AUTHOR:

TITLE:

YEAR:

OpenAIR citation:

This work was submitted to- and approved by Robert Gordon University in partial fulfilment of the following degree:
_____

The design, synthesis and evaluation of selective, non invasive imaging agents for atherosclerotic plaque

Jesintha Rose Jesudoss

A thesis submitted in partial fulfilment of the requirements of The Robert Gordon University for the degree of Master of Philosophy

September 2017

Declaration

I declare that this thesis was composed entirely by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Signed:

Date:

Dedication

I dedicate my work to each and every one who believed in me and strengthened me with their prayers, love and support

---

Especially to my parents Mr. Jesudoss and Mrs. Jayamani Regina; my brother Mr. Justin and my friend Mr. Samuel Joseph.

Acknowledgement

All praise, honour and glory to my Father God in Heaven, for giving me the wisdom, guidance and strength to handle everything in life.

I owe a great many thanks to a great many people who helped and supported me during my journey as a research student at RGU.
I'd like to express my deep and sincere gratitude to Dr. Graeme Kay my principal supervisor and the person who encouraged and believed in me throughout this research program. He was an excellent tutor and his guidance and support helped me a lot and especially during my thesis writing. Besides my advisor, I would also like to thank the rest of my research team: Dr. Alberto Di Salvo for NMR analysis and Prof. Cherry Wainwright and the Institute for Health and Welfare Research for providing the funding for my PhD. I would like to acknowledge the support and expertise provided by Prof. Paul Kong and Dr. Gemma Barron during the cell culture work.

I also want to express my never failing love and gratitude to Dr. David Mincher and Dr. Agnes Turnbull my former MSc tutors and project supervisors, without whom I probably may not have had the opportunity to do my MPhil in this field. I appreciate all the advice that I received from them over the years and not to forget the lovely discussions and celebrations outside work.

I further want to express my gratitude to all my colleagues who were instrumental in helping me progress through the years. I always enjoyed our constructive conversions about lab, work and the life after MPhil. Special thanks to Andrea MacMillan, Dr.Radisti, Evie, Dr.Amina, Dr. Gemma and Dr. Kyari Yates for always being there to talk, even when I just needed a coffee and sometime away from lab.

I am thankful to all the technical staff specially John, Kirstein, Beth, Moira, Marie, Tina, Laura and Lauren for providing an amazing, inspiring and friendly working environment.

Last but always the best part of my life- my family. Love you loads Appa Amma and Anna. The flawless love and endless sacrifices that you guys did for me shaped my personality and you always remain the sole owners of every goal I achieve. Thanks for all the wonderful family times and for helping me to grow into the kind of person who I like to be.

Finally, to my best friend and partner in crime, Joe. My life would never have been complete without you. 'Thanks' would be a very small expression for all that you have done for me, starting from the delicious curries that you cook for me right till the moments when you wiped my tears and said that you still believed in me. Time spent with you was a harmony to my soul and I'm very much looking forward to the new phase in our life.

God is good all the time.

Abstract

The aim of this research study was the synthesis of an imaging agent, which could be used to non-invasively image atherosclerotic plaques *via* MRI. In order to achieve this, a library of peptide substrates designed to be substrate specific for either legumain or MMP-2/9 were synthesized. This was achieved *via* a combination of both solid and solution phase peptide synthesis. The first phase of the research study focused on the synthesis of a targeted MRI contrast agent designed to image legumain expressing atherosclerotic plaques. This was achieved by using a Fluorescent Resonance Energy Transfer (FRET) active system to evaluate the substrate specificity of synthesized peptide sequences for legumain. The self quenched FRET substrates were engineered to carry a legumain sensitive bond in between the donor (5(6)-carboxyfluorescein) and quencher (aminoanthraquinone) moiety. Legumain mediated activation of the FRET substrates was measured *via* fluorescence spectroscopic analysis which showed the release in fluorescence that was initially quenched by the anthraquinone moiety. Optimized peptide sequences were then covalently coupled to the gadolinium chelate (Gd-DOTA) so that, upon enzymatic cleavage, the contrast agent could be released at the site of activation. Mass spectrometric analysis on the post enzyme digested samples (incubated with recombinant human legumain) showed that the substrates were fragmented by the activated enzyme *in vitro*. The targeted MRI contrast agent and FRET substrates were successfully synthesized and then characterized by low/high resolution mass spectrometry. Additionally cytotoxic studies on selected compounds and a model of the enzyme-digested product [Lys(DOTA-Gd)-Sp-AQ] *via* MTT assay demonstrated that the compounds did not affect the cell viability of MDA-MB-231 cell line.

The second phase of this project was the synthesis of an aminoanthraquinone peptide substrate coupled to a contrast agent (CA). This substrate was designed to be specific to MMP-2/9. Although the synthetic strategy was successful, subsequent reaction optimization will be required (increasing yield) before *in vitro* CA release could be evaluated. Thus, the targeted MRI CAs synthesized during this research study provide a starting point for novel imaging agents. These agents, which upon successful *in vitro* and *in vivo* trials could potentially be used to image legumain expression or MMP (2/9) expression or atherosclerotic plaques.

Keywords: Atherosclerosis; MMPs; Contrast agent; FRET; Legumain; Gd-DOTA.

Abbreviations

^{1}H NMR	Proton nuclear magnetic resonance
2-ME	2-Methoxyestradiol
5(6)-CBF	5,6-Carboxyfluorescein
ACE	Angiotensin converting enzyme
Ado	12-Aminododecanoic acid
AEP	Asparaginyl endopeptidase
Ala	Alanine
AQ	Anthraquinone
Arg	Arginine
Asp	Aspartic acid
BBB	Blood brain barrier
bFGF	Basic fibroblast growth factor
CAD	Coronary artery disease
CAs	Contrast agents
CDCl$_3$	Deuterated chloroform
CEST	Chemical exchange saturation transfer agents
CHD	Coronary heart disease
CPR	Cysteine proteases
CT	Computed tomography
DCM	Dichloromethane
DIM	3,3′-Diindolylmethane
DIPEA	*N,N*-Diisopropylethylamine
DMF	Dimethylformamide
dMRI	Diffusion MRI
DMSO	Dimethyl sulfoxide
DOTA	1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetraacetic acid
DTPA	Diethylenetriaminepentaacetic acid
ECM	Extracellular matrix
EDT	1,2-Ethanedithiol
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
Em	Emission
ESMS	Electrospray mass spectrometry
EST	Expressed sequence tags
EtOH	Ethanol
FDG	2-[^{18}F]Fluoro-2-deoxy-d-glucose
FITC	Fluorescein isothiocyanate
FLUVACS	Flu vaccination acute coronary syndrom
Fmoc	9-Fluorenylmethyloxycarbonyl
Fmoc-OSu	Fmoc *N*-hydroxysuccinimide ester
FP	Fluorescent protein
FRET	Fluorescence resonance energy transfer
Gd	Gadolinium
Glu	Glutamic acid

Gln	Glutamine
Gly	Glycine
HOBt	1-Hydroxy or N-hydroxy benzotriazole
HPLC	High performance liquid column chromatography
Hrs	Hours
ICAM-1	Intercellular adhesion molecule-1
IL	Interleukin
Ile	Isoleucine
i-Pr-OH	Isopropyl alcohol
IVUS	Intravascular ultrasound
kDA	Kilo Daltons
LDL	Low density lipids
Leu	Leucine
LSAM	Legumain stabilization and activity modulation
m/z	Mass/charge
MeOH	Methanol
Met	Methionine
MI	Molecular imaging
MMP	Matrix metalloproteinase
MRI	Magnetic resonance imaging
MT-MMPs	Membrane type MMPs
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NIRF	Near infrared fluorescence
Nva	Norvaline
NSF	Nephrogenic systemic fibrosis
OCT	Optical coherence tomography
PAMAM	Polyamidoamine
PDGF	Platelet-derived growth factor
PEG	Polyethylene glycol
PET	Positron emission tomography
Phe	Phenylalanine
ppm	Parts per million
Pro	Proline
PS-SPL	Positional scanning-synthetic peptide library.
PSV	Protein storage vacuole
PVD	Peripheral vascular disease
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphoniumhexafluoro phosphate
RER	Rough endoplasmic reticulum
Rt	Room temperature
Ser	Serine
SMC	Smooth muscle cells
SP	Signal peptide
SPECT	Single-photon-emission computed tomography
SPPS	Solid phase peptide synthesis
TA	Thioanisole
tBoc	Tertiarybutoxycarbonyl

TBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluroniumtetrafluoro borate
TETA	1,4,8,11-Tetraazacyclotetradocane-N, N', N'', N'''-tetraacetic acid
TFA	Trifluoroacetic acid
Thr	Threonine
TIMPs	Tissue inhibitors of metalloproteinase
TLC	Thin layer chromatography
TNF	Tumor necrosis factor
Tyr	Tyrosine
US	Ultra sound
Val	Valine
VCAM-1	Vascular cell adhesion molecule-1
VPE	Vacuole processing enzyme
VSMCs	Vascular smooth muscle cells

Contents

Abstract

Declaration

Dedication

Acknowledgement

Abbreviations

1. Introduction

2. Results and Discussion

1. Introduction

1.1. Atherosclerosis

Atherosclerosis can be defined as a disease related to the arterial networks characterised by the deposition of lipids, calcium, matrix components, inflammatory cells and smooth muscle cells inside the intima of a blood vessel (Homeister and Willis 2010). These deposits are known as plaques and are prone to cause several complications, for example:

- Plaques block the space inside the arteries and interrupt the flow of blood either partially or completely.
- The rupture of the plaque, due to thinning of the fibrous cap leading to formation of a blood clot or thrombus (Cipollone *et al*. 2005).
- They can also cause the arteries to weaken leading to formation of a localized blood-filled sac from the dilation of blood vessels, which is defined as an aneurysm (Ross 1993). The rupture of this sac leads to severe internal bleeding.

Cullen *et al.,* (2007) stated that the term 'atherosclerotic artery' was first described by Leonardo da Vinci (1452-1519) as a broadening of vessel walls due to 'excessive nourishment' drawn from blood. The phrase arteriosclerosis, which is now often used interchangeably for atherosclerosis, was first coined by a German-born French surgeon and pathologist Johann G.C.F.M Lobstein in 1833, but the actual term 'atherosclerosis' was introduced by Felix J. Marchand, a German physician. Lobstein related atherosclerosis to a hardened arterial wall formation due to alterations in tissues in correlation with factors such as age, stress and metabolic dysfunction.

Research into atherosclerosis started as early as the 18th century, but most of the results were not supported by experimental data. In 1856, the insudation theory by Virchow provided information on atherosclerotic lesions, but this work was also purely based on microscopic pictures of plaques at autopsy (Steinberg 2005). Approximately 50 years later, it was Anitschkow who demonstrated that the very first visible sign of a lesion would be seen only when the space between endothelium and the underlying inner elastic membrane was filled with lipid substances, which were polyblactic or monocytic in nature (Anitschkow and

Cowdry 1933). Experiments conducted by Timothy Leary in 1934, in lesions of rabbits with a high cholesterol diet, showed structural similarities to the lesion of human coronary arteries (Leary 1943). In 1951, it was found that the contents of atherosclerotic lesions in both animal models and human extracts were mostly of lipid origin and cholesterol (Duff and McMillan 1951).

1.1.1. Morphology of atherosclerosis

The first distinguishable form of the lesion is known as 'intimal xanthoma' or 'fatty streak' which generally is a very flat non-destructive collection or build-up of foam cells. These foam cells are normally present in most individuals in their middle age, but may not cause any health related problems. However, there are certain types of lesions, normally known as progressive or developing lesions, which tend to occur during the advanced stages of the disease.
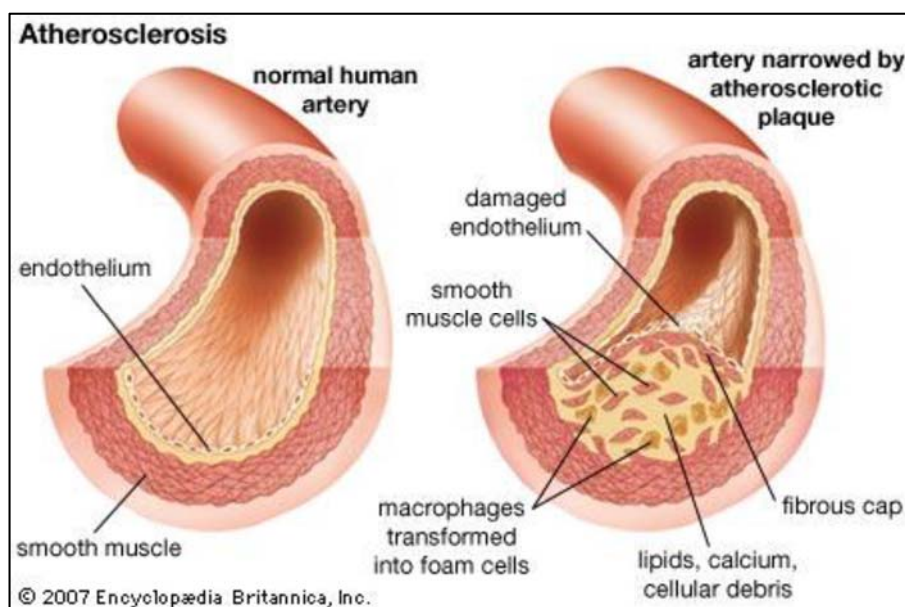
Figure 1: Pictorial representation of a normal artery and an artery damaged by atherosclerotic plaque (Hicks 2009).

These are characterized by a much more diverse and convoluted morphology, which may cause life-threatening problems (figure 1). Thus, based on various studies and knowledge about atherosclerotic plaques they can be broadly classified as (Virmani *et al*. 2000):

- **Pathologic intimal thickening lesions**: These lesions contain proliferative smooth muscle cells (SMC) within an extracellular lipid environment rich in a proteoglycan matrix.
- **Fibrocalcific plaques**: These lesions have fewer inflammatory cells but are generally rich in collagen and have a necrotic centre.

- **Fibrous cap atheromas**: These are similar to pathologic intimal thickening lesions, but they also contain macrophages derived from foam cells and inflammatory cells. They are characterized with a fibrous cap, rich in collagen, followed by a necrotic centre filled with extracellular lipid and apoptotic cellular debris. Thus, they are said to present features of instability such as a thin fibrous cap and a profuse amount of inflammatory cells (Homeister and Willis 2010). These lesions may be associated with lethal or non-lethal clinical manifestations.

