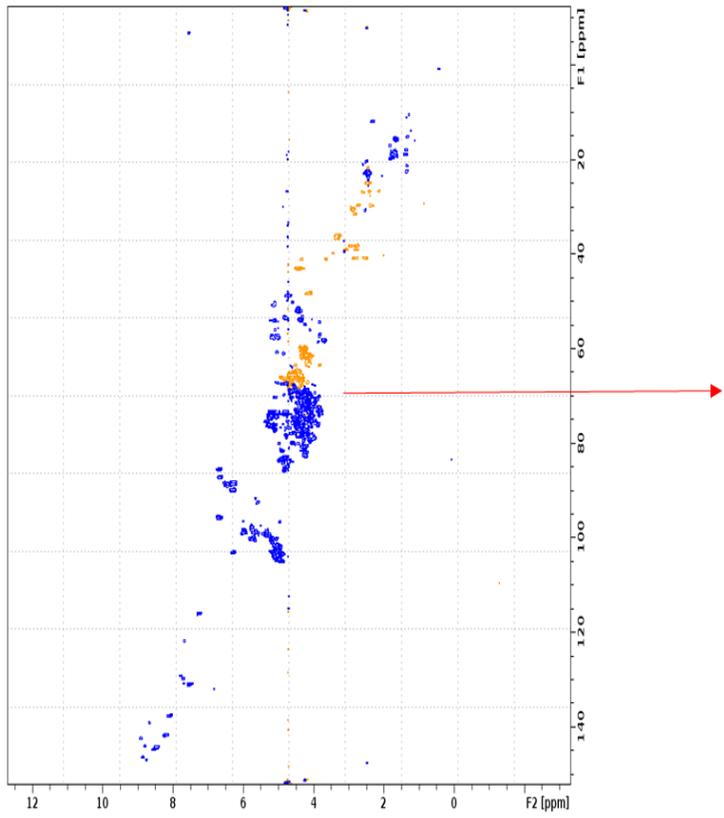
The data in one dimensional NMR (<sup>1</sup>D) is usually presented with x axis corresponding to the frequency axis which is the chemical shifts measured in ppm and the Y axis corresponds to the intensity, while in <sup>2</sup>D-NMR there are two frequency axes and the intensities present in third axis, therefore, it is present as plots.

The aim of this experiment was to identify the structure of the complex compounds on the atomic level and their interactions with other compounds. Sample preparation and solvent exchange used in this experiment is as shown in section (2.12).

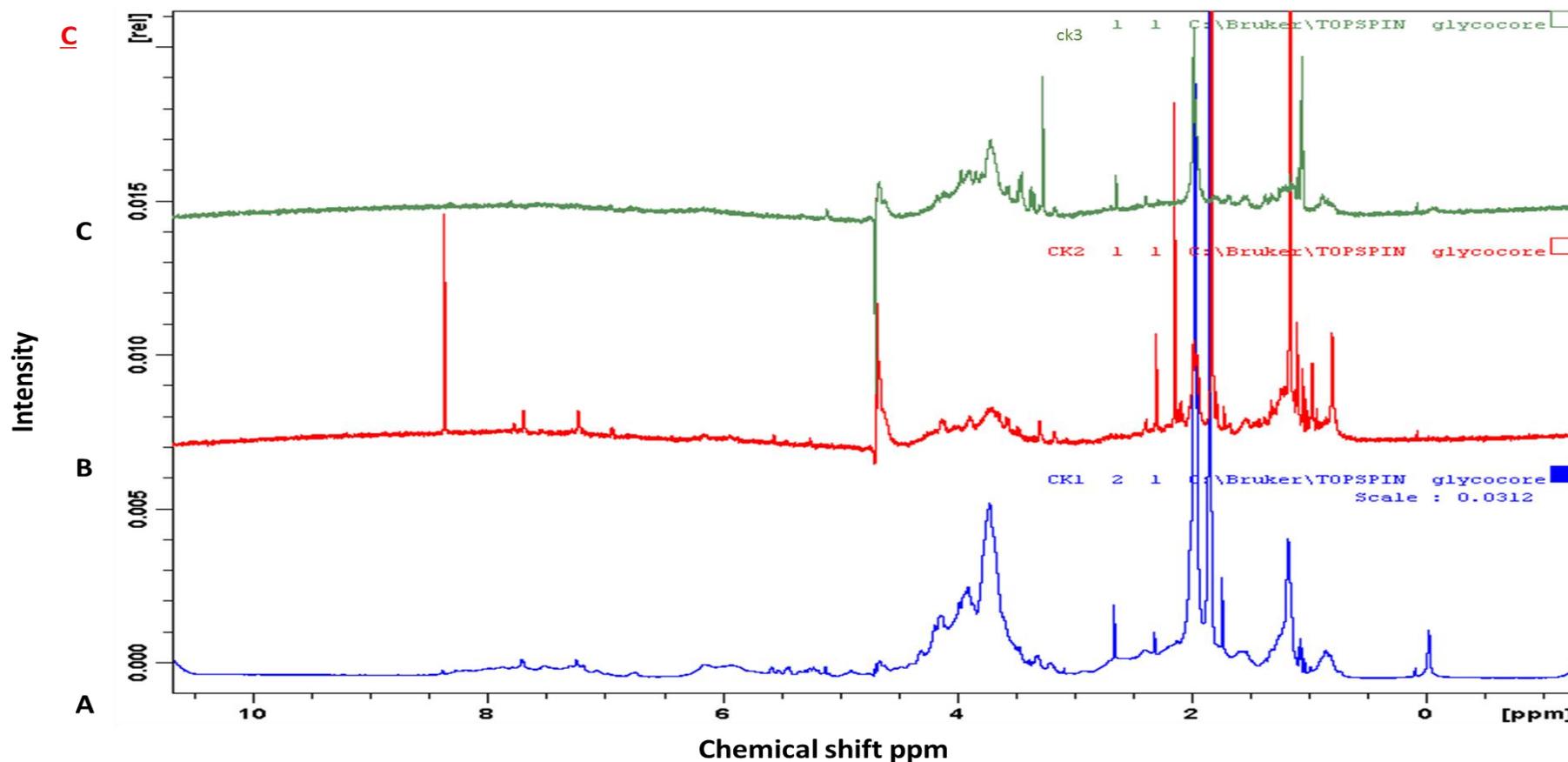
Figure (4-11) shows the signal in both <sup>1</sup>D/<sup>2</sup>D-NMR of intact whelk-GAG samples. Broad peaks were observed from <sup>1</sup>H-NMR of intact whelk-GAG and CK1, CK2, and CK3 fractions. Measurement of multiple coupling constants directly was not possible from all spectra signals due to peaks overlap. Unfortunately, the complexity of the mixtures in each sample made it impossible to make unambiguous assignments of the glycan constituents within the samples. More figures provided in appendix V.



**B**



Chemical shift ppm



**Figure 4-11: Brucker NMR spectrometer at  $^1\text{D}$  and  $^2\text{D}$ , frequency of 800MHz at 70 °C and 600MHz at 37 °C.**

**Panel A:**  $^1\text{H}$  spectrum of CK1 extracts at frequency of 800MHz at 70 °C. **Panel B:** HSQC spectrum of crude GAG extracts at frequency of 800MHz at 70 °C. **Panel C:**  $^1\text{H}$  spectrum of fractions CK1, CK2 and CK3 at frequency of 600MHz at 37 °C.

### 4.3 Discussion

The aim of this chapter was to analyse the structure of GAGs isolated from whelks and to begin to link biological activity with structural fine detail. This is important in the progression of designing bioactive compounds from natural products and to ultimately move towards the development of potential therapeutic approaches.

The main finding of this chapter were that the structure of the whelk GAG is significantly different from the structure of mammalian GAGs, which lies mainly in the abundance of different disaccharide units and the sulphation patterns presents in these molecules.

Initially, the monosaccharide composition of the whelk GAG was investigated. The amount, and type, of monosaccharides was determined using highly sensitive techniques, most of the monosaccharides found in the whelk GAG samples are also typically present in different glycoconjugates such as glycoproteins and proteoglycans from other organisms.

The monosaccharide composition analysis is an important method used to identify the structure of complex GAGs, this methods has been adopted in several studies (Gao *et al.*, 2011). Monosaccharide composition analysis by HPAEC-PAD suggests the presence of specific non-GAG related oligosaccharide classes such as *N*-glycans or *O*-glycans within the crude GAG samples.

High abundance of glucose, fucose and galactose monosaccharide are seen in the whelk GAG samples, this in agreement with the finding by Tanoue & Handa, (1987). According to KerheRvé *et al.*, (1995), monosaccharide composition analysis of different marine species revealed a high abundance of glucosamine, mannose, ribose, rhamnose, fucose, and xylose. Most of these monosaccharides are also detected in the crude whelk GAG samples in this study.

The results showed lower amounts of uronic acid distributed among whelk GAGs chain. This finding has been confirmed by the result obtained from GC-MS. Some marine organisms

have been reported to express small amounts of uronic acid whether iduronic or glucouronic acid comparing to the amount of other monosaccharides found in mammalian GAG chains (Jiao *et al.*, 2011). This finding concurs with another study that showed the chemical composition of green and red seaweeds *Codium fragile* and *Chondrus crispus*, which are reported to have 31% neutral sugars within their structure compared to only 0.6% uronic acid (Kulshreshtha *et al.*, 2015).

The most likely explanation of the unclear findings from the size exclusion chromatography experiments is the incomplete digestion of GAG substrate by the enzymes. This resulted in the presence of broad peaks observed from elution profiles. Unfortunately, an extensive period of sample incubation with lyases showed no further degradation, which might indicate that whelk GAGs are unfavourable substrate of either enzyme. Many factors should be taken into consideration when broad peaks are observed from size exclusion chromatography, such as sample impurities. This makes purification of heterogeneous GAG samples even more difficult and increases the complexity within the samples elution profile. In addition, the main limitation of this technique is the low resolution profile with carbohydrate comparing to proteins. As carbohydrate is heterogeneous and larger than other macromolecules.

The above analysis does not enable us to determine comprehensively the fine structure of Whelk GAG chains but the disaccharides or oligosaccharides generated from enzymatical degradation does give us an insight into the reasons why the biological properties of Whelk GAGs differ so significantly from mammalian GAG chains.

Ion exchange chromatography enabled the fractionation the GAG sample according to their net negatively charged groups using stepwise adding of sodium chloride with three different buffer strengths. The result show that the fraction eluted with 1 M NaCl yielded approximately 90% of total sample applied to the column, followed by smaller amounts of

fractions eluted with 2M or 3M NaCl salt concentrations. This indicates that most pure polysaccharide of the hexuronic acid and neutral sugars composition of whelk GAG is elute at low salt concentrations according to their net charge at the selected pH. The above finding was confirmed by monosaccharide composition analysis of each fractions. In addition, monosaccharide quantification for all fractions (1M, 2M and 3M NaCl) enabled the calculation of the corresponding amounts of each enzyme required per substrate for complete enzyme digestion for further structure elucidation process.

The results in this chapter show for the first time the abundance of different disaccharide units presents in whelk GAGs along with their sulphation patterns differ significantly from mammalian GAGs, using the GRIL-LC/MS technique. The results are broadly consistent with some reports in the literature, such as the structural analysis of compound with reduced anti-coagulant activity isolated from the shrimp *L. vannamei* by Brito *et al.*, (2008), suggesting that some GAG isolated from marine organisms differ in their chemical composition than mammalian GAGs. The finding is in line with another study that shows the content of unique markers for the antithrombin-binding region of heparin isolated from the two marine clam species *Anomalocardia brasiliana* and *Tivela mactroides* are present in large amounts (up to 25-30%) of the total disaccharide units. This also demonstrated the occurrence of a novel saccharide sequence identified as (-GlcNSO<sub>3</sub>)-IdA-GlcNSO<sub>3</sub>(3,6-di-OSO<sub>3</sub>), that has not previously been found in heparin or related polysaccharides in mammalian GAGs.

This finding might explain, at least in part the incomplete digestion of the whelk GAGs by heparinase enzymes. Since each heparinase enzyme is known to act on specific substrate, changes in the amount of specific disaccharide, sulphation pattern and/or their configuration within the intact chain might result in resistance toward enzymes activity.

Reasonable amounts of GAGs detected from the retained fraction of the Nanosep centrifugal devices contain omega membrane (10K filter Eppendorf tube), which again could be the result of incomplete digestion by heparinase enzymes. This fraction has proven biological activity toward cancer unlike the flow through fractions, which demonstrated no activity toward cancer cells (further explained in chapter 5). This finding was unexpected and suggests that retained fragments contained the key heparinase resistant sequences that may be responsible for the biological activity and as such their structure is going to be important for future studies.

This study has a number of possible limitations, including the main factor, which is the degree of complexity of whelk GAG structure that will take an extensive study to identify biologically active sequences within the whelk GAG chains.

These findings contribute considerably to the evaluation and plans for further structure analysis of whelk GAG in the development of potential therapeutic approaches, which may have an important role as future anticancer and antimalarial agents.

Further studies are required to further analyse the structure of whelk GAGs in particular the bioactive compounds detected in the retained fractions. More research in this area is necessary to answer several questions that remained unresolved in terms of structure-function relationship.

# **Chapter 5 : The effect of whelk GAG on triple negative breast cancer and elucidation of their mechanisms of action**

## 5.1 Introduction

Recent concerns about triple negative breast cancer (TNBC) have generated a considerable body of research. Triple negative breast cancer is an aggressive disease subtype that has the poorest outcomes and accounts for as many as one in five cases. Compared to positive breast cancer, TNBC is a faster growing type of breast cancer. Treatment options for patients with TNBC are limited to the use of conventional medicine, including a combination of therapies such as surgery, radiotherapy and chemotherapy. Hormone-based therapies cannot be used to treat TNBC as these cells do not have receptors for oestrogen, progesterone or HER2 proteins. Chemotherapy or radiotherapy is more toxic to the body's immune system and in most cases creates resistive cancer cells. In addition, there is a high potential for reoccurrence of cancer and metastasis. Many studies have been conducted to expand our understanding of TNBC and to find alternative treatment options; as a consequence, many inhibitors are currently under development to treat TNBC (Cleator *et al.*, 2007).

Treatment directed against proliferative pathways such as cytokines and growth factors that activate receptors are promising therapeutic options for TNBC. Many compounds in these pathways are either overexpressed or mutated in cancer cells which provide potential therapeutic targets in the treatment of cancer (Adjei & Hidalgo, 2005).

Recently, there has been growing interest in developing potential drugs and biomaterials from marine organisms for their particular physical and chemical properties. Researchers have identified a wide variety of active compounds including GAGs, which have the potential to inhibit the growth of human tumour cells *in vitro* and *in vivo*.

Many marine polysaccharides are subjected to chemical modifications to yield derivatives that are degraded, semi-synthetic and smaller in size, which have been found to exhibit extensive and diverse biological action, such as treating and even preventing cancer.

In addition to their ability to induce programmed cell death (apoptosis) or anti-proliferation, as small molecules they have industrial value because they can increase drug design and drug discovery, since large fragments are unlikely to be good drug candidates. Their availability, lower toxicity and suitability for human application must also be taken into account.

A study by Saravanan & Shanmugam, (2010), characterised crude GAGs extracted from the bivalve mollusc *Amussium pleuronectus*, which were fractionated by ion exchange chromatography to a low molecular weight, resulting in derivative exposed anticoagulant activity. This study suggested that GAG from bivalve molluscs could be an alternative source of heparin.

Research has suggested that applying diverse extraction methods might yield different compounds that vary not only in their structural composition but also in their biological composition. For example, using two different methods for isolating brown seaweed *Adenocystis utricularis* at room temperature and 70 °C yielded two different compounds, each with totally different activities against herpes simplex virus 1 and 2 (Ponce *et al.*, 2003). Another study by Ye *et al.*, (2008), suggested that the biological activity of polysaccharide isolated from *Sargassum pallidum* was significantly increased by reducing the chain size and increasing the sulphation pattern.

Several studies have investigated the biological activity of marine GAG extracts. Researchers have identified GAGs with anticoagulant, antibacterial, antiviral and anticancer properties. However, these studies have not addressed the treatment of aggressive cancers such as TNBC. Furthermore, while a few attempts have been made to characterise active GAGs from marine animal bodies, whelk has received little attention

This chapter focuses on four objectives, the first of which is to examine the biological activity of crude whelk GAG extraction. Secondly, examine the biological activity of whelk GAG derivatives from different analytical techniques on two subtypes of TNBC, in order to

identify the most significant active fraction. The third objective is to evaluate the anti-proliferation activity of whelk GAGs on cancer stem cell. Finally, determine the mechanism of action of whelk GAG on TNBC cell lines using label-free proteomic approaches.

There are two types of proteomic, labelling approaches. The first is simply to label the peptides with isotopes; the second is to fluoresce labels then mix and analyse all the samples together in a single LC-MS/MS run. The peptides originating from different samples can then be distinguished in a mass spectrometer.

By contrast, label-free approaches simply compare the peptides originating from different samples in different LC-MS/MS runs. Recently this approach has been widely used, especially as the sensitivity and resolution of LC-MS/MS has increased and with advances in computational algorithms (Nahnsen *et al.*, 2013). Moreover, label-free quantitative proteomics is a powerful tool with which to estimate overall changes in protein abundances among a variety of different samples (Megger *et al.*, 2013).

One of the most effective tools for drug target identification is based on the comparison of protein expression between drug-treated and untreated tumour cells. This promising approach helps to characterise proteins enriched or isolated due to interacting or binding to bioactive molecules, causing what are called drug–protein interactions (Drewes, 2012). In pharmacology, proteomics is a promising tool for determining the specificity of drugs and for identifying new targets for interaction with known drugs or those that are sensitive to newly designed drugs. Proteomics also can provide direct information about side effects and toxicities to minimize failure in clinical trials (Sutton, 2012).

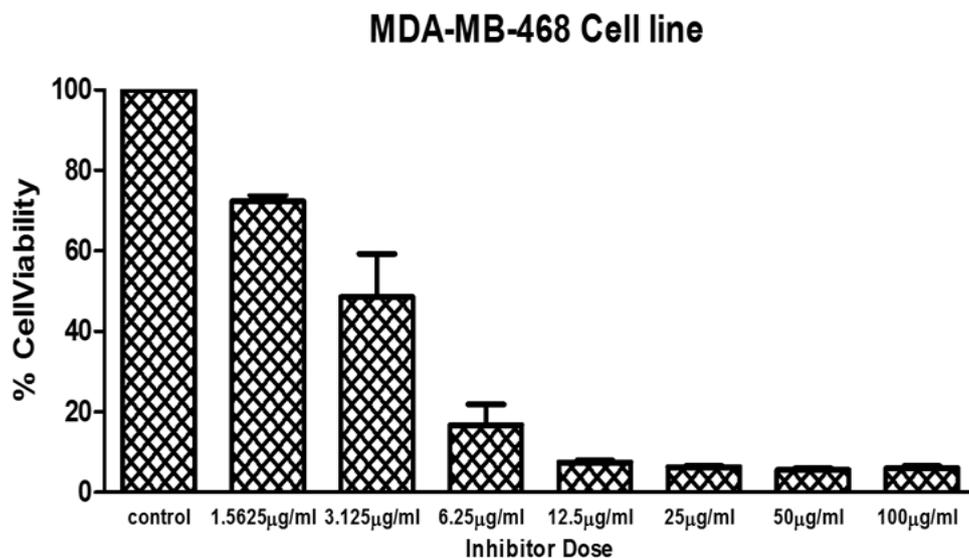
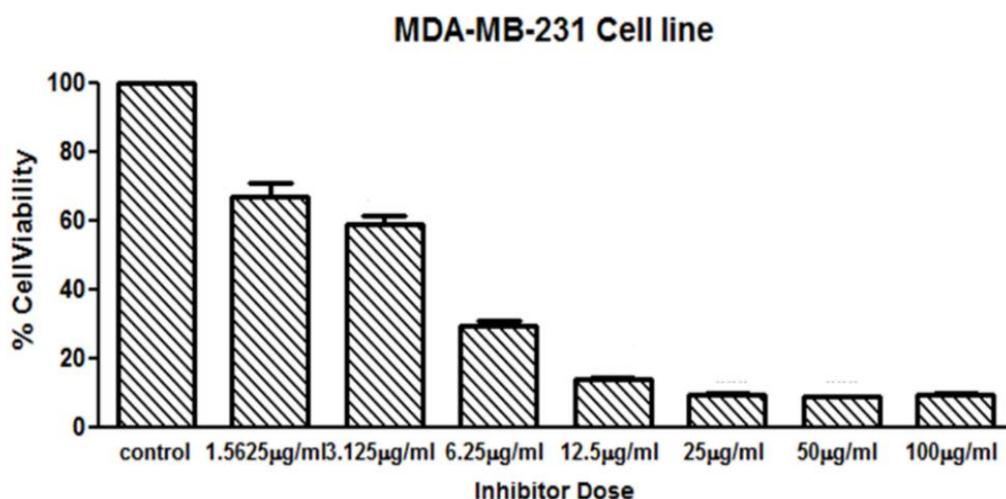
It is important to examine the capability of whelk GAGs to induce cell death among small populations of cells known as cancer stem cells (CSC), and this chapter addresses this aspect of the current study. Initial bio screening of the whelk GAG crude extraction was carried out on mammosphere generated from MDA-MB-468, MDA-MB-231 and MCF-7. Promising

results were reported from all cell lines, further suggesting the importance of whelk GAGs in treating cells with cancer and stem cell properties among TNBC cells.

## **5.2 Results**

### **5.2.1 Effect of intact GAG on triple negative breast cancer**

Both triple negative breast cancer cells (MDA-MB-468 and MDA-MB-231) were plated at a density of  $3 \times 10^3$  in 96 well plates. Cells were incubated for 24 hours to allow confluent monolayer formation then washed and withdrawn from the effects of trypsin. Cells were incubated in the presence and absence of whelk GAG crude extractions. Serial dilution of whelk GAG extraction was used to treat the cells then incubated for 5 days, each condition was in triplicate. Cell proliferation was evaluated by measuring the absorbance of MTT dye (purple formazan crystals) using a plate reader. The percentage of viable cells was represented as mean  $\pm$ SD, and IC50 was calculated for each cell line using GraphPad Prism 5.0 software. The effect of crude whelk GAG extracts on both cell lines is illustrated in Figure 5-1. In the presence of whelk GAG, cell proliferation significantly decreased in both cell lines as the dosage increased. IC50 values of  $3.48 \pm 0.02$   $\mu$ g/ml and  $2.76 \pm 0.01$   $\mu$ g/ml were observed from MDA-MB-231 and MDA-MB-468 respectively.

**A****B**

**Figure 5-1: Inhibition activity of intact whelk GAG sample on negative breast cancer cells.**

**Panel A:** Cell proliferation inhibition of whelk GAG on MDA-MB-468. **Panel B:** Cell proliferation inhibition of whelk GAG on MDA-MB-231. Data represent the percentage of viable cells for sample size (n=3) as mean  $\pm$  SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay were performed using an MTT colorimetric assay. The IC<sub>50</sub> values were calculated using nonlinear regression analysis (GraphPad Prism 5.0).

### **5.2.2 Effect of HS/CS-disaccharide observed from digestion of whelk GAG and commercial porcine HS/CS on triple negative breast cancer**

To compare the bioactivity of whelk GAG extracts and commercial porcine HS/CS on triple negative breast cancer, both cell lines MDA-MB-468 and MDA-MB-231 were cultured in a monolayer in 96 well plates and dosed with a serial dilution of all compounds over a 5-day incubation at 37 °C and 5% CO<sub>2</sub>, and each condition was in triplicate. In addition, whelk GAG extraction and commercial porcine HS/CD were subjected to enzyme depolymerisation using a mixture of Hep I, II, III and ABC lyase. The results from this experiment demonstrate the effect of whelk GAG extraction and the whelk GAG fragments library generated by Hep I, II, III/ABC lyase depolymerisations on both cell lines. This experiment has been repeated for commercial porcine HS/CS and its depolymerisation derivatives on TNBC.

The results clearly suggest significant activity of crude extraction and fractions yielded from enzymatic digestion of Hep I, II, III and a combination of all Hep lyase. In addition, the results from the fragments generated by ABC lyase showed a dramatic anti-proliferation activity. However, the anti-proliferation activity reported from Hep depolymerisation fragments is slightly different from the anti-proliferation activity shown with fragments generated from ABC lyase. A noticeable resistance has been shown from TNBC cells toward whelk GAG fragments generated by Hep lyases, whereas a similar inhibition profile is observed from whelk GAG crude extraction before and after ABC digestion (Table 5-1).

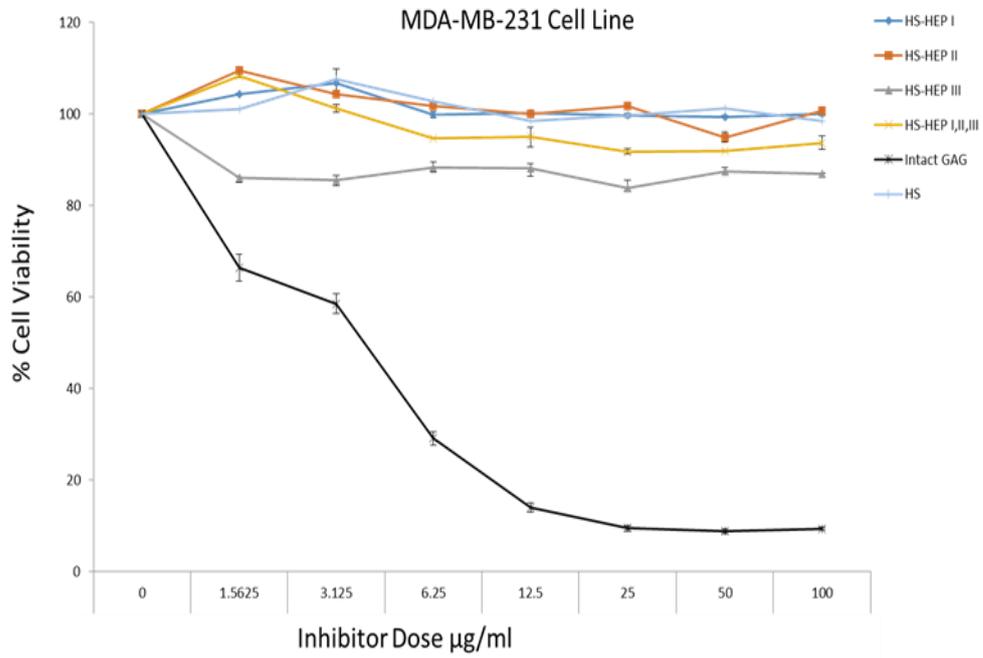
**Table 5-1: IC50 values of whelk GAG fragments libraries from Hep lyase and ABC lyase on TNBC cell lines.**

Cell line	IC50 µg/ml Crude whelk GAG	IC50 µg/ml whelk GAG + Hep I	IC50 µg/ml whelk GAG + Hep II	IC50 µg/ml whelk GAG + Hep III	IC50 µg/ml whelk GAG + Hep I, II, III	IC50 µg/ml whelk GAG +
1 MDA-MB-468	2.76	2.61	42.32	37.89	33.15	3.03
2 MDA-MB-231	3.48	13.97	24.92	26.03	27.61	4.60

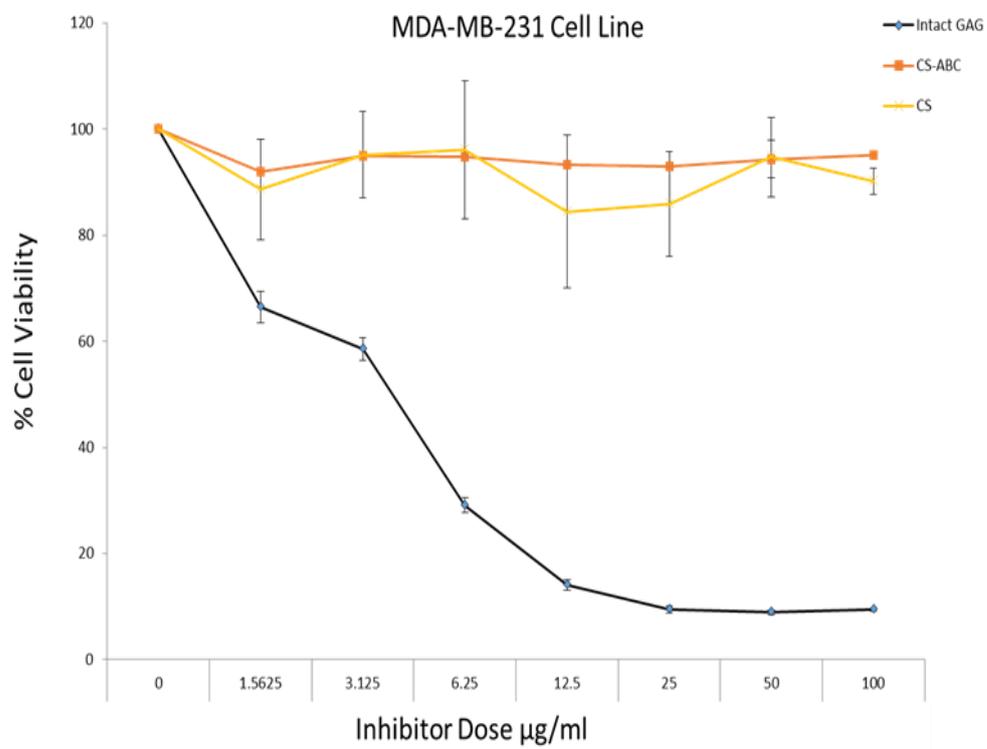
Commercial porcine HS/CS have no inhibition activity on the MDA-MB-468 and MDA-MB-231 cell line either before or after enzymatic degradation compared to the activity of whelk GAG crude extraction.

The results shown in Figure 5-2 indicates the inhibition activity of whelk GAG crude extraction on the MDA-MB-468 and MDA-MB-231 cell line, in comparison to mammalian GAGs and their fragments generated by Hep lyase and ABC lyase degradation. This leads to conclude that mammalian GAG has no inhibition activity on TNBC cell lines, as compared to whelk GAG extraction and its derivatives from enzymatic degradation.

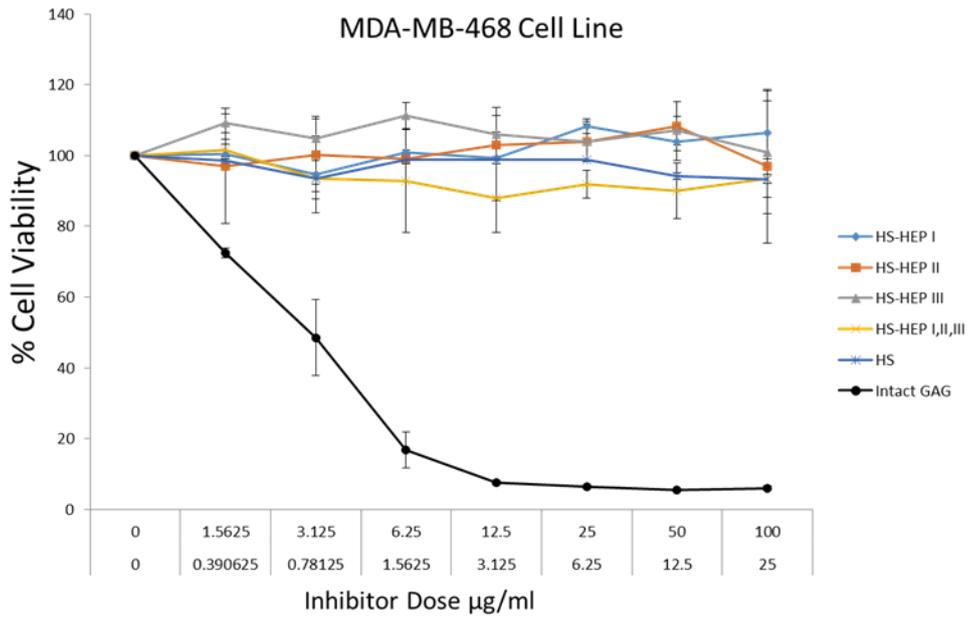
A1



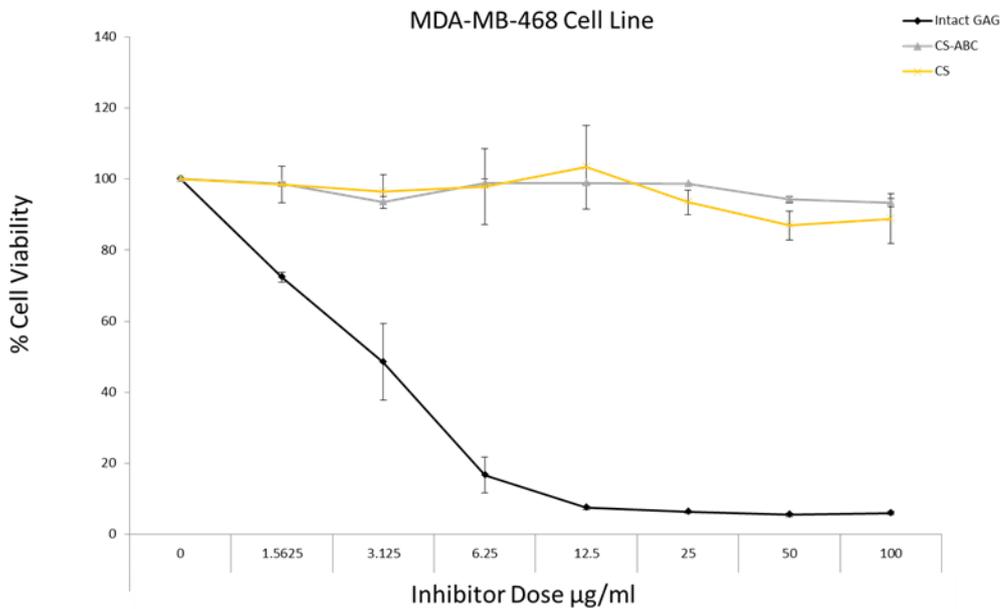
B1



A2



B2



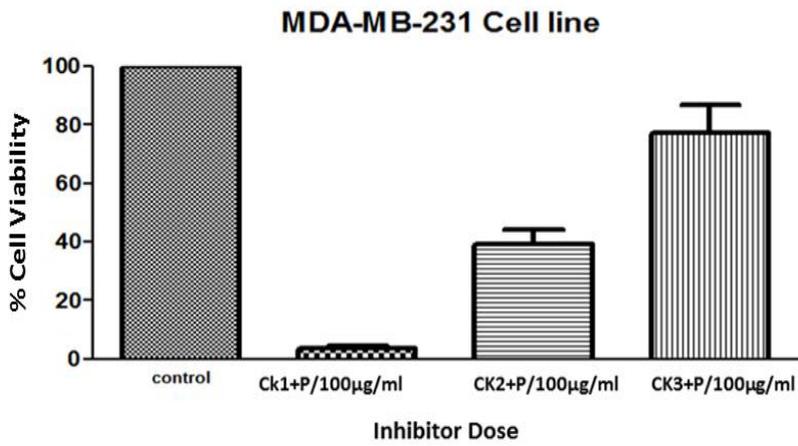
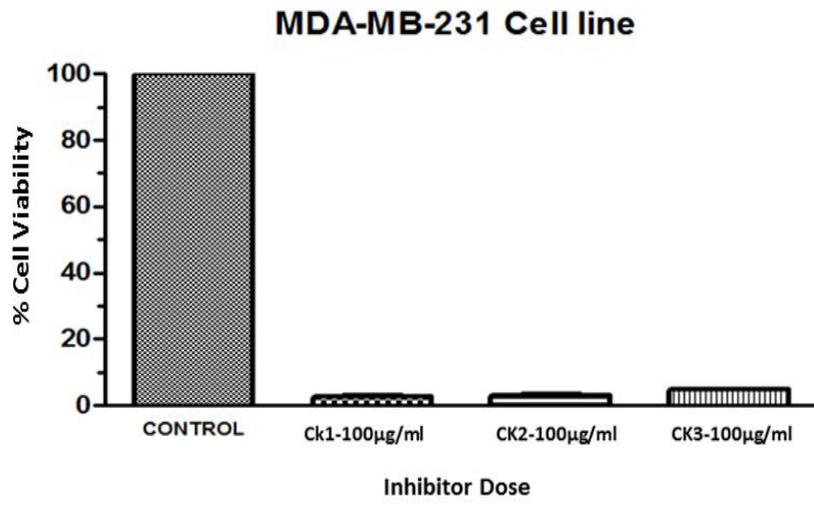
**Figure 5-2: Anti-proliferation activity of crude whelk GAG extracts compared to the activity of commercial mammalian HS/CS before and after enzymatic depolymerisation.**

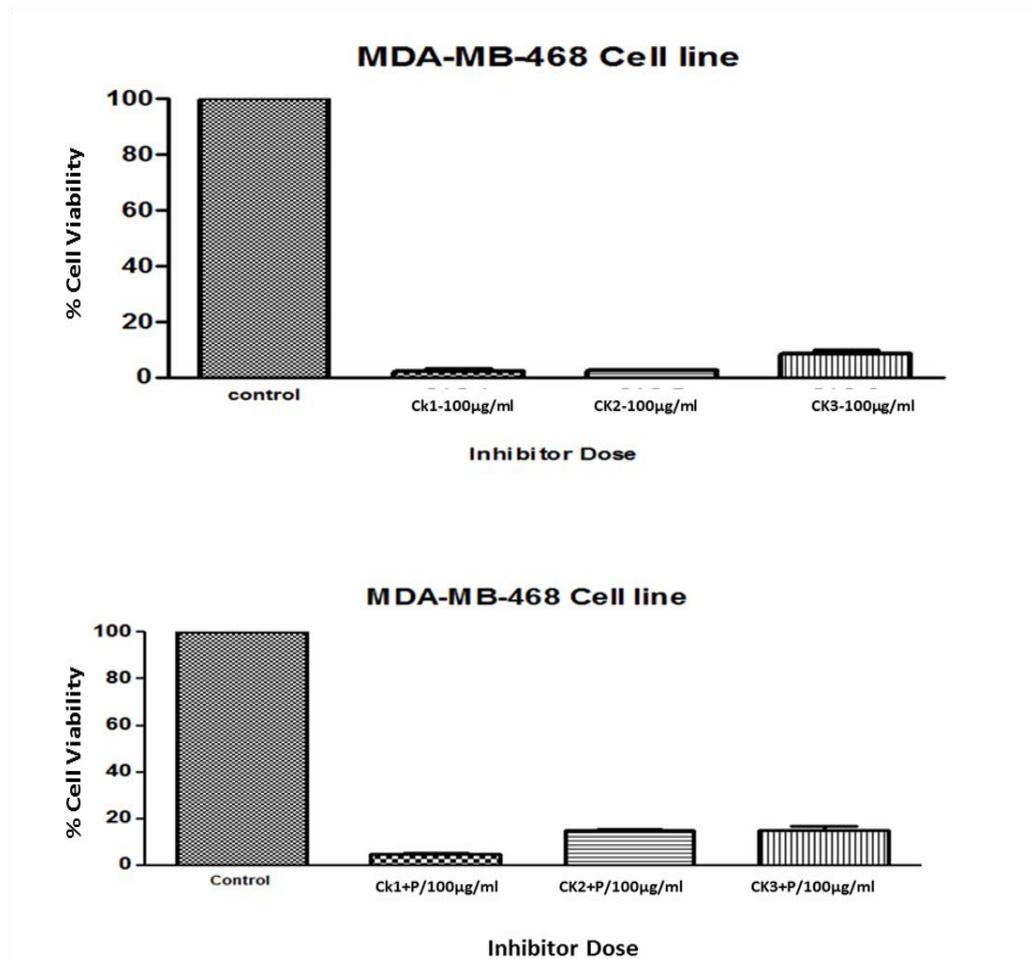
**Panel A1** and **B1**: Shown the effect toward MDA-MB-231 cell line. **Panel A2** and **B2**: Shown the effect toward MDA-MB-468 cell line. The data represent the percentage of viable cells for sample size (n=3) as mean ± SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay were performed using an MTT colorimetric assay.