Pathologic intimal thickening lesions and fibrocalcific plaques are more stable when compared to fibrous cap atheromas and are therefore rarely associated with a thrombosis that may lead to a heart attack or stroke. Conditions of chronic angina pectoris and claudication are the clinical indicators of fibrocalcific plaques, which are caused by the reduction in the flow of blood due to the unusual protrusion of these plaques into the lumen of the blood vessel. Sometimes the inner walls of pathologic initimal thickening lesions undergo minor damage (fissures) or total ulceration, which assists in acute thrombosis (Homeister and Willis 2010). Most of the fatal consequences of atherosclerotic conditions are caused by fibrous cap atheromas, which are otherwise called vulnerable or unstable plaques. The disruption of the plaque wall is mainly due to an increase in inflammatory cell proteases being released continuously or due to plaque neovascularisation triggering intraplaque haemorrhage (O'Brien *et al*. 2004). Further studies on plaque growth shed more light on the remodelling strategy of these plaques into positive or negative remodelling (figure 2). When the inner space of the artery remains uncompromised with the arterial wall swelling outwards it is defined as positive remodelling (Glagov phenomenon). Such plaques do not cause angina and are generally progressive in development. They normally don't intrude the luminal space until 40% of the cross sectional area is being used and only after which they start to restrict blood flow. Thus they grow for years and ultimately form the bulk of unstable plaques that normally rupture causing several complications in humans (Boudi and Ahsan 2010).
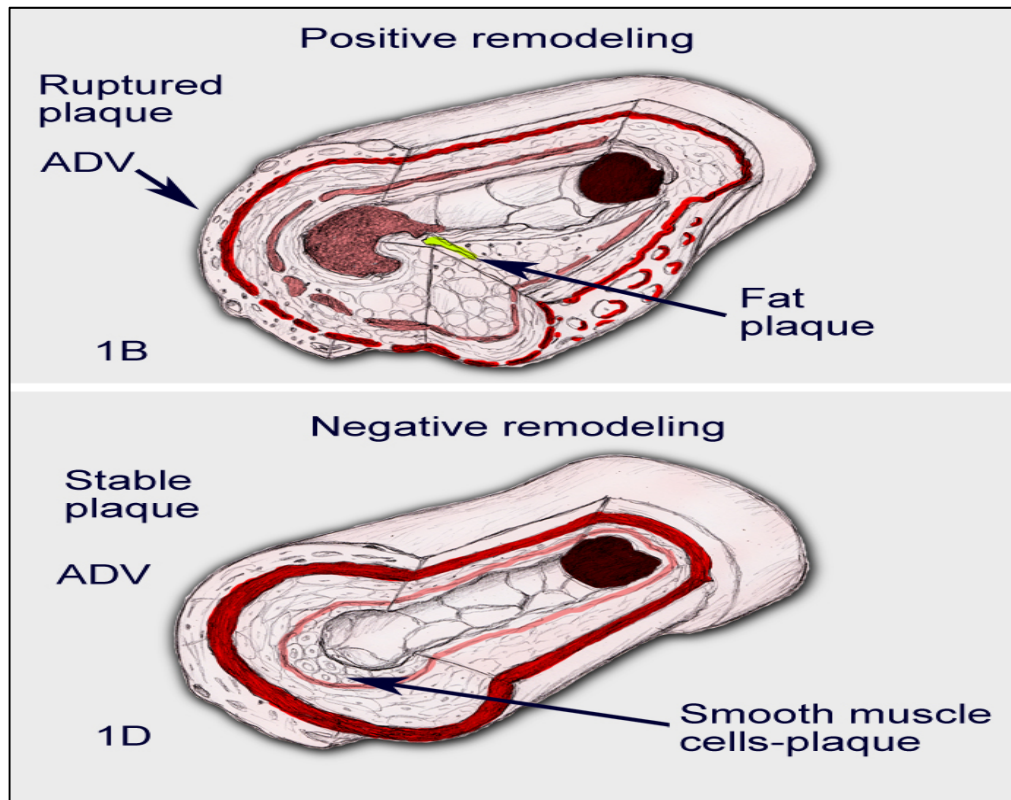
Figure 2: Pictorial representation of positive and negative remodelling of plaques (Boudi and Ahsan 2010).

Most of these positive remodelled plaques, on reaching maximum outward dilation, start to grow inward into the luminal space thus progressing into negative remodelling phase. These negative alterations develop steady angina and also cause rupture of unstable plaques and thrombosis (Boudi and Ahsan 2010).

1.1.2. Clinical impacts of atherosclerosis

Several pathological findings in both human and animal samples have described atherosclerosis as 'response-to-injury' due to endothelial dysfunction (Cullen *et al.* 2005). These abnormalities in the endothelium caused due to injuries trigger imbalanced reactions, which change the regular homeostatic conditions of the endothelium. These injuries initiate the inflow of leukocytes or platelets and also influence the endothelium to concentrate procoagulant properties instead of anti-coagulants forming cytokines, growth factors and vaso-active molecules. If the response to the endothelial dysfunction is not neutralized or controlled it increases the concentration of these inflammatory molecules leading to the formation of plaques or lesions, which gradually thicken the walls of the arteries (Ross 1999). These lesions further mature by developing into fibrous plaques structured by

SMCs and connective tissues enclosing a lipid-rich centre in patients who are more vulnerable to develop a plaque.
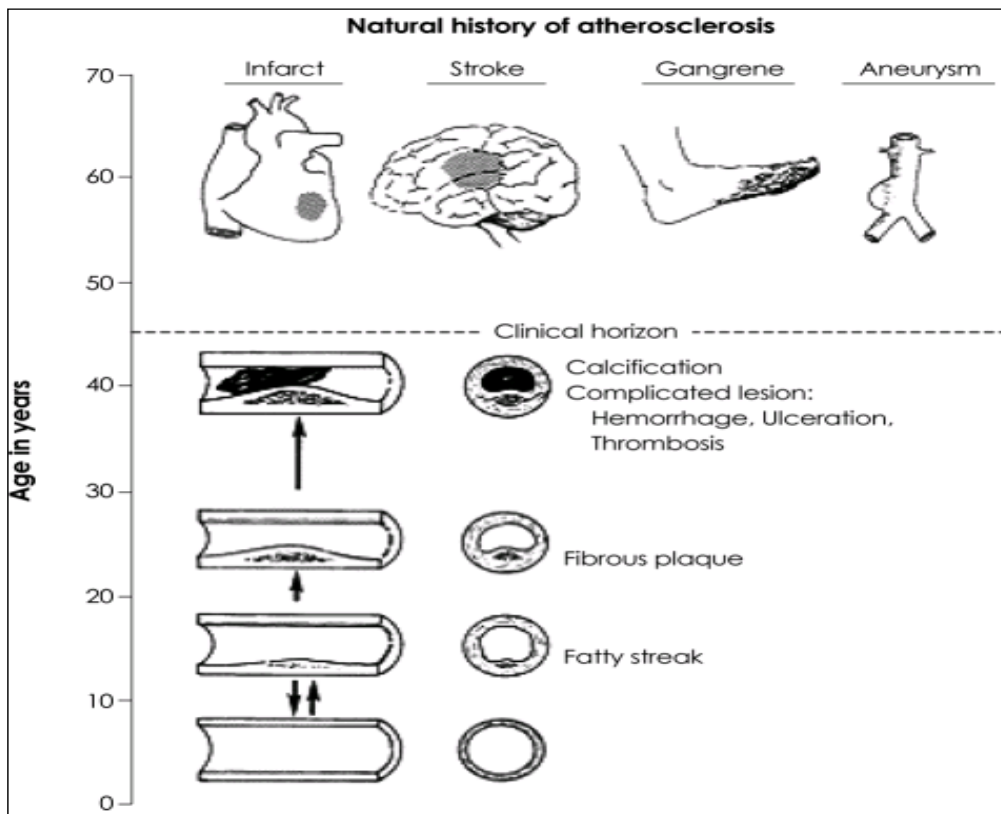
Figure 3: Progress and the types of atherosclerotic effect in humans (Hong 2010).

The plaques then undergo vascularisation, haemorrhage, and ulceration ultimately leading to several disorders in humans such as heart attack, stroke and peripheral vascular diseases depending upon the organ they affect as shown in figure 3. Thus the characteristic features of plaques, which are prone to rupture, are said to be large, showing active remodelling along with a large lipid centre almost covering 40% of the plaque space (Ross 1999).

1.1.3. Risk factors developing atherosclerosis

Historically, the major causes or risk factors for atherosclerosis have been thought to be mainly smoking, alcohol in-take, age, sex, diet, diabetes, high blood pressure or obesity (Kannel *et al*. 1961). However, with the vast improvement in our knowledge about the pathogenesis of the disease more interest in understanding the novel risk factors of the disease has emerged (Ross 1999). In an article by Hackam and Anand (2003) about risk factors in atherosclerosis, they have described four factors, which are said to be predictors of this disease even in people with very few chances of developing atherosclerosis.

Atherosclerosis is a complex multifactorial disease that presents a lot of risk factors, both of genetic and environment origin. These factors can be classified as modifiable and non-modifiable, factors such as sex, age and family history are grouped as non-modifiable factors whereas smoking, alcohol intake, diabetes, obesity, hypertension are classified as modifiable factors (Tolfrey 2002). The probability of occurrence of atherosclerosis remains high with the high occurrence of these risk factors. Understanding these risk factors helps prevent the severity of the disease and also serves as a guide to treat susceptible patients at the early stage of the disease (Kullo and Ballantyne 2005).

1.1.4. Current therapies to treat atherosclerosis

The aim in treating atherosclerosis is to decrease morbidity and mortality and to deter any further complications (Boudi and Ahsan 2014). The prevention and management of this condition requires control over adaptable life style, as well as monitoring and treating certain clinical conditions such as hypertension, hyperlipidemia and diabetes mellitus (Boudi and Ahsan 2014). Since the vascular characterization and anatomy of atherosclerosis is well studied, this has led to the development of novel therapies that identify the potential targets of this disease. Therapies, such as amino acid supplementation, antiplatelet agents, angiotensin-receptor blockers, angiotensin converting enzymes (ACE) and antioxidant therapies have shown to slow the development of atherosclerosis (Boudi and Ahsan 2014).

A number of pharmacological agents haven been used to treat angina. Agents such as beta-blockers restrain sympathetic stimulation of the heart inhibiting heart rate and heart muscle contractility. This ultimately reduces myocardial oxygen consumption and hence relieves angina. In the same way, nitrates such as glyceryl trinitrate, isosorbide mononitrate and erythrityl tetranitrate also work to reduce myocardial oxygen uptake by decreasing systemic vascular resistance (Boudi and Ahsan 2014). Calcium channel blockers are substances that restrict calcium entry into myocytes and vascular smooth muscle cells. These agents increase systemic vasodilation which in turn increases coronary blood flow. They also bring down arterial pressure and systemic vascular constriction and produce a negative inotropic effect (Boudi and Ahsan 2014). Fibric acid derivatives such as fenofibrate, benzafibrate and gemfibrozil work by increasing lipoprotein lipase activity and boost the metabolic breakdown of triglyceride-rich lipoproteins (figure 4).
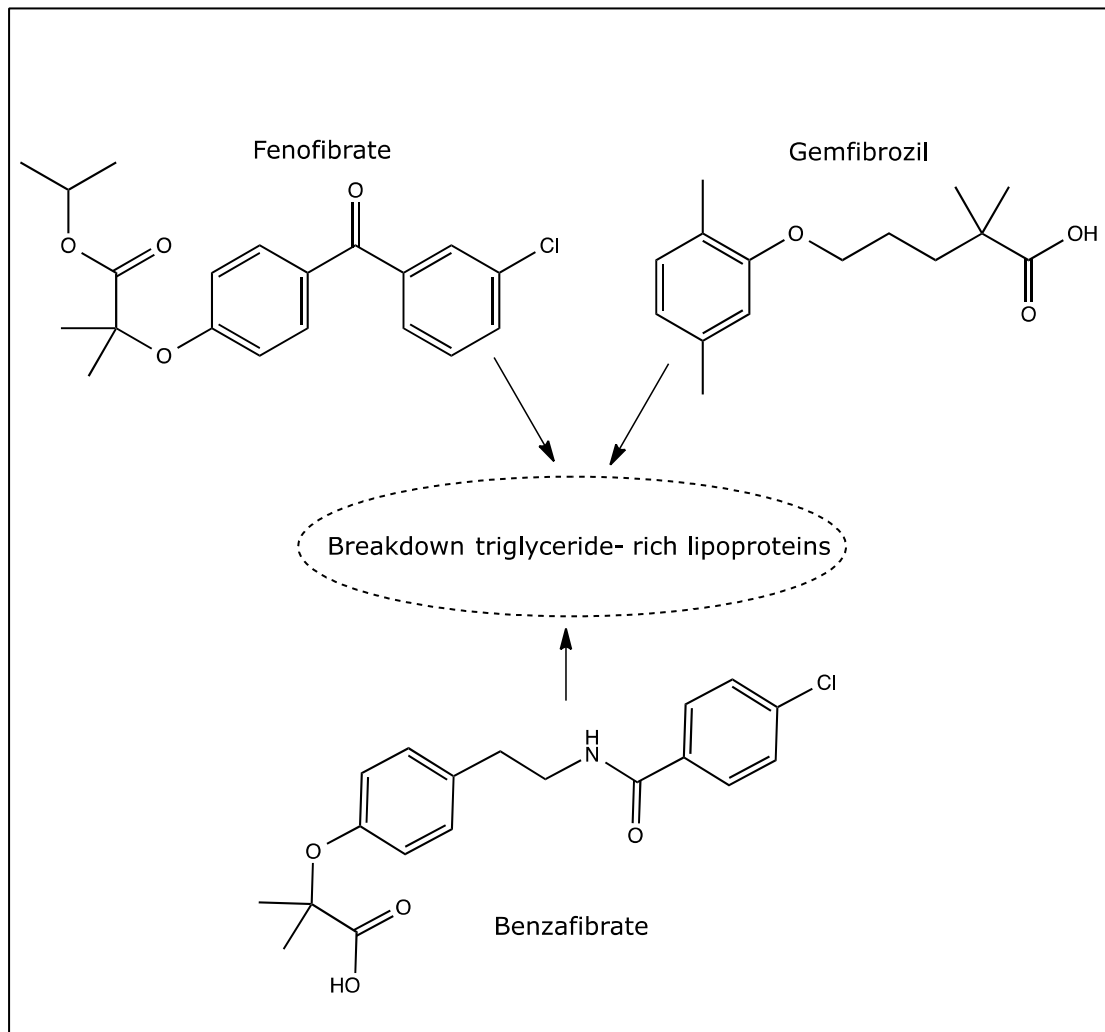
Figure 4: Chemical structure of fibrates.

These proteins are normally required for a rise in HDL concentration. It has been stated that fibrates usually decrease triglycerides volume by 20-50% and enhance HDL concentration by 10-15% (Boudi and Ahsan 2014). Other drugs that are normally taken by atherosclerotic patients minimise the concentration of lipid contents in blood. These drugs, such as statins and niacin, work by controlling the production of cholesterol and triglycerides. Other antithrombotic drugs such as warfarin, aspirin (lower strength), clopidogrel and prasugrel work by thinning the blood, therefore preventing any further deposit of plaques, reducing the injuries triggered by blood clots due to atherosclerosis.

However, these drugs might not be effective for people who are allergic to any of these chemical compounds and hence there is always a demand for complementary treatment (Baker 2008). Studies on hormonal therapies in clinical and animal analysis have demonstrated that the occurrence of coronary atherosclerosis can be highly reduced by estrogen, which works by inhibiting ACE (Brosnihan *et al*. 1999). Moreover, this hormone also increases the production of

endothelium-derived nitric oxide, which inhibits the proliferation of vascular smooth muscle cells (VSMCs) and prevents adhesion of leukocytes. 17-β-Estradiol (figure 5) is a naturally occurring estrogen hormone in humans which methylates to form 2-methoxyestradiol (2-ME) and this compound inhibits VSMC proliferation and thus prevents the progression of atherosclerosis (Barchiesi *et al*. 2006).
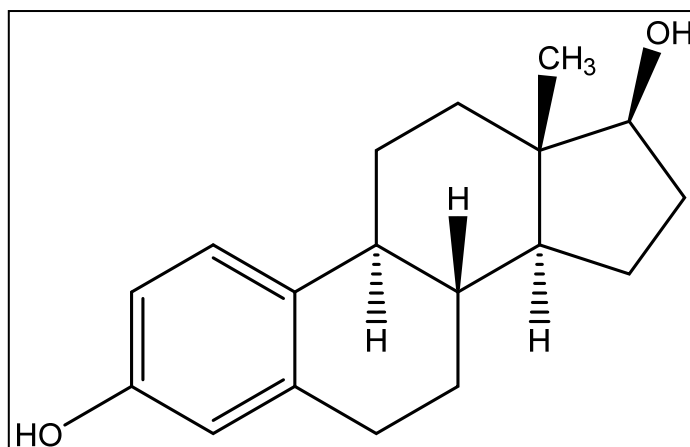
Figure 5: Structure of 17-β-Estradiol.