### **5.2.3 Cytotoxic effect of ion-exchange chromatography fractions of whelk GAG before and after oxidation with periodic acid on triple negative breast cancer cell lines**

Whelk GAG crude extraction was fractionated by ion exchange chromatography. The sample was eluted in a salt (ionic strength) gradient in order to group the GAG according to charge. As described in detail in Chapter 4, three fractions were obtained and examined for their activity on TNBC cells, MDA-MB-231 and MDA-MB-468. The results were compared with the activity obtained from GAG sample fractions that had been oxidised with periodic acid. The results shown in Figure 5-3 compare the percentage of viable cells ( $\pm$ SD) for each cell line after dosing the cells with 100  $\mu$ g/ml fractions of CK1, CK2 and CK3 in both conditions (oxidised and non-oxidised with periodic acid) over a 5-day incubation period, with each condition performed in triplicate. The results shown in Panel A indicate highly significant inhibition activity of whelk GAG fragments CK1, CK2 and CK3 toward the MDA-MB-231 cell line compared with untreated cells (negative control). However, the activity of these fractions on the same cell line varies after oxidation carried out by periodic acid. Treating the cell lines with an equal amount of whelk GAG fragments that is oxidised by periodic acid over the same period of time results in a high level of significant activity of fragments CK1 and CK2, whereas fraction CK3 was found to be less significant. In addition, the percentage of viable cells obtained was higher from CK3 than from other fractions. This experiment was repeated with the MDA-MB-468 cell line. As shown in Figure 5-3 (Panel B), the activity of whelk GAG fractions oxidised with periodic acid remained similar to that of the non-oxidised GAG, with a significant level of activity being obtained from both conditions toward this cell line. The percentage of viable cells was significantly lower than that of corresponding untreated cells.

**A**



**B**

**Figure 5-3: Inhibition activity of whelk GAG fractions (CK1, CK2, and CK3) from anion-exchange comparing with the biological activity of fractions oxidised with periodic acid on triple negative breast cancer cells.**

**Panel A** shows highly significant activity of whelk GAG fractions CK1, CK2, and CK3 on MDA-MB-231 breast cancer cells from all fragments. This compares to the data shown in the second graph, which illustrates the activity of whelk GAG fragments CK1, CK2, and CK3 treated with periodic acid oxidation (P), which shows a slightly different effect from each fragment CK2, CK2 and CK3.

**Panel B** shows high significant activity of whelk GAG fragments CK1, CK2 and CK3 treated and untreated with periodic acid oxidation on MDA-MB-468 breast cancer cells. Data represent the percentage of viable cells for sample size (n=3) as mean  $\pm$  SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay were performed using an MTT colorimetric assay.

#### **5.2.4 Cytotoxic activity of whelk GAG fractions CK1, CK2 and CK3 before and after enzymatic degradation on triple negative breast cancer cell lines**

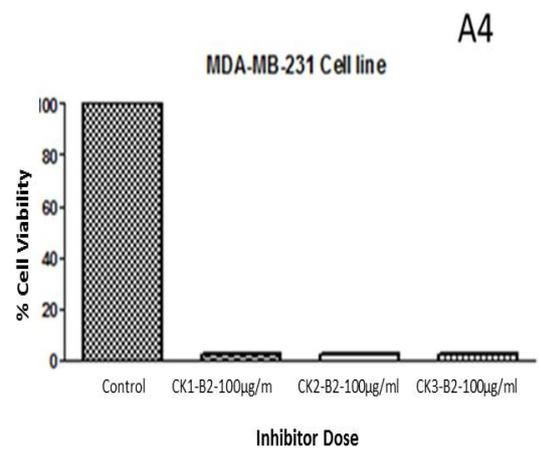
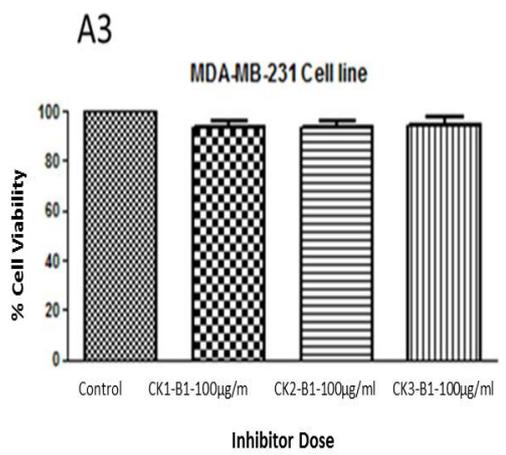
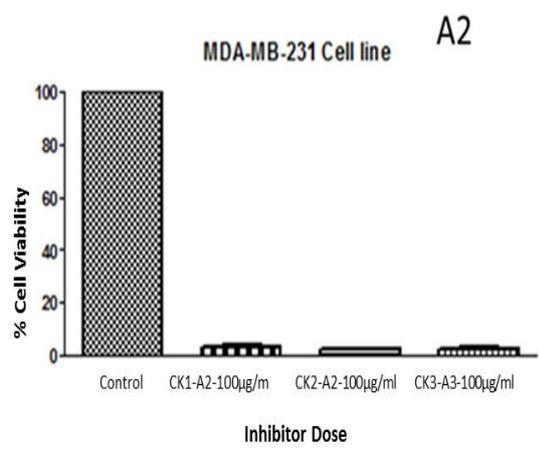
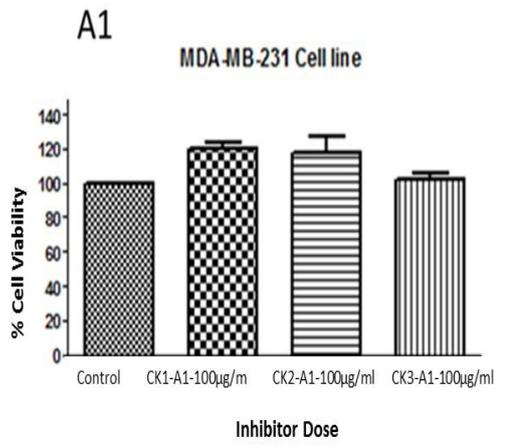
Whelk GAG crude extraction was analysed using a variety of analytical techniques in order to identify the structure of the GAG in relation to the active part of the polymer responsible for their activity toward many cancer cell lines, including triple negative breast cancer. The inhibition activity remained for the retained fraction (R) which did not pass through the centrifuge filtration device following enzymatic degradation of all fragments obtained from whelk GAG by ion exchange chromatography.

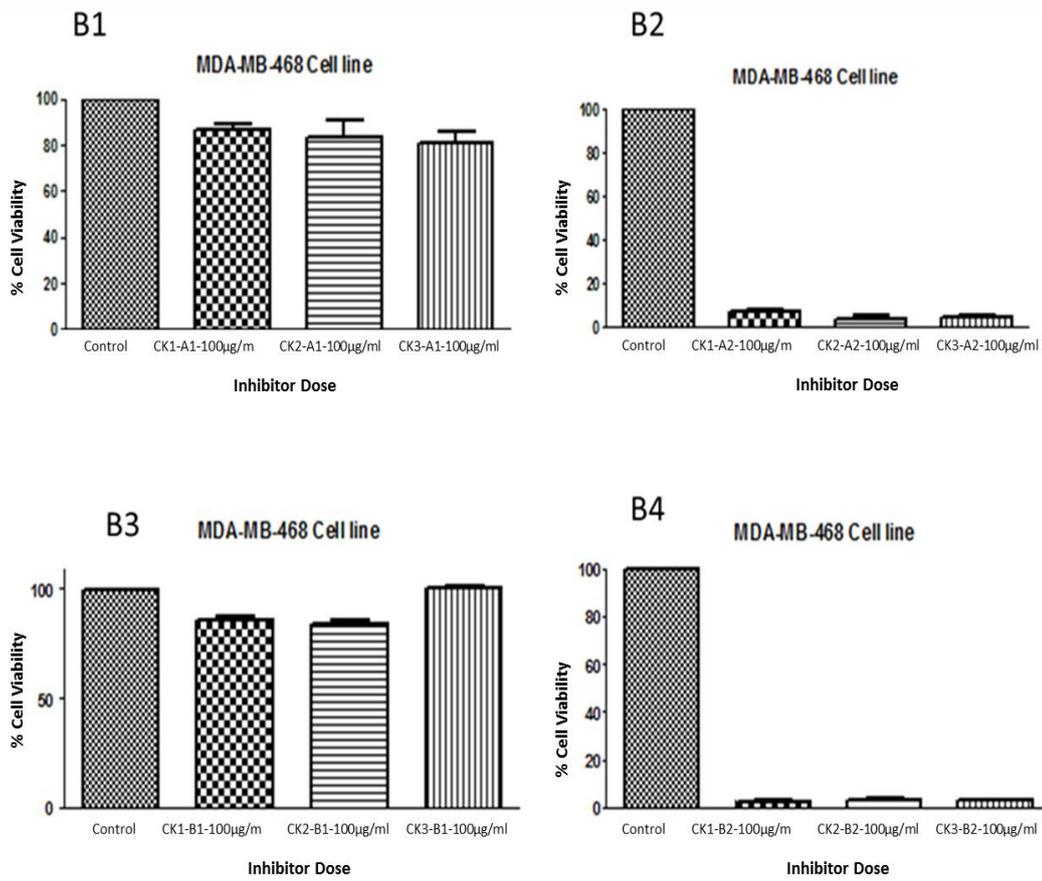
The aim of this experiment was to determine the activity of whelk GAGs that has been fractionated by ion-exchange chromatography and de-polymerised by enzymes. As described in detail in chapter 4, these compounds are passed through a 10K filter (Nanosep centrifugal devices with omega membrane). The monosaccharide and disaccharide composition of both the flow-through (FT) and the retained fractions (R) was then thoroughly analysed. After determining the molecules involved in the structural backbone of these two fractions (FT and R), the dried material was examined for its biological activity on the MDA-MB-231 and MDA-MB-468 breast cancer cell lines.

An equal amount (100 µg/ml) of each fraction of FT and R of all fragments (CK1, CK2 and CK3) were incubated with both cell lines for five days at 37 °C and 5% CO<sub>2</sub>. The cells were cultured in monolayer in 96 well plates and incubated for 24 hours, then the medium was changed and a fresh medium was added to all wells then dosed with 100 µg/ml of each fraction (FT and R) in triplicate. The cytotoxic activity was obtained using an MTT assay. The data were validated using GraphPad Prism 5.0 software. The results shown in Figure 5-4 demonstrate the activity of whelk GAG which has been depolymerised by Hep lyase or ABC lyases and which has passed through the filter. This flow through (FT) is more likely to contain di or oligosaccharide that is smaller in size than the pore size of the filter and

therefore is capable of passing through the device. The other fraction, which is often larger in size than the filter and unable to pass through it, is called the retained fraction (R). The results shown in panel A1 and panel B1 indicate no activity of the FT fractions from Hep lyase on either the MDA-MB-231 or MDA-MB-468 cell line. Moreover, no inhibition activity was seen for the FT fraction of ABC lyase fragments against either cell line, as shown in Panels A3 and B3. On the other hand, highly significant inhibition activity was observed from the retained fractions (R) of Hep lyase and ABC lyase fragments on both cell lines.

**A**



**B**

**Figure 5-4: Anti-proliferation activity of whelk GAG fractions CK1, CK2 and CK3 before and after enzymatic degradation on triple negative breast cancer.**

**Panel A:** Inhibition activity of FT and R from Hep lyase and ABC lyase fragments on the MDA-MB-231 cell line. **A1** shows the activity of FT fraction from a 10K filter of Hep lyase degradation of whelk GAG. **A2** shows the activity of R fraction from a 10K filter of Hep lyase degradation of whelk GAG. **A3** shows the activity of FT fraction from a 10K filter of ABC lyase degradation of whelk GAG. **A4** shows the activity of R fraction from a 10K filter of ABC lyase degradation of whelk GAG for each fragments CK1, CK2 and CK3 from anion exchange chromatography.

**Panel B:** Inhibition activity of FT and R from Hep lyase and ABC lyase on the MDA-MB-468 cell line. **B1** shows the inhibition activity of FT fraction from a 10K filter of Hep lyase degradation of whelk GAG. **B2** shows the inhibition activity of R fraction from a 10K filter of Hep lyase degradation of whelk GAG. **B3** shows the inhibition activity of FT fraction from a 10K filter of ABC lyase degradation of whelk GAG. **B4** shows the inhibition activity of R fraction from a 10K filter of ABC lyase degradation of whelk GAG for each fragments CK1, CK2 and CK3 from anion exchange chromatography. Data represent the percentage of viable cells for sample size (n=3) as mean  $\pm$  SD. Cells were cultured in monolayers and maintained at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. Cytotoxicity assay were performed using an MTT colorimetric assay.

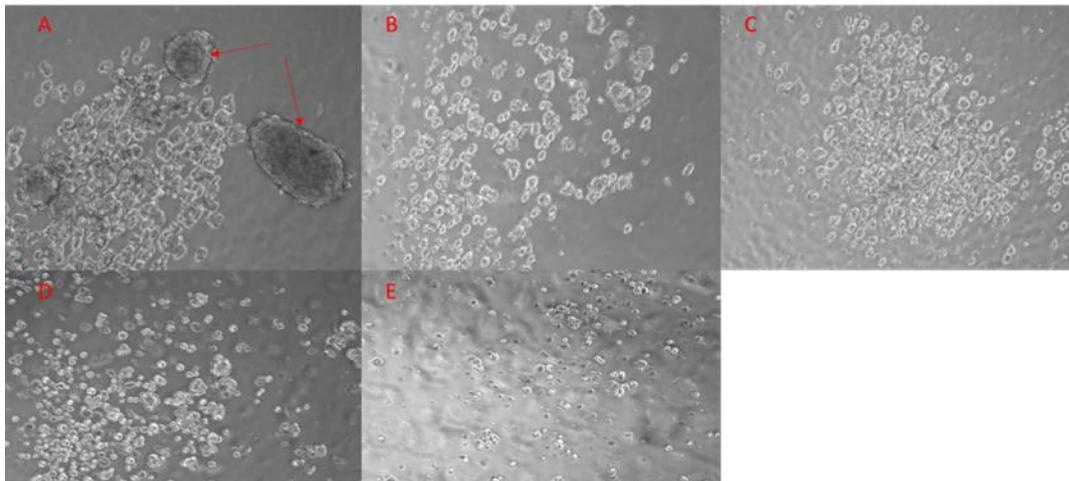
### **5.2.5 Cytotoxic effect of crude whelk GAG extracts on mammosphere formation in breast cancer**

It is critical to examine the capability of the new therapeutic candidate to induce cell death even among small populations of cells that has self-renewable properties. Herein, we examined the cytotoxic effect of crude whelk GAG extraction on mammosphere formation in three breast cancer cell lines'. Culture methods and materials are described in detail in chapter 2 (section 2.15).

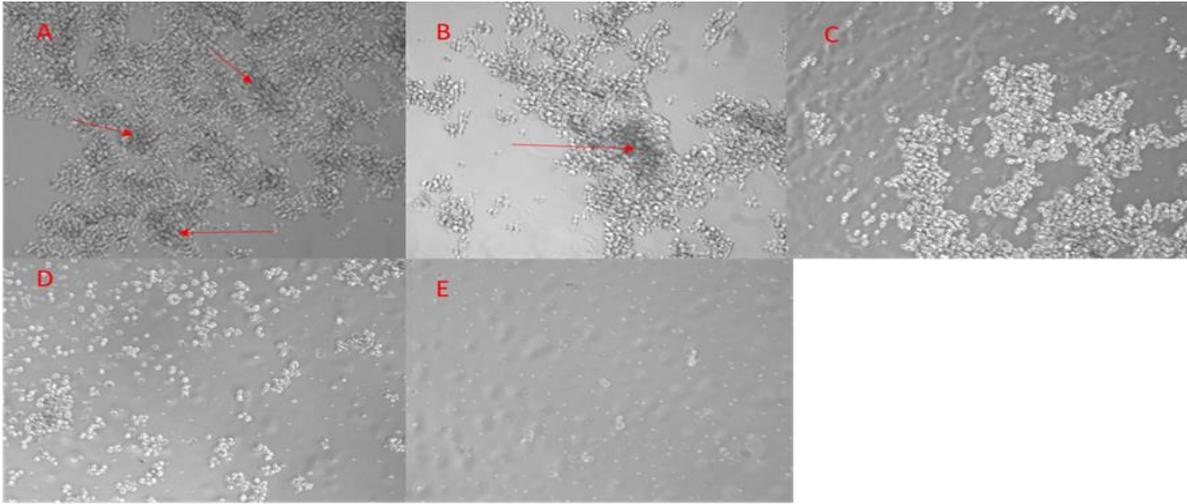
This experiment aimed first to determine the ability of different subtypes of breast cancer to induce mammosphere formation and then to highlight the potential inhibition activity of whelk GAG extraction on the tumourspheres formed in each cell line. Cells were plated in culture plates coated with poly-HEMA in triplicate and dosed with ideal IC<sub>50</sub> values of whelk GAG crude extraction corresponding to each cell line, in serial concentrations of 20 µg, 50 µg, and 100 µg/ ml. Plates were incubated for 5 days at 37 °C and 5 % CO<sub>2</sub>. In designing this experiment with different concentrations of whelk GAG alone in addition to the IC<sub>50</sub> values corresponding to each cell line, the aim was to compare the cytotoxic effect from each concentration on the mammosphere formation as these populations of cells usually are known for their resistance, which raises the possibility of inducing inhibition among them using different concentrations from the IC<sub>50</sub> values of whelk GAGs used for primary breast cancer cells.

The results obtained from culturing mammospheres clearly indicate that the capability of MCF-7 to form mammosphere is greater than that of either MDA-MB-231 or MDA-MB-468. Figure 5-5 (Panel A) illustrates the mammosphere formation in the MCF-7 cell line. Image A depicts the untreated mammosphere; images B, C, D and E show clusters of cells clumped together with no obvious signs of mammosphere formation. However, TNBC cell lines survived the non-adherent culture conditions and were able to grow a high number of cells in the suspension form, which confirms the ability of all subtypes of breast cancer cells to

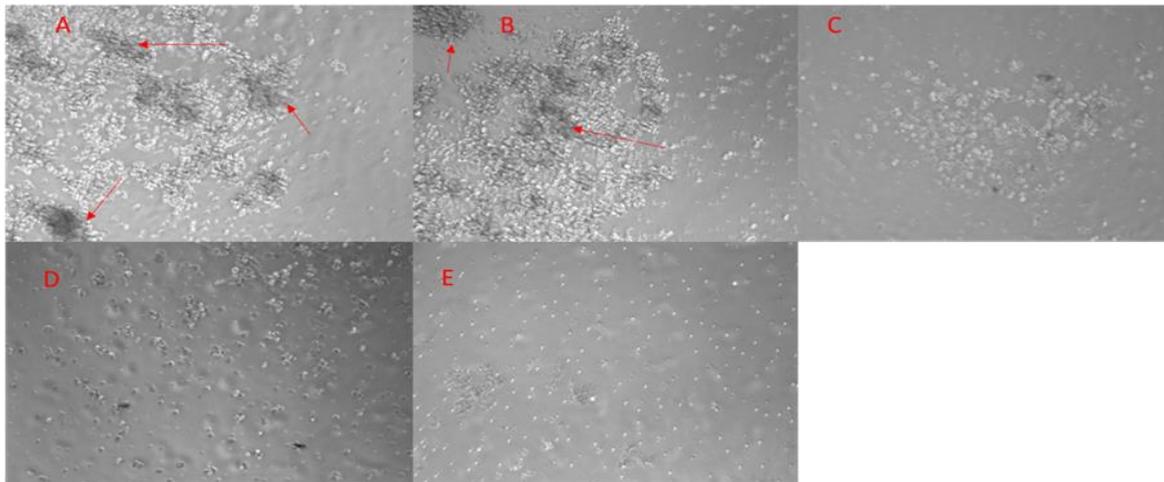
develop malignancy and the high risk of metastasis. The mammosphere and clumped cells decreased when the cells treated with the IC50 values corresponded to each cell line. The images show that increasing the dose from 20  $\mu\text{g/ml}$  to 50  $\mu\text{g/ml}$  and 100  $\mu\text{g/ml}$  induced significant cell death. Given that CSC possess the cell-surface phenotype CD44<sup>+</sup>/CD24<sup>-</sup>/low cells, further investigation is necessary into the CD44<sup>+</sup>/CD24<sup>-</sup>/low cell subpopulation on the mammosphere formation in MCF-7, MDA-MB-231 and MDA-MB-468 using fluorescence-activated cell sorting (FACS).



A: MCF-7 Cell line



B: MDA-MB-468 Cell line



C: MDA-MB-231 Cell line

**Figure 5-5: Mammosphere formation in three breast cancer cell lines.**

**Panel A:** Mammosphere formation in the MCF-7 cell line. Image A (untreated) shows clear mammosphere formation; image B depicts cells treated with the desirable IC50 value; images C, D and E show the results of treatment for 5 days with 20 µg/ml, 50 µg/ml and 100 µg/ml GAG respectively. **Panel B:** Mammosphere formation in the MDA-MB-468 cell line. Image A (untreated) shows clumped cell formation; image B shows the results of treatment with desirable IC50 value; images C, D, and E depict cells treated for 5 days with 20 µg/ml, 50 µg/ml and 100 µg/ml GAG respectively. **Panel C:** Mammosphere formation in the MDA-MB-231 cell line. Cells survived and grew in suspension condition (image A, untreated); image B shows the result of treatment with desirable IC50 value; images C, D, and E depict cells treated for 5 days with 20 µg/ml, 50 µg/ml and 100 µg/ml GAG respectively.

## **5.2.6 Label free protein quantification**

### **5.2.6.1 Quantitative proteomics analysis**

Quantitative proteomics is one of the most powerful methods for identifying the regulation or dysregulation of biological systems that can result from changes in protein levels, changes in subcellular localization or changes in post-translational modifications (PTMs), which cause changes in protein-protein interactions. There are two types of quantitative analysis: relative quantification and absolute quantification. Relative quantification identifies changes in the expression level of protein between samples conditions, while absolute quantification determines the number of copies of the protein per cell and measures the expression level of different proteins within a single sample or between different samples. This approach is used in the pharmaceutical and biomedical industries for quality control, therapeutic analysis and drug testing purposes.

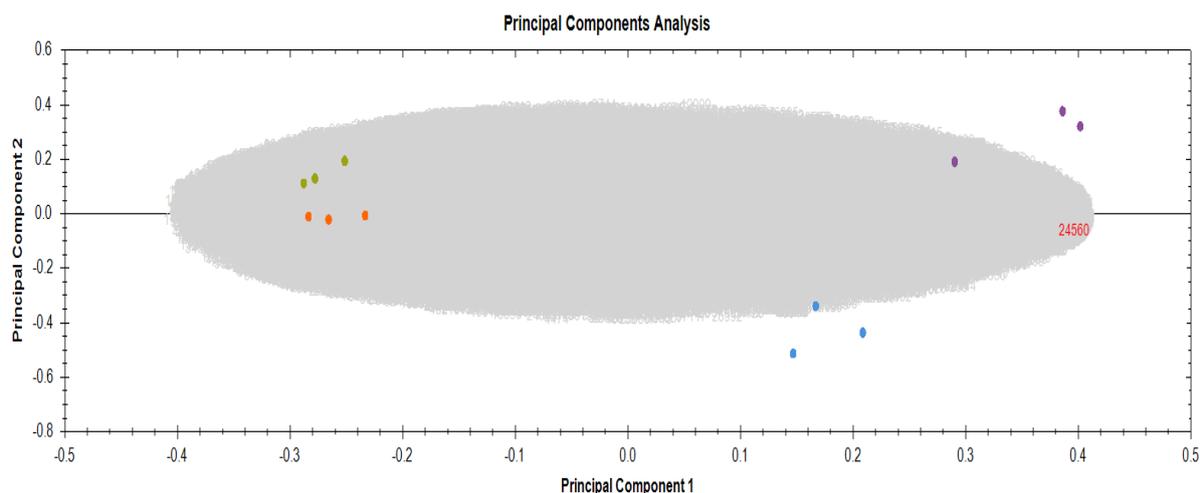
Both relative and absolute quantification can be performed using one of two methods, labelling or label-free. In the labelling approach, a sample is labelled with either a metabolic label such as SILAC or  $^{15}\text{N}$  while in culture, or a chemical label such as like iTRAQ, ICAT, TMT and dimethyl after extraction and digestion of the protein from the cells or tissue. In the label-free approach, protein quantification usually is done by counting the spectra or ion intensities of individual samples run separately and the expression levels or protein abundance is calculated by comparison between different samples runs (Zhu *et al.*, 2009).

Here, the complex peptide mixture from triple negative breast cancer (TNBC) replicates were analysed. Digested samples were analysed by LC-MS/MS using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a QE HF (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. The complex spectra were further analysed using PROGENESIS Qi software, which performed the peak identification, followed by protein identification and protein quantification by searching with the web-based UniProt mapping tool. Two different conditions of the total proteome of breast tumours was

analysed in each cell line: untreated cells (control) and cells treated with whelk GAG crude extraction over 24 hours.

The cut-off parameters used were p-value < 0.05 for determining the significant difference, and maximum fold  $\geq 1.5$  of the increase/decrease in protein levels. Analysis of the MS data identified 846 distinct proteins differentially expressed between treated and untreated samples in the MDA-MB-468 cell line. A total of 553 proteins were identified and were differentially expressed between treated and untreated sample in the MDA-MB-231 cell line.

Normalised abundance is calculated as the sum of the normalised area under the curve for all the quantitative peptides for each sample, whereas raw abundance is the sum of the raw (real) area under the curve for all the quantitative peptides for each sample. A comparison of all data is shown in Figure 5-6. Principal component analysis (PCA), which was performed to obtain a simplified data visualisation, shows that sample groups cluster reasonably well and that the greatest variation in the data is attributable to the cell line when followed by the treatment. The data clusters for each individual cell line, which are also shown, indicate that the greatest change is most likely caused by the treatment applied to the cells. In these cases, the p-value obtained from statistical analysis of the samples most likely accounts for the differences between the control and treated cells and allows a reasonable interpretation of the data.



**Figure 5-6: Principal Component Analysis PCA of Triple Negative Breast Cancer TNBC cells proteome at protein level.**

PCA of both cell lines shown in colour variation as follows: control (blue) and Whelk GAG treated (purple) in MDA-MB-468, and control (orange) and whelk GAG treated (green) in MDA-MB-231, all at protein levels.

### 5.2.6.2 Mapping identified proteins to functional clusters in two triple negative breast cancer cell lines

Gene ontology analysis of all the proteins identified from two cell lines was obtained by Protein ANalysis THrough Evolutionary Relationships (PANTHER version 11.1) (Mi *et al.*, 2013). Firstly, existing ANOVA p-values were extracted from the data from the two label-free proteomic analyses and imported into R. Benjamini-Hochberg, p-values then were derived from the existing p-values for each analysis using the ‘p.adjust’ method and volcano plot images were obtained using R (Figure 5-7). Protein groups which were under the 5% significance threshold were selected for further analysis.

The functional classification was summarised using PANTHER biological process (BP) gene ontologies, molecular function (MF) gene ontology and cellular component (CC) gene ontology, using the list of all identified proteins from each cell line as data input. Proteins were classified into several categories according to the protein database available in PANTHER.

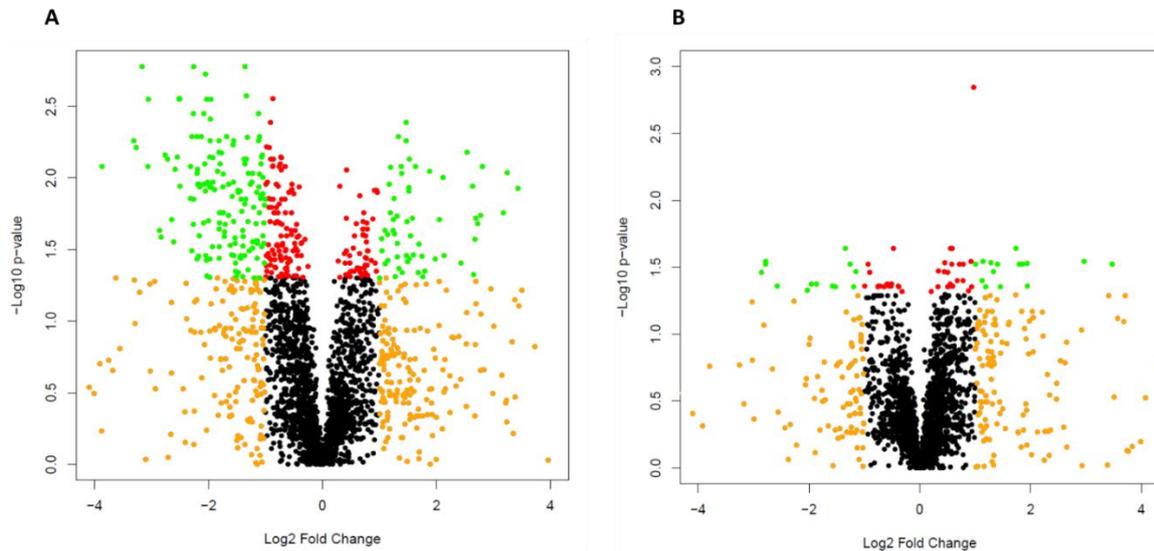
Gene ontology (GO) molecular function categories according to the PANTHER database in both TNBC cell lines (Figures 5-8 (A) and 5-9 (A)) showed that the majority of the identified proteins are categorised being catalytically active and protein having binding properties. The percentage of each functionally overrepresented cluster varies between cell lines. For instance, metabolic activity represents the majority of molecular function in MDA-MB-468 for GTPase activity and pyrophosphatase activity (39%), followed by protein binding (37%) and structural molecular activity (10%), while the majority of molecular function in MDA-MB-231 is categorised as protein binding (47%), metabolic activity (21%) and structural molecular activity (15%). The majority of protein binding is related to cytoskeletal protein binding, for example actin binding and microtubule binding, followed by receptor binding such as growth factor activity and hormone activity.

The remaining identified proteins in both cell lines contributed to transporter activity, signal transduction activity and receptor activity. A closer look at the BP ontologies of these proteins (Figure 5-8 (B) and 5-9 (B)) showed 29% and 28% of the identified proteins were related to the cellular process in MDA-MB-468 and MDA-MB-231 respectively, while 27% in MDA-MB-468 and 18% in MDA-MB-231 related to the metabolic process.

The distribution of protein localisation in MDA-MB-468 (Figure 5-8 (C)) showed that the majority of the identified platelet proteins were located in parts of the cell, including cytoplasmic proteins (43%), organelles (32%) and the macromolecular complexes (16%). The remaining proteins contributed to the membrane (7%) and extracellular region (1%) respectively. The distribution of protein localisation in MDA-MB-231 (Figure 5-9 (C)) was restricted to the parts of the cell (33%) and cell organelle (29%); the remainder of the identified proteins were mapped to the macromolecular complex (16%), membrane (8%), extracellular region (8%) and extracellular matrix (6%).

Furthermore, gene functional annotation in a network context was performed using online tools including Database for Annotation, Visualization and Integrated Discovery (DAVID). This tool has developed functions to cluster related GO terms or genes, an approach that allows the enrichment analysis to progress from gene-centric to biological module-centric analysis. These methods take into account the networked nature of biological annotation content in order to concentrate on building the larger biological picture from gene group-based analysis rather than focusing on an individual term or gene. Like other methods, this method has obvious limitations, as genes that have a weak relationship with other genes or terms will not be included in the analysis.

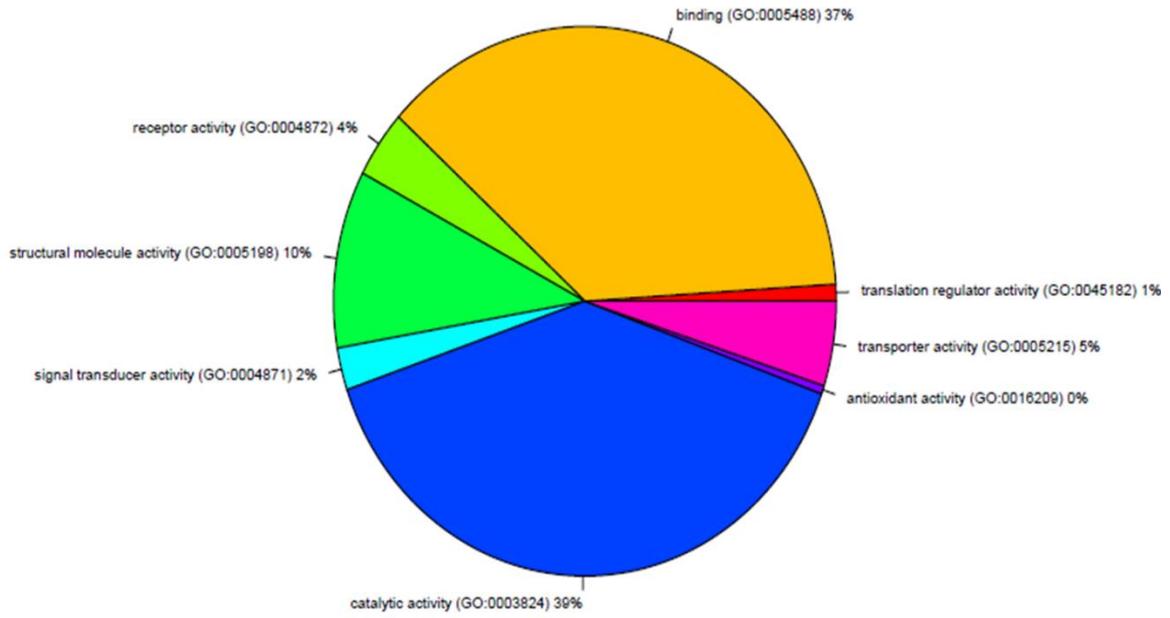
Proteomic data were analysed to generate biological functional clusters, which were verified by the significant enrichment of the online tool, DAVID v6.8 (<https://david.ncifcrf.gov>). For both MDA-MB-468 and MDA-MB-231, experiments with proteins with adjusted p-values of less than 0.05 generated functional annotation clustering. The ontologies were ranked by the significance of enrichment calculated using the DAVID online tool. A false discovery rate (FDR) corrected p-value was calculated for each protein for multiple hypothesis comparisons by the Benjamini-Hochberg method (Benjamini & Hochberg, 1995). Briefly, this uses a strategy to identify groups of genes with high connectivity; these groups contain genes defined by similar functional properties and can be considered as functional clusters. Lists of proteins with FDR-corrected p-values < 0.05 were generated.