Studies on respiratory infection have shown that they might also enhance the condition of atherosclerosis in susceptible people (Meier *et al*. 1999). Evidence states that the occurrence of myocardial infarction is quite high during the cold seasons (The Eurowinter Group 1997) and during the spread of the influenza A virus. It has been stated that the virus causing influenza A also helps in the process of atherosclerosis by incubating and infecting the lesions (Haidaria *et al*. 2010). In a study conducted by the FLUVACS study group (Saskia and Kuiper 2011), it has been stated that mortality was reduced in people with cardiovascular conditions when they received the vaccine for the influenza A virus for two years. However, it has also been stated, that the use of vaccines to control cholesterol, oxidized LDL, endothelial cells, cytokines/growth factors and metabolic factors which generally reduces atherosclerotic progression could at the same time also increase plaque progression (Saskia and Kuiper 2011). The potential role of the immune system, in the pathogenesis of atherosclerotic plaques, has been studied by Chyu and colleagues (2011) where they explained the activity of lymphocytes and dendritic cells and their role in atherosclerotic plaque formation.

Moreover, clinical procedures involving the administration of immunoglobulin intravenously to treat immune mediated inflammatory conditions such as autoimmune disease suggest the idea of administering polyclonal immunoglobulin to treat atherosclerotic inflammations and pre-clinical evaluations on the same have provided promising results (Nicoletti *et al*. 1998). Surgical procedures such as partial ileal bypass have been practised since the 1960s to treat hyperlipidemia.

In this procedure, the length of ileum is shortened to reduce circulating cholesterol volume, which can also be thought of as a means of treating the atherosclerotic condition (Boudi and Ahsan 2014). In a recent study, Rajoria and colleagues (2011) demonstrated the down regulatory effect of MMP-2 and MMP-9 (which are normally overexpressed in atherosclerotic plaques) by a natural dietary complex 3,3'-diindolylmethane (DIM) present in cruciferous vegetables. This finding suggests that DIM could be used for reducing atherosclerosis plaque progression. Another means of treating the atherosclerotic condition is by the use of prodrugs. Huttunen and colleagues (2011) defined prodrugs as a chemically caged derivative of the active drug which is programmed to be released via a chemical or enzymatic reaction *in vivo*. The rationale behind this concept was to enhance the physicochemical properties of the drug, such as drug stability and solubility; to enhance drug pharmacokinetic/pharmacologic characteristic, such as oral absorption and selectivity and also to prevent unwanted side effects or non-specific toxicity to healthy cells/tissues (Rautio *et al*. 2008). In this manner, prodrugs are designed to target proteases as they are involved in many diseases. Proteases present an important and fundamental role in several biological and pathological functions by their regulatory mechanism, proteolysis (Choi *et al*. 2012). By their highly restricted functioning, proteases control DNA replication and transcription, cell growth and differentiation, neurogenesis, angiogenesis, stem cell mobilization, immunity, necrosis and apoptosis (Lopez-Otin and Bond 2008). Hence any abnormal modifications in the proteolytic function cause diseases such as cancer, atherosclerosis and neurodegenerative disorders (Choi *et al*. 2012) and therefore making these attractive targets for prodrugs. For protease activated prodrugs (PAPs) to be efficient the bond between the therapeutic/prognostic agent and the promoiety must be stable while in the blood stream and then degrade upon reaching the target protease (Rautio *et al*. 2008). This can be achieved by using peptide sequences which remain stable in blood but are highly specific towards the protease of interest have been developed (Choi *et al*. 2012). In their study on PAPs Choi and colleagues (2012) demonstrated that the synthesized peptide prodrugs were 97% stable after 24hrs in serum and plasma. They also suggested that the stability of the peptide substrates can be enhanced by the use of macromolecular structures such as dextran.

MMPs are one such target protease for prodrugs and are also one of the well-studied protease family. Mainly because of their pericelluar and extracellular location on cells they serve as valuable targets for delivering prodrugs to atherosclerotic and cancer microenvironments. Prodrugs have been synthesized

which are cleaved *in vivo* by MMPs to deliver compounds to atherosclerotic sites. Some of the known prodrugs that are specifically cleaved by MMP-2 are: melitin, TNF and Dox micelle whereas FITC based prodrugs and SeV are cleaved by MMP-9 (Edward 2008). However, with the growing number of atherosclerotic incidences there is a need for a transition from the therapeutic state of primarily treating the cardiovascular defects to one which identifies plaques even before they could progress to the stage of rupture, and this can be achieved by molecular imaging (Hwang *et al*. 2011). As discussed in previous sections, atherosclerosis can be described as lipid-driven chronic inflammatory progression involving the remodelling of the ECM within blood vessels, which subsequently result in plaque formation (Hansson *et al*. 2006). Moreover, it has been demonstrated that MMPs were over expressed in atherosclerotic plaques and were also involved in the degradation of all components of the ECM (Wagsater *et al*. 2011).

1.2. Matrix metalloproteinase (MMPs)

Matrix metalloproteinases, normally known as MMPs, belong to the family of proteins dependent on zinc and calcium (Borkakoti 2004). MMPs are involved in physiological functions such as healing of wounds, ovulation, development of the skeletal system and the development of the mammary gland involutions in adult reproductive tissues (McCawley and Matrisian 2001). One major function of MMPs is the production and degradation of the extracellular matrix generally known as ECM. This is done under tight control of the, naturally available endogenous inhibitors: α-2-macroglobulins (Nagase and Woessner 1999) and tissue inhibitors of metalloproteinases (TIMPs) (Hansen *et al*. 1993). MMPs are broadly classified into four groups depending on their sequence homology and substrate specificity:
- Collagenases
- Gelatinases
- Stromelysins and
- Membrane type MMPs (MT-MMPs)

MMP-1 (interstitial collagenase), MMP-8 (neutrophil collagenase) and MMP-13 (collagenase III) are the group of MMPs that belong to the collagenases. These MMPs are capable of cleaving fibrillar collagens I, II and III. MMP-2 and 9 belong to the second group representing gelatinases, and are known for their ability to disintegrate gelatins. Type IV collagen is also degraded by these gelatinases in the basement membrane. MMP-3, 10 and 11 belong to the third group of MMPs denoted as stromelysins. These MMPs are capable of degrading a broad range of ECM component, taking into account proteoglycans, laminis, fibronectin,

vitronectin and certain classes of collagenases as well. The final group of MMPs-the MT-MMPs are capable of degrading many of the ECM components and additionally they also have the capacity to activate other MMP types (Creemers *et al.* 2001). Some distinctive features of MMPs as elucidated by Denis and Verweij (1997); Nagase and Woessner 1999; Hansen and co-workers (1993) and Westermarck and Kahari (1999) are:

- MMPs are generally in their inactive form. They are produced as pre pro-enzymes and released as pro-enzymes.
- They have a distinct sequence made of 80 amino acids in the propeptide domain-PRCG(V/N)PD.
- MMPs have two catalytic zinc atoms of which one is present at the active site.
- The stability of the enzyme is sustained by two calcium ions.
- There is a highly conserved region between the catalytic domain and the propeptide *N*-terminal domain in the primary region.
- α-macroglobulin and TIMPs are the natural inhibitors of MMPs.

1.2.1. Parameters for the regulation of MMPs

Natural inhibitors such as α-macroglobulin and TIMPs regulate the functioning of this family of enzymes and when these regulations are disrupted they are said to cause various disorders in living organisms (McGeehan *et al.* 1994). Such disorders include rheumatism and osteoarthritis, tumor growth, angiogenesis and atherosclerosis leading to brain injuries, periodontal diseases, acute lung injury and acute respiratory distress syndrome (McCawley and Matrisian 2001; Westermarck and Kahari 1999).