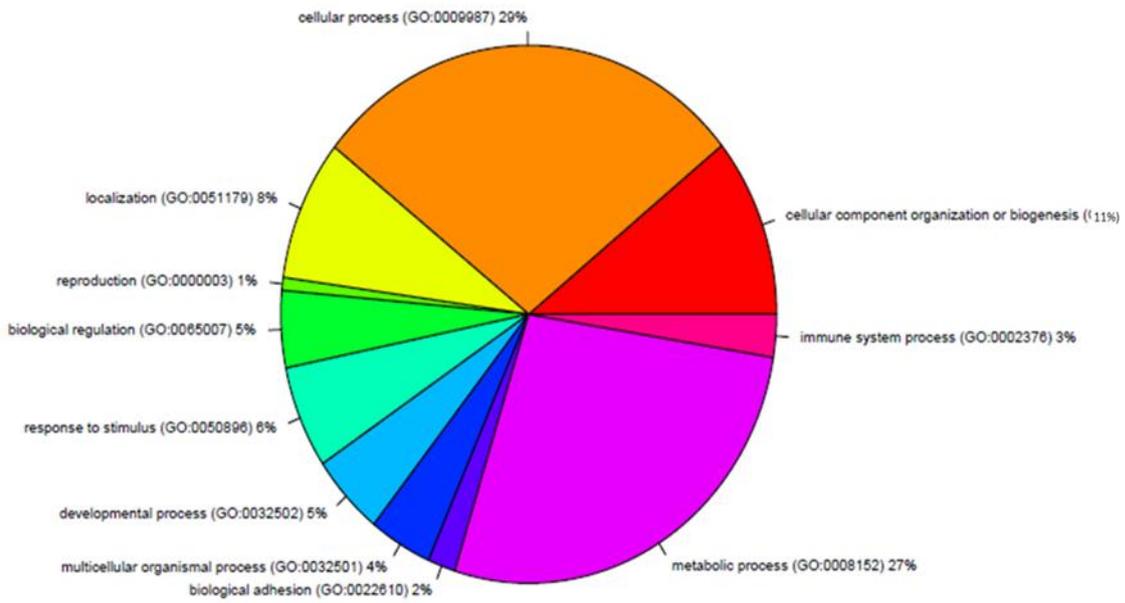
**Figure 5-7: Volcano plot images were produced using R, for each analysis, for corrected p-values.**

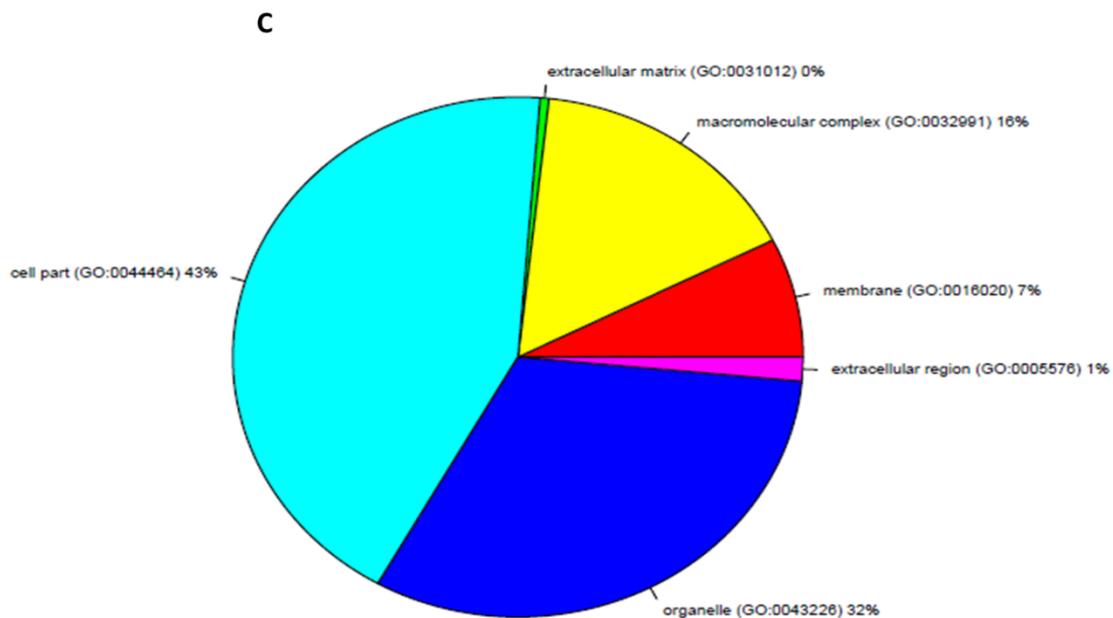
**Panel A** presents data for wheik GAG treated vs control samples in the MDA-MB-468 cell line. **Panel B** presents data for treated vs control samples in the MDA-MB-231 cell line. High values on the y-axis of a volcano plot indicate low p-values, so higher  $-\log_{10}$  (p-values) are the most significant. The x-axis (fold change) is  $\log_2$ -scaled, so a value of 0 indicates no 'fold change'. The fold change corresponds to the ratio of C / T (control / treated) for each protein (group). Points are coloured as follows: absolute fold changes larger than 1 is orange, a p-value less than 0.05 is red, if the p-values is less than 0.05 and the absolute fold change is larger than 1 the point is green. All other points are black

**A**



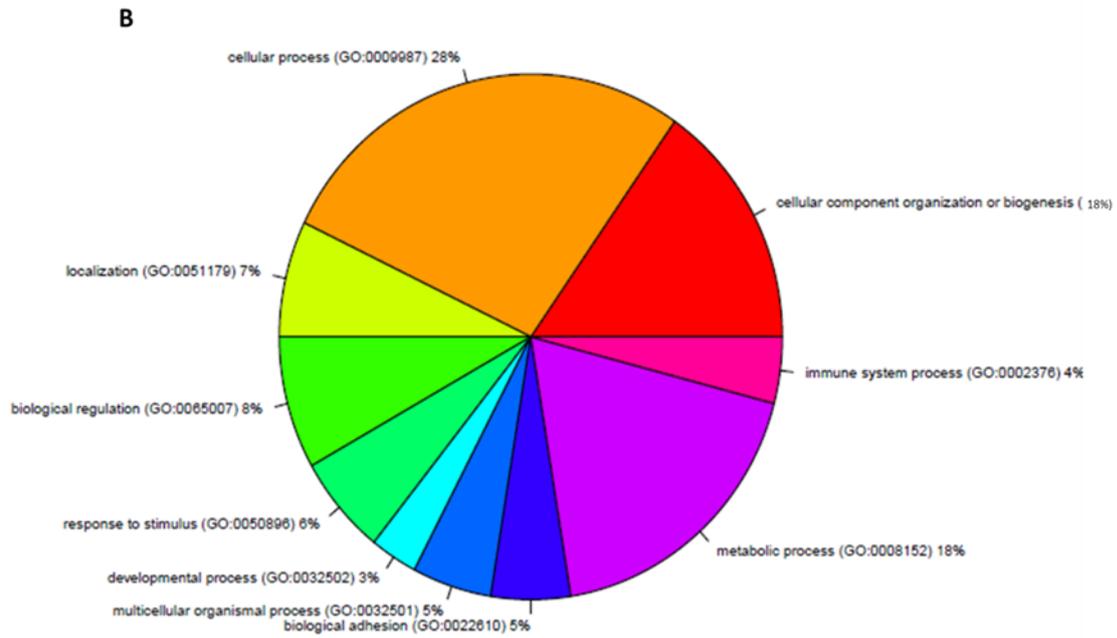
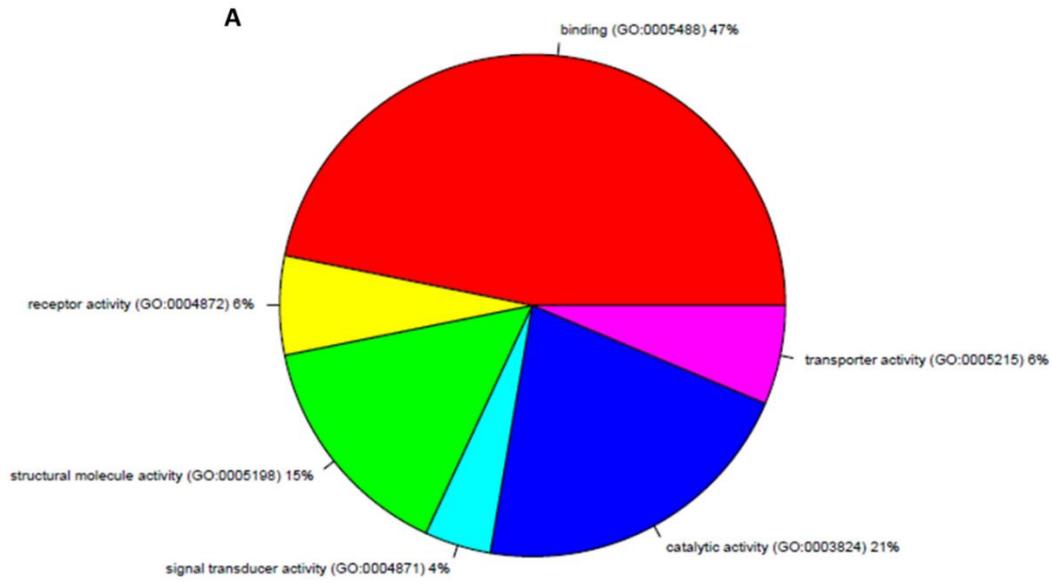
**B**



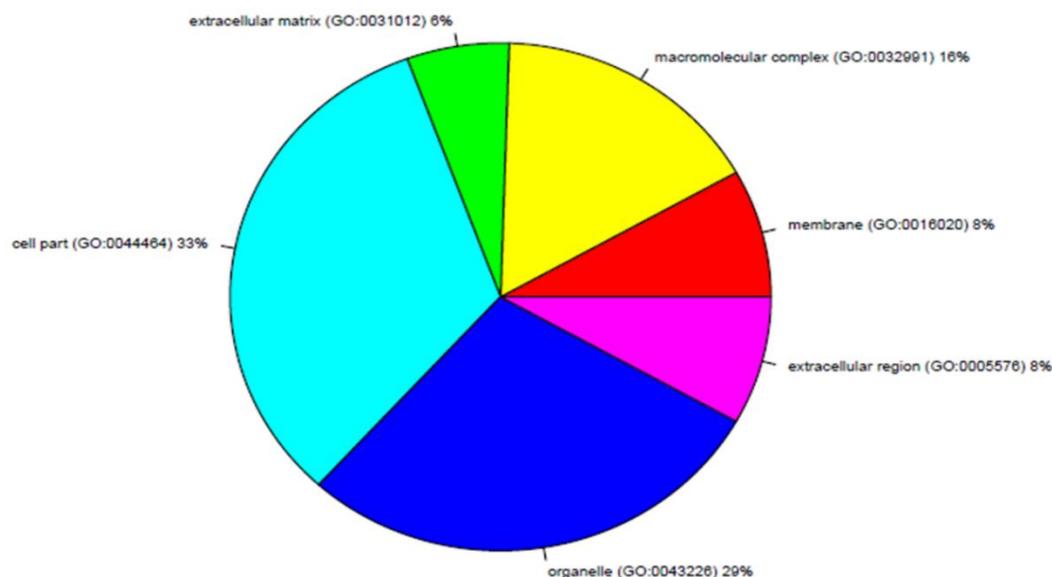


**Figure 5-8: Function classification from PANTHER online tool analysis of differential proteins expressed in the MDA-MB-468 cell line upon GAG treatment.**

**Panel A:** Functionally enriched GO: Molecular function categories. **Panel B:** Functionally enriched GO: Biological process categories. **Panel C:** Functionally enriched GO: Cellular composition categories.



C



**Figure 5-9: Function classification from PANTHER online tool analysis of differential proteins expressed in the MDA-MB-231 cell line upon GAG treatment.**

**Panel A:** Functionally enriched GO: Molecular function categories. **Panel B:** Functionally enriched GO: Biological process categories. **Panel C:** Functionally enriched GO: Cellular composition categories.

For every functional category from a number of sources (e.g. GO, KEGG, INTERPRO) there are associated proteins, counts in experimental/reference data, fold enrichment values, p-values and Bonferroni/Benjamini/FDR corrected p-values. Here the most enriched functional categories in the experimental data are visualised. However, the functional categories are clustered together based on the proteins/genes annotated in each functional category. An enrichment score is given for each cluster. Based on the results of this analysis shown in Tables 5-2 and 5-3, highest enrichment scores were obtained for cadherin binding involved in cell-cell adhesion, and the cell-cell adherent's junction and extracellular matrix in both TNBC cell lines. Cell adhesion is an important process on the tissue level for maintaining morphologies and their functions in an integrated manner. It is known as an interaction between a cell and a surface, substrate or another cell that occurs from the action of

transmembrane glycoproteins, called cell adhesion molecules, such as receptors. Such processes have a crucial role in tumour progression, in particular during invasion and metastasis (Cavallaro & Christofori, 2004).

Cadherins (calcium-dependent adhesion) are defined as any cadherin binding that occurs as part of the process of cell-cell adhesion. A class of type-1 transmembrane protein, they play important roles in cell adhesion, forming adherens junctions to bind cells within tissues together, and therefore play a critical role during tissue differentiation. They are dependent on calcium ( $\text{Ca}^{2+}$ ) ions to function. There are 18 proteins involved in this biological function in the human proteome and the cytosol-facing domains of these proteins are usually connected to elements of the cytoskeleton (Lodish *et al.*, 1995).

This finding is in line with the data generated from PANTHER. The molecular function and biological process annotations were enriched in both TNBC cell lines for protein binding, which involved cytoskeletal protein binding, receptor binding and calcium ion binding.

Another annotation cluster shown in Table 5-2 relates to the MAPK cascade signalling pathway. This is probably the most important in instigating the cellular response to the whelk GAG. Mitogen-activated protein kinases (MAPKs) are a highly conserved family of serine/threonine protein kinases involved in a variety of fundamental cellular processes such as proliferation, differentiation, motility, stress response, apoptosis and survival. Other proteins that are enriched contribute to cellular processes such as signal transduction/cell communication that initiate by binding to cell receptors (Seger & Krebs, 1995). However, the proteins related to nucleic acid binding transcription factor activity appeared to be less important during cell growth inhibition by whelk GAG.

The finding indicated that whelk GAG contributed to the interaction of the protein with an integrin and growth factor. Neuregulin (NRG) is an example of a specific growth factor that has a unique heparin-binding domain; binding to their receptors leads to rapid receptor-

tyrosine phosphorylation and activation of downstream signalling pathways. The heparin-binding domain binds specific sulphate groups on negatively charged HSPGs, leading to a highly specific tissue distribution of NRG (Pankonin *et al.*, 2005).

Other proteins that were found to be enriched in the MDA-MB-231 cell line are involved in important pathways, such as PI3K-Akt signalling pathway. It is known to contribute to the biology of various types of cancer including proliferation, adhesion, angiogenesis and metastasis. These findings suggest the influence of whelk GAG to downregulate crucial pathways such as proliferation, adhesion, migration and metastasis of breast cancer cells (Vara *et al.*, 2004).

**Table 5-2: Adjusted functional annotation clustering of significant proteins from DAVID online analysis of label-free proteomic data from the MDA-MB-468 cell line.**

Functional annotation clustering was conducted using DAVID analysis tool using subset data from quantitative proteomics analysis. Fold enrichment measures the magnitude of enrichment compared to human genome. Fold enrichment > 2 was considered as significant. The population percentage is the total number of proteins involved in a given term divided by the total number of input proteins. P-values examine the significance of gene-term enrichment. P-value <0.05 was considered significant. Existing ANOVA p-values were extracted from the two label-free proteomic analyses and imported into R. Benjamini-Hochberg p-values and were derived from the existing p-values for each analysis using the ‘p. adjust’ method. Protein groups which were under the 5% significance threshold for corrected p-values were selected for further analysis. False discovery rate (FDR) is the corrected over-representation probability, helps estimating the false positives and it is calculated using the Benjamini-Hochberg approach.

**Annotation Cluster 1                      Enrichment Score: 11.88**

Category	Term	%	pValue	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_MF_DIRECT	GO:0098641~cadherin binding involved in cell-cell adhesion	8.95	1.57E-13	5.33	8.84E-11	2.95E-11	2.30E-10
GOTERM_CC_DIRECT	GO:0005913~cell-cell adherens junction	8.95	5.24E-13	5.11	2.36E-10	7.85E-11	7.43E-10
GOTERM_BP_DIRECT	GO:0098609~cell-cell adhesion	7.80	2.69E-11	5.05	4.98E-08	4.98E-08	4.59E-08

**Annotation Cluster 2      Enrichment Score: 5.14**

Category	Term	%	pValue	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_DIRECT	GO:0031145~anaphase-promoting complex-dependent catabolic process	4.62	3.18E-11	10.27	5.87E-08	2.94E-08	5.41E-08
GOTERM_BP_DIRECT	GO:0006521~regulation of cellular amino acid metabolic process	3.17	4.69E-08	10.94	8.67E-05	1.44E-05	7.99E-05
GOTERM_MF_DIRECT	GO:0004298~threonine-type endopeptidase activity	2.31	1.09E-07	19.02	6.12E-05	1.53E-05	1.59E-04
GOTERM_BP_DIRECT	GO:0033209~tumor necrosis factor-mediated signaling pathway	3.46	2.11E-05	5.15	0.03	0.002	0.03
GOTERM_BP_DIRECT	GO:0002223~stimulatory C-type lectin receptor signaling pathway	3.17	4.15E-05	5.31	0.07	0.004	0.07
GOTERM_BP_DIRECT	GO:0050852~T cell receptor signaling pathway	3.17	6.97E-04	3.77	0.72	0.052	1.18
GOTERM_BP_DIRECT	GO:0000165~MAPK cascade	4.33	7.09E-04	2.90	0.73	0.051	1.20
GOTERM_CC_DIRECT	GO:0000932~cytoplasmic mRNA processing body	1.73	0.015461702	4.09	0.99	0.160	19.83

**Annotation Cluster 3      Enrichment Score: 4.55**

Category	Term	%	pValue	Fold Enrichment	Bonferroni	Benjamini	FDR
UP_KEYWORDS	mRNA processing	6.93	1.10E-08	4.29	3.66E-06	4.57E-07	1.49E-05
UP_KEYWORDS	mRNA splicing	5.78	9.24E-08	4.57	3.09E-05	2.81E-06	1.26E-04
GOTERM_BP_DIRECT	GO:0000398~mRNA splicing, via spliceosome	5.20	1.96E-06	4.11	0.003	3.01E-04	0.003
UP_KEYWORDS	Spliceosome	3.46	9.86E-06	5.62	0.003	1.94E-04	0.013
GOTERM_CC_DIRECT	GO:0005681~spliceosomal complex	3.17	1.04E-05	6.23	0.004	3.59E-04	0.014
GOTERM_CC_DIRECT	GO:0071013~catalytic step 2 spliceosome	2.60	3.29E-04	5.21	0.13	0.007	0.465
GOTERM_BP_DIRECT	GO:0008380~RNA splicing	3.17	0.001	3.36	0.95	0.09	2.807
KEGG_PATHWAY	hsa03040:Spliceosome	3.17	0.003	2.97	0.56	0.09	4.562
GOTERM_BP_DIRECT	GO:0006397~mRNA processing	2.60	0.025	2.55	1	0.49	35.476

**Table 5-3: Adjusted functional annotation clustering of significant proteins from DAVID online analysis of label-free proteomic data from MDA-MB-231 cell line.**

Functional annotation clustering was conducted using DAVID analysis tool using subset data from quantitative proteomics analysis. Fold enrichment measures the magnitude of enrichment compared to human genome. Fold enrichment > 2 was considered as significant. The population percentage is the total number of proteins involved in a given term divided by the total number of input proteins. P-values examine the significance of gene-term enrichment. P-value <0.05 was considered significant. Existing ANOVA p-values were extracted from the two label-free proteomic analyses and imported into R. Benjamini-Hochberg p-values and were derived from the existing p-values for each analysis using the ‘p.adjust’ method. Protein groups which were under the 5% significance threshold for corrected p-values were selected for further analysis. False discovery rate (FDR) is the corrected over-representation probability, helps estimating the false positives and it is calculated using the Benjamini-Hochberg approach.

**Annotation Cluster 1            Enrichment Score: 2.36**

Category	Term	%	pValue	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_CC_DIRECT	GO:0031012~extracellular matrix	10.14	8.21E-04	6.24	0.13	0.04	1.00
GOTERM_BP_DIRECT	GO:0007155~cell adhesion	10.14	0.009	3.76	0.99	0.35	14.00
GOTERM_BP_DIRECT	GO:0010811~positive regulation of cell-substrate adhesion	4.34	0.010	19.49	0.99	0.34	14.22

**Annotation Cluster 2****Enrichment Score: 2.25**

<b>Category</b>	<b>Term</b>	<b>%</b>	<b>pValue</b>	<b>Fold Enrichment</b>	<b>Bonferroni</b>	<b>Benjamini</b>	<b>FDR</b>
GOTERM_MF_DIRECT	GO:0031072~heat shock protein binding	5.79	6.14E-04	23.64	0.10	0.02	0.75
GOTERM_BP_DIRECT	GO:0006457~protein folding	7.24	0.005	6.85	0.98	0.26	8.44
GOTERM_MF_DIRECT	GO:0051082~unfolded protein binding	5.79	0.009	9.02	0.81	0.16	11.07
UP_KEYWORDS	Chaperone	5.79	0.028	5.93	0.99	0.28	30.22

**Annotation Cluster 3      Enrichment Score: 2.21**

Category	Term	%	PValue	Fold Enrichment	Bonferroni	Benjamini	FDR
UP_SEQ_FEATURE	region of interest:Heparin-binding	5.79	4.07E-05	58.15	0.01	0.01	0.05
GOTERM_MF_DIRECT	GO:0005178~integrin binding	8.69	6.01E-05	14.18	0.01	0.01	0.07
UP_KEYWORDS	Heparin-binding	7.24	2.06E-04	16.94	0.03	0.007	0.25
GOTERM_BP_DIRECT	GO:0030198~extracellular matrix organization	8.69	0.001	7.55	0.55	0.15	1.66
GOTERM_BP_DIRECT	GO:0033627~cell adhesion mediated by integrin	4.34	0.001	49.38	0.69	0.15	2.39
UP_SEQ_FEATURE	domain:VWFC	4.34	0.001	48.46	0.39	0.22	2.21
KEGG_PATHWAY	hsa04145:Phagosome	8.69	0.002	6.30	0.25	0.25	2.52
GOTERM_MF_DIRECT	GO:0008201~heparin binding	7.24	0.003	7.75	0.48	0.10	4.49
GOTERM_MF_DIRECT	GO:0050840~extracellular matrix binding	4.34	0.004	28.64	0.56	0.09	5.66
GOTERM_BP_DIRECT	GO:0007160~cell-matrix adhesion	5.79	0.005	10.97	0.98	0.27	8.13
SMART	SM00214:VWC	4.34	0.006	23.29	0.31	0.31	6.57
GOTERM_BP_DIRECT	GO:0007155~cell adhesion	10.14	0.009	3.76	0.99	0.35	14.00
INTERPRO	IPR001007: von Willebrand factor, type C	4.34	0.009	19.68	0.89	0.42	11.92
SMART	SM00209:TSP1	4.34	0.020	13.26	0.68	0.43	18.25

UP_KEYWORDS	Cell adhesion	8.69	0.021	3.73	0.97	0.24	22.92
INTERPRO	IPR000884:Thrombospondin, type 1 repeat	4.34	0.023	12.41	0.99	0.66	26.41
KEGG_PATHWAY	hsa05205:Proteoglycans in cancer	7.24	0.032	4.01	0.98	0.88	31.88
KEGG_PATHWAY	hsa04512:ECM-receptor interaction	4.34	0.097	5.54	0.99	0.89	69.87
KEGG_PATHWAY	hsa04510:Focal adhesion	5.79	0.128	3.12	0.99	0.86	80.03
KEGG_PATHWAY	hsa04151:PI3K-Akt signaling pathway	7.24	0.155	2.32	1	0.86	86.15
GOTERM_CC_DIRECT	GO:0009897~external side of plasma membrane	4.34	0.188	3.71	1	0.84	92.32

### 5.2.6.3 Cell signalling pathway analysis

The online PANTHER tool makes it possible to identify components of functional pathways from lists of dysregulated proteins. This allows candidate dysregulated pathways to be selected for more detailed investigation. PANTHER uses binomial statistics tools to compare a given gene list to a reference list (NCBI: Homo sapiens genes) to determine statistically significant overrepresentation of functional groups of genes. The approach employed in this chapter uses the quantitative proteomic data derived from label-free quantitative proteomic analysis from two triple negative breast cancer cell lines, MDA-MB-468 and MDA-MB-231. Each identified protein was automatically annotated according to the pathways in which it is involved. The respective pathways were then statistically analysed for their overrepresentation, and a list of the possible pathways through which whelk GAG inhibits cancer growth was generated. The resulting list was then filtered using a minimum 3-fold enrichment and p-value < 0.05 for the enrichment of data output. The pathways enriched in both cell lines are listed in tables 5-4 and 5-5. When subset data from the MDA-MB-468 cell line was used as data input, the result showed five pathways are significantly involved in the whelk GAG modes of action. Four pathways were significantly enriched in MDA-MB-231 cell line.

**Table 5-4: Pathways enriched in MDA-MB-468 from PANTHER online analysis.**

<b>Pathway</b>	<b>Homo sapiens (REF)</b>	<b>Total Protein</b>	<b>Fold Enrichment</b>	<b>P-value</b>
Glycolysis	22	8	14.98	1.54E-05
DNA replication	35	12	14.13	1.90E-08
Ubiquitin proteasome pathway	66	11	6.87	1.51E-04
EGF receptor signaling pathway	139	14	4.15	1.78E-03
FGF signaling pathway	124	12	3.99	1.08E-02

**Table 5-5: Pathways enriched in MDA-MD-231 from PANTHER online analysis.**

<b>Pathway</b>	<b>Homo sapiens (REF)</b>	<b>Total Protein</b>	<b>Fold Enrichment</b>	<b>P-value</b>
De novo pyrimidine deoxyribonucleotide biosynthesis	13	4	19.04	1.07E-02
DNA replication	35	7	12.37	3.37E-04
Cytoskeletal regulation by Rho GTPase	83	10	7.45	2.21E-04
Integrin signalling pathway	192	12	3.87	1.38E-02

According to the results observed from PANTHER and REACTOME from MDA-MB-468 cell line, when GAG have an impact on EGF receptor signalling pathway and FGF signalling pathway. When GAG might bind to EGFR and block RAS signalling which eventually inhibits phosphatidylinositol 3-kinase PI3K-AKT- mTOR, this further leads to cell arrest and anti-proliferation. RAS is a chain of proteins in the cell that communicates a signal from a receptor on the surface of the cell to the DNA in the nucleus of the cell. It is normally activated in response to the binding of extracellular ligands to various receptors such as epidermal growth factor (EGF) binding to its cognate receptor EGFR, Upon EGF binding to the extracellular domain of EGFR, the intracellular domain of EGFR forms an asymmetric dimer in the cytosol. EGFR and its ERBB receptor family members can form homo- or hetero-dimers. Downstream signalling proceeds through Ras in the Raf-MEK-ERK which is a key Ras effector pathway, known as the mitogen-activated protein kinase (MAPK), Raf-MEK- ERK pathway, and/or PI3K-Akt-mTOR pathways. The signalling is involved in numerous cellular functions, including cell proliferation, apoptosis, migration, and differentiation (DerMardirossian & Bokoch, 2001).

PANTHER pathway analysis suggests that DNA replication signalling pathways are affected in MDA-MB-468 and MDA-MB-231 as a consequence of GAG treatment. This pathway is a complex network of interacting proteins and enzymes that are required for DNA replication. The role for DNA replication stress in human tumorigenesis has been extensively studied. The early stages of cancer development, before genomic instability and malignant conversion commonly express markers of an activated DNA damage response to delay cancer. These include phosphorylated kinases ATM and Chk2, phosphorylated histone H2AX and p53. However, mutations in some pathways such as ATM or p53 might allow cell proliferation (Bartkova *et al.*, 2005).

REACTOME (<http://reactome.org>) is a free bioinformatics online tool for visualisation and analysis of pathways to support genome analysis of large gene lists derived from high-throughput genomic, proteomic approaches. Pathway analysis is a powerful tool for extracting meaning from a long list of differentially expressed genes and proteins. REACTOME offers a set of pathway analysis tools which aim to deal with the large size of the data samples, and provides reliable and accurate results as well as additional information. The analysis algorithm conducted in REACTOME is comprised of four steps. The first involves finding matches between the protein/gene identifiers in the user's sample and the background data contained in REACTOME. The second step involves grouping the identified proteins in sets or species projections. This is followed by aggregating the found identifiers in the pathways and finally, performing statistical testing to calculate the likelihood that the association between the sample identifiers and the found pathway is due to random chance (Fabregat *et al.*, 2017).

According to García-Campos *et al.*, (2015), the simple hypothesis in an over-representation analysis is that relevant pathways can be detected if the proportion of differentially expressed genes within a given pathway exceeds the proportion of genes that could be randomly

expected. The method used by Reactome to calculate statistical significance is the Binomial Test. Together with the p-value, the false discovery rate (FDR) helps to estimate false positives and is calculated using the Benjamini-Hochberg approach (Benjamini & Hochberg, 1995).

Here, a set of protein IDs from proteomic data for each cell line was uploaded in REACTOME, which served as an additional tool with which to verify the previous pathway analysis results from PANTHER and functional annotation clustering from DAVID. The output was filtered by searching for pathways with an entities p-value < 0.05. The over-represented pathways generated from this tool are listed in Tables 5-6 and 5-7.

**Table 5-6: Pathways enriched from REACTOME in MDA-MB-468 cell line.**

<b>Pathway identifier</b>	<b>Pathway name</b>	<b>Entities pValue</b>	<b>Entities FDR</b>
<b>R-HSA-216083</b>	Integrin cell surface interactions	< 0.01	< 1%
<b>R-HSA-1474244</b>	Extracellular matrix organization	< 0.01	< 1%
<b>R-HSA-5687128</b>	MAPK6/MAPK4 signalling	< 0.01	< 1%
<b>R-HSA-70171</b>	Glycolysis	< 0.01	< 1%
<b>R-HSA-8866910</b>	TFAP2 (AP-2) family regulates transcription of growth factors and their receptors	< 0.01	< 1%
<b>R-HSA-3371497</b>	HSP90 chaperone cycle for steroid hormone receptors (SHR)	< 0.01	< 1%
<b>R-HSA-140342</b>	Apoptosis induced DNA fragmentation	< 0.01	< 1%
<b>R-HSA-211227</b>	Activation of DNA fragmentation factor	< 0.01	< 1%
<b>R-HSA-179419</b>	APC:Cdc20 mediated degradation of cell cycle proteins prior to satisfaction of the cell cycle checkpoint	< 0.01	< 1%
<b>R-HSA-392517</b>	Rap1 signalling (cell cycle)	< 0.01	< 1%
<b>R-HSA-8866911</b>	TFAP2 (AP-2) family regulates transcription of cell cycle factors	< 0.01	< 1%
<b>R-HSA-69473</b>	G2/M DNA damage checkpoint	< 0.05	< 1%
<b>R-HSA-351906</b>	Apoptotic cleavage of cell adhesion proteins	< 0.05	< 1%
<b>R-HSA-70326</b>	Glucose metabolism	< 0.05	< 1%
<b>R-HSA-5626467</b>	RHO GTPases activate IQGAPs	< 0.05	< 1%
<b>R-HSA-351906</b>	Apoptotic cleavage of cell adhesion proteins	< 0.05	< 1%
<b>R-HSA-2262752</b>	Cellular responses to stress	< 0.05	< 1%
<b>R-HSA-5638302</b>	Signalling by Overexpressed Wild-Type EGFR in Cancer	< 0.05	< 1%
<b>R-HSA-5638303</b>	Inhibition of Signalling by Overexpressed EGFR	< 0.05	< 1%
<b>R-HSA-5637812</b>	Signalling by EGFRvIII in Cancer	< 0.05	< 1%
<b>R-HSA-5673001</b>	RAF/MAP kinase cascade	< 0.05	< 1%

**Table 5-7: Pathway enriched from REACTOME in MDA-MB-231 cell line.**

<b>Pathway identifier</b>	<b>Pathway name</b>	<b>Entities pValue</b>	<b>Entities FDR</b>
<b>R-HSA-216083</b>	Integrin cell surface interactions	< 0.01	< 1%
<b>R-HSA-1474244</b>	Extracellular matrix organization	< 0.01	< 1%
<b>R-HSA-5683057</b>	MAPK family signalling cascades	< 0.01	< 1%
<b>R-HSA-5625900</b>	RHO GTPases activate CIT	< 0.01	< 1%
<b>R-HSA-351906</b>	Apoptotic cleavage of cell adhesion proteins	< 0.01	< 1%
<b>R-HSA-351906</b>	Apoptotic cleavage of cell adhesion proteins	< 0.01	< 1%
<b>R-HSA-2644603</b>	Signalling by NOTCH1 in Cancer	< 0.01	< 1%
<b>R-HSA-5626467</b>	RHO GTPases activate IQGAPs	< 0.05	< 1%
<b>R-HSA-69205</b>	G1/S-Specific Transcription	< 0.05	< 1%
<b>R-HSA-68886</b>	M Phase	< 0.05	< 1%
<b>R-HSA-8956320</b>	Nucleobase biosynthesis	< 0.05	< 1%
<b>R-HSA-170834</b>	Signalling by TGF-beta Receptor Complex	< 0.05	< 1%
<b>R-HSA-176814</b>	Activation of APC/C and APC/C:Cdc20 mediated degradation of mitotic proteins	< 0.05	< 1%
<b>R-HSA-176408</b>	Regulation of APC/C activators between G1/S and early anaphase	< 0.05	< 1%
<b>R-HSA-8875878</b>	MET promotes cell motility	< 0.05	< 1%
<b>R-HSA-8852276</b>	The role of GTSE1 in G2/M progression after G2 checkpoint	< 0.05	< 1%
<b>R-HSA-174143</b>	APC/C-mediated degradation of cell cycle proteins	< 0.05	< 1%
<b>R-HSA-453276</b>	Regulation of mitotic cell cycle	< 0.05	< 1%
<b>R-HSA-8874081</b>	MET activates PTK2 signalling	< 0.05	< 1%
<b>R-HSA-75153</b>	Apoptotic execution phase	< 0.01	< 1%

The analysis revealed several pathways and sub-pathways, of which the integrin cell surface interactions pathway was the most overrepresented.

Integrins are the receptors that mediate cell adhesion to the extracellular matrix ECM. Integrins consists of one alpha and one beta subunit forming a non-covalently bound heterodimer. Eighteen alpha and eight beta subunits have been identified in humans that combine to form 24 different receptors. The control of integrin function occurs via regulatory signals that originate within the cell cytoplasm and are then transmitted to the external ligand-binding domain of the receptor. As previously indicated by DAVID annotation analysis, REACTOME pathway analysis showed that whelk GAG demonstrated a direct inhibition towards the signalling of these integrins through their receptors.

ECM composition is highly heterogeneous and dynamic, modulated largely by matrix metalloproteinases (MMPs) and growth factors. Collagen is the most abundant fibrous protein within the ECM, constituting up to 30% of total protein in multicellular animals. It provides tensile strength and the ability to recover after stretching (Daley *et al.*, 2008). Whelk GAG affected the collagen through their signalling events in both cell lines.

ECM performs many functions such as proliferation, adhesion and migration, and regulates cell differentiation and death (Hynes, 2009). Within the pathways related to extracellular matrix organization, whelk GAG was found to influence deregulated cell proliferation and invasion, potentially attributed to the inhibition of the epidermal growth factor (EGF) signalling cascade. Thus suggesting that whelk GAG specifically influence cell proliferation and adhesion, as also was shown by the data obtained from PANTHER.

As described in the previous section, our study has demonstrated that whelk GAG reduced the level of glycolysis in the MDA-MB-468 cell line following 24 hours of treatment. In line with this finding, REACTOME pathway analysis showed that whelk GAG interfered with the glycolysis metabolic complex signalling cascade that converts glucose 6-phosphate to

pyruvate, which then controls rearrangement of cancer cells metabolism. Glucose via glycolysis provides the carbon skeletons, NADPH and ATP, to build new cancer cells, which persist in hypoxia that in turn rewires metabolic pathways for cell growth and survival (Dang, 2012).

Interestingly, this effect has not been reported from MDA-MB-231. This cell line looks to have been effected less by the GAG treatment (figure 5-5). It may be the glycolysis effects are a later effect of wheelk GAG treatment as the cells begin to die.

Moreover, pathway analysis using subset data from the MDA-MB-468 and MDA-MB-231 cells showed direct effect of wheelk GAG on Rho GTPase (guanine-nucleotide-exchange factors (GEFs) or GTPases). The latest are molecular switches that control a wide variety of signal transduction pathways in all cells. They are known principally for their pivotal role in regulating the actin cytoskeleton, but their ability to influence cell polarity, microtubule dynamics, membrane transport pathways and cell cycle progression is probably just as significant (Bremm *et al.*, 2008; Mateus *et al.*, 2007). Specific members of the Rho family and their downstream targets contribute to the regulation of key elements from the core cell cycle machinery, mostly involved in the G1/S transition (Villalonga *et al.*, 2006).

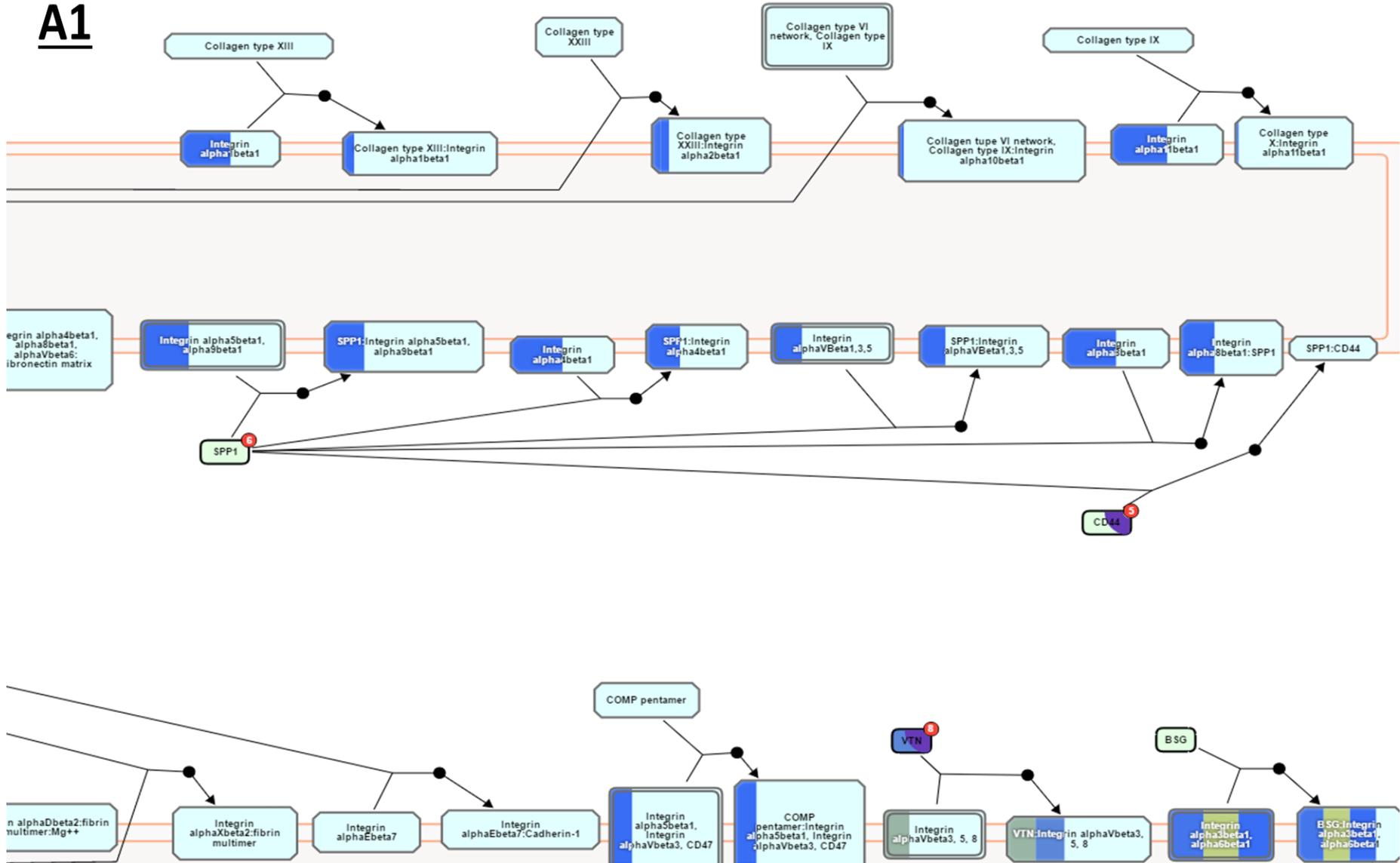
Moreover, pathway analysis using subset data from the MDA-MB-231 cell line showed selective effect of wheelk GAG on Notch signalling pathway. Notch signalling is a highly conserved signal transduction pathway that regulates stem cell maintenance and differentiation in several organ systems. Upon activation, the Notch receptor is proteolytically processed, its intracellular domain (NICD) translocates into the nucleus and activates expression of target genes. A crosstalk with other conserved signalling pathways such as the Wnt, Hedgehog, hypoxia and TGF $\beta$ /BMP pathways can affect Notch signalling output (Borggreffe *et al.*, 2016).

This regulation can happen by regulation of ligand, receptor or transcription factor expression, regulation of protein stability of intracellular key components, usage of the same cofactors or coregulation of the same key target genes (Borggrefe *et al.*, 2016). Since carcinogenesis is often dependent on at least two of these pathways, a better understanding of their molecular crosstalk is pivotal. There are various examples of elevated Notch signalling being associated with a particular subtype of breast cancer and response to targeted therapy. For example, elevated Notch1 expression is found in HER2 +ve (Zardawi *et al.*, 2010) and triple negative/basal breast tumours (Reedijk *et al.*, 2008). Recently, a study has identified a series of activating mutations within the PEST domain of Notch1, Notch2, and Notch3, these mutations were enriched in triple negative breast cancers (Wang *et al.*, 2015).

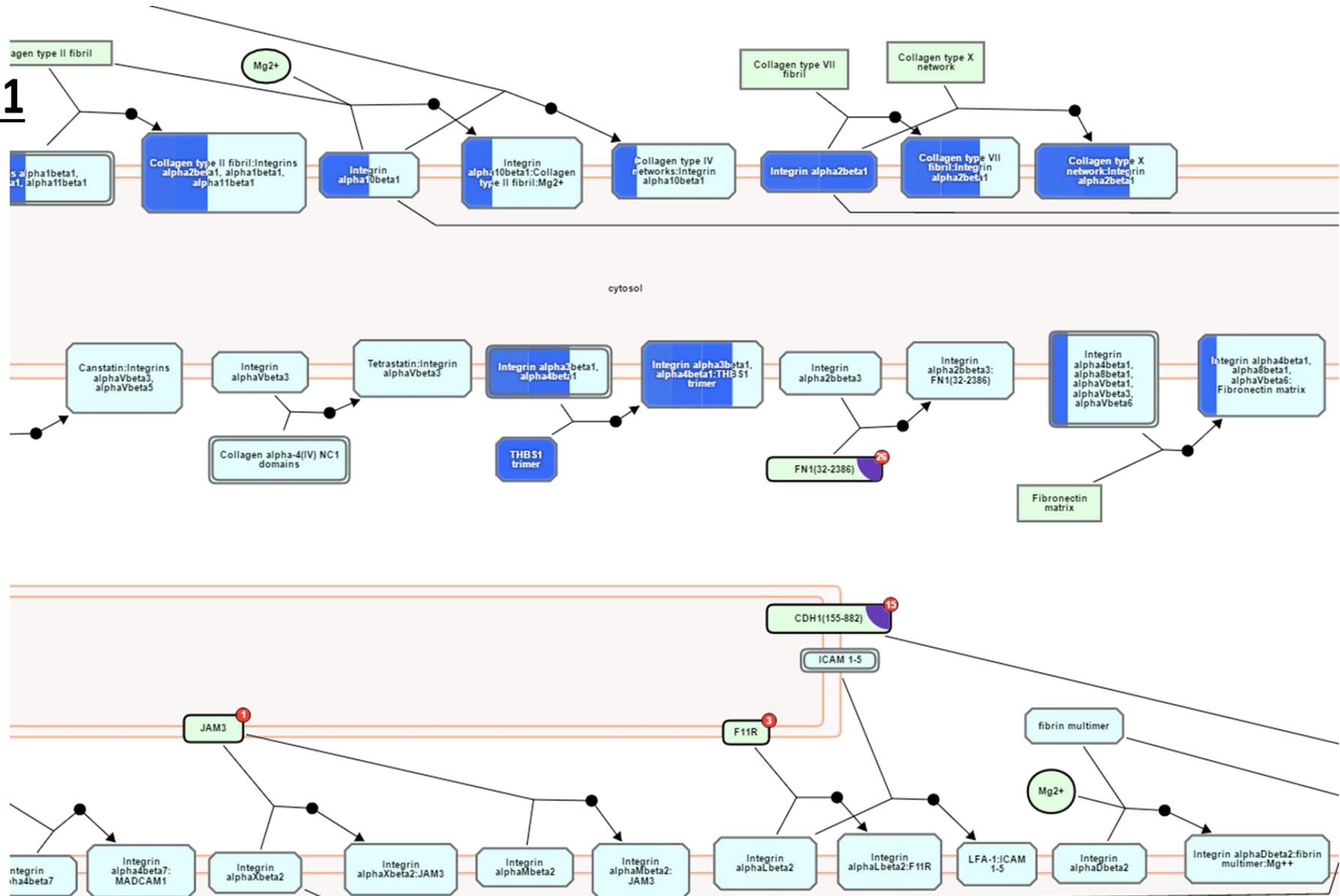
Notch signalling is known to regulate many cellular processes of breast cancer including proliferation by upregulating cyclinA, cyclinB, and cyclinD1 expression (Rizzo *et al.*, 2008), angiogenesis, by activate Notch signalling in neighbouring endothelial cells to promote angiogenesis (Dufraine *et al.*, 2008; Zeng *et al.*, 2005), metastasis, as shown by Sethi *et al.*, (2011), the breast cancer cell line MDA-MB-231 was forced to express the Notch ligand Jagged1, and this was found to significantly increase bone metastasis.

Finally, Notch signalling was found to regulate cancer stem cell activity, by regulating the self-renewal of breast cancer stem cells (Harrison *et al.*, 2010; Harrison *et al.*, 2013). Altogether, it is very clear that activation of Notch plays a key role in breast cancer. Therefore it represents a very attractive therapeutic target for many breast cancer subtypes, such as ER-positive tumours, as they frequently develop resistance to endocrine treatments. Several recent studies have suggested that ER mutations may occur during endocrine treatment (Robinson *et al.*, 2013). More recently, pharmacologic and genetic inhibition of the Notch signalling was found to reduce breast CSC activity *in vitro* and tumour formation *in vivo* (Harrison *et al.*, 2010), in addition to their roles in treating TNBC.

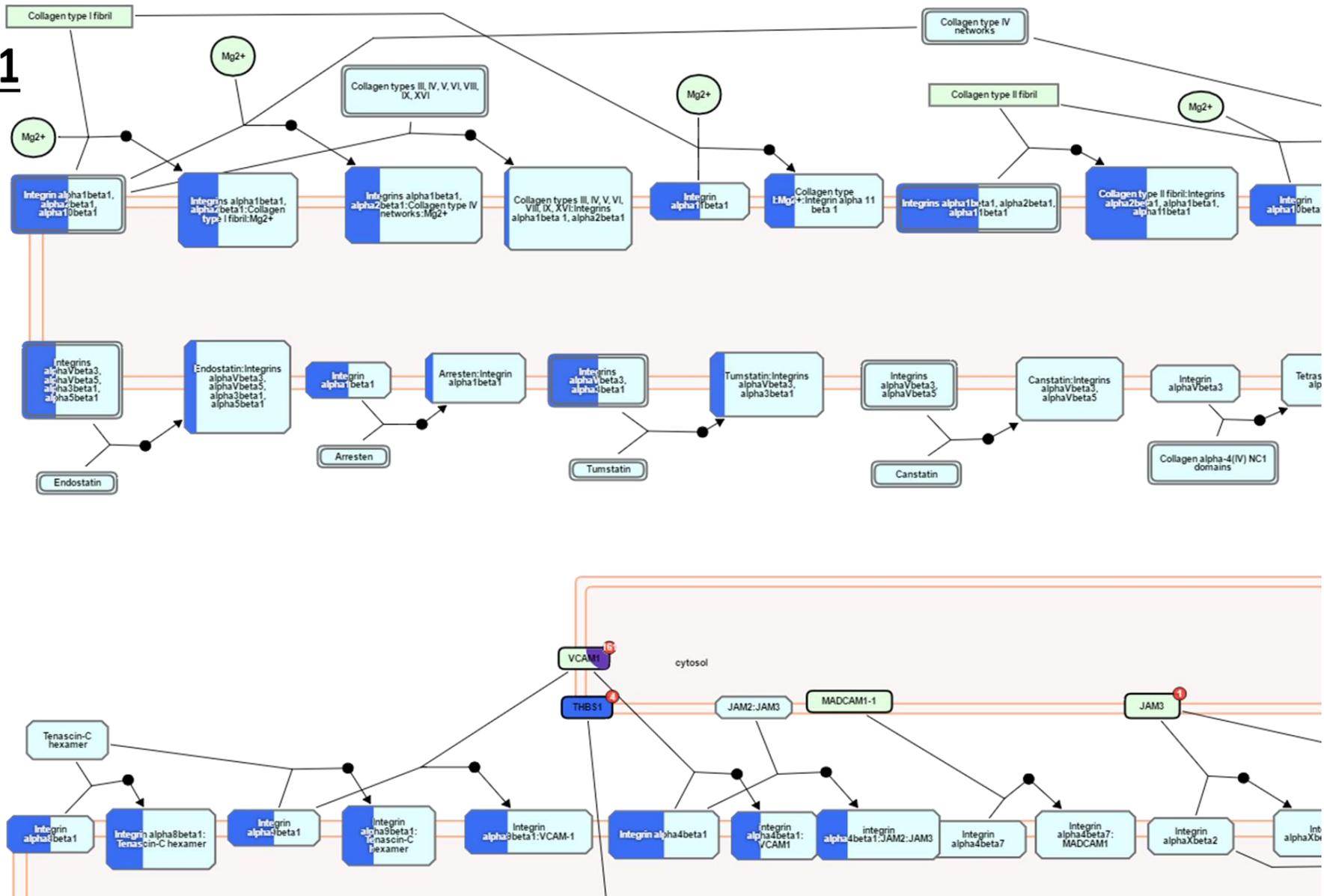
# A1



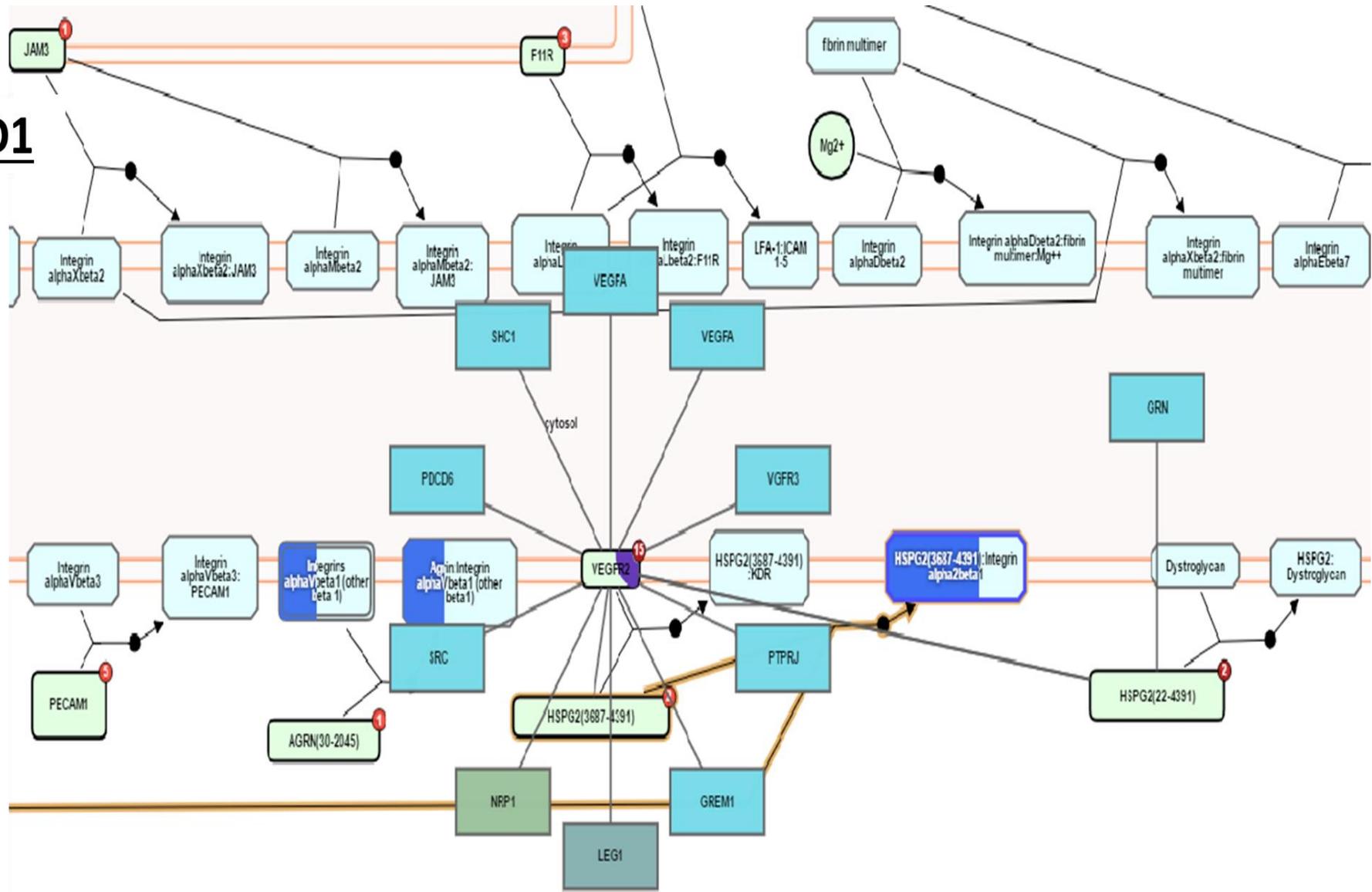
**B1**



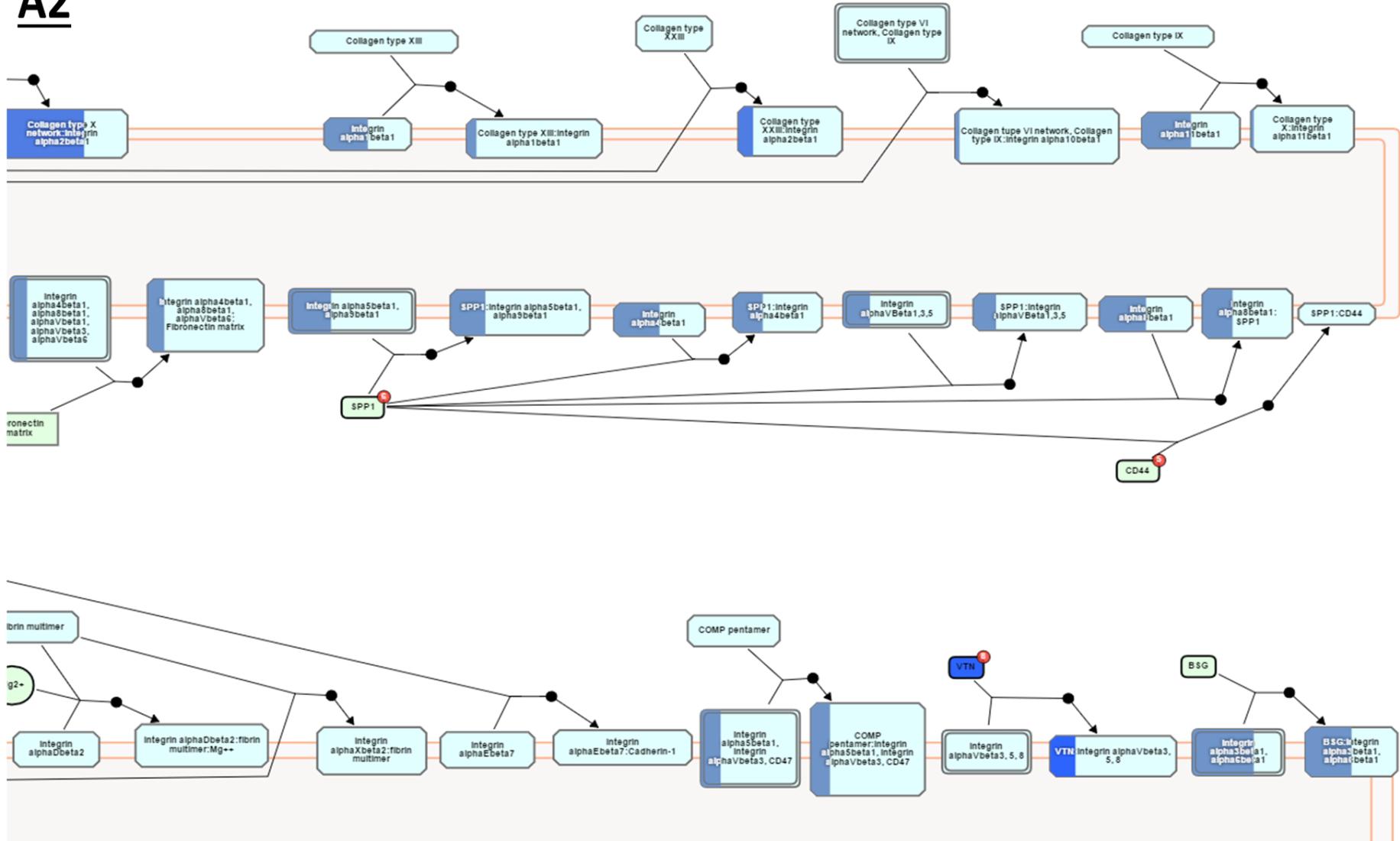
**C1**



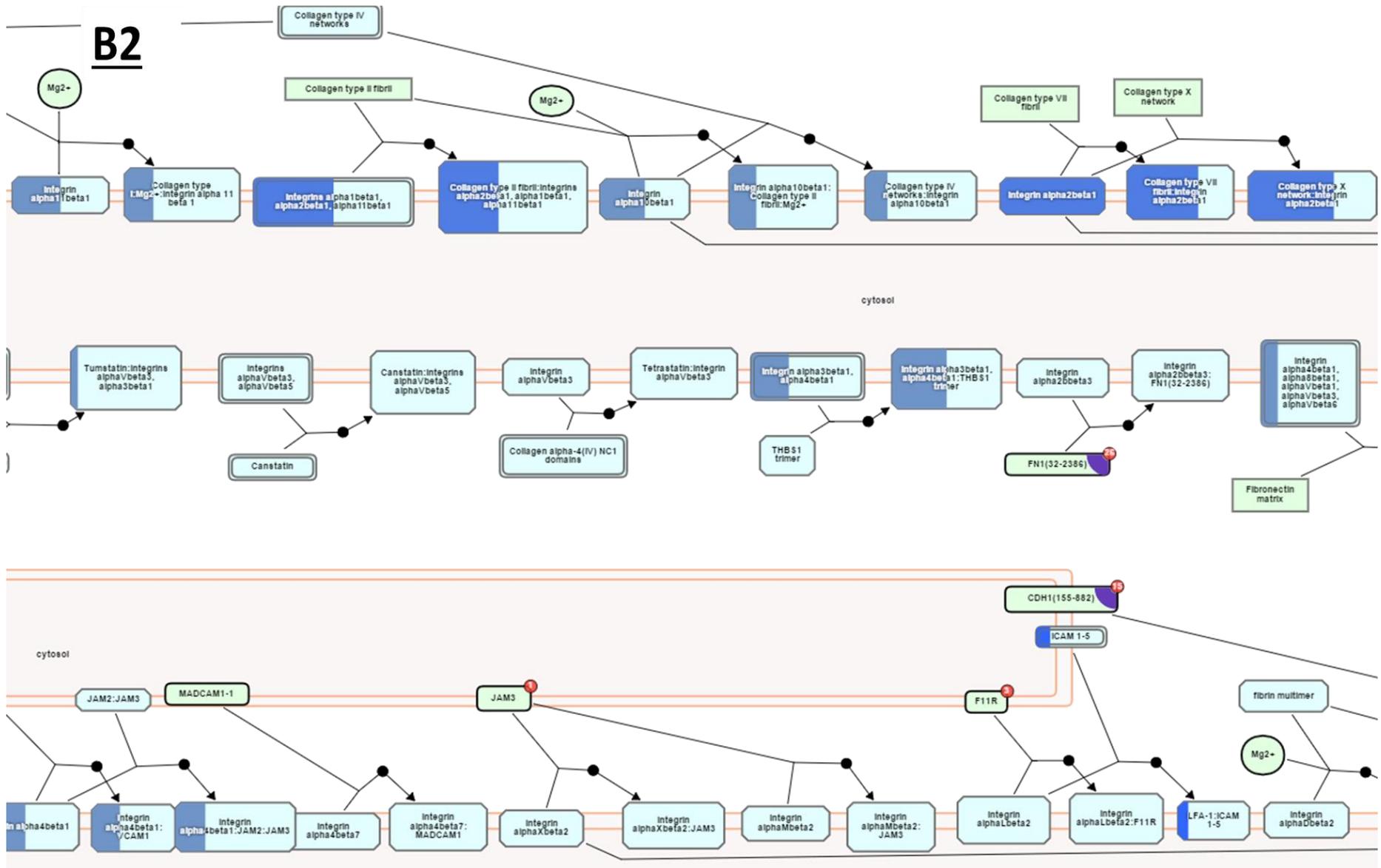
**D1**



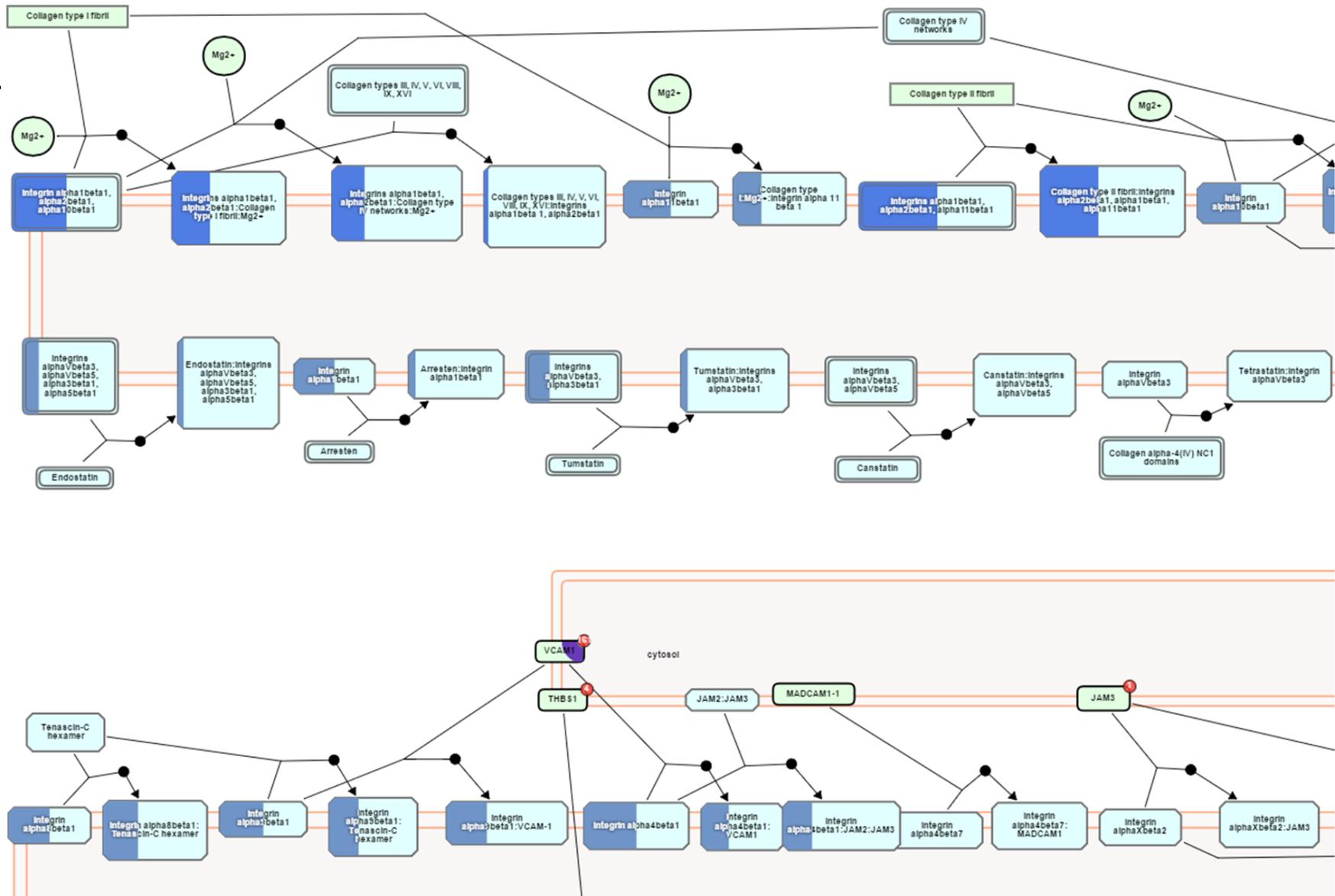
# A2

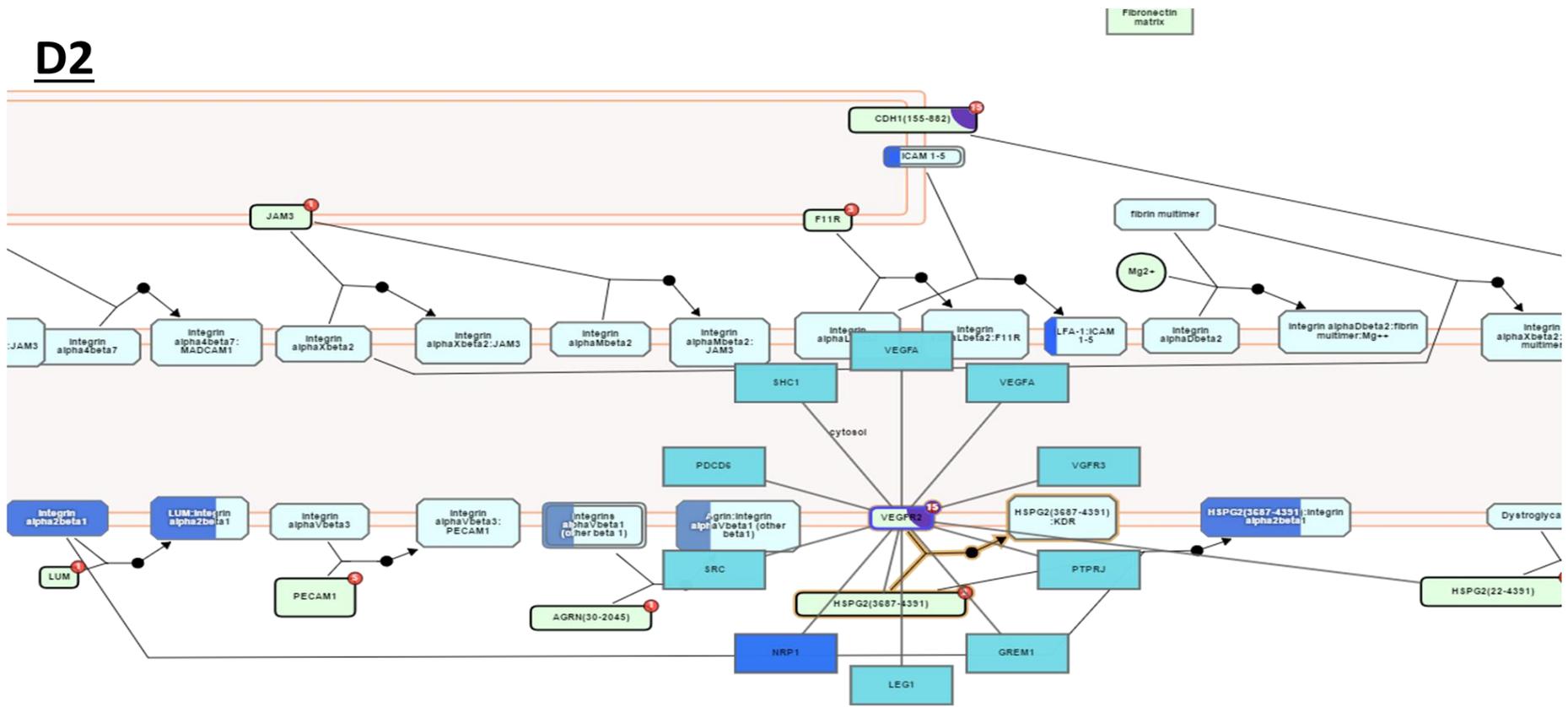


# B2



**C2**





**Figure 5-10: The integrin cell surface interactions pathway.**

Downregulation (showing in a dark blue) in the majority of integrin signalling events in the cytosol in the MDA-MB-468 cell line (Panel A1, B1, C1 and D1) and MDA-MB-231 cell line (Panel A2, B2, C2 and D2).

### 5.3 Discussion

This chapter set out to assess the impact of the crude whelk GAG extracts and their fragments on TNBC cell proliferation and to investigate their mode of action on TNBC using label-free proteomic approaches. In addition, this chapter has described the results from experiments designed to examine the activity of whelk GAGs toward cancer stem cells by using mammosphere assays on different breast cancer subtypes.

The main finding of this chapter is that whelk GAG crude extracts have significant inhibitory activity on the triple negative breast cancer cell lines MDA-MB-231 and MDA-MB-468. Remarkable IC<sub>50</sub> values were reported from both cell lines, with nearly  $3.48 \pm 0.02$   $\mu\text{g/ml}$  and  $2.76 \pm 0.01$   $\mu\text{g/ml}$  observed from MDA-MB-231 and MDA-MB-468 respectively.

A comparison between the bioactivity of whelk GAG extracts and commercial mammalian HS/CS on triple negative breast cancer has been reported under two conditions, before and after enzymatic degradation of each source of GAG (whelk and commercial mammalian HS/CS). The results indicated no growth inhibitory activity from commercial mammalian HS/CS either before or after enzymatic degradation on TNBC cell lines. By contrast, significant inhibition was observed following treatment with whelk GAG extracts before and after enzymatic degradation with HS and CS lyases.

The results reported in Chapter 3 suggested incomplete enzymatic digestion of whelk GAG, according to the elution profile from size exclusion chromatography. On the evidence from high-performance liquid chromatography, neither ABC nor Hep lyases completely breakdown the whelk GAG chain, as evidenced by the overlapping peaks that were obtained after enzymatic depolymerisation, which revealed partially depolymerised compounds.

In line with these results, the biological activity of the fragments generated from Heparinase lyase has slightly reduced. This could be as a result of losing fraction of the Hep/HS unit

from the chain that might play an important role in cell growth inhibition. The activity from fragments generated by ABC enzyme seemed to be similar to the crude whelk GAG extracts. This study has shown that whelk GAG fractions from high performance anion exchange chromatography (HPAEC) exhibit significant inhibition activity toward TNBC cell lines. These fragments were oxidised by periodic acid, which acts on glycoside bonds. The aim was to break down the sugar rings and evaluate the structure-function activity. Remarkably, the biological activity remained in these fragments and fucose units were the only monosaccharide that have been oxidised by periodic acid as indicated in previous chapter. This finding was unexpected and suggests that the substrate undergoes partial oxidation process even after a long-term oxidation period. In addition, losing of fucose units did not affect the biological activity of whelk GAGs. This indicates that fucose units that attached to the GAG chain in whelk extracts seem less important in the biological activity.

Nevertheless, a study has shown that the C-2-C-3 glycol group in the uronic acid moiety of the GAGs has low reactivity towards periodic acid oxidation. The reactivity of D-glucuronic acid in chondroitin 4-sulphate and 6-sulphate was reported to be relatively low, which might oxidize 100 fold more slowly compared to the fast oxidation of the C-5 epimer L-iduronic acid in heparan sulphate (Scott & Tigwell, 1978). This might explain the incomplete oxidation process.

Fragments CK1, CK2 and CK3 from ion exchange chromatography were degraded by enzymatic action. Each fragment was then separated by a 10K filter into two fractions, a flow-through fraction (FT) and a retained fraction (R). Each fraction from each fragment (CK1, CK2 and CK3) was tested *in vitro*. No inhibition activity toward either the MDA-MB-231 or MDA-MB-468 cell line was obtained from the FT fraction after Hep and ABC lyase degradation. This result clearly suggests that fragments with a size equal to or smaller than 10K tended to be weaker and their anti-proliferative effects were lost.