The pro-MMPs are said to be in an inactive form due to the electrostatic interaction between the free thiol in the prodomain and the zinc atom, which is ligated to histidine in the catalytic pocket as shown in figure 6. This electrostatic interaction can be disrupted either by proteases or by non-proteolytic means ultimately activating the MMPs. During this process of activation the prodomain may not be removed completely, however the enzyme still remains active because the presence of the prodomain does not interfere with the activation of the enzyme and can be removed later either by autolysis intra or intermolecularly (Fu *et al.* 2008). This process of activation was termed as 'cysteine-switch' by Van Wart and Hansen (1990), which makes the MMPs catalytically active by any of the following three mechanisms:

- Alteration of the free thiol by physiological or non-physiological compounds.
- Activation by interacting with another proteinase leading to disruption of the prodomain.
- Allosteric interference causing inter or intramolecular autolytic cleavage of the prodomain.
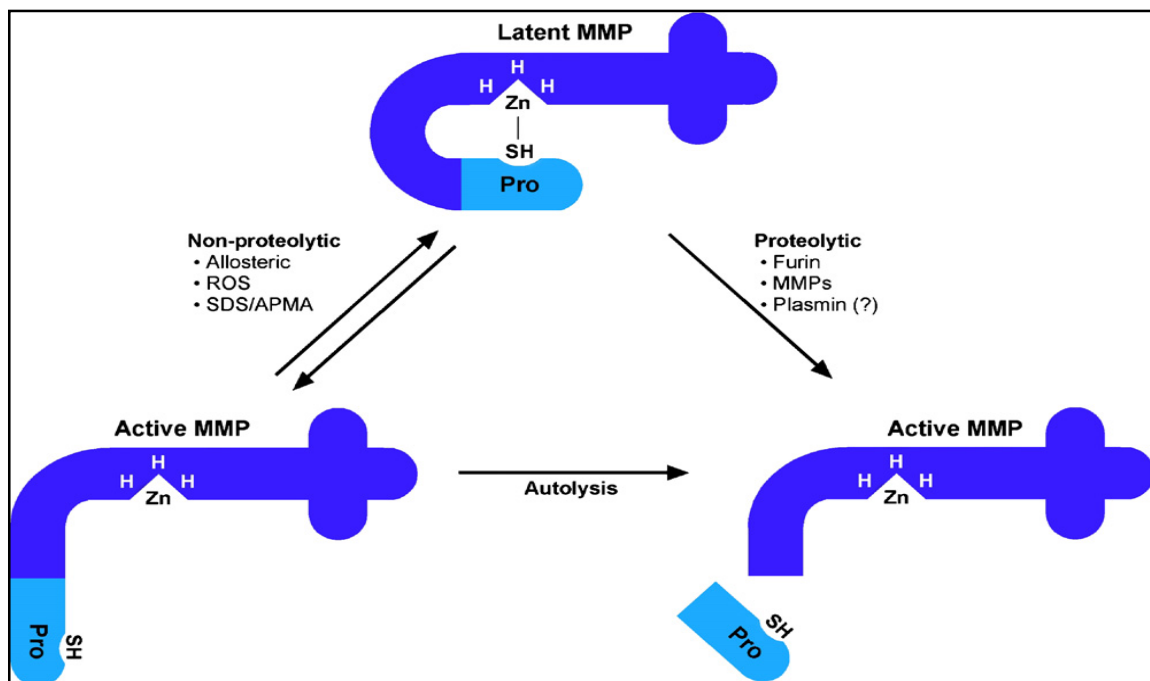
Figure 6: Activation of a latent MMP by proteolytic or non-proteolytic mechanism (Fu *et al*. 2008).

In healthy tissues and normal cells the concentration of MMPs are generally low and are said to be over expressed when stimulated by various factors like cell matrix, growth factors, cytokines, cell-cell interactions, oncogeneic transformations, physical stress or tumor promoters (Westermarck and Kahari 1999). Once the MMPs are activated, they continuously degrade the ECM and hence it is very important that their functioning is kept under tight control.

1.2.2. Structure of a MMP

MMPs normally have three domains (figure 7), the signal peptide (1) which operates as the signalling tool for secretion, a pro-peptide domain (2), which detaches itself when the enzyme is activated, and finally a catalytic domain (3) (Nguyen *et al*. 2001). The MMPs remain inactive because of the presence of the cysteine switch which is ~80 amino acids long (Himelstein *et al*. 1994-95). The MMPs are activated by the removal of the amino terminal domain from the enzyme in a proteolytic fashion (Denis and Verweil 1997).
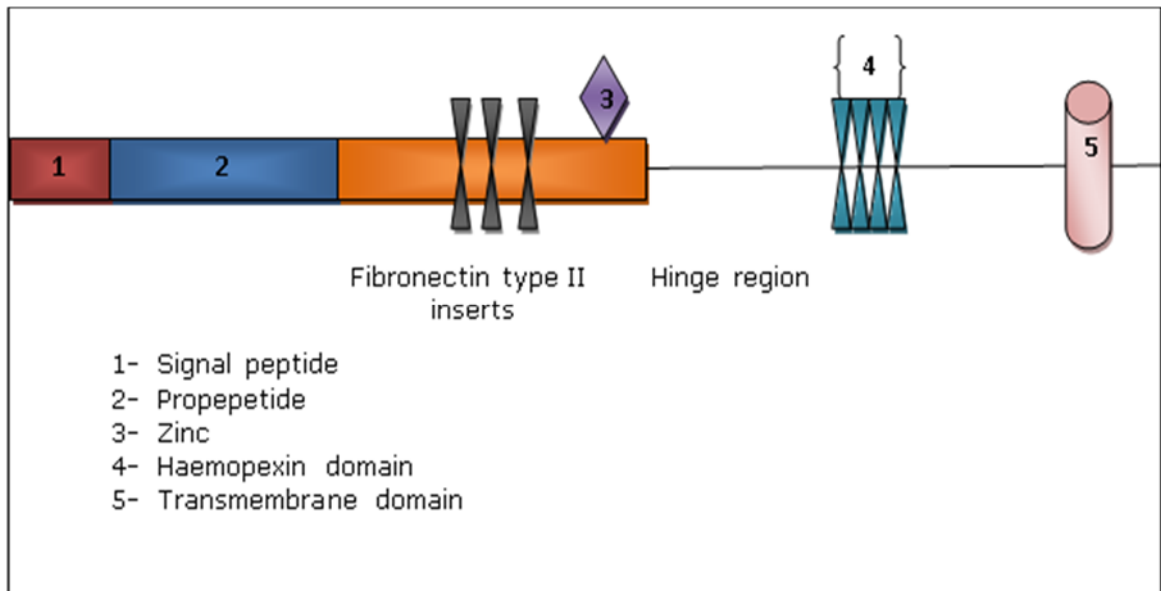
Figure 7: Structure of a MMP (adapted from Hidalgo and Eckhardt 2001).

It has been demonstrated that the reduction in the components of the matrix, (particularly that of fibrillar collagens in the fibrous cap) lead to thinning of the cap wall (Shah 2003). Increased breakdown of the matrix is a characteristic feature of a particular group of MMPs involving MMP-1, MMP-2, MMP-3 and MMP-9. These MMPs are expressed by the inflammatory cells and, to a smaller quantity, by SMCs and endothelial cells in atherosclerotic plaques (Galis *et al*. 1995). Since most of the properties of these proteases are well studied with a clear knowledge about their chemistry in most diseases they can be efficiently used as targets for drug design. With progress in the fields of molecular imaging and the use of fluorescent probes such as near infrared fluorescent probes (NIRF) (Zucker and Cao 2001), these proteases (MMPs) can be characterised in their natural environment of action and may also be exploited as diagnostic tools to detect atherosclerosis or tumour at a much early stage of progression (Neefjes and Dantuma 2004).