However, massive inhibition activity on both cell lines was reported from the retained fraction of all fragments. These results indicate that the active fragments are bigger in size than 10K (the size of the filter pores), and therefore could not pass through the filter.

These findings concur with other studies that show different sizes of marine GAG fragments seem to exhibit different biological activity. Many polysaccharides isolated from marine organisms have been chemically modified or degraded and are involved directly or indirectly in inhibiting cancer growth or influence different stages of carcinogenesis (Stonik & Fedorov, 2011).

Fragments with different molecular weight of fucosylated chondroitin sulphate are reported to exhibit anticoagulant activity, but this activity has been shown to reduce with the decrease in molecular size of the polysaccharide obtained, and a molecular size above 6 kDa is required to achieve thrombin inhibition (Wu *et al.*, 2010).

The results are broadly consistent with those of another study of unfractionated chains of Fucoidans which is isolated from many marine organisms. These are reported to exhibit activity in several aspects, such as inducing apoptosis in human colon cancer cells as well as breast adenocarcinoma MCF-7 (Hyun *et al.*, 2009; Kim *et al.*, 2010; Yamasaki-Miyamoto *et al.*, 2009).

Other studies reported that naturally over-sulphated and high molecular weight fucoidans from *U. pinnatifida* exhibit significant activity as anticancer agents and prevent angiogenesis due to their molecular size (Cho *et al.*, 2010; Narazaki *et al.*, 2008).

In contrast, some reports in the literature indicate that low molecular weight oligosaccharides isolated and prepared from *Chondrus ocellatus* may be more promising as anticancer agents than high molecular forms of the same polysaccharide (Lins *et al.*, 2009).

This is the first study, to our knowledge, to examine the bioactivity of crude whelk GAG crude extracts on mammospheres formation in MCF-7, MDA-MB-468 and MDA-MB-231,

although the mammospheres formed from MDA-MB-468 and MDA-MB-231 were not as clear as those obtained from MCF-7. However, it is clear that some population of each cell line survived the conditions of undetected culture and proliferated over a 5-day incubation period. The crude whelk GAG extracts reduced the number of mammospheres in all cell lines after treatment with the IC50 values corresponding to each cell line but did not completely inhibit mammosphere growth. The inhibition activity was increased by increasing the dose of whelk GAG. This important finding shows the capability of crude whelk GAG extracts to induce cell death among fewer cell populations from different subtypes of breast cancer which have the ability to develop metastasis or act as cancer stem cells (CSC).

In order to gain a better understanding of the mechanism of action of the whelk GAG extracts and highlight potential targets for the active compounds from whelk GAG in TNBC, this study has employed a label-free proteomic approach.

The results from these experiments suggested many pathways that might be involved in the whelk GAG inhibition mechanisms on both cell lines. The integrin cell surface interaction pathway was identified by REACTOME as the first route of whelk GAG to induce growth inhibition. The further downstream signalling transduction of this pathway leads to cell cycle arrest and anti-proliferation.

The results demonstrated that whelk GAG impedes FGF and FGFRs pathways in MDA-MB-468 cells, which eventually inhibit cell proliferation and tumour progression. The results also suggested the ability of whelk GAG to bind to EGFR which may block or inhibit RAS signalling which eventually inhibit PI3K-AKT- mTOR, this further leads to cell arrest and anti-proliferation in MDA-MB-468 cells.

The results showed that whelk GAGs affected the extracellular organization in both cell lines. The pathways related to extracellular matrix organization consist of several proteins that are responsible for many cellular functions such as cell proliferation, cell differentiation and cell

adhesion. Each protein has several domains, and each domain has its own specific structure and function. For example, the extracellular domain of E-cadherin that was significantly enriched in both cell lines upon whelk GAG treatment, is essential for cell-to-cell adhesion. The cytoplasmic domain of E-cadherin interacts with the catenins ( $\alpha$ -,  $\beta$ -,  $\gamma$ - and p120 catenin) anchored to the actin cytoskeleton, establishing cadherin-catenin complexes (Gumbiner & McCrea, 1993).

It has been reported that the conformation of E-cadherin is only stable upon  $\text{Ca}^{2+}$  binding to its highly conserved, negatively charged extracellular motifs. E-cadherin forms an adherens junction with its binding partner  $\beta$ -catenin and actin filaments. This complex is critical to inhibiting individual epithelial cell motility and to providing homeostatic tissue architecture (Cavallaro & Christofori, 2004; Kemler & Ozawa, 1989; Wen *et al.*, 2015). E-cadherin is also involved in a number of signalling pathways in carcinogenesis. One of the pathways that is frequently overexpressed in breast cancer involves Rho GTPases, among which Rho A, Rac1 and Cdc42 have been extensively studied (Chen *et al.*, 2014; Li *et al.*, 2014; Menezes *et al.*, 2014). They contribute to the regulation of key elements from the core cell cycle machinery, mostly involved in the G1/S transition. This pathway was significantly affected by whelk GAG treatment in MDA-MB-231 cells.

The results demonstrate downregulation in the glycolysis metabolism pathway in MDA-MB-468, which is important in cancer survival. It may be the glycolysis effects are a later effect of whelk GAG treatment as the cells begin to die. This pathway seems to be less important in whelk GAG modes of action on the MDA-MB-231 cell line.

However, selective effect of whelk GAG on Notch signalling pathway has been reported from MDA-MB-231 cell line. Recently, a study has identified a series of activating mutations within the PEST domain of Notch1, Notch2, and Notch3, these mutations were enriched in triple negative breast cancers including MDA-MB-231 cell. Notch signalling is known to

regulate many cellular processes of breast cancer including proliferation, metastasis, cancer stem cells, angiogenesis. Therefore, this result describes a strong potential target for whelk GAG and demonstrates one of the most important mechanisms of action on this cell line.

In conclusion, the crude whelk GAG extracts were extensively active on both cell lines, whereas no activity was reported from commercial mammalian HS or CS on either cell line.

In this study, a library of fragments was generated from whelk GAG using ion-exchange chromatography, Heparinase lyase, and ABC lyase. Incomplete enzymatic digestion of whelk GAG was suggested by the elution profile of size exclusion chromatography. The results from MTT assays from both cell lines show an obvious resistance from both TNBC cell lines towards fragments generated from Hep lyase. This could be as a result of losing some Hep/HS units from the chain sequence that may play an important role in inhibiting the growth of TNBC cells. The results from fragments generated from ABC lyases showed high inhibition activity toward both cell lines with no resistance as previously noticed with Hep fragments.

Moreover, fragments from ion-exchange chromatography have shown significant inhibition activity on both cell lines. The oxidised form of these fractions showed noticeable inhibition activity on both cell lines. Generally, D-glucuronic acid moiety of the GAG shows low reactivity towards periodic acid oxidation, hence the whelk GAG sample is reported to have a high abundance of glucuronic acid, which might explain the resistance to the oxidation process. In addition, losing of fucose units because of oxidation process did not affect the biological activity of whelk GAG. This suggests that fucose is less important in the whelk GAG function against cancer growth.

Each fraction (FT and R) from each fragment (CK1, CK2 and CK3) was monitored for its biological activity. No inhibition activity on TNBC cell lines was obtained from the FT fraction of Heparinase lyase and ABC lyase degradation. However, significant inhibition

activity on both cell lines was observed from the retained fractions from all fragments. The inhibition activity of the obtained purified GAG (FT fraction) on TNBC cell lines tended to be weaker and its anti-proliferative effects were lost to some extent. This strongly suggests that the active fragment is clearly longer in size than 10K.

In addition, crude whelk GAG extracts have shown a remarkable inhibition activity on mammospheres obtained from MCF-7, MDA-MB-468 and MDA-MB-23 cell lines. This important finding shows the capability of whelk GAGs to induce cell death among cells that might develop metastasis or act as CSC.

An overall understanding of the mechanism of action of whelk GAG toward TNBC cell lines has emerged from this study. An important finding demonstrated that whelk GAG selectively impedes specific pathways in each cell line studied, such as FGF and FGFRs pathways in MDA-MB-468 cells, leading to the inhibition of cell growth, cell arrest and decreased cell survival. In addition to their selective activity on Notch signalling pathway that has been reported from MDA-MB-231 cell line. Furthermore, whelk GAG affected the extracellular organisation, which is important in cell proliferation and survival. In addition to their influence on glycolysis metabolism in MDA-MB-468 and their potential attribute in downregulating the DNA replication process, whelk GAGs induce cell arrest and anti-proliferation among cells. Taken together, the results of this study have confirmed the anti-proliferation activity of whelk GAG and explained its coherent modes of action, which support the application of whelk GAG as a naturally sourced anticancer agent.

# **Chapter 6 : Conclusion and Future Works**

This study has evaluated the role of whelk GAGs in the growth inhibition of several cancers and malaria *in vitro*. An additional aim was to investigate the structure of whelk GAGs and to highlight the key differences between these GAGs and mammalian GAGs, and to associate structural differences with their biological activity.

Many naturally derived active compounds have been discovered because of intensive research by pharmaceutical companies and academia. Studies have addressed the contribution of marine GAGs in the development of new bioactive candidates with anti-proliferation activity toward different cancers and malaria. However, despite great interest in the investigation of natural products from marine organisms, many marine species and their constituents remain undiscovered.

Data from advanced analytical techniques has enabled significant new insights to be gained into the role that sulphation patterns and the organisation/abundance of different disaccharide units in molecules may play in influencing oligosaccharide conformation and, more importantly, function. These have been shown to be significant factors and to play a major role in biological functions, with differences identified between inactive and active oligosaccharides in whelk GAGs in comparison to mammalian GAGs.

For the first time, this study has shown that crude whelk GAG extracts inhibit the growth of HepG2 cells, MCF-7, SKBR-3, K562, A549 cells and plasmodial growth. Although the exact mechanism of action of whelk GAG toward these cancers and towards malaria is not fully understood, the most likely explanation for this finding is that whelk GAG extracts act as an anti-proliferation agent. Administration of exogenous GAGs from whelk with active structural features in effective doses may lead to a competition process in which the functional endothelial GAGs bind to the growth factors that are commonly seen in the cell differentiation event. This could therefore give rise to an anti-proliferation outcome. Administration of some marine GAGs led to many cellular changes such as reduction of

cancer cell adhesion, suppression of cell migration and tube formation in those cells, suppression of angiogenesis, inhibition of cell proliferation and tumour invasion. These findings can further contribute to the development of anticancer and antimalarial agents. Further studies are required in future to investigate their mode of action.

Our data indicates distinct differences in the structure of whelk GAG and mammalian GAGs, which lie mainly in the abundance of different disaccharide units and the sulphation patterns present in these molecules. This finding might explain, at least in part, the incomplete digestion of whelk GAGs by heparinase enzymes. Our data has indicated that a reasonable amount of GAG in the retained fraction of the Nanosep centrifugal devices contain omega membrane (10K filter Eppendorf tube), which again could be the result of incomplete digestion by heparinase enzymes. Unlike the flow-through fractions, the retained fraction has proven but reduced, in comparison to intact GAG chains, biological activity toward cancer cells. This suggests that retained fractions contain the key heparinase-resistant sequences, which are essential for their biological activity. The structure of these fractions is still relatively unknown, and further research is necessary to fully explore the structure of this fraction.

This study assessed the impact of crude whelk GAG extracts and its fragments on triple negative breast cancer (TNBC) cell proliferation and investigated their mode of action on TNBC using label-free proteomic approaches. In addition, it examined the activity of whelk GAGs toward cancer stem cells by using mammosphere assays on different breast cancer subtypes.

Significant inhibition activity toward the TNBC cell lines MDA-MB-231 and MDA-MB-468 and the mammosphere formation of several subtypes of breast cancer has been indicated in whelk GAGs relative to inactive GAG fragments from mammalian sources.

The label-free proteomic data has enabled significant new insights to be gained into the mechanism of action of whelk GAG activities toward TNBC. Its role has suggested many pathways that might be involved in whelk GAG inhibition mechanisms on MDA-MB-468 and MDA-MB-231 cell lines. An important finding demonstrated that whelk GAG selectively impedes the FGF and FGFRs pathways, leading to inhibited cell growth and decreased cell survival. In addition to their selective activity on Notch signalling pathway that has been reported from MDA-MB-231 cell line. Notch signalling is known to regulate many cellular processes of breast cancer including proliferation, metastasis, cancer stem cell regulation and angiogenesis, it is very clear that activation of Notch plays a key role in breast cancer. Therefore, it represents a very attractive therapeutic target for many breast cancer subtypes. Furthermore, whelk GAG affected extracellular organisation and many pathways that are important in cell proliferation and survival. The study also demonstrated the influence of whelk GAG on glycolysis metabolism in MDA-MB-468, also shown potential attributes of whelk GAG in down-regulating the DNA replication process, which further induces cell arrest and anti-proliferation among cells.

These findings contribute considerably to the evaluation and plans for further structure analysis of whelk GAG in the development of potential therapeutic approaches, which may have an important role as future anticancer and antimalarial agents.

Finally, further studies are required to further analyse the structure of whelk GAGs in particular the bioactive compounds detected in the retained fractions. More research in this area is necessary to answer several questions that remained unresolved in terms of structure-function relationship. In addition, more research require on screening the bioactivity of whelk GAGs toward TNBC cells that have resistance toward chemotherapy, this resistance toward chemotherapeutic agent could be developed *in vitro* by treating cells with low doses of chemotherapy agents such as Cisplatin.

## References

- Adamo, M., Qiu, D., Dick, L. W., Zeng, M., Lee, A.-H., & Cheng, K.-C. (2009). Evaluation of oligosaccharide methods for carbohydrate analysis in a fully human monoclonal antibody and comparison of the results to the monosaccharide composition determination by a novel calculation. *Journal of pharmaceutical and biomedical analysis*, *49*(2), 181-192.
- Adjei, A. A., & Hidalgo, M. (2005). Treating cancer by blocking cell signals. *Journal of Clinical Oncology*, *23*(23), 5279-5280.
- Ai, X., Do, A.-T., Lozynska, O., Kusche-Gullberg, M., Lindahl, U., & Emerson, C. P. (2003). QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. *The Journal of cell biology*, *162*(2), 341-351.
- Aminin, D. L., Menchinskaya, E. S., Pisiagin, E. A., Silchenko, A. S., Avilov, S. A., & Kalinin, V. I. (2015). Anticancer activity of sea cucumber triterpene glycosides. *Marine Drugs*, *13*(3), 1202-1223.
- Antoniou, A., Pharoah, P., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J., . . . Borg, Å. (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *The American Journal of Human Genetics*, *72*(5), 1117-1130.
- Arasta, M. (2015). Importance of different growth factors in oestrogen receptor-negative breast cancer.
- Badve, S., Dabbs, D. J., Schnitt, S. J., Baehner, F. L., Decker, T., Eusebi, V., . . . Lakhani, S. R. (2011). Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. *Modern pathology*, *24*(2), 157-167.
- Bancroft, J. D., & Gamble, M. (2008). *Theory and practice of histological techniques*: Elsevier Health Sciences.
- Bartkova, J., Hořejší, Z., Koed, K., Krämer, A., Tort, F., Zieger, K., . . . Lukas, C. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature*, *434*(7035), 864-870.
- Bautista, J. M., Marín-García, P., Diez, A., Azcárate, I. G., & Puyet, A. (2014). Malaria proteomics: insights into the parasite–host interactions in the pathogenic space. *Journal of proteomics*, *97*, 107-125.
- Bender, E. (2014). Epidemiology: the dominant malignancy. *Nature*, *513*(7517), S2-S3.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289-300.
- Bernfield, M., Götte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J., & Zako, M. (1999). Functions of cell surface heparan sulfate proteoglycans. *Annual review of biochemistry*, *68*(1), 729-777.
- Berteau, O., & Mulloy, B. (2003). Sulfated fucans, fresh perspectives: structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. *Glycobiology*, *13*(6), 29R-40R.
- Bilan, M. I., Vinogradova, E. V., Shashkov, A. S., & Usov, A. I. (2007). Structure of a highly pyruvylated galactan sulfate from the Pacific green alga *Codium yezoense* (Bryopsidales, Chlorophyta). *Carbohydrate Research*, *342*(3), 586-596.
- Bixler, H. (1994). The carrageenan connection IV. *British Food Journal*, *96*(3), 12-17.
- Borggreffe, T., Lauth, M., Zwijsen, A., Huylebroeck, D., Oswald, F., & Giaimo, B. D. (2016). The Notch intracellular domain integrates signals from Wnt, Hedgehog, TGFβ/BMP and hypoxia pathways. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1863*(2), 303-313.
- Boyle, M. J., Richards, J. S., Gilson, P. R., Chai, W., & Beeson, J. G. (2010). Interactions with heparin-like molecules during erythrocyte invasion by *Plasmodium falciparum* merozoites. *Blood*, *115*(22), 4559-4568.

- Boyle, P., & Ferlay, J. (2005). Cancer incidence and mortality in Europe, 2004. *Annals of oncology*, 16(3), 481-488.
- Bremm, A., Walch, A., Fuchs, M., Mages, J., Duyster, J., Keller, G., . . . Langer, R. (2008). Enhanced activation of epidermal growth factor receptor caused by tumor-derived E-cadherin mutations. *Cancer research*, 68(3), 707-714.
- Bringmann, G., Gulder, T. A., Lang, G., Schmitt, S., Stöhr, R., Wiese, J., . . . Imhoff, J. F. (2007). Large-scale biotechnological production of the antileukemic marine natural product sorbicillactone A. *Marine Drugs*, 5(2), 23-30.
- Brito, A. S., Arimatéia, D. S., Souza, L. R., Lima, M. A., Santos, V. O., Medeiros, V. P., . . . Justo, G. Z. (2008). Anti-inflammatory properties of a heparin-like glycosaminoglycan with reduced anticoagulant activity isolated from a marine shrimp. *Bioorganic & medicinal chemistry*, 16(21), 9588-9595.
- Brueggemeier, R. W. (1994). Aromatase inhibitors—mechanisms of steroidal inhibitors. *Breast cancer research and treatment*, 30(1), 31-42.
- Cao, R., Bråkenhielm, E., Pawliuk, R., Wariaro, D., Post, M. J., Wahlberg, E., . . . Cao, Y. (2003). Angiogenic synergism, vascular stability and improvement of hind-limb ischemia by a combination of PDGF-BB and FGF-2. *Nature medicine*, 9(5), 604-613.
- Carey, D. J. (1997). Syndecans: multifunctional cell-surface co-receptors. *Biochemical Journal*, 327(1), 1-16.
- Cassaró, C., & Dietrich, C. P. (1977). Distribution of sulfated mucopolysaccharides in invertebrates. *Journal of Biological Chemistry*, 252(7), 2254-2261.
- Cavallaro, U., & Christofori, G. (2004). Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nature Reviews Cancer*, 4(2), 118-132.
- Cesaretti, M., Luppi, E., Maccari, F., & Volpi, N. (2004). Isolation and characterization of a heparin with high anticoagulant activity from the clam *Tapes philippinarum*: evidence for the presence of a high content of antithrombin III binding site. *Glycobiology*, 14(12), 1275-1284.
- Chavaroche, A. A., van den Broek, L. A., & Eggink, G. (2013). Production methods for heparosan, a precursor of heparin and heparan sulfate. *Carbohydrate Polymers*, 93(1), 38-47.
- Chen, M., Knifley, T., Subramanian, T., Spielmann, H. P., & O'Connor, K. L. (2014). Use of synthetic isoprenoids to target protein prenylation and Rho GTPases in breast cancer invasion. *PLoS One*, 9(2), e89892.
- Chevolot, L., Foucault, A., Chaubet, F., Kervarec, N., Siquin, C., Fisher, A.-M., & Boisson-Vidal, C. (1999). Further data on the structure of brown seaweed fucans: relationships with anticoagulant activity. *Carbohydrate Research*, 319(1), 154-165.
- Chlebowski, R. T., Manson, J. E., Anderson, G. L., Cauley, J. A., Aragaki, A. K., Stefanick, M. L., . . . Chen, C. (2013). Estrogen plus progestin and breast cancer incidence and mortality in the Women's Health Initiative Observational Study. *Journal of the National Cancer Institute*, 105(8), 526-535.
- Cho, M. L., Lee, B.-Y., & You, S. G. (2010). Relationship between oversulfation and conformation of low and high molecular weight fucoidans and evaluation of their in vitro anticancer activity. *Molecules*, 16(1), 291-297.
- Christensen, S. B., & Kharazmi, A. (2001). Antimalarial natural products. *Bioactive Compounds from Natural Sources*, 379.
- Clark, G. M., Osborne, C. K., & McGuire, W. L. (1984). Correlations between estrogen receptor, progesterone receptor, and patient characteristics in human breast cancer. *Journal of Clinical Oncology*, 2(10), 1102-1109.
- Cleator, S., Heller, W., & Coombes, R. C. (2007). Triple-negative breast cancer: therapeutic options. *The lancet oncology*, 8(3), 235-244.
- Coombe, D. R., Parish, C. R., Ramshaw, I. A., & Snowden, J. M. (1987). Analysis of the inhibition of tumour metastasis by sulphated polysaccharides. *International Journal of Cancer*, 39(1), 82-88.

- Cotran, R., Kumar, V., & Robbins, S. (2010). *Pathologic Basis of Disease*. Philadelphia, PA: Saunders: Elsevier.
- Daley, W. P., Peters, S. B., & Larsen, M. (2008). Extracellular matrix dynamics in development and regenerative medicine. *Journal of cell science*, *121*(3), 255-264.
- Dang, C. V. (2012). Links between metabolism and cancer. *Genes & development*, *26*(9), 877-890.
- DerMardirossian, C., & Bokoch, G. M. (2001). Regulation of cell function by Rho GTPases. *Drug News Perspect*, *14*(7), 389.
- Dietrich, C. P., de Paiva, J., Moraes, C. T., Takahashi, H. K., Porcionatto, M. A., & Nader, H. B. (1985). Isolation and characterization of a heparin with high anticoagulant activity from *Anomalocardia brasiliensis*. *Biochimica Et Biophysica Acta (BBA)-General Subjects*, *843*(1-2), 1-7.
- Dietrich, C. P., Nader, H. B., de Paiva, J. F., Santos, E. A., Holme, K. R., & Perlin, A. S. (1989). Heparin in molluscs: chemical, enzymatic degradation and <sup>13</sup>C and <sup>1</sup>H nmr spectroscopical evidence for the maintenance of the structure through evolution. *International Journal of Biological Macromolecules*, *11*(6), 361-366.
- Dong, Y., Ren, W., Qi, J., Jin, B., Li, Y., Tao, H., . . . Han, B. (2016). EGFR, ALK, RET, KRAS and BRAF alterations in never-smokers with non-small cell lung cancer. *Oncology letters*, *11*(4), 2371-2378.
- Drewes, G. (2012). Chemical proteomics in drug discovery. *Chemical Proteomics: Methods and Protocols*, 15-21.
- Dufraine, J., Funahashi, Y., & Kitajewski, J. (2008). Notch signaling regulates tumor angiogenesis by diverse mechanisms. *Oncogene*, *27*(38), 5132-5137.
- Duncan, G., McCormick, C., & Tufaro, F. (2001). The link between heparan sulfate and hereditary bone disease: finding a function for the EXT family of putative tumor suppressor proteins. *The Journal of clinical investigation*, *108*(4), 511-516.
- Earl, W. L., & VanderHart, D. (1982). Measurement of <sup>13</sup>C chemical shifts in solids. *Journal of Magnetic Resonance (1969)*, *48*(1), 35-54.
- Edwards, J. C. (2009). *Principles of NMR*. Process NMR Associates LLC, 87A Sand Pit Rd, Danbury CT, 6810.
- Eroles, P., Bosch, A., Pérez-Fidalgo, J. A., & Lluch, A. (2012). Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. *Cancer treatment reviews*, *38*(6), 698-707.
- Esko, J. D., Kimata, K., & Lindahl, U. (2009). Proteoglycans and sulfated glycosaminoglycans.
- Fabbro, D., & Garcia-Echeverria, C. (2002). Targeting protein kinases in cancer therapy. *Current opinion in drug discovery & development*, *5*(5), 701-712.
- Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnau, V., . . . Hermjakob, H. (2017). Reactome pathway analysis: a high-performance in-memory approach. *BMC bioinformatics*, *18*(1), 142.
- Farias, E. H., Pomin, V. H., Valente, A.-P., Nader, H. B., Rocha, H. A., & Mourão, P. A. (2008). A preponderantly 4-sulfated, 3-linked galactan from the green alga *Codium isthmocladum*. *Glycobiology*, *18*(3), 250-259.
- Fritz, T. A., Gabb, M. M., Wei, G., & Esko, J. D. (1994). Two N-acetylglucosaminyltransferases catalyze the biosynthesis of heparan sulfate. *Journal of Biological Chemistry*, *269*(46), 28809-28814.
- Galliher, P., Cooney, C., Langer, R., & Linhardt, R. (1981). Heparinase production by *Flavobacterium heparinum*. *Applied and environmental microbiology*, *41*(2), 360-365.
- Gallo, L. H., Nelson, K. N., Meyer, A. N., & Donoghue, D. J. (2015). Functions of Fibroblast Growth Factor Receptors in cancer defined by novel translocations and mutations. *Cytokine & growth factor reviews*, *26*(4), 425-449.
- Gao, Q., Nilsson, U., Ilag, L. L., & Leck, C. (2011). Monosaccharide compositional analysis of marine polysaccharides by hydrophilic interaction liquid chromatography-tandem mass spectrometry. *Analytical and bioanalytical chemistry*, *399*(7), 2517-2529.

- Gierisch, J. M., Coeytaux, R. R., Urrutia, R. P., Havrilesky, L. J., Moorman, P. G., Lowery, W. J., . . . Sanders, G. D. (2013). Oral contraceptive use and risk of breast, cervical, colorectal, and endometrial cancers: a systematic review. *Cancer Epidemiology Biomarkers & Prevention*, 22(11), 1931-1943.
- Gill, V. L., Aich, U., Rao, S., Pohl, C., & Zaia, J. (2013). Disaccharide analysis of glycosaminoglycans using hydrophilic interaction chromatography and mass spectrometry. *Analytical chemistry*, 85(2), 1138-1145.
- Goodger, S. J., Robinson, C. J., Murphy, K. J., Gasiunas, N., Harmer, N. J., Blundell, T. L., . . . Gallagher, J. T. (2008). Evidence that heparin saccharides promote FGF2 mitogenesis through two distinct mechanisms. *Journal of Biological Chemistry*, 283(19), 13001-13008.
- Goupille, R. E., Smith, A. E., Toida, T., & Linhardt, R. J. (1997). Preparation and structure of heparin lyase-derived heparan sulfate oligosaccharides. *Glycobiology*, 7(2), 231-239.
- Griffith, K. S., Lewis, L. S., Mali, S., & Parise, M. E. (2007). Treatment of malaria in the United States: a systematic review. *Jama*, 297(20), 2264-2277.
- Group, E. B. C. T. C. (2011). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. *The Lancet*, 378(9804), 1707-1716.
- Gumbiner, B. M., & McCrea, P. D. (1993). Catenins as mediators of the cytoplasmic functions of cadherins. *J Cell Sci*, 1993(Supplement 17), 155-158.
- Habuchi, H., Tanaka, M., Habuchi, O., Yoshida, K., Suzuki, H., Ban, K., & Kimata, K. (2000). The Occurrence of Three Isoforms of Heparan Sulfate 6-O-Sulfotransferase Having Different Specificities for Hexuronic Acid Adjacent to the Targeted N-Sulfoglucosamine. *Journal of Biological Chemistry*, 275(4), 2859-2868.
- Hamai, A., Hashimoto, N., Mochizuki, H., Kato, F., Makiguchi, Y., Horie, K., & Suzuki, S. (1997). Two distinct chondroitin sulfate ABC lyases an endoeliminase yielding tetrasaccharides and an exoeliminase preferentially acting on oligosaccharides. *Journal of Biological Chemistry*, 272(14), 9123-9130.
- Haroun-Bouhedja, F., Ellouali, M., Sinquin, C., & Boisson-Vidal, C. (2000). Relationship between sulfate groups and biological activities of fucans. *Thrombosis research*, 100(5), 453-459.
- Harrison, H., Farnie, G., Howell, S. J., Rock, R. E., Stylianou, S., Brennan, K. R., . . . Clarke, R. B. (2010). Regulation of breast cancer stem cell activity by signaling through the Notch4 receptor. *Cancer research*, 70(2), 709-718.
- Harrison, H., Simões, B. M., Rogerson, L., Howell, S. J., Landberg, G., & Clarke, R. B. (2013). Oestrogen increases the activity of oestrogen receptor negative breast cancer stem cells through paracrine EGFR and Notch signalling. *Breast cancer research*, 15(2), R21.
- Hawa, I., Zulaikah, M., Jamaludin, M., Zainal Abidin, A., Kaswandi, M., & Ridzwan, B. (1999). The potential of the coelomic fluid in sea cucumber as an antioxidant. *Malaysian Journal of Nutrition*, 5(1&2), 55-59.
- Herschkowitz, J. I., Simin, K., Weigman, V. J., Mikaelian, I., Usary, J., Hu, Z., . . . Chandrasekharan, S. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome biology*, 8(5), 1.
- Holliday, D. L., & Speirs, V. (2011). Choosing the right cell line for breast cancer research. *Breast cancer research*, 13(4), 215.
- Honya, M., Mori, H., Anzai, M., Araki, Y., & Nisizawa, K. (1999). *Monthly changes in the content of fucans, their constituent sugars and sulphate in cultured Laminaria japonica*. Paper presented at the Sixteenth International Seaweed Symposium.
- Huo, D., Ikpat, F., Khramtsov, A., Dangou, J.-M., Nanda, R., Dignam, J., . . . Oluwasola, O. (2009). Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer. *Journal of Clinical Oncology*, 27(27), 4515-4521.
- Hynes, R. O. (2009). The extracellular matrix: not just pretty fibrils. *Science*, 326(5957), 1216-1219.

- Hyun, J.-H., Kim, S.-C., Kang, J.-I., Kim, M.-K., Boo, H.-J., Kwon, J.-M., . . . Yoo, E.-S. (2009). Apoptosis inducing activity of fucoidan in HCT-15 colon carcinoma cells. *Biological and Pharmaceutical Bulletin*, 32(10), 1760-1764.
- Jamieson, A. G., Boutard, N., Sabatino, D., & Lubell, W. D. (2013). Peptide Scanning for Studying Structure – Activity Relationships in Drug Discovery. *Chemical biology & drug design*, 81(1), 148-165.
- Jastrebova, N., Vanwildemeersch, M., Rapraeger, A. C., Giménez-Gallego, G., Lindahl, U., & Spillmann, D. (2006). Heparan sulfate-related oligosaccharides in ternary complex formation with fibroblast growth factors 1 and 2 and their receptors. *Journal of Biological Chemistry*, 281(37), 26884-26892.
- Jemal, A., Center, M. M., DeSantis, C., & Ward, E. M. (2010). Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiology Biomarkers & Prevention*, 19(8), 1893-1907.
- Jemth, P., Kreuger, J., Kusche-Gullberg, M., Sturiale, L., Giménez-Gallego, G., & Lindahl, U. (2002). Biosynthetic Oligosaccharide Libraries for Identification of Protein-binding Heparan Sulfate Motifs EXPLORING THE STRUCTURAL DIVERSITY BY SCREENING FOR FIBROBLAST GROWTH FACTOR (FGF) 1 AND FGF2 BINDING. *Journal of Biological Chemistry*, 277(34), 30567-30573.
- Jiao, G., Yu, G., Zhang, J., & Ewart, H. S. (2011). Chemical structures and bioactivities of sulfated polysaccharides from marine algae. *Marine Drugs*, 9(2), 196-223.
- Johnston, S. R., & Yeo, B. (2014). The Optimal Duration of Adjuvant Endocrine Therapy for Early Stage Breast Cancer—With What Drugs and for How Long? *Current oncology reports*, 16(1), 1-8.
- Jordan, R. E., & Marcum, J. A. (1986). Anticoagulant active heparin from clam (*Mercenaria mercenaria*). *Archives of biochemistry and biophysics*, 248(2), 690-695.
- Kakuta, Y., Sueyoshi, T., Negishi, M., & Pedersen, L. C. (1999). Crystal structure of the sulfotransferase domain of human heparan sulfate N-deacetylase/N-sulfotransferase 1. *Journal of Biological Chemistry*, 274(16), 10673-10676.
- Kantarjian, H. M., Deisseroth, A., Kurzrock, R., Estrov, Z., & Talpaz, M. (1993). Chronic myelogenous leukemia: a concise update. *Blood* 1993 Aug, 1(82), 3.
- Karl, S., Wong, R. P., St Pierre, T. G., & Davis, T. M. (2009). A comparative study of a flow-cytometry-based assessment of in vitro Plasmodium falciparum drug sensitivity. *Malaria journal*, 8(1), 294.
- Katajamaa, M., & Orešič, M. (2005). Processing methods for differential analysis of LC/MS profile data. *BMC bioinformatics*, 6(1), 1.
- Kayser, O., Kiderlen, A., & Croft, S. (2003). Natural products as antiparasitic drugs. *Parasitology research*, 90(2), S55-S62.
- Kemler, R., & Ozawa, M. (1989). Uvomorulin - catenin complex: Cytoplasmic anchorage of a Ca<sup>2+</sup> - dependent cell adhesion molecule. *Bioessays*, 11(4), 88-91.
- KerheRvé, P., Charrière, B., & Gadel, F. (1995). Determination of marine monosaccharides by high-pH anion-exchange chromatography with pulsed amperometric detection. *Journal of Chromatography A*, 718(2), 283-289.
- Key, T. J., Verkasalo, P. K., & Banks, E. (2001). Epidemiology of breast cancer. *The lancet oncology*, 2(3), 133-140.
- Kim, B.-T., Kitagawa, H., Tamura, J.-i., Saito, T., Kusche-Gullberg, M., Lindahl, U., & Sugahara, K. (2001). Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode  $\alpha$ 1, 4-N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/heparin biosynthesis. *Proceedings of the National Academy of Sciences*, 98(13), 7176-7181.
- Kim, E. J., Park, S. Y., Lee, J.-Y., & Park, J. H. Y. (2010). Fucoidan present in brown algae induces apoptosis of human colon cancer cells. *BMC gastroenterology*, 10(1), 96.
- Kim, Y. S., Jo, Y. Y., Chang, I. M., Toida, T., Park, Y., & Linhardt, R. J. (1996). A new glycosaminoglycan from the giant African snail *Achatina fulica*. *Journal of Biological Chemistry*, 271(20), 11750-11755.

- Kinoshita, A., Yamada, S., Haslam, S. M., Morris, H. R., Dell, A., & Sugahara, K. (1997). Novel tetrasaccharides isolated from squid cartilage chondroitin sulfate E contain unusual sulfated disaccharide units GlcA (3-O-sulfate)  $\beta$ 1–3GalNAc (6-O-sulfate) or GlcA (3-O-sulfate)  $\beta$ 1–3GalNAc (4, 6-O-disulfate). *Journal of Biological Chemistry*, *272*(32), 19656-19665.
- Kiviranta, I., Tammi, M., Jurvelin, J., Säämänen, A.-M., & Helminen, H. (1985). Demonstration of chondroitin sulphate and glycoproteins in articular cartilage matrix using periodic acid-Schiff (PAS) method. *Histochemistry*, *83*(3-4), 303-306.
- Kjellen, L. (2003). Glucosaminyl N-deacetylase/N-sulphotransferases in heparan sulphate biosynthesis and biology. *Biochemical Society Transactions*, *31*(2), 340-342.
- Knutsen, S., Myslabodski, D., Larsen, B., & Usov, A. (1994). A modified system of nomenclature for red algal galactans. *Botanica marina*, *37*(2), 163-170.
- Kobayashi, M., Habuchi, H., Yoneda, M., Habuchi, O., & Kimata, K. (1997). Molecular cloning and expression of Chinese hamster ovary cell heparan-sulfate 2-sulfotransferase. *Journal of Biological Chemistry*, *272*(21), 13980-13985.
- Koyanagi, S., Tanigawa, N., Nakagawa, H., Soeda, S., & Shimeno, H. (2003). Oversulfation of fucoidan enhances its anti-angiogenic and antitumor activities. *Biochemical pharmacology*, *65*(2), 173-179.
- Kreuger, J., Spillmann, D., Li, J.-p., & Lindahl, U. (2006). Interactions between heparan sulfate and proteins: the concept of specificity. *The Journal of cell biology*, *174*(3), 323-327.
- Kulshreshtha, G., Burlot, A.-S., Marty, C., Critchley, A., Hafting, J., Bedoux, G., . . . Prithiviraj, B. (2015). Enzyme-assisted extraction of bioactive material from *Chondrus crispus* and *Codium fragile* and its effect on Herpes simplex virus (HSV-1). *Marine Drugs*, *13*(1), 558-580.
- Lahaye, M. (2001). Developments on gelling algal galactans, their structure and physico-chemistry. *Journal of applied phycology*, *13*(2), 173-184.
- Lamanna, W. C., Baldwin, R. J., Padva, M., Kalus, I., Ten Dam, G., Van Kuppevelt, T. H., . . . Merry, C. L. (2006). Heparan sulfate 6-O-endosulfatases: discrete in vivo activities and functional cooperativity. *Biochemical Journal*, *400*(1), 63-73.
- Lauder, R. M. (2009). Chondroitin sulphate: a complex molecule with potential impacts on a wide range of biological systems. *Complementary therapies in medicine*, *17*(1), 56-62.
- Lawrence, R., Lu, H., Rosenberg, R. D., Esko, J. D., & Zhang, L. (2008). Disaccharide structure code for the easy representation of constituent oligosaccharides from glycosaminoglycans. *Nature methods*, *5*(4), 291-292.
- Lawrence, R., Olson, S. K., Steele, R. E., Wang, L., Warrior, R., Cummings, R. D., & Esko, J. D. (2008). Evolutionary differences in glycosaminoglycan fine structure detected by quantitative glycan reductive isotope labeling. *Journal of Biological Chemistry*, *283*(48), 33674-33684.
- Leblond, C., Glegg, R., & Eidinger, D. (1957). Presence of carbohydrates with free 1, 2-glycol groups in sites stained by the periodic acid-Schiff technique. *Journal of Histochemistry & Cytochemistry*, *5*(5), 445-458.
- Legg, J., Jensen, U. B., Broad, S., Leigh, I., & Watt, F. M. (2003). Role of melanoma chondroitin sulphate proteoglycan in patterning stem cells in human interfollicular epidermis. *Development*, *130*(24), 6049-6063.
- Li, H., Peyrollier, K., Kilic, G., & Brakebusch, C. (2014). Rho GTPases and cancer. *Biofactors*, *40*(2), 226-235.
- Li, X., Lewis, M. T., Huang, J., Gutierrez, C., Osborne, C. K., Wu, M.-F., . . . Chamness, G. C. (2008). Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *Journal of the National Cancer Institute*, *100*(9), 672-679.
- Lindahl, U., Kusche-Gullberg, M., & Kjellén, L. (1998). Regulated diversity of heparan sulfate. *Journal of Biological Chemistry*, *273*(39), 24979-24982.
- Linker, A. (1979). Structure of heparan sulphate oligosaccharides and their degradation by exoenzymes. *Biochemical Journal*, *183*(3), 711-720.

- Lins, K. O., Bezerra, D. P., Alves, A. P. N., Alencar, N., Lima, M. W., Torres, V. M., . . . Costa - Lotufo, L. V. (2009). Antitumor properties of a sulfated polysaccharide from the red seaweed *Champia feldmannii* (Diaz - Pifferer). *Journal of Applied Toxicology*, *29*(1), 20-26.
- Liu, J., Shworak, N. W., Fritze, L. M., Edelberg, J. M., & Rosenberg, R. D. (1996). Purification of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. *Journal of Biological Chemistry*, *271*(43), 27072-27082.
- Liu, J., Shworak, N. W., Sinay, P., Schwartz, J. J., Zhang, L., Fritze, L. M., & Rosenberg, R. D. (1999). Expression of heparan sulfate D-glucosaminyl 3-O-sulfotransferase isoforms reveals novel substrate specificities. *Journal of Biological Chemistry*, *274*(8), 5185-5192.
- Liu, L., & Rein, K. S. (2010). New peptides isolated from *Lyngbya* species: a review. *Marine Drugs*, *8*(6), 1817-1837.
- Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Matsudaira, P., & Darnell, J. (1995). *Molecular cell biology* (Vol. 3): Scientific American Books New York.
- López-Terrada, D., Cheung, S. W., Finegold, M. J., & Knowles, B. B. (2009). Hep G2 is a hepatoblastoma-derived cell line. *Human pathology*, *40*(10), 1512-1515.
- Ludwig, R. J., Boehme, B., Podda, M., Henschler, R., Jager, E., Tandi, C., . . . Gille, J. (2004). Endothelial P-selectin as a target of heparin action in experimental melanoma lung metastasis. *Cancer research*, *64*(8), 2743-2750.
- Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., . . . Haluska, F. G. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *New England Journal of Medicine*, *350*(21), 2129-2139.
- Mabeau, S., Kloareg, B., & Joseleau, J.-P. (1990). Fractionation and analysis of fucans from brown algae. *Phytochemistry*, *29*(8), 2441-2445.
- Maeda, N. (2015). Proteoglycans and neuronal migration in the cerebral cortex during development and disease. *Frontiers in neuroscience*, *9*.
- Maheswaran, S., & Haber, D. A. (2010). Circulating tumor cells: a window into cancer biology and metastasis. *Current opinion in genetics & development*, *20*(1), 96-99.
- Malhotra, G. K., Zhao, X., Band, H., & Band, V. (2010). Histological, molecular and functional subtypes of breast cancers. *Cancer biology & therapy*, *10*(10), 955-960.
- Malumbres, M., & Barbacid, M. (2009). Cell cycle, CDKs and cancer: a changing paradigm. *Nature Reviews Cancer*, *9*(3), 153-166.
- MarinLit, V. S. (2003). A marine literature database produced and maintained by the Department of Chemistry. *University of Canterbury, New Zealand*.
- Marques, J., Vilanova, E., Mourão, P. A., & Fernández-Busquets, X. (2016). Marine organism sulfated polysaccharides exhibiting significant antimalarial activity and inhibition of red blood cell invasion by *Plasmodium*. *Scientific reports*, *6*.
- Mateus, A. R., Seruca, R., Machado, J. C., Keller, G., Oliveira, M. J., Suriano, G., & Luber, B. (2007). EGFR regulates RhoA-GTP dependent cell motility in E-cadherin mutant cells. *Human molecular genetics*, *16*(13), 1639-1647.
- Mathews, M. (1974). Connective tissue. Macromolecular structure and evolution. *Molecular biology, biochemistry, and biophysics*(19), 1-318.
- Medeiros, G. F., Mendes, A., Castro, R. A., Baú, E. C., Nader, H. B., & Dietrich, C. P. (2000). Distribution of sulfated glycosaminoglycans in the animal kingdom: widespread occurrence of heparin-like compounds in invertebrates. *Biochimica Et Biophysica Acta (BBA)-General Subjects*, *1475*(3), 287-294.
- Megger, D. A., Bracht, T., Meyer, H. E., & Sitek, B. (2013). Label-free quantification in clinical proteomics. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1834*(8), 1581-1590.

- Menezes, S., Kovacevic, Z., & Richardson, D. (2014). The Metastasis Suppressor NdrG1 Down-regulates The Expression Of Snail And Slug To Promote E-cadherin Expression In Pancreatic Cancer. *Asia-pacific Journal of Clinical Oncology*, *10*, 37-38.
- Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nature protocols*, *8*(8), 1551-1566.
- Mian, A. J., & Percival, E. (1973). Carbohydrates of the brown seaweeds *Himantalia lorea*, *Bifurcaria bifurcata*, and *Padina pavonia*: Part I. extraction and fractionation. *Carbohydrate Research*, *26*(1), 133-146.
- Miertus, S., & Fassina, G. (2005). *Combinatorial chemistry and technologies: methods and applications*: CRC Press.
- Mita, T., Tanabe, K., & Kita, K. (2009). Spread and evolution of *Plasmodium falciparum* drug resistance. *Parasitology international*, *58*(3), 201-209.
- Moehler, T. M., Seckinger, A., Hose, D., Andrulis, M., Moreaux, J., Hielscher, T., . . . Jauch, A. (2013). The glycome of normal and malignant plasma cells. *Plos One*, *8*(12), e83719.
- Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of immunological methods*, *65*(1-2), 55-63.
- Mourão, P. A. (2004). Use of sulfated fucans as anticoagulant and antithrombotic agents: future perspectives. *Current pharmaceutical design*, *10*(9), 967-981.
- Mulloy, B., Khan, S., & Perkins, S. J. (2011). Molecular architecture of heparin and heparan sulfate: Recent developments in solution structural studies. *Pure and Applied Chemistry*, *84*(1), 65-76.
- Munagala, R., Aqil, F., & Gupta, R. C. (2011). Promising molecular targeted therapies in breast cancer. *Indian journal of pharmacology*, *43*(3), 236.
- Nader, H. B., Lopes, C. C., Rocha, H. A., Santos, E. A., & Dietrich, C. P. (2004). Heparins and heparinoids: occurrence, structure and mechanism of antithrombotic and hemorrhagic activities. *Current pharmaceutical design*, *10*(9), 951-966.
- Nahnsen, S., Bielow, C., Reinert, K., & Kohlbacher, O. (2013). Tools for label-free peptide quantification. *Molecular & Cellular Proteomics*, *12*(3), 549-556.
- Nakato, H., & Kimata, K. (2002). Heparan sulfate fine structure and specificity of proteoglycan functions. *Biochimica Et Biophysica Acta (BBA)-General Subjects*, *1573*(3), 312-318.
- Narazaki, M., Segarra, M., & Tosato, G. (2008). Sulfated polysaccharides identified as inducers of neuropilin-1 internalization and functional inhibition of VEGF165 and semaphorin3A. *Blood*, *111*(8), 4126-4136.
- Newman, D. J., Cragg, G. M., & Snader, K. M. (2000). The influence of natural products upon drug discovery. *Natural product reports*, *17*(3), 215-234.
- Nielsen, T. O., Hsu, F. D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., . . . Dressler, L. (2004). Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clinical cancer research*, *10*(16), 5367-5374.
- Nuijen, B., Bouma, M., Manada, C., Jimeno, J., Schellens, J. H., Bult, A., & Beijnen, J. (2000). Pharmaceutical development of anticancer agents derived from marine sources. *Anti-Cancer Drugs*, *11*(10), 793-811.
- Nur Hanim Zainudin, K. N. S. S. a. F. C. G. (2014). Marine Sourced Glycosaminoglycans 'GAGS'. *Journal of Advanced Laboratory Research in Biology*(ISSN 0976-7614 ).
- O'Brien, C., Henrich, P. P., Passi, N., & Fidock, D. A. (2011). Recent clinical and molecular insights into emerging artemisinin resistance in *Plasmodium falciparum*. *Current opinion in infectious diseases*, *24*(6), 570.
- Ogundipe, O. (2015). *Isolation and characterization of a glycosaminoglycan with anticancer activity*. University of Salford.
- Ohkawa, T., Naomoto, Y., Takaoka, M., Nobuhisa, T., Noma, K., Motoki, T., . . . Shirakawa, Y. (2004). Localization of heparanase in esophageal cancer cells: respective roles in prognosis and differentiation. *Laboratory investigation*, *84*(10), 1289-1304.

- Oldham, R. K. (1984). Biologicals and biological response modifiers: fourth modality of cancer treatment. *Cancer treatment reports*, 68(1), 221-232.
- Ornitz, D. M. (2000). FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays*, 22(2), 108-112.
- Osborne, C. K. (1998). Tamoxifen in the treatment of breast cancer. *New England Journal of Medicine*, 339(22), 1609-1618.
- Osborne, C. K., & Schiff, R. (2011). Mechanisms of endocrine resistance in breast cancer. *Annual review of medicine*, 62, 233.
- Pancake, S., Holt, G., Mellouk, S., & Hoffman, S. (1993). Malaria sporozoites and circumsporozoite protein bind sulfated glycans: carbohydrate binding properties predicted from sequence homologies with other lectins. *Parassitologia*, 35, 77-80.
- Pankonin, M. S., Gallagher, J. T., & Loeb, J. A. (2005). Specific structural features of heparan sulfate proteoglycans potentiate neuregulin-1 signaling. *Journal of Biological Chemistry*, 280(1), 383-388.
- Parker, J. S., Mullins, M., Cheang, M. C., Leung, S., Voduc, D., Vickery, T., . . . Hu, Z. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of Clinical Oncology*, 27(8), 1160-1167.
- Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., & Negishi, M. (2000). Heparan/chondroitin sulfate biosynthesis structure and mechanism of human glucuronyltransferase I. *Journal of Biological Chemistry*, 275(44), 34580-34585.
- Pejler, G., Danielsson, A., Björk, I., Lindahl, U., Nader, H., & Dietrich, C. (1987). Structure and antithrombin-binding properties of heparin isolated from the clams *Anomalocardia brasiliana* and *Tivela mactroides*. *Journal of Biological Chemistry*, 262(24), 11413-11421.
- Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., & Blundell, T. L. (2000). Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature*, 407(6807), 1029-1034.
- Percival, E., & Ross, A. (1950). 145. Fucoidin. Part I. The isolation and purification of fucoidin from brown seaweeds. *Journal of the Chemical Society (Resumed)*, 717-720.
- Pereira, M. G., Benevides, N. M., Melo, M. R., Valente, A. P., Melo, F. R., & Mourão, P. A. (2005). Structure and anticoagulant activity of a sulfated galactan from the red alga, *Gelidium crinale*. Is there a specific structural requirement for the anticoagulant action? *Carbohydrate Research*, 340(12), 2015-2023.
- Pereira, M. S., Vilela-Silva, A.-C. E., Valente, A.-P., & Mourão, P. A. (2002). A 2-sulfated, 3-linked  $\alpha$ -L-galactan is an anticoagulant polysaccharide. *Carbohydrate Research*, 337(21), 2231-2238.
- Perou, C. M., Sørli, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., . . . Akslen, L. A. (2000). Molecular portraits of human breast tumours. *Nature*, 406(6797), 747-752.
- Perrimon, N., & Bernfield, M. (2000). Specificities of heparan sulphate proteoglycans in developmental processes. *Nature*, 404(6779), 725-728.
- Pikas, D. S., Eriksson, I., & Kjellén, L. (2000). Overexpression of different isoforms of glucosaminyl N-deacetylase/N-sulfotransferase results in distinct heparan sulfate N-sulfation patterns. *Biochemistry*, 39(15), 4552-4558.
- Pinhal, M. A., Smith, B., Olson, S., Aikawa, J.-i., Kimata, K., & Esko, J. D. (2001). Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. *Proceedings of the National Academy of Sciences*, 98(23), 12984-12989.
- Pomin, V. H. (2009). Review: An overview about the structure–function relationship of marine sulfated homopolysaccharides with regular chemical structures. *Biopolymers*, 91(8), 601-609.
- Pomin, V. H. (2012a). Fucanomics and galactanomics: Current status in drug discovery, mechanisms of action and role of the well-defined structures. *Biochimica Et Biophysica Acta (BBA)-General Subjects*, 1820(12), 1971-1979.
- Pomin, V. H. (2012b). Fucanomics and galactanomics: marine distribution, medicinal impact, conceptions, and challenges. *Marine Drugs*, 10(4), 793-811.

- Pomin, V. H. (2015). A Dilemma in the Glycosaminoglycan - Based Therapy: Synthetic or Naturally Unique Molecules? *Medicinal research reviews*, 35(6), 1195-1219.
- Pomin, V. H., & Mourão, P. A. (2008). Structure, biology, evolution, and medical importance of sulfated fucans and galactans. *Glycobiology*, 18(12), 1016-1027.
- Ponce, N. M., Pujol, C. A., Damonte, E. B., Flores, L., & Stortz, C. A. (2003). Fucoidans from the brown seaweed *Adenocystis utricularis*: extraction methods, antiviral activity and structural studies. *Carbohydrate Research*, 338(2), 153-165.
- Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., . . . Perou, C. M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research*, 12(5), 1.
- Proksch, P., Edrada, R., & Ebel, R. (2002). Drugs from the seas—current status and microbiological implications. *Applied Microbiology and Biotechnology*, 59(2-3), 125-134.
- Prydz, K., & Dalen, K. T. (2000). Synthesis and sorting of proteoglycans. *Journal of cell science*, 113(2), 193-205.
- Pye, D. A., Vivès, R. R., Hyde, P., & Gallagher, J. T. (2000). Regulation of FGF-1 mitogenic activity by heparan sulfate oligosaccharides is dependent on specific structural features: differential requirements for the modulation of FGF-1 and FGF-2. *Glycobiology*, 10(11), 1183-1192.
- Pye, D. A., Vives, R. R., Turnbull, J. E., Hyde, P., & Gallagher, J. T. (1998). Heparan sulfate oligosaccharides require 6-O-sulfation for promotion of basic fibroblast growth factor mitogenic activity. *Journal of Biological Chemistry*, 273(36), 22936-22942.
- Rabenstein, D. L. (2002). Heparin and heparan sulfate: structure and function. *Natural product reports*, 19(3), 312-331.
- Rakha, E. A., Reis-Filho, J. S., & Ellis, I. O. (2008). Basal-like breast cancer: a critical review. *Journal of Clinical Oncology*, 26(15), 2568-2581.
- Read, M., & Hyde, J. E. (1993). Simple in vitro cultivation of the malaria parasite *Plasmodium falciparum* (erythrocytic stages) suitable for large-scale preparations. *Protocols in molecular parasitology*, 43-55.
- Reedijk, M., Pinnaduwege, D., Dickson, B. C., Mulligan, A. M., Zhang, H., Bull, S. B., . . . Andrulis, I. L. (2008). JAG1 expression is associated with a basal phenotype and recurrence in lymph node-negative breast cancer. *Breast cancer research and treatment*, 111(3), 439-448.
- Rieckmann, T., Kotevic, I., & Trueb, B. (2008). The cell surface receptor FGFR1 forms constitutive dimers that promote cell adhesion. *Experimental cell research*, 314(5), 1071-1081.
- Rinehart Jr, K. L., Gloer, J. B., Cook Jr, J. C., Mizzak, S. A., & Scahill, T. A. (1981). Structures of the didemnins, antiviral and cytotoxic depsipeptides from a Caribbean tunicate. *Journal of the American Chemical Society*, 103(7), 1857-1859.
- Ringvall, M., Ledin, J., Holmborn, K., van Kuppevelt, T., Ellin, F., Eriksson, I., . . . Forsberg, E. (2000). Defective heparan sulfate biosynthesis and neonatal lethality in mice lacking N-deacetylase/N-sulfotransferase-1. *Journal of Biological Chemistry*, 275(34), 25926-25930.
- Rizzo, P., Miao, H., D'Souza, G., Osipo, C., Yun, J., Zhao, H., . . . Hao, L. (2008). Cross-talk between notch and the estrogen receptor in breast cancer suggests novel therapeutic approaches. *Cancer research*, 68(13), 5226-5235.
- Robinson, C. J., Mulloy, B., Gallagher, J. T., & Stringer, S. E. (2006). VEGF165-binding sites within heparan sulfate encompass two highly sulfated domains and can be liberated by K5 lyase. *Journal of Biological Chemistry*, 281(3), 1731-1740.
- Robinson, D. R., Wu, Y.-M., Vats, P., Su, F., Lonigro, R. J., Cao, X., . . . Hodges, L. (2013). Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nature genetics*, 45(12), 1446-1451.
- Roby, J. F. (1998). Transformations *Essentials of Carbohydrate Chemistry* (pp. 48-75): Springer.
- Roden, L., & Schwartz, N. (1975). Biosynthesis of connective tissue proteoglycans. *Biochem Ser One*.
- Rostand, K. S., & Esko, J. D. (1997). Microbial adherence to and invasion through proteoglycans. *Infection and immunity*, 65(1), 1.

- Russo, J., & Russo, I. H. (2004). The breast as a developing organ *Molecular Basis of Breast Cancer* (pp. 11-48): Springer.
- Saad, O. M., Ebel, H., Uchimura, K., Rosen, S. D., Bertozzi, C. R., & Leary, J. A. (2005). Compositional profiling of heparin/heparan sulfate using mass spectrometry: assay for specificity of a novel extracellular human endosulfatase. *Glycobiology*, *15*(8), 818-826.
- Saad, O. M., & Leary, J. A. (2003). Compositional analysis and quantification of heparin and heparan sulfate by electrospray ionization ion trap mass spectrometry. *Analytical chemistry*, *75*(13), 2985-2995.
- Salmivirta, M., Lidholt, K., & Lindahl, U. (1996). Heparan sulfate: a piece of information. *The FASEB Journal*, *10*(11), 1270-1279.
- Sanderson, R. D., Yang, Y., Suva, L. J., & Kelly, T. (2004). Heparan sulfate proteoglycans and heparanase—partners in osteolytic tumor growth and metastasis. *Matrix Biology*, *23*(6), 341-352.
- Saravanan, R., & Shanmugam, A. (2010). Isolation and characterization of low molecular weight glycosaminoglycans from marine mollusc *Amussium pleuronectus* (Linne) using chromatography. *Applied biochemistry and biotechnology*, *160*(3), 791-799.
- Schlessinger, J., Plotnikov, A. N., Ibrahimi, O. A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., . . . Mohammadi, M. (2000). Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Molecular cell*, *6*(3), 743-750.
- Schmidt-Kittler, O., Ragg, T., Daskalakis, A., Granzow, M., Ahr, A., Blankenstein, T. J., . . . Müller, P. (2003). From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proceedings of the National Academy of Sciences*, *100*(13), 7737-7742.
- Schmitz, F. (1994). *Cytotoxic compounds from sponges and associated microfauna*. Paper presented at the Sponges in time and space: biology, chemistry, paleontology. Proceedings of the fourth international Porifera congress. Rotterdam.
- Scott, J., & Harbinson, R. (1969). Periodate oxidation of acid polysaccharides. *Histochemistry and Cell Biology*, *19*(2), 155-161.
- Scott, J. E., & Tigwell, M. J. (1978). Periodate oxidation and the shapes of glycosaminoglycuronans in solution. *Biochemical Journal*, *173*(1), 103-114.
- Sedita, J., Izvolsky, K., & Cardoso, W. V. (2004). Differential expression of heparan sulfate 6-O-sulfotransferase isoforms in the mouse embryo suggests distinctive roles during organogenesis. *Developmental dynamics*, *231*(4), 782-794.
- Seeger, R., & Krebs, E. G. (1995). The MAPK signaling cascade. *The FASEB Journal*, *9*(9), 726-735.
- Sethi, N., Dai, X., Winter, C. G., & Kang, Y. (2011). Tumor-derived jagged1 promotes osteolytic bone metastasis of breast cancer by engaging notch signaling in bone cells. *Cancer cell*, *19*(2), 192-205.
- Sharkey, L. C., Radin, M. J., Heller, L., Rogers, L. K., Tobias, A., Matise, I., . . . McCune, S. A. (2013). Differential cardiotoxicity in response to chronic doxorubicin treatment in male spontaneous hypertension-heart failure (SHHF), spontaneously hypertensive (SHR), and Wistar Kyoto (WKY) rats. *Toxicology and applied pharmacology*, *273*(1), 47-57.
- Silva, T. H., Alves, A., Ferreira, B., Oliveira, J. M., Reys, L., Ferreira, R., . . . Reis, R. (2012). Materials of marine origin: a review on polymers and ceramics of biomedical interest. *International Materials Reviews*, *57*(5), 276-306.
- Soeda, Ohmagari, Y., Shimeno, H., & Nagamatsu, A. (1993). Preparation of oversulfated fucoidan fragments and evaluation of their antithrombotic activities. *Thrombosis research*, *72*(3), 247-256.
- Soeda, S., Kozako, T., Iwata, K., & Shimeno, H. (2000). Oversulfated fucoidan inhibits the basic fibroblast growth factor-induced tube formation by human umbilical vein endothelial cells: its possible mechanism of action. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1497*(1), 127-134.

- Stecklein, S., Elsarraj, H., Valdez, K., Paul, A., & Behbod, F. (2013). Breast Cancer Invasion and Metastasis *Experimental Metastasis: Modeling and Analysis* (pp. 27-56): Springer.
- Stonik, V., & Fedorov, S. (2011). Cancer preventive marine natural product. *Cellular and Genetic Practices for Translational Medicine*, 1-36.
- Sugahara, K., & Kitagawa, H. (2000). Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Current opinion in structural biology*, 10(5), 518-527.
- Sugahara, K., Mikami, T., Uyama, T., Mizuguchi, S., Nomura, K., & Kitagawa, H. (2003). Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Current opinion in structural biology*, 13(5), 612-620.
- Sugahara, K., Ohi, Y., Harada, T., De Waard, P., & Vliegthart, J. (1992). Structural studies on sulfated oligosaccharides derived from the carbohydrate-protein linkage region of chondroitin 6-sulfate proteoglycans of shark cartilage. I. Six compounds containing 0 or 1 sulfate and/or phosphate residues. *Journal of Biological Chemistry*, 267(9), 6027-6035.
- Sutton, C. W. (2012). The role of targeted chemical proteomics in pharmacology. *British journal of pharmacology*, 166(2), 457-475.
- Sweeley, C. C., Bentley, R., Makita, M., & Wells, W. (1963). Gas-liquid chromatography of trimethylsilyl derivatives of sugars and related substances. *Journal of the American Chemical Society*, 85(16), 2497-2507.
- Synytsya, A., Kim, W.-J., Kim, S.-M., Pohl, R., Synytsya, A., Kvasnička, F., . . . Park, Y. I. (2010). Structure and antitumour activity of fucoidan isolated from sporophyll of Korean brown seaweed *Undaria pinnatifida*. *Carbohydrate Polymers*, 81(1), 41-48.
- Tanoue, E., & Handa, N. (1987). Monosaccharide composition of marine particles and sediments from the Bering Sea and northern North Pacific. *Oceanologica Acta*, 10(1), 91-99.
- Turnbull, J., Powell, A., & Guimond, S. (2001). Heparan sulfate: decoding a dynamic multifunctional cell regulator. *Trends in cell biology*, 11(2), 75-82.
- Turner, N., Lambros, M. B., Horlings, H. M., Pearson, A., Sharpe, R., Natrajan, R., . . . Mackay, A. (2010). Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. *Oncogene*, 29(14), 2013-2023.
- Urbán, P., Estelrich, J., Cortés, A., & Fernández-Busquets, X. (2011). A nanovector with complete discrimination for targeted delivery to *Plasmodium falciparum*-infected versus non-infected red blood cells in vitro. *Journal of controlled release*, 151(2), 202-211.
- Uyama, T., Kitagawa, H., & Sugahara, K. (2007). Biosynthesis of glycosaminoglycans and proteoglycans. *Comprehensive glycoscience*, 3, 79-104.
- Valastyan, S., & Weinberg, R. A. (2011). Tumor metastasis: molecular insights and evolving paradigms. *Cell*, 147(2), 275-292.
- van de Velde, F., Pereira, L., & Rollema, H. S. (2004). The revised NMR chemical shift data of carrageenans. *Carbohydrate Research*, 339(13), 2309-2313.
- Vangapandu, S., Jain, M., Kaur, K., Patil, P., Patel, S. R., & Jain, R. (2007). Recent advances in antimalarial drug development. *Medicinal research reviews*, 27(1), 65-107.
- Vara, J. Á. F., Casado, E., de Castro, J., Cejas, P., Belda-Iniesta, C., & González-Barón, M. (2004). PI3K/Akt signalling pathway and cancer. *Cancer treatment reviews*, 30(2), 193-204.
- Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., . . . Etzler, M. E. (2009). *Methods and Applications*.
- Viale, G. (2012). The current state of breast cancer classification. *Annals of oncology*, 23(suppl 10), x207-x210.
- Vilarino-Varela, M., Chin, Y. S., & Makris, A. (2009). *Current indications for post-mastectomy radiation*. Paper presented at the International Seminars in Surgical Oncology.
- Villalonga, P., Villalonga, P., & Ridley, A. J. (2006). *Rho GTPases and cell cycle control*: Taylor & Francis.

- Wang, K., Zhang, Q., Li, D., Ching, K., Zhang, C., Zheng, X., . . . Wang, H. (2015). PEST domain mutations in Notch receptors comprise an oncogenic driver segment in triple-negative breast cancer sensitive to a  $\gamma$ -secretase inhibitor. *Clinical cancer research*, 21(6), 1487-1496.
- Wang, R., Lv, Q., Meng, W., Tan, Q., Zhang, S., Mo, X., & Yang, X. (2014). Comparison of mammosphere formation from breast cancer cell lines and primary breast tumors. *Journal of thoracic disease*, 6(6), 829-837.
- Wei, Z., Swiedler, S. J., Ishihara, M., Orellana, A., & Hirschberg, C. B. (1993). A single protein catalyzes both N-deacetylation and N-sulfation during the biosynthesis of heparan sulfate. *Proceedings of the National Academy of Sciences*, 90(9), 3885-3888.
- Weigelt, B., Peterse, J. L., & Van't Veer, L. J. (2005). Breast cancer metastasis: markers and models. *Nature reviews. Cancer*, 5(8), 591.
- Weitzhandler, M., Kadlecak, D., Avdalovic, N., Forte, J., Chow, D., & Townsend, R. (1993). Monosaccharide and oligosaccharide analysis of proteins transferred to polyvinylidene fluoride membranes after sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Journal of Biological Chemistry*, 268(7), 5121-5130.
- Wells, T. N., Alonso, P. L., & Gutteridge, W. E. (2009). New medicines to improve control and contribute to the eradication of malaria. *Nature reviews Drug discovery*, 8(11), 879-891.
- Wen, L., Sun, L., Xi, Y., Chen, X., Xing, Y., Sun, W., . . . Cai, L. (2015). Expression of calcium sensing receptor and E - cadherin correlated with survival of lung adenocarcinoma. *Thoracic cancer*, 6(6), 754-760.
- Wiederschain, G. Y. (2009). Essentials of glycobiology. *Biochemistry (Moscow)*, 74(9), 1056-1056.
- Willför, S., Pranovich, A., Tamminen, T., Puls, J., Laine, C., Suurnäkki, A., . . . Hemming, J. (2009). Carbohydrate analysis of plant materials with uronic acid-containing polysaccharides—A comparison between different hydrolysis and subsequent chromatographic analytical techniques. *Industrial Crops and Products*, 29(2), 571-580.
- Wu, M., Xu, S., Zhao, J., Kang, H., & Ding, H. (2010). Physicochemical characteristics and anticoagulant activities of low molecular weight fractions by free-radical depolymerization of a fucosylated chondroitin sulphate from sea cucumber *Thelenata ananas*. *Food Chemistry*, 122(3), 716-723.
- Wu, Z. L., Zhang, L., Yabe, T., Kuberan, B., Beeler, D. L., Love, A., & Rosenberg, R. D. (2003). The involvement of heparan sulfate (HS) in FGF1/HS/FGFR1 signaling complex. *Journal of Biological Chemistry*, 278(19), 17121-17129.
- Xiao, L., Yang, C., Patterson, P. S., Udhayakumar, V., & Lal, A. A. (1996). Sulfated polyanions inhibit invasion of erythrocytes by plasmodial merozoites and cytoadherence of endothelial cells to parasitized erythrocytes. *Infection and immunity*, 64(4), 1373-1378.
- Yamada, S., Sugahara, K., & Özbek, S. (2011). Evolution of glycosaminoglycans: Comparative biochemical study. *Communicative & integrative biology*, 4(2), 150-158.
- Yamagata, T., Saito, H., Habuchi, O., & Suzuki, S. (1968). Purification and properties of bacterial chondroitinases and chondrosulfatases. *Journal of Biological Chemistry*, 243(7), 1523-1535.
- Yamasaki-Miyamoto, Y., Yamasaki, M., Tachibana, H., & Yamada, K. (2009). Fucoidan induces apoptosis through activation of caspase-8 on human breast cancer MCF-7 cells. *Journal of agricultural and food chemistry*, 57(18), 8677-8682.
- Ye, H., Wang, K., Zhou, C., Liu, J., & Zeng, X. (2008). Purification, antitumor and antioxidant activities in vitro of polysaccharides from the brown seaweed *Sargassum pallidum*. *Food Chemistry*, 111(2), 428-432.
- Yuan, J., Hashii, N., Kawasaki, N., Itoh, S., Kawanishi, T., & Hayakawa, T. (2005). Isotope tag method for quantitative analysis of carbohydrates by liquid chromatography–mass spectrometry. *Journal of Chromatography A*, 1067(1), 145-152.
- Zardawi, S. J., Zardawi, I., McNeil, C. M., Millar, E. K., McLeod, D., Morey, A. L., . . . Lopez - Knowles, E. (2010). High Notch1 protein expression is an early event in breast cancer development and is associated with the HER - 2 molecular subtype. *Histopathology*, 56(3), 286-296.

- Zeng, Q., Li, S., Chepeha, D. B., Giordano, T. J., Li, J., Zhang, H., . . . Wang, C.-Y. (2005). Crosstalk between tumor and endothelial cells promotes tumor angiogenesis by MAPK activation of Notch signaling. *Cancer cell*, 8(1), 13-23.
- Zhang, Z., Coomans, C., & David, G. (2001). Membrane Heparan Sulfate Proteoglycan-supported FGF2-FGFR1 Signaling EVIDENCE IN SUPPORT OF THE "COOPERATIVE END STRUCTURES" MODEL. *Journal of Biological Chemistry*, 276(45), 41921-41929.
- Zhu, W., Smith, J. W., & Huang, C.-M. (2009). Mass spectrometry-based label-free quantitative proteomics. *BioMed Research International*, 2010.
- Zierer, M. S., & Mourão, P. A. (2000). A wide diversity of sulfated polysaccharides are synthesized by different species of marine sponges. *Carbohydrate Research*, 328(2), 209-216.

## Appendices

## **Appendix I: Recipes for solutions and buffers**

### **Enzymatic digestion:**

#### **Heparin Lyase Buffer: 100 mM Sodium Acetate/0.1 mM Calcium Acetate pH 7.0**

- 3.4 g NaOAc·3H<sub>2</sub>O
- 4.4 mg Ca(OAc)<sub>2</sub>·H<sub>2</sub>O.
- Dissolve in 200 mL of Milli-Q water, adjust pH to 7.0 and bring volume to 250 mL.
- GE-H0001-Heparinase I (EC 4.2.2.7), 2 IU, GE-H 0002- Heparinase II, 0.25 IU, GE-H 0003-Heparinase III (EC 4.2.2.8), 0.5 IU from Grampian Enzymes.
- 

#### **Chondroitin ABC Lyase Buffer : 50 mM Tris-Cl/50 mM Sodium Acetate pH 8.0**

- 1.5 g Tris Base
- 3.4 g NaOAc·3H<sub>2</sub>O.
- Dissolve in 200 mL of Milli-Q water, adjust pH to 8.0 and bring volume to 250 mL
- Chondroitinase AC-50 µL- / 0.5 IU- 50-013, Chondroitinase B-50 µL- / 5 µg- 50-018 from IBEX Technologies Inc-Canada.

### **Analysis of Monosaccharides by HPAEC-PAD**

#### **1-Preparation of 1L of 100 mM NaOH with 5 mM NaOAc:**

- 0.7 g of Sodium Acetate Trihydrate
- 5.2 mL of 50% NaOH
- Dissolve the Sodium Acetate Trihydrate in Milli-Q water that has been filtered through a 0.2 µm filter. Then add the 50% NaOH and bring the volume to 1 L.

#### **2-Preparation of 1L of 100 mM NaOH with 250 mM NaOAc:**

- 34 g of Sodium Acetate Trihydrate
- 5.2 mL of 50% NaOH

- Dissolve the Sodium Acetate Trihydrate in Milli-Q water that has been filtered through a 0.2  $\mu\text{m}$  filter. Then add the 50% NaOH and bring the volume to 1 L.

**DEAE Pre-Wash Buffer:**

**50 mM NaOAc/150 mM NaCl with 0.1% Triton X-100 pH 6.0**

- 6.8 g NaOAc·3H<sub>2</sub>O
- 4.4 g NaCl
- 0.5 g Triton X-100
- Dissolve in 400 mL Milli-Q water, adjust pH to 6.0 and bring volume to 500 mL

**DEAE Wash Buffer:**

**50 mM NaOAc/150 mM NaCl with pH 6.0**

- 6.8 g NaOAc·3H<sub>2</sub>O
- 4.4 g NaCl
- Dissolve in 400 mL Milli-Q water, adjust pH to 6.0 and bring volume to 500 mL

**DEAE Elution Buffer:**

**1- 50 mM NaOAc/1M NaCl pH 6.0**

- 6.8 g NaOAc·3H<sub>2</sub>O
- 29.22 g NaCl
- Dissolve in 400 mL Milli-Q water, adjust pH to 6.0 and bring volume to 500 mL

**2- 50 mM NaOAc/2M NaCl pH 6.0**

- 6.8 g NaOAc·3H<sub>2</sub>O
- 58.44g NaCl
- Dissolve in 400 mL Milli-Q water, adjust pH to 6.0 and bring volume to 500 mL

### **3- 50 mM NaOAc/3M NaCl pH 6.0**

- 6.8 g NaOAc·3H<sub>2</sub>O
- 87.660g NaCl
- Dissolve in 400 mL Milli-Q water, adjust pH to 6.0 and bring volume to 500 mL

### **Reagents of TMS Derivatization and Analysis by GC-MS**

- 3 M Methanolic-HCl (Supelco, 33355).
- Methanol (Sigma, 322415).
- Acetic Anhydride (Sigma, 539996).
- Pyridine (Sigma, 270970).
- Tri-Sil HTP Reagent (Pierce, TS-48999).