1.2.3. MMPs and their role in atherosclerosis

MMPs play a vital role in the development, decay and rupture of atherosclerotic plaques (Galis *et al*. 1994). Aspects of which are listed below:

- The damage of the intimal ECM is the main reason for the rupture of the plaque initiated by proteases.
- MMPs are normally found to be active at neutral pH, which is the most desirable condition for vascular ECM remodelling to occur.
- MMPs have at least one component of the ECM as a recognized substrate.

- The thickening and growth of the injured blood vessels cause the accumulation of new cells derived by VSMCs (Newby 2005).

Atherosclerotic plaques are said to be in a dynamic state undergoing constant remodelling of the ECM, to maintain the stability of the plaque (Loftus *et al*. 2000). There is evidence to show the over expression of MMPs in such plaques along with their involvement in aneurismal and degradation process (Thompson *et al*. 1995). Overexpression of these MMPs in atherosclerotic lesions leads to unstable plaques, which are prone to rupture resulting in clinical conditions such as thrombosis, occlusion, heart attack or stroke (Galis and Khatri 2002). One major source of MMPs is the foam cells, which are derived from macrophages. These macrophages are normally said to be highly concentrated in human and animal atherosclerotic plaques (Galis *et al*. 1995). The early stage of the growth of the plaque is initiated by VSMCs, which cause the migration, and over expression of MMP-2 at the site of injury (Galis *et al*.1994). Over expression of MMP-2 and 9 is detected in injured arteries as they play a vital role in remodelling (Chesler *et al*. 1999). There is also accumulation of other MMPs including MMP-1, 3 and 9 as a result of the signalling of inflammatory cytokines and tumour necrosis factor-α at the site of injury (Yanagi *et al*. 1992). Furthermore immune activation also leads to increased production of MMP-1, 3, 8 and 9 in the vascular SMCs by the addition of T-lymphocytes and recombinant CD40 ligands (Schonbeck *et al*. 1999).

Work carried out by Loftus and colleagues (2000), has shown an increase in the expression of MMP-9 in most of the unstable carotid plaques and has also demonstrated that most of the plaques with the characteristic features of instability such as the presence of cerebral particulate embolization, plaque rupture and intraplaque haemorrhage have all demonstrated increased concentration of MMP-9. It has been previously stated that the site of plaque burst is concentrated with inflammatory infiltrates, which are mostly characterized by macrophages, foam cells and T-lymphocytes (Moreno *et al*. 1994). Buja and Willerson (1994) have shown that these inflammatory infiltrates play a vital role in plaque rupture.

Welgus and his colleagues (1990) explained that macrophages were effective producers of MMP-9 and also the increased expressions of MMP-3 and MMP-1 in carotid plaques was related with macrophage and mast cell infiltration (Nikkari *et al*. 1995). All these studies demonstrate that MMPs are up regulated at the site of injury to enhance the repairing process of the injured site. Thus the occurrence of MMPs does not only have a role in the pathogenesis of atherosclerosis, but it is also a strong indication of the stability of the plaques (Galis 2004). With increased

knowledge about the structural chemistry of MMPs and available peptide substrates specific for particular MMPs, they can be effectively used as diagnostic tools for studying atherosclerotic plaques by using them as markers for molecular imaging in MRI or fluorescent probes techniques. Hence MMPs can be used as targets for both therapy and imaging (Lancelot *et al.* 2008).

1.2.4. MMP-2 and MMP-9

Several research studies have demonstrated that MMPs particularly MMP-2 and MMP-9 have a potential involvement in angiogenic lesions of both atherosclerotic and tumour origin and in their progression (Kurizaki *et al.* 1998). MMP-2 and MMP-9 are generally termed as gelatinases because they can cleave gelatin, which is a denatured form of collagen. The one and main difference between gelatinases and other MMPs is the presence of three fibronectin type II repeats which is the key reason for enhanced binding affinity towards collagen (Kridel *et al.* 2001). Gelatinases are similar in structure having the three-fibronectin units placed in the catalytic domain. They also have similar specificities for substrates like collagen, type X, VII, V and IV, fibronectin, fibrin, interleukin-I and α-1 proteinase inhibitor (Xu *et al.* 2005).

The major difference between the two MMPs is their mechanism of activation (table 1):

- The mesenchymal cells are responsible for the activation of MMP-2 (Lozito *et al.* 2014) and they are finally activated by membrane type MMP (MT-MMP) and more recently it has been found that they are also activated by legumain (Mattock *et al.* 2010).
- Cells such as neutrophils and eosinophils are said to produce MMP-9 and they are ultimately activated by MMPs such as MMP-3 (Xu *et al.* 2005).

Similar to other MMPs, both MMP-2 and 9 have a hemopexin domain separated from the catalytic domain by a fibronectin motif. Moreover, MMP-9 has a domain similar to collagen. As mentioned previously this hemopexin unit at the C- terminal region favours binding of MMP to substrates and may also be involved with TIMP activities (Nguyen *et al.* 2001).

Common names	Gelatinase A,	Gelatinase B,
Nomenclature	MMP-2	MMP-9
Substrate specificity	Gelatin type I, IV, V, collagen, elastin	Gelatin type I, IV, V, collagen, elastin
Molecular mass	72kDa	92kDa
Molecular mass of active units	64kDa, 62kDa	82kDa, 67kDa
Physiological activators	Type I collagen, lipopolysaccharide, hepatocyte growth factor, thrombin, APC, legumain and MT1-MMP.	Serine proteases
Latent form binds to TIMP	TIMP-2	TIMP-1

Table1: Features of gelatinase A and B (Nguyen *et al*. 2001).

1.2.5 MMP-9/2 and its substrate specificity

Overexpression of MMPs in vascular and cardiac tissues has been associated in the pathogenesis of several cardiovascular disorders such as atherosclerosis. Development of atherosclerotic plaques, also known as negative remodelling, requires the production and migration of smooth muscle cells (SMCs) within the arterial wall, and this progression is largely dependent on MMPs, particularly MMP-2 and MMP-9, which can degrade reticular type IV and V collagen surrounding SMC (Guo *et al*. 2009). Experiments conducted on *Ldlr$^{-/-}$Apob$^{100/100}$* mice demonstrated that MMP-2 and MMP-9 were the two main proteases overexpressed in atherosclerotic mice model (Wagsater *et al*. 2011). Based on this evidence part of this project involved the synthesis of MMP-2 and MMP-9 specific peptide substrates. The rationale was to exploit the peptide hydrolysis ability of MMP-2/9 to deliver a prognostic chemical agent in the atherosclerotic microenvironment, which in this project was a gadolinium-DOTA chelate. In order to avoid any false signals or unspecific loading of the chemical agent disturbing the healthy cells, it is essential for the imaging agent/ therapeutic drug to be selectively deposited at the site of plaque formation. Hence, the stability and selectivity of the peptide substrate is vital for the design of such prodrugs. Thereby the peptide substrate