### **Cell lines purchase**

- The human lung adenocarcinoma epithelial A549 cell line (passage 9) (86012804, Sigma Aldrich(UK).
- 92020424/ MDA-MB-231 Human Caucasian breast Adenocarcinoma-Frozen  
In Stock, Public health England-cell collections.
- 86012803/ MCF7- Human Caucasian breast adenocarcinoma- Frozen  
In Stock, Public health England-cell collections.
- ATCC- HTB-132/ MDA-MB-468 Mammary gland/breast; derived from metastatic site: pleural effusion black female, Frozen in Stock, ATCC/LGC Promochem, Middlesex,UK.
- ATCC- HTB-30/ Caucasian female, mammary gland/breast; derived from metastatic site: pleural effusion, Frozen in Stock, ATCC/LGC Promochem, Middlesex,UK.

- ATCC- CCL-243/ Homo sapiens, human bone marrow, lymphoblast, Frozen in Stock, ATCC/LGC Promochem, Middlesex,UK.

### **Mammosphere assay buffers**

- DMEM/F12 (Thermofisher, product code 11320-033).
- Basic Fibroblast Growth Factor (bFGF), Human Recombinant (Corning, Product code 354060).
- Epidermal Growth Factor (EGF), Mouse Natural (Culture Grade), (Corning, Product code 354001).
- ITS Liquid Media Supplement (100x), liquid, sterile-filtered, BioReagent, suitable for cell culture (Sigma, Product code I3146-5ML).
- B-27 Supplement (50X), serum free (Thermofisher, Product code 17504-044).
- Poly (2-hydroxyethyl methacrylate) (Sigma, Product code P3932).

### **Reagents Supplier and Instruments Details**

- PBS tablets (Phosphate Buffered Saline), 1 x tablet dissolved in 200mL water yields 0.01M phosphate buffer, 0.0027M KCl, and 0.137M NaCl, pH 7.4 at 25°C for HPLC from Fisher Scientific.
- Acetone
- Sodium Carbonate Sigma Aldrich-1.06392.0500
- Alcalase Enzyme
- trichloroacetic acid Sigma Aldrich -T6399
- potassium acetate Fisher Scientific- 11376798

- cetylpyridinium chloride Sigma Aldrich -C0732
- Sodium chloride Sigma Aldrich-S7653
- Dialysis tube Spectrum Labs.com
- Bovine Serum Albumin BSA Sigma Aldrich A2153.
- Sodium acetate NaOAc $\cdot$ 3H $_2$ O Sigma Aldrich- S2889.
- Ca(OAc) $_2$  $\cdot$ H $_2$ O.
- Tris Base Sigma Aldrich -697737.
- Sodium Acetate Trihydrate Fisher Scientific-12350000.
- Sodium hydroxide NaOH Sigma Aldrich -71687.
- Triton X-100 Sigma Aldrich.

### **The Giemsa staining solution**

Gurr's Giemsa stain solution (BDH/VWR international limited, UK) prepared at 1:10 as below:

- one tablet pH 6.4 (BDH laboratory supplies, England).
- 1 litre of freshly distilled water.

### **Sample preparation for LC-MS analysis using filter-aided sample preparation (FASP)**

#### **1- Materials and Reagents:**

- Microcon-30kDa Centrifugal Filter Unit with Ultracel-30 membrane (MRCF0R030) and Direct Detect® spectrometer (DDHW00010-WW) from Millipore,
- ammonium bicarbonate (ambic, 09830) from Sigma Aldrich,
- Trypsin (Promega, cat. no. V528A),
- LysC MS grade (Wako Chemicals, cat. no. 129-02541),
- Dithiothreitol (DTT, MB105) and Tris Base (B2005) from Melford laboratories,
- Iodoacetamide (IAM; Sigma-Aldrich, cat. no. I6125)
- Urea (Fisher, cat. no. U/0500/53),
- POROS R3 beads (1-339-03),
- 96-well with 0.2  $\mu$ M PVDF membrane (3504) from Corning.
- Microtube -15 from Covaris (520145),
- Formic acid (FA; Fluka, cat. no. 94318),
- Acetonitrile (MeCN; Fisher, cat. no. A955-1),

- Acetonitrile with 0.1% formic acid (Sigma Aldrich, 34688),
- Water with 0.1% formic acid, (Sigma-Aldrich, 34673),

## 2- Buffers:

- SDS buffer stock = 2% SDS in 50 mM Tris-HCl pH 7.4.
- UA1 = 8M urea, 0.1M Tris-HCl (pH 8.5).
- UA2 = 8M urea, 0.1 Tris-HCl (pH 8.5) with 15 mM DTT.
- UA3 = 75% UA1 buffer + 25% deionised water (=6M Urea).
- Wet solution = 50% acetonitrile.
- Wash solution = 0.1% formic acid.
- Elute solution = 50% acetonitrile, 0.1% formic acid

## 3- Measurement of protein concentration for proteomic experiment.

Replicate number	Concentration (ug.uL <sup>-1</sup> )	%CV coefficient of variation	Average concentration (ug.uL <sup>-1</sup> )	To digest 25 ug add x uL to each filter tube
MDA-MB-468/ C1 – 1,2,3	0.432	16%	0.38	67
MDA-MB-468/C1 – 4,5,6	0.313	4%		
MDA-MB-231/ C1 – 7,8,9	0.393	5%	0.37	68
MDA-MB-231/ C1 – 10,11,12	0.344	27%		
MDA-MB-468/ T1 – 13,14,15	0.590	6%	0.62	40
MDA-MB-468/ T1 – 16,17,18	0.657	8%		
MDA-MB-231/T1 – 18,19,20	0.214	18%	0.29	87
MDA-MB-231/T1 – 21,22,23	0.360	4%		

**Appendix II: A: LC-MS spectra of aniline-tagged standards and whelk-GAG flow throw (FT) sample fractions. B: MS spectra of the unknown peak in CK1 fraction (CK-1 sample in negative mode direct injection MS), Standard CS in negative mode direct injection MS, Standard GlcA in negative mode direct injection MS, and Reagent blank run on MS**

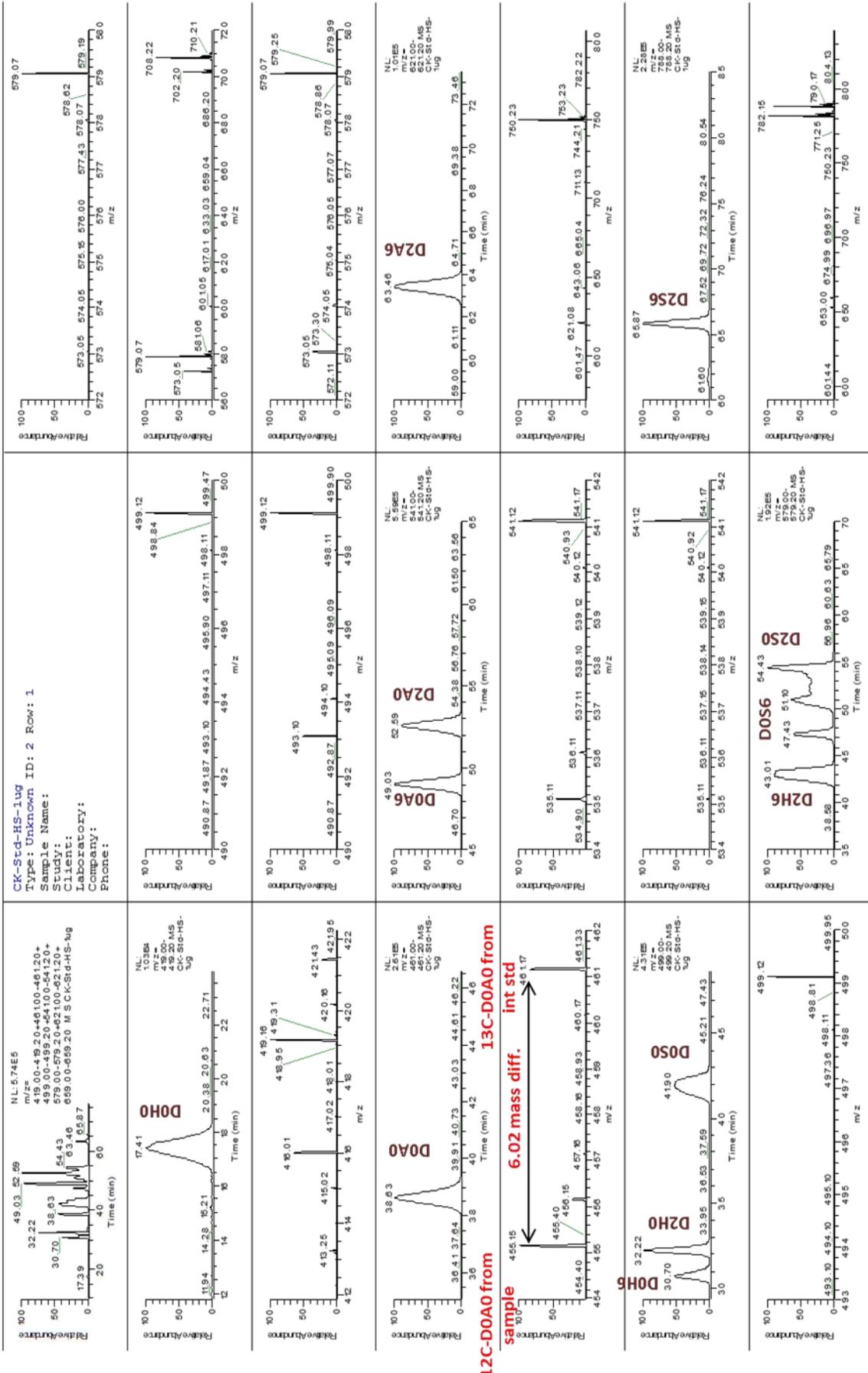
The retention times and mass spectra of aniline-tagged standards and whelk-GAG flow throw (FT) sample fractions from GRIL-LC-MS. A: Represent HS-Std and sample fractions digested with HS lyase, A1-Standard disaccharide mixture tagged with 12C-Aniline, A2-Standard Heparin digested with Hep-lyase, 12C-aniline tagged and mixed with known amount of 13C-aniline tagged internal standard, A3-CK-1 GAG sample isolated in 1M NaCl fraction, digested with Hep-lyase and GRIL tagged with 12C-aniline and spiked with 13C-aniline tagged disaccharide mixture as int std, A4-CK-2 GAG sample isolated in 2M NaCl fraction, digested with Hep-lyase and GRIL tagged with 12C-aniline and spiked with 13C-aniline tagged disaccharide mixture as int std, A5-CK-3 GAG sample isolated in 3M NaCl fraction, digested with Hep-lyase and GRIL tagged with 12C-aniline and spiked with 13C-aniline tagged disaccharide mixture as int std.

B: Represent CS-Std and sample fractions digested with ABC lyase, B1- Standard disaccharide mixture tagged with 12C-Aniline, B2- Standard chondroitin digested with CS-ABC, 12C-aniline tagged and mixed with known amount of 13C-aniline tagged internal standard, B3- CK-1 GAG sample isolated in 1M NaCl fraction, digested with CS-ABC and GRIL tagged with 12C-aniline and spiked with 13C-aniline tagged disaccharide mixture as int std, B4- CK-2 GAG sample isolated in 2M NaCl fraction, digested with CS-ABC and GRIL tagged with 12C-aniline and spiked with 13C-aniline tagged disaccharide mixture as int std, B5- CK-3 GAG sample isolated in 3M NaCl fraction, digested with CS-ABC and GRIL tagged with 12C-aniline and spiked with 13C-aniline tagged disaccharide mixture as int std.



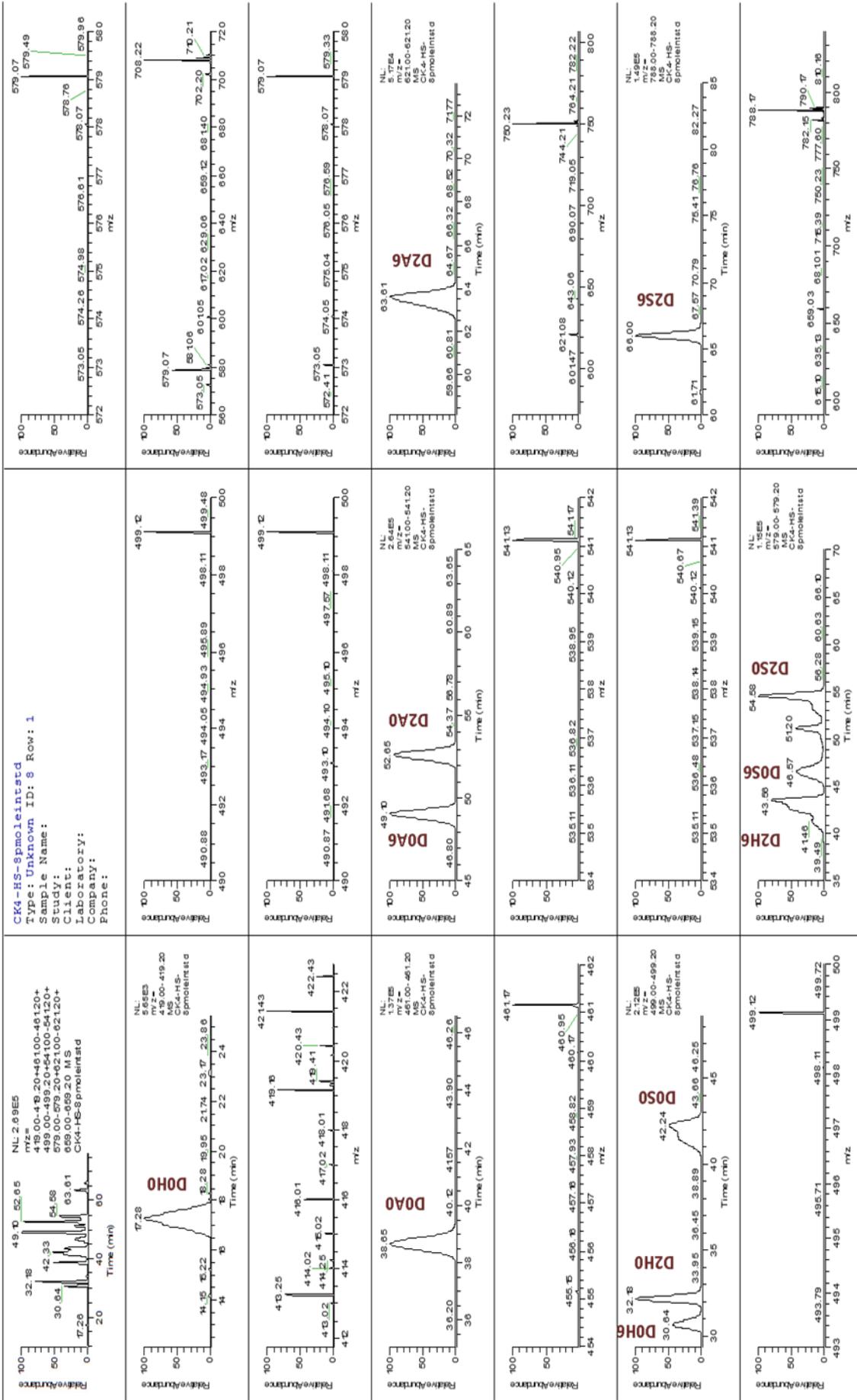
A2

C:\Xcalibur\...\CK-Std-HS-1ug



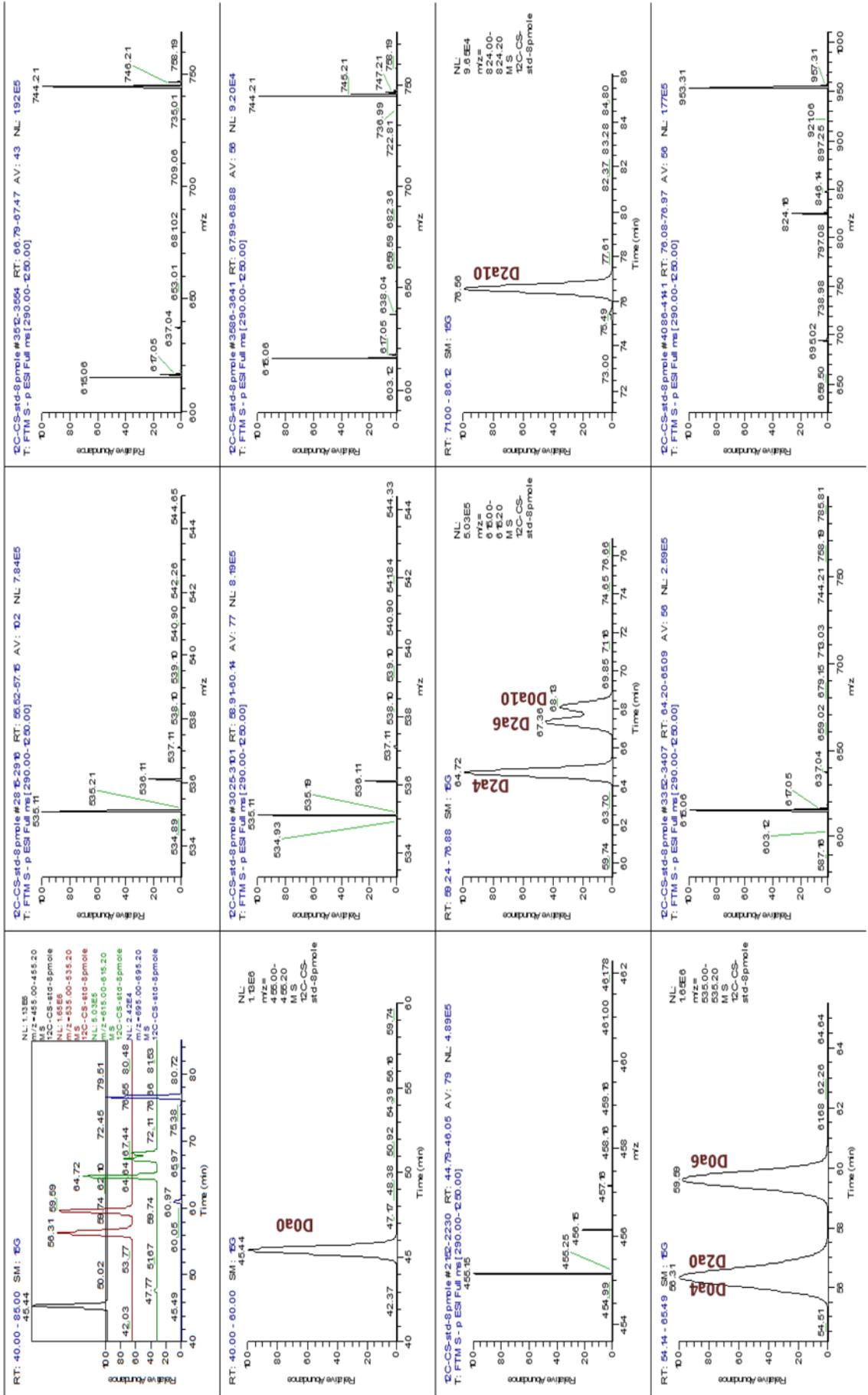






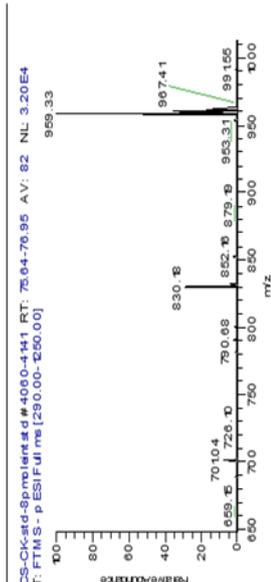
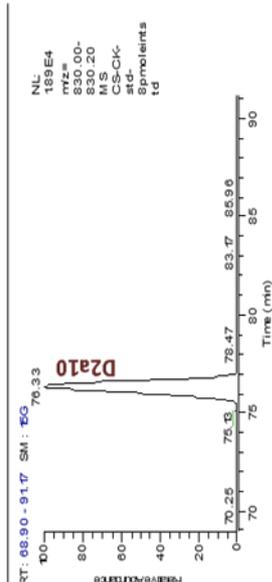
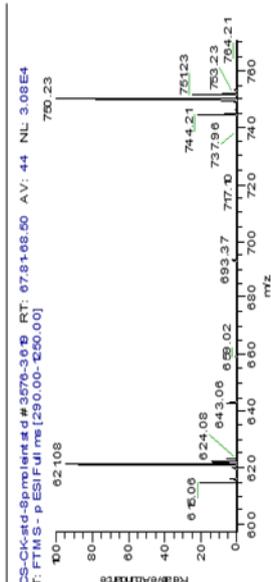
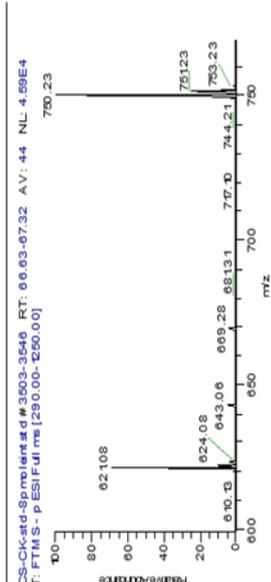
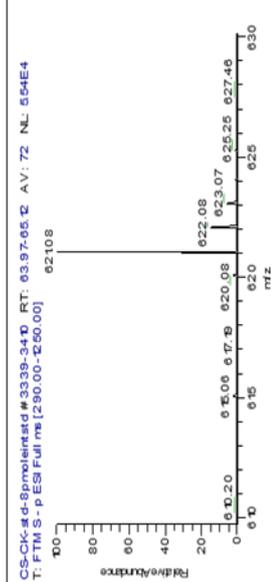
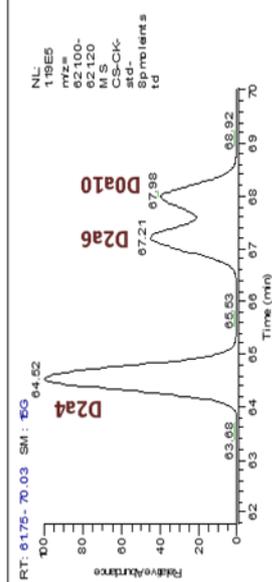
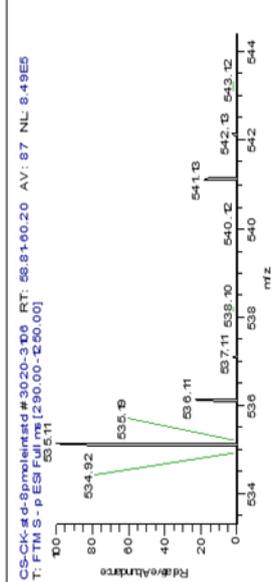
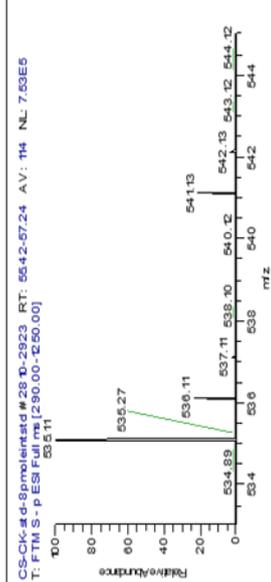
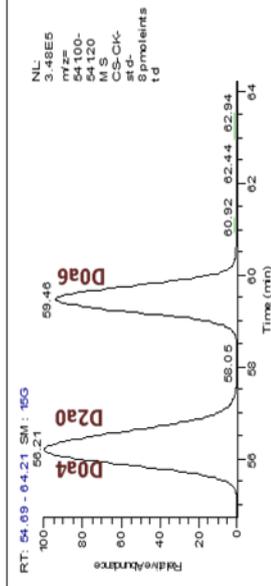
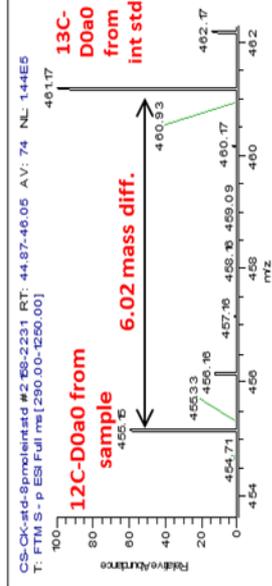
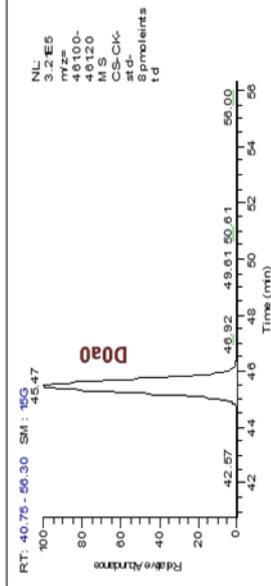
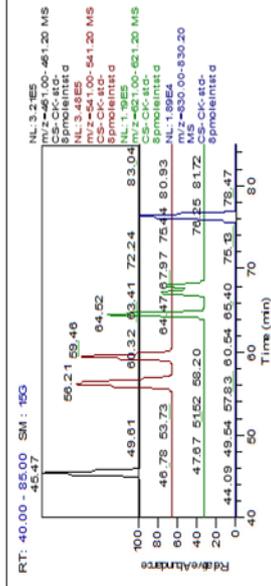
B1

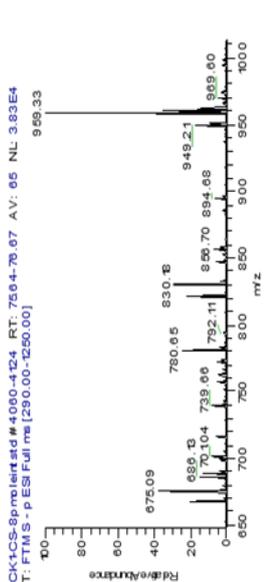
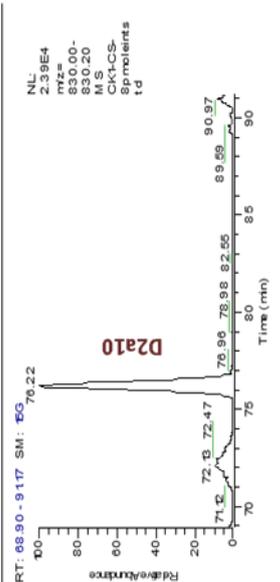
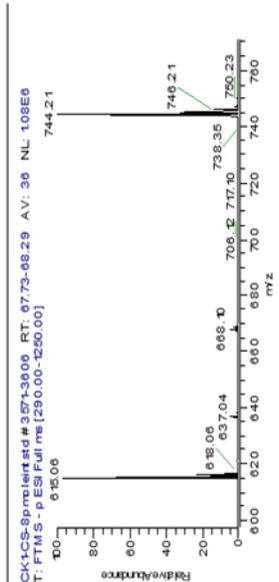
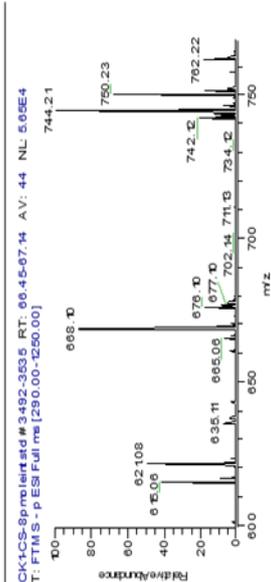
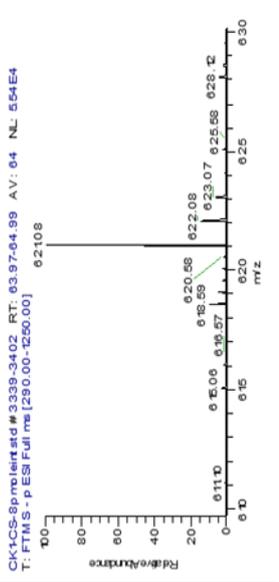
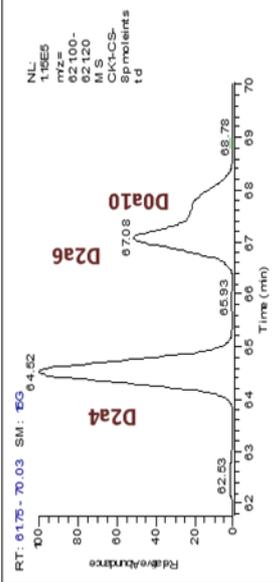
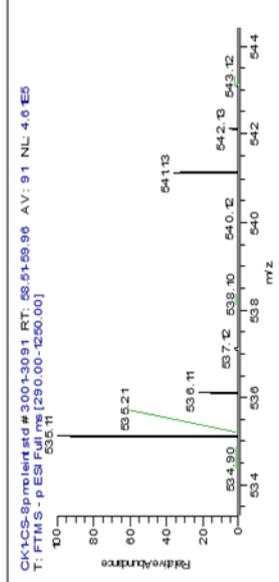
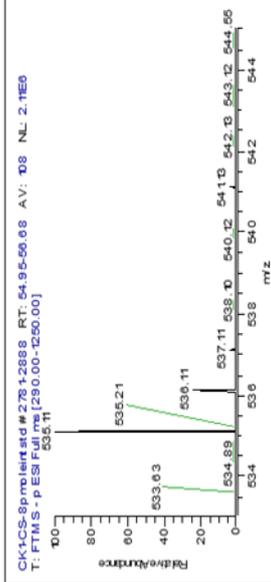
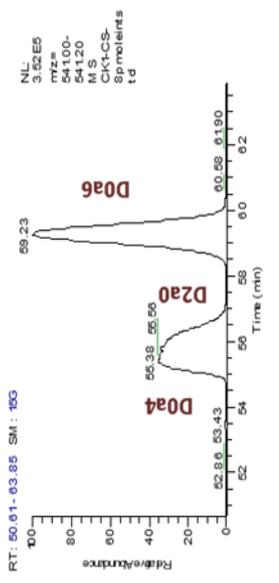
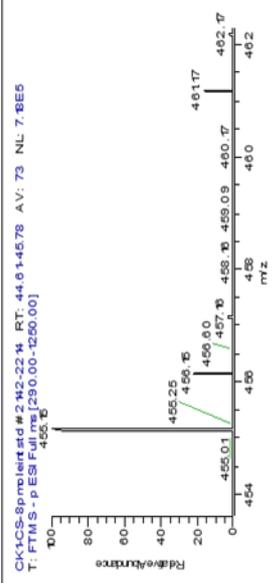
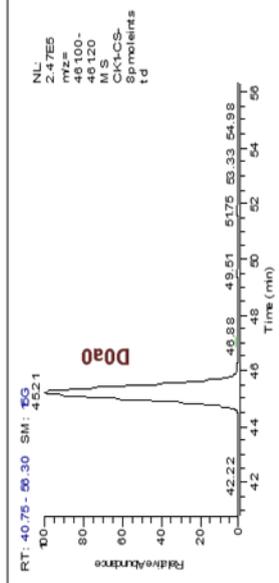
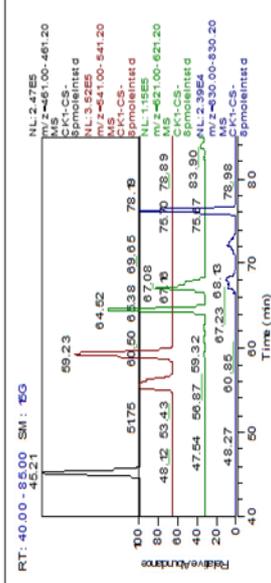
C:\Xcalibur\...12C-CS-std-8pmole



# B2

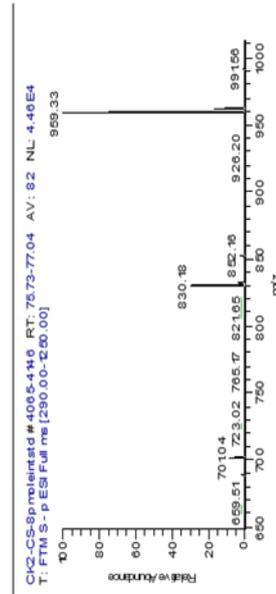
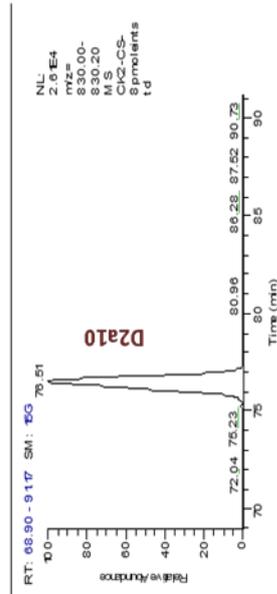
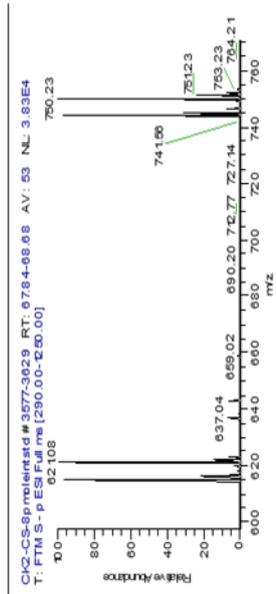
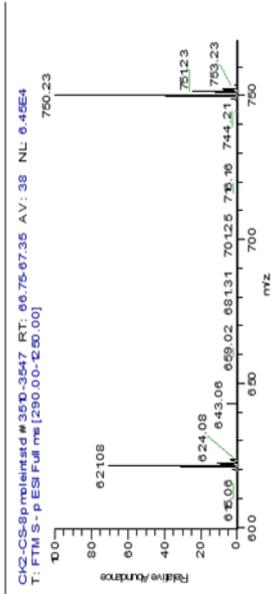
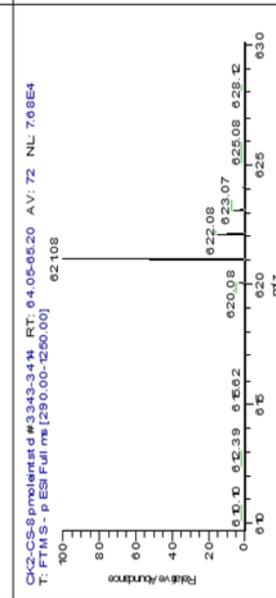
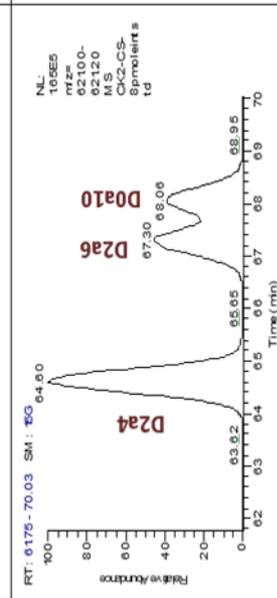
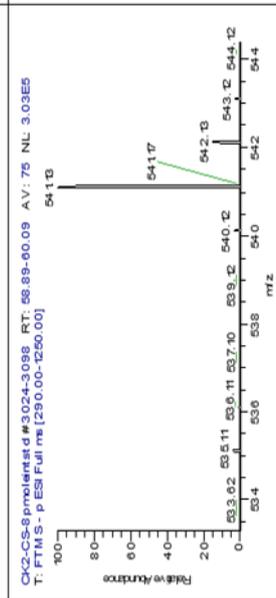
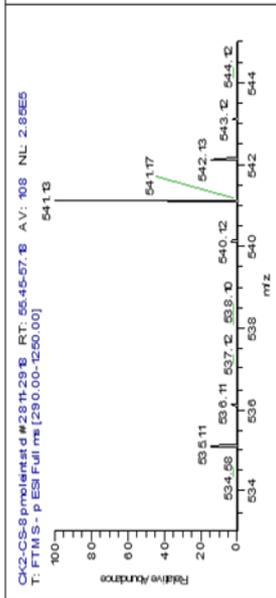
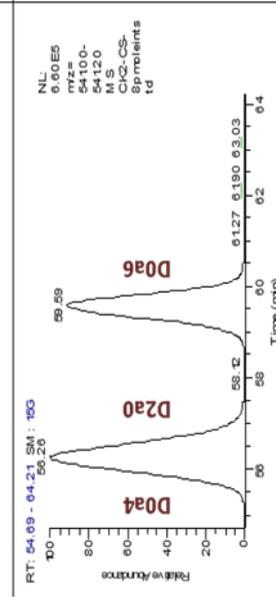
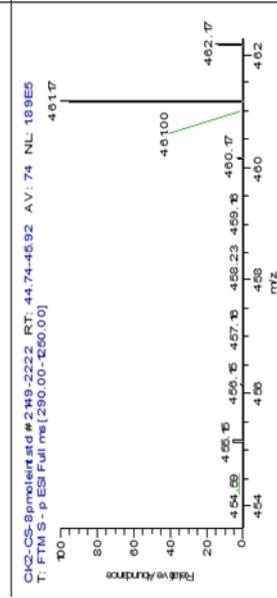
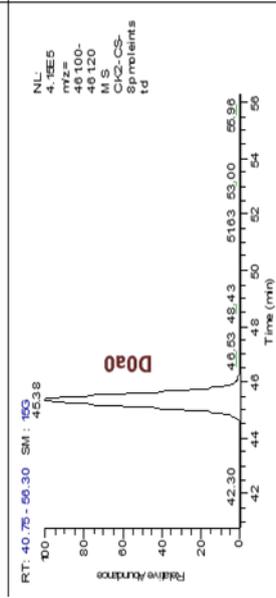
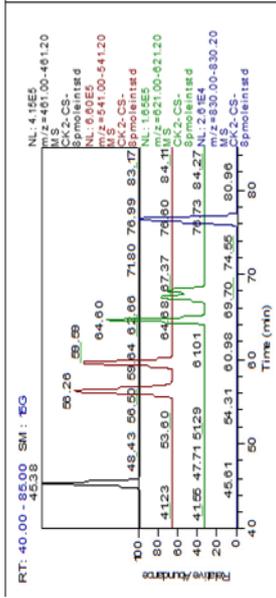
C:\xcalibur\...CS-CK-std-8pm oleints.d





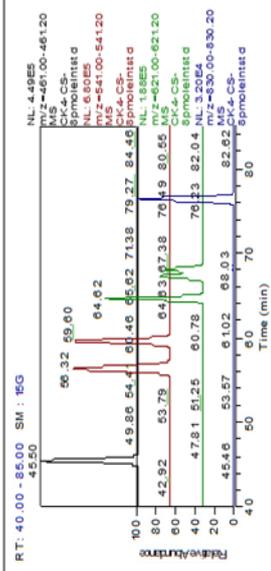
B4

C:\xcalibur\...ck2-cs-8pmoleints



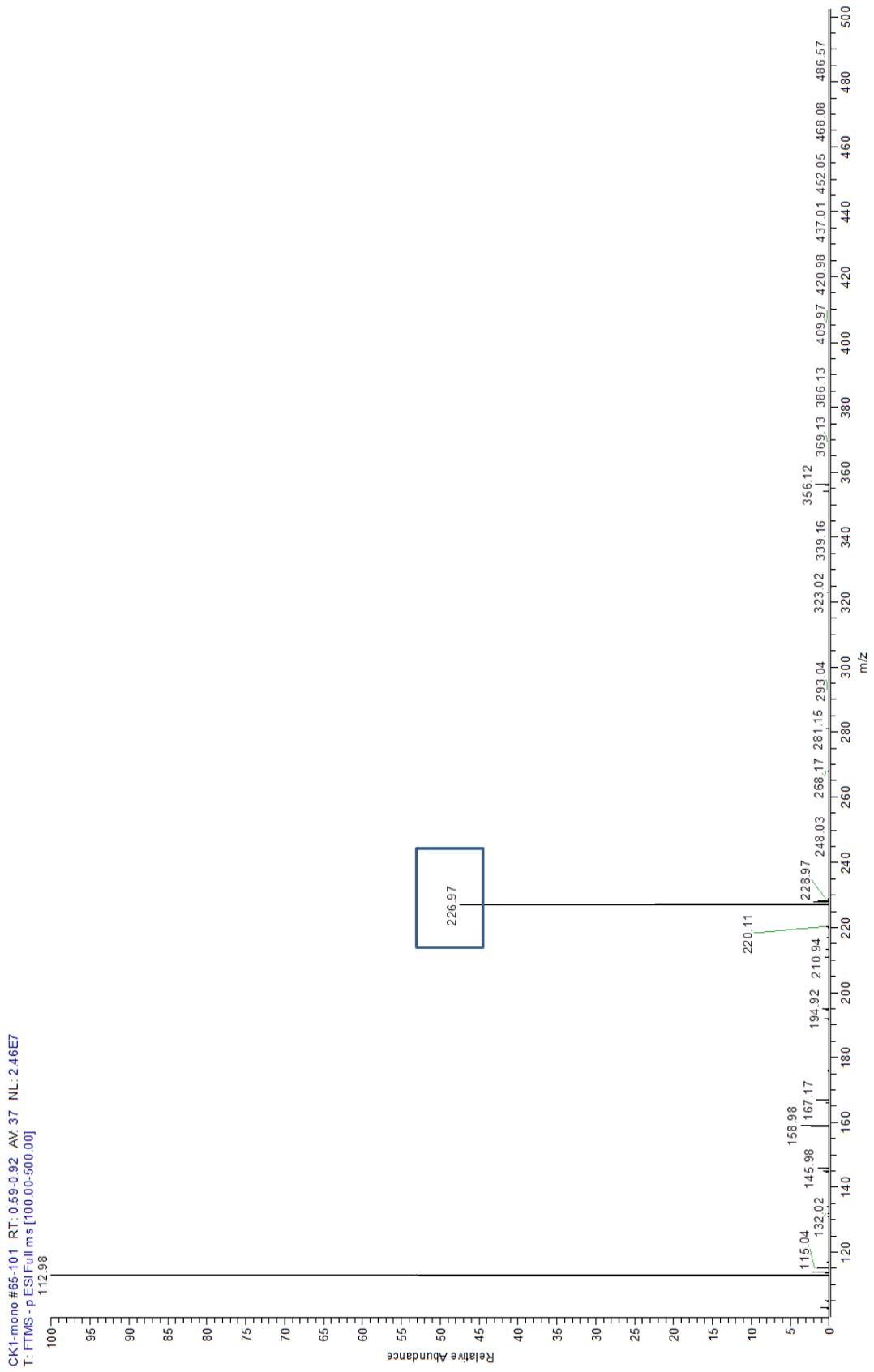
# B5

C:\Xcalibur\...CK4-CS-8pmoleintstd

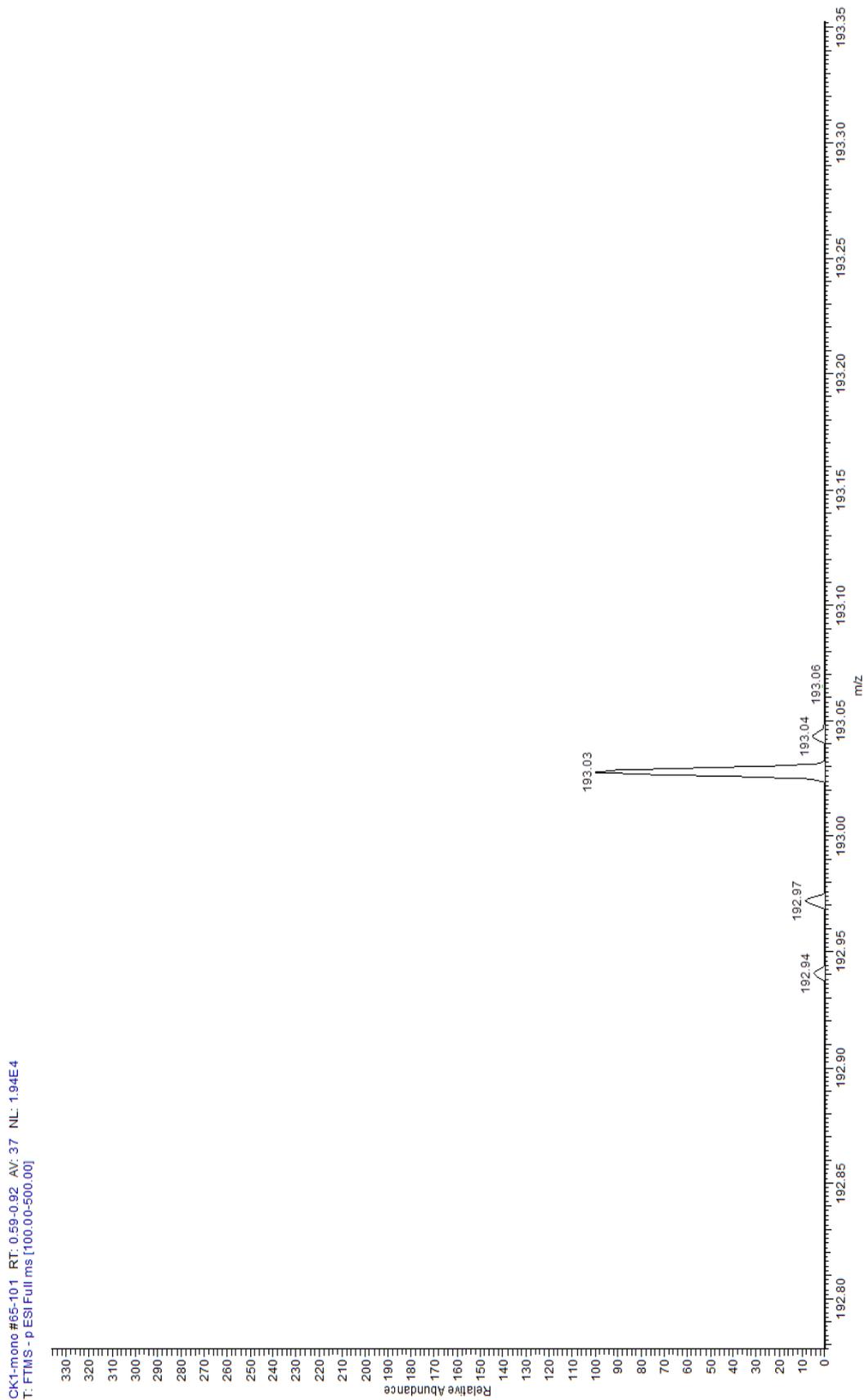


## CK-1 sample in negative mode direct injection ms

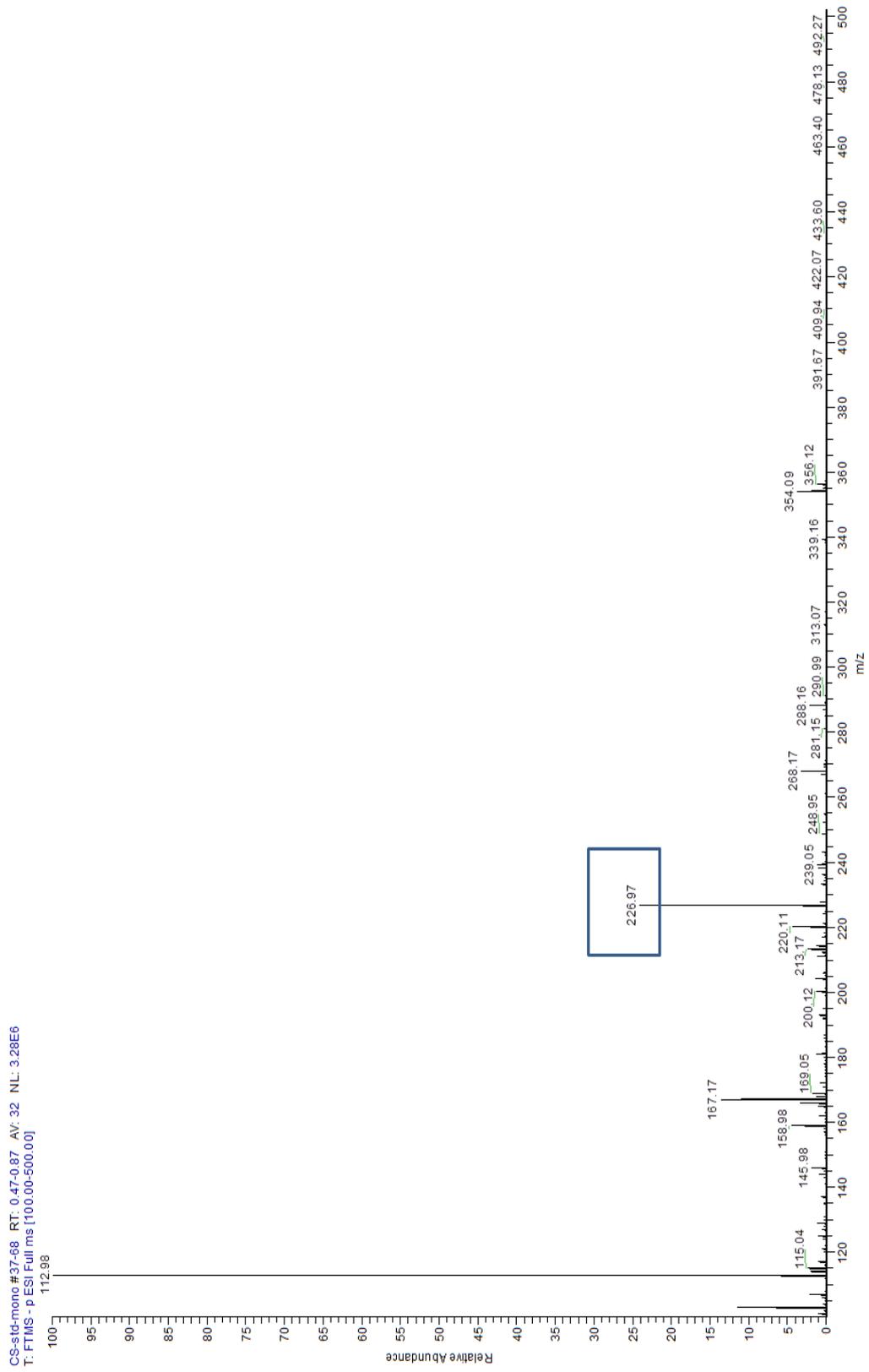
The ion shown in box are also present in std CS, however I could not assign any structure for those ions



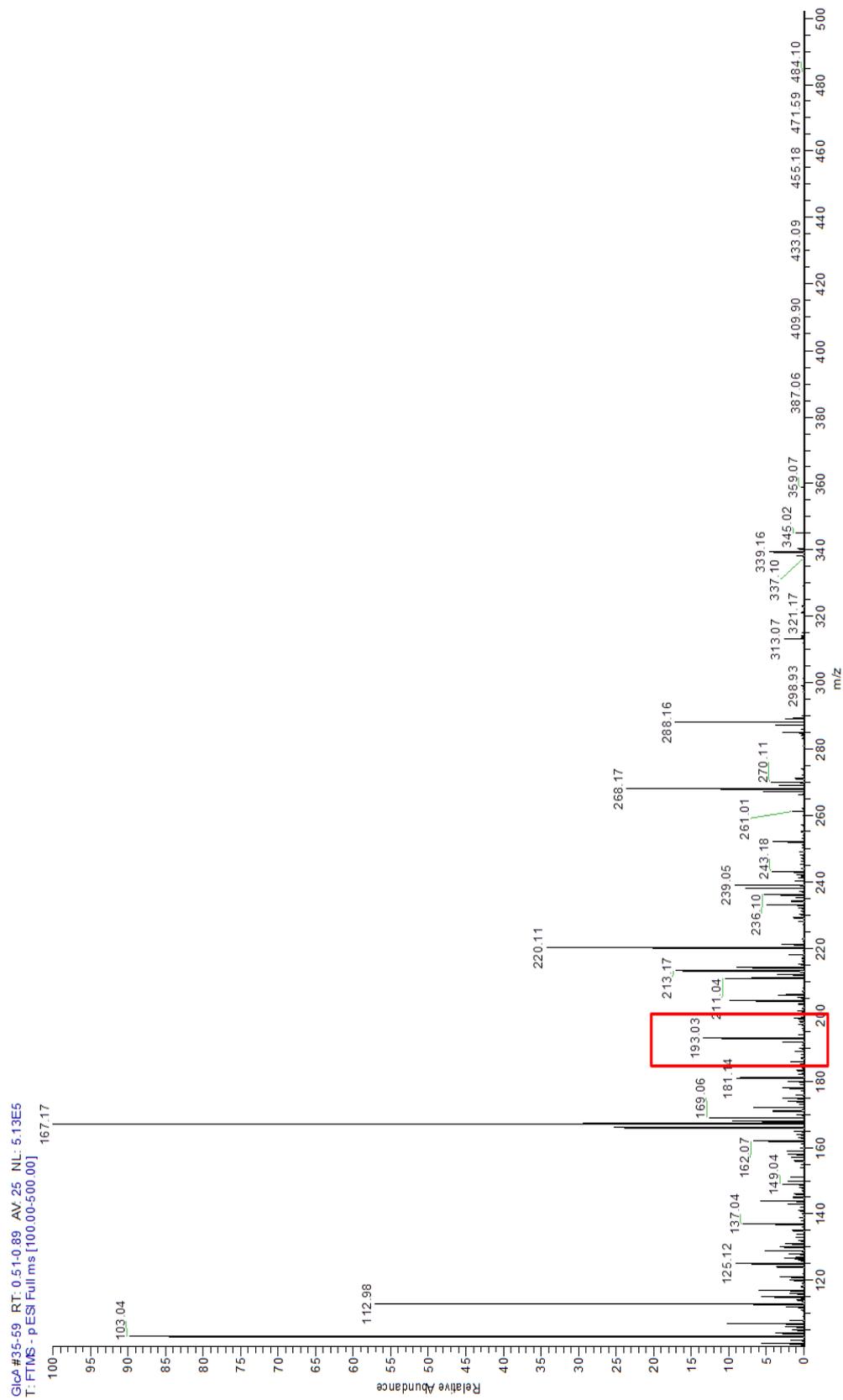
# CK-1 sample in negative mode direct injection ms (expanded region showing the presence of GlcA ion (M-H)<sup>-</sup>)



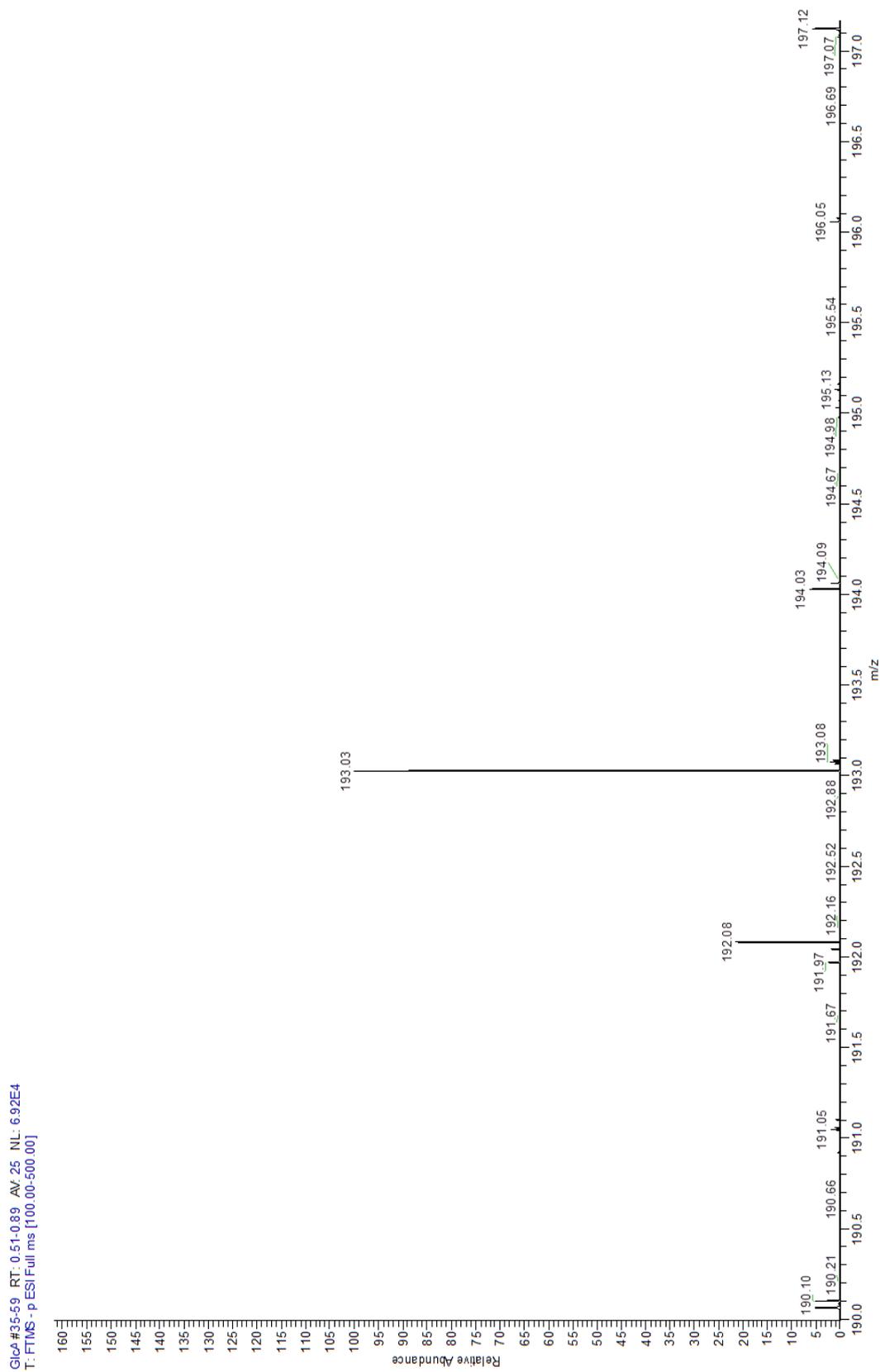
# Standard CS in negative mode direct injection ms



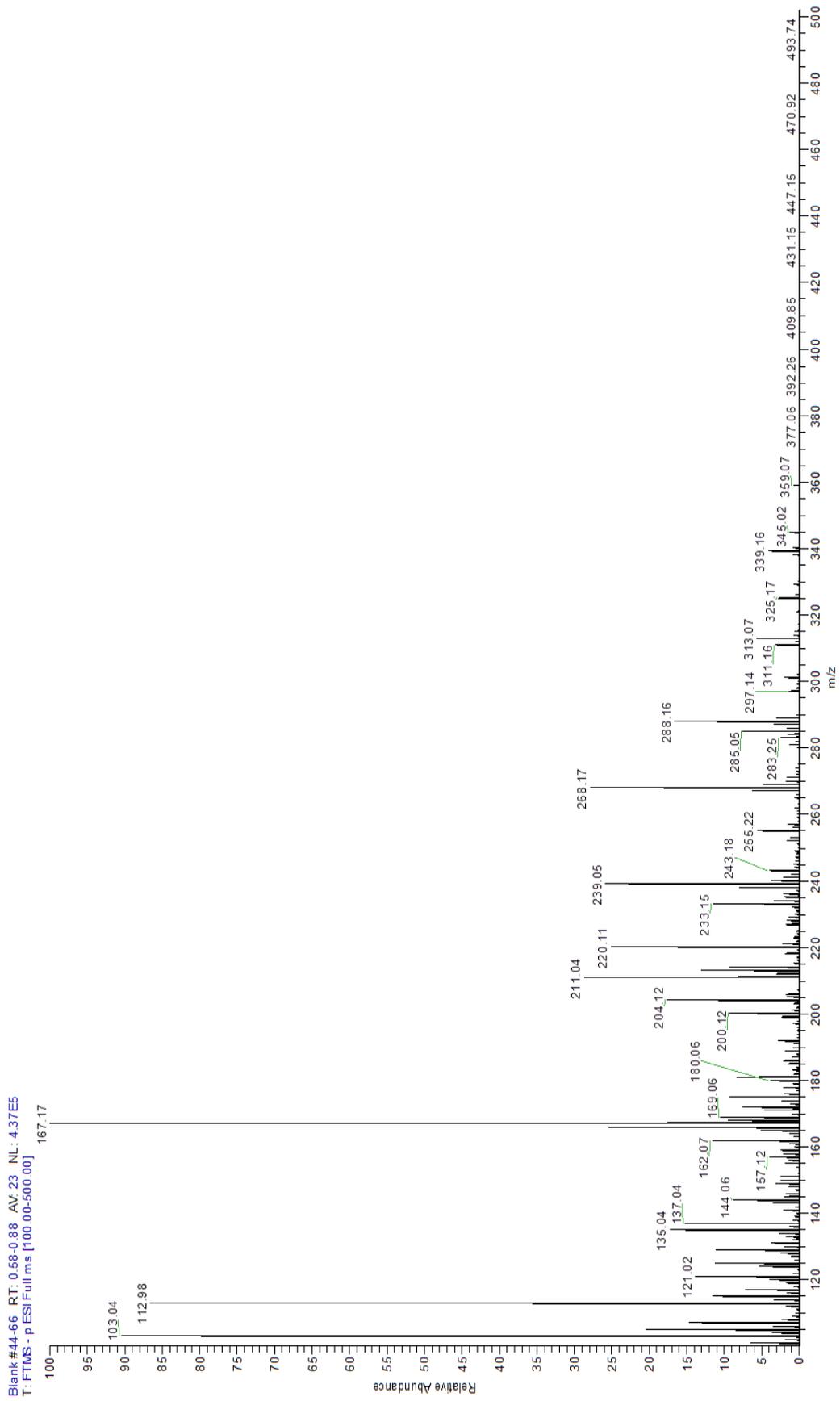
# Standard GlcA in negative mode direct injection ms



# Standard GlcA in negative mode direct injection ms (expanded)



# Reagent blank run on MS



### **Appendix III: Monosaccharide quantification**

---

2N-TFA hydrolysis 5h on 10% material and 7% injected

2N-TFA hydrolysis 5h on 10% material and 7% injected

Amount of Monosaccharide in Injected Sample (nmole):

Sample name	Fuc	GalNAc	GlcNAc	Gal	Glc	Man	GalA	GlcA
CK-1 (1M NaCl fr)	11.2293	9.2713	7.6965	28.0277	6.6071	8.0659	0.0000	0.6863
CK-2 (2M NaCl fr)	0.6905	1.7117	1.0513	3.3548	1.0016	-	0.0000	0.0000
CK3 (3M NaCl fr)	1.2842	3.6896	1.9991	4.5115	2.6818	0.7679	0.0000	0.3780
CK-4 (FT on PD10)			0.1050	0.3386		0.5791	0.0000	0.0000

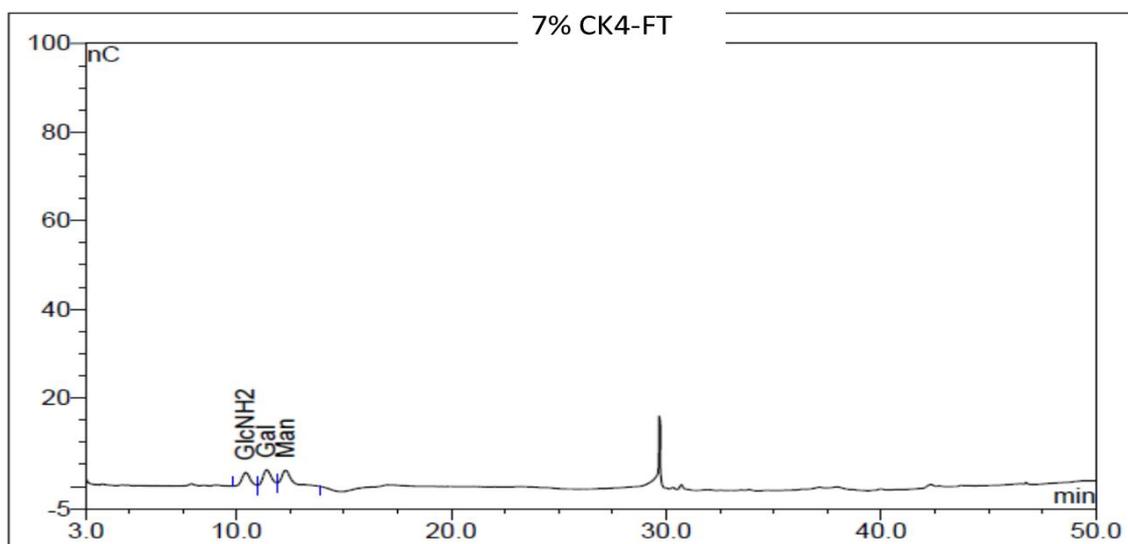
**Amount of Monosaccharide in Injected Sample (ng):**

Sample name	Fuc	GalNAc	GlcNAc	Gal	Glc	Man	GalA	GlcA	Total Carb
CK-1 (1M NaCl fr)	1843.402	2050.904	1702.543	5049.470	1190.335	-	0.000	133.238	11969.893
CK-2 (2M NaCl fr)	113.352	378.645	232.558	604.401	180.448	-	0.000	0.000	1509.405
CK3(3M NaCl fr)	210.814	816.176	442.221	812.792	483.153	138.345			2903.501
CK-4 (FT on PD10)	0.000	0.000	23.227	61.002	0.000	104.331	0.000	0.000	188.560

**Amount of Monosaccharide per  $\mu$ L (ng/ $\mu$ L)\*:**

Sample name	Fuc	GalNAc	GlcNAc	Gal	Glc	Man	GalA	GlcA	Total Carb	Amt taken for GRIL-LCMS(ml)
CK-1 (1M NaCl fr)	26.334	29.299	42.564	126.237	29.758	-	0.000	1.903	256.095	10
CK-2 (2M NaCl fr)	1.619	5.409	5.814	15.110	4.511	-	0.000	0.000	32.464	62.5
CK3(3M NaCl fr)	3.012	11.660	11.056	20.320	12.079				58.125	62.5
CK-4 (FT on PD10)	0.000	0.000	2.323	6.100	0.000	10.433	0.000	0.000	18.856	not taken

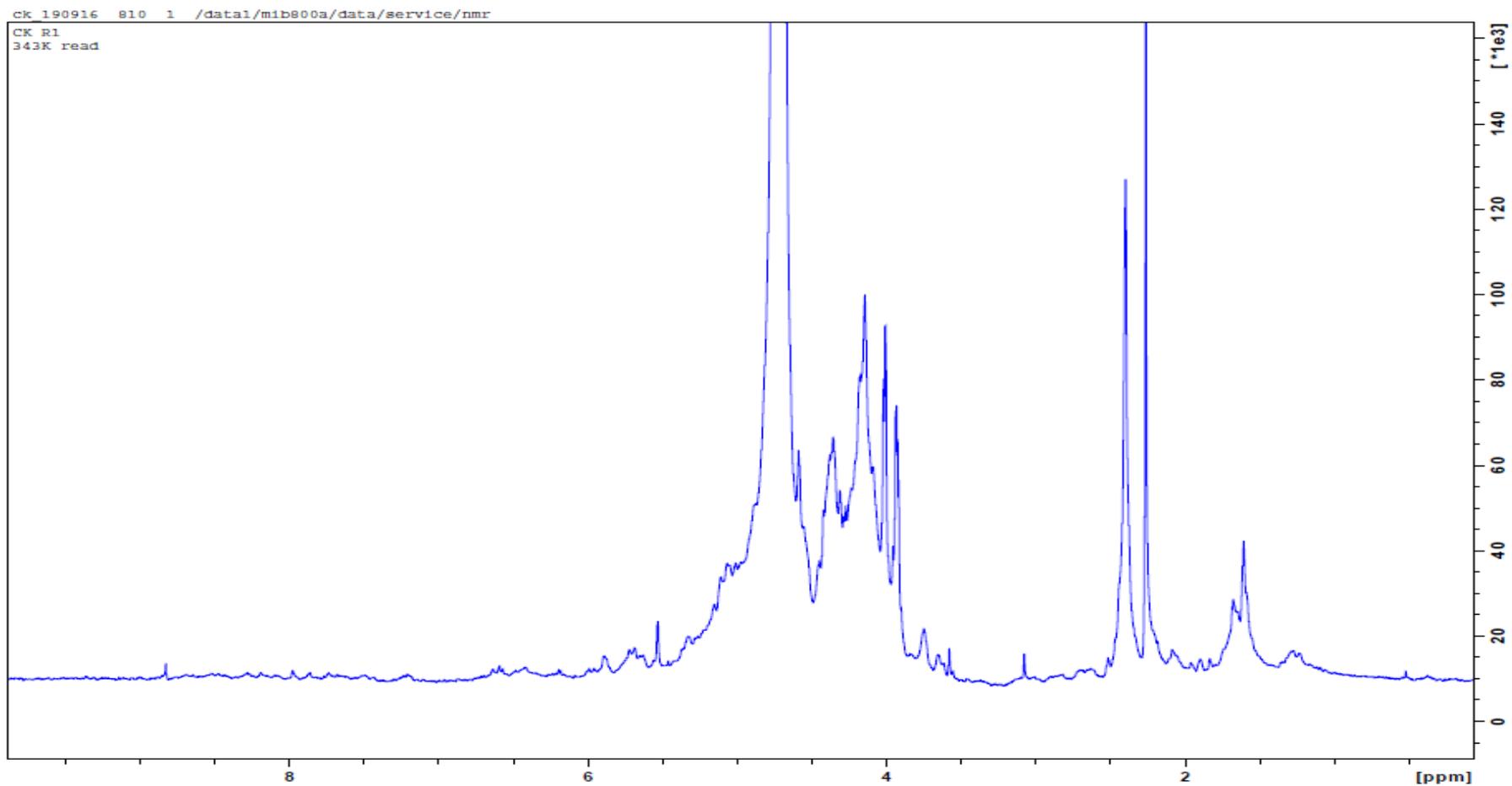
**Appendix IV: Monosaccharide analysis of the flow through fraction of the DEAE ion exchange column.**



Peak No.	Component Name	Retention Time min	Area nC*min	Rel.Area %	Height nC	Amount nmol
1	GlcNH2	10.43	1.373	26.61	3.097	0.1050
2	Gal	11.41	1.730	33.52	3.715	0.3386
3	Man	12.29	2.058	39.87	3.595	0.5791

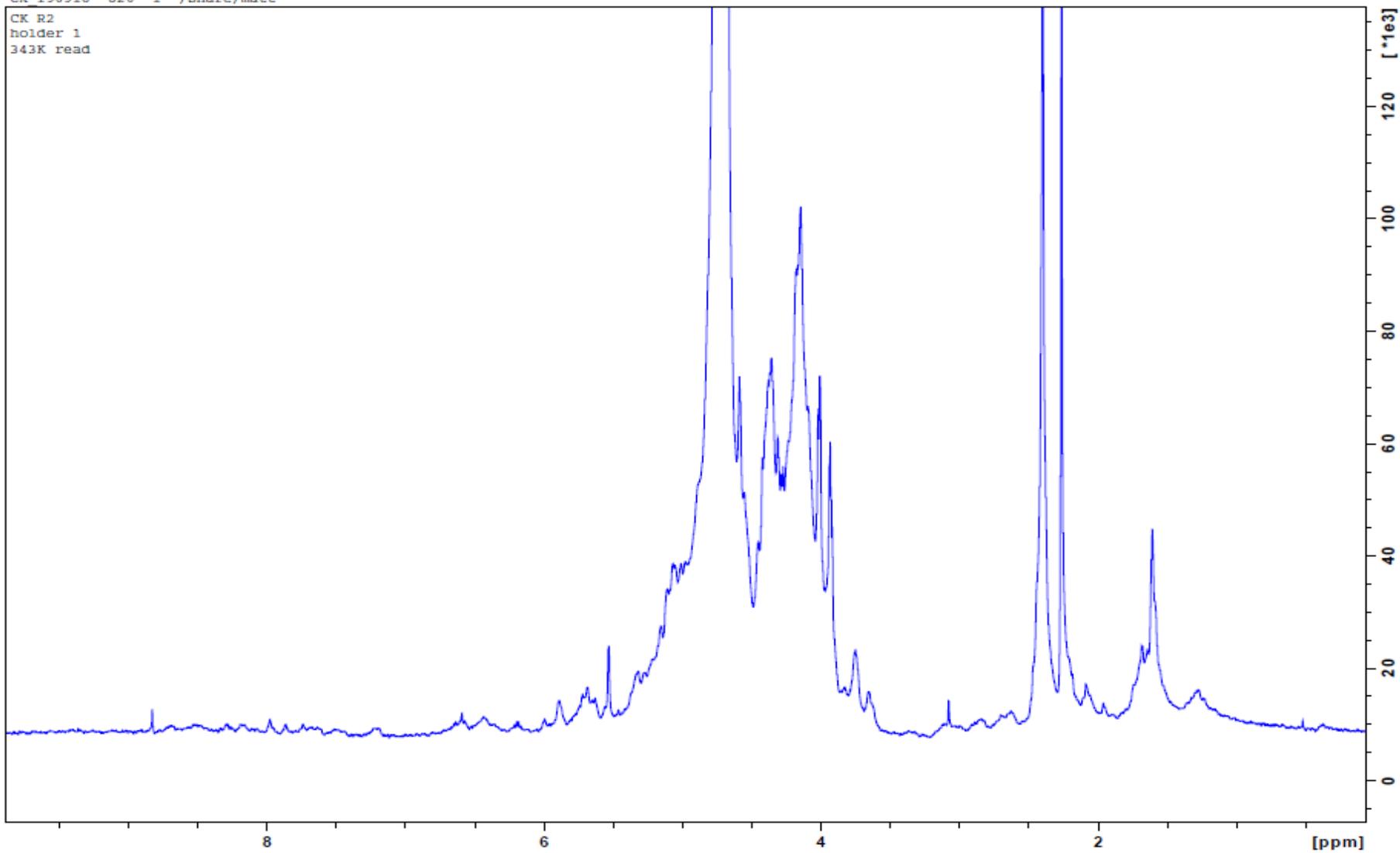
**Appendix V: NMR spectra of whelk GAG.**

1-  $^1\text{H}$  NMR spectra of the retrained fraction of 10K filter from fractions CK1, CK2 and CK3 after digesting with Hep lyase, spectra recorded at frequency of 800MHz at 70 °C.



ck\_190916\_820\_1 /share/matt

CK R2  
holder 1  
343K read



ck 190916 830 1 /share/matt

CK R3  
holder 2  
343K read