should have the potential to carry the prognostic/therapeutic agent to the desired site while still being stable in the circulatory system.
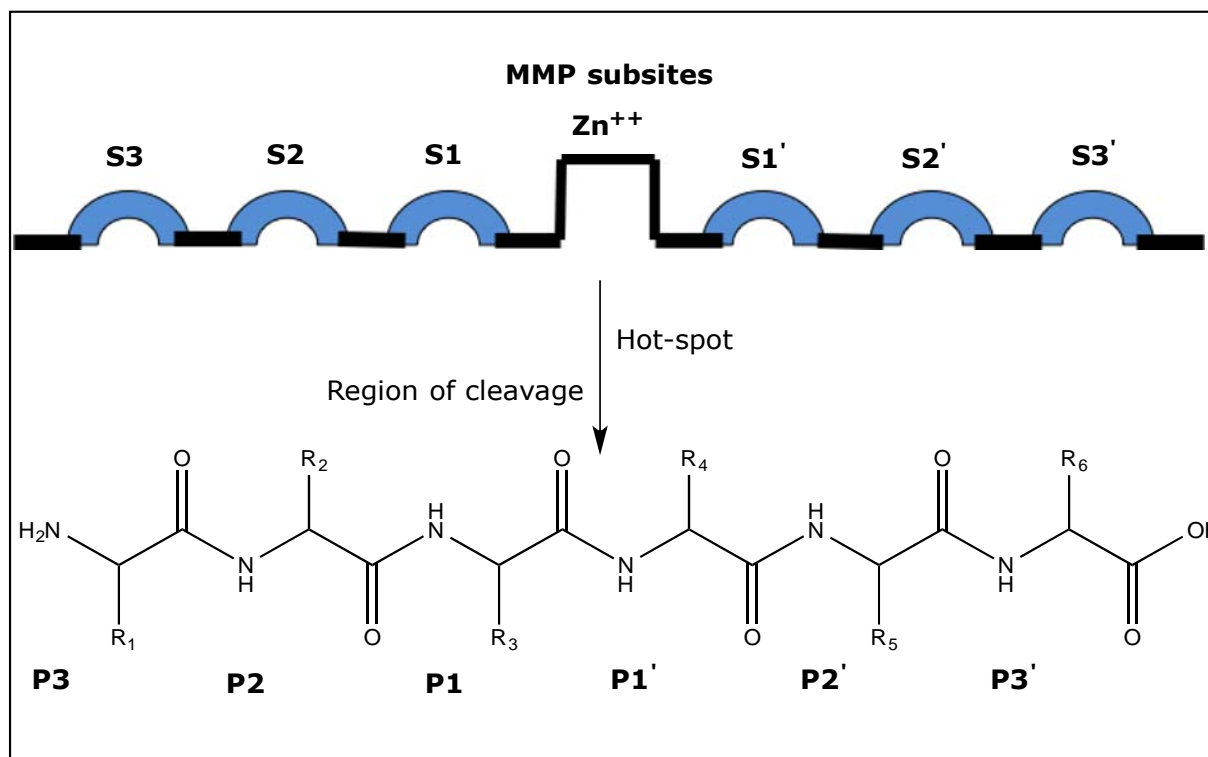
Figure 8: A representation of scissile bond.

The synthesis of peptide substrates activated by MMP-9/2 to release the required therapeutic drug or imaging agent can be done, by mimicking the available resources of natural substrates (figure 8). The peptide substrates are constructed in a way to mirror the subsites/pockets of the corresponding enzyme into which the side chain residues of the relevantly positioned amino acid projects (Atlas *et al*. 1970). A peptide sequence around the active site would be labelled as P3-P2-P1-P1′-P2′-P3′ and the region of cleavage or 'hot-spot' is present between P1-P1′ positions (Lim *et al*. 2010). The symbol P denotes the position of amino acids in the sequence. On drawing the peptide structure in a conventional way as in a scissile bond, the arrangement will have the *N*-terminus to the left followed by labelling the amino acids as P1 and those amino acids towards the right are labelled as P1′ (figure 8). The major amino acids for proteolytic hydrolysis by the enzyme (MMP-2/9) are placed within P3-P3′ position.

In MMP-2 and MMP-9, the S1′ subsite has a deep hydrophobic pocket, that prefers a leucine or isoleucine residue at this location. Normally MMP-2 tolerates hydrophobic amino acids on the carboxyl side of the sequence (right side of scissle bond) much better than on the other side (Massova *et al*. 1997; Murphy and Nagase 2008). Synthetic hydrophobic amino acid residues such as norvaline or norleucine with straight and long side chains and also naturally available branched

amino acids such as valine, leucine and isoleucine are favoured at the S1′ subsite. Moreover MMP-9 can tolerate further branching at the β and γ-carbon. Compared to S1′, the S1 subsite is shallow in nature with limited preference for amino acids residues at this position. However, it was demonstrated that glycine was the most preferred amino acid at this position (McGeehan *et al*. 1994).

The S2′ subsite has a small hydrophobic surface, and with MMP-9 this position accepts amino acids with a small hydrophobic side chain or with a large hydrophilic side chain. The S3 subsite is characterized with a deep hydrophobic pocket and also creates a deviation from linearity into the active site cleft. Hence an amino acid similar to proline; which has a five membered ring structure, is well tolerated at this position (Kridel *et al*. 2001). Chen and co-workers demonstrated that the S2 position was the main area of difference in selectivity between MMP-2 and MMP-9.

MMP	P3	P2	P1	P1′	P2′
	Hydrophobic pocket	Long side chain amino acids	Small subsite	Amino acids with hydrophobic side chains.	Very broad specificity for amino acids residues
MMP-2	Proline	Serine, lysine, alanine and glutamic acid	Alanine and glycine	Leucine and isoleucine	Valine, leucine and isoleucine
MMP-9	Proline	Arginine and aspartic acid	Glycine	Leucine, isoleucine and norvaline	Serine and threonine

Table 2: Summary of preferred amino acids for MMP-2/9.

For example, if glutamic acid in the sequence was substituted for aspartic acid then the substrate specificity for MMP-2 would be switched to MMP-9 (Chen *et al*. 2002). Normally the selectivity and sensitivity of cleavage by MMP-9 can be enhanced by placing amino acids with large side chain groups at position P2. The presence of arginine at P2 position was found to be well tolerated by MMP-9.

Furthermore the positon of arginine at this location improved the proteolytic effect of MMP-9 whereas; this had unfavourable hydrolysis for MMP-2 (Murphy *et al*. 2002; Chen *et al*. 2002). Additional studies have shown the successful cleavage of a prodrug (EV1-FITC) by MMP-9 with a straight chain residue at the α-carbon of norvaline which was found to perfectly occupy the S1′ pocket of the enzyme (Van Valckenborgh *et al*. 2005) (Table 2).

1.2.6. Legumain in MMP activation and plaque progression

Papaspyridonos and colleagues (2006), in their findings on gene expression as a novel tool for detecting unstable atherosclerotic plaques, demonstrated the up regulation of MMPs in atherosclerotic lesions. Moreover, it lead to important information about a novel gene, termed legumain, as a potent activator of MMPs at protein level (Papaspyridonos *et al*. 2006). RT-PCR, western blot and affymetrix analysis demonstrated consistent up regulation of legumain in vulnerable plaques at both transcript and protein levels. Although this enzyme has not been correlated with atherosclerosis previously, studies by Hashimoto and colleagues (1999) has highlighted the expression of legumain by macrophages and its involvement in activation of MMPs and cathepsins as vital to the role of this enzyme in atherosclerosis (Shirahama-Noda *et al*. 2003). It was also shown that the occurrence of legumain was associated, significantly, with the expression of CD68-macrophage markers providing clear evidence that macrophages largely expressed legumain (Papaspyridonos *et al*. 2006).

Liu and colleagues (2003) also demonstrated the involvement of legumain with the invasion and metastasis of tumour; largely by increased ECM destabilization. This function could have been possible by the initiation of progelatinase A (MMP-2), which already has a vital role in ECM degradation. Moreover, the activation of MMP-2 involves hydrolysis of an asparaginyl bond which is the universal hot-spot for legumain recognition (Nagase 1997). Liu also confirmed that there was an increase in the concentration of the active form of MMP-2 (62kDa) when the concentration of legumain was increased while incubation with the inactive 72kDa MMP-2. It was also demonstrated that there was very little or zero response when pro MMP-9 were incubated with legumain even though MMP-9 is a functionally similar protease to MMP-2 in tumour and atherosclerotic plaque progression (Liu *et al.* 2003; Chen *et al*. 2001). Moreover, it was demonstrated that normal cells expressing legumain were not affected by legumain-based prodrugs used in treating cancer, indicating that these kinds of therapeutics or diagnostics require conditions that are not similar to normal tissues. Thus there is every possibility of

legumain effecting the remodelling of ECM through processing of other proteases specifically MMPs (Morita *et al*. 2007) explaining their involvement in human atherosclerosis.

1.3. Legumain

Asparaginyl Endopeptidase (AEP), otherwise known as legumain (EC 3.4.22.34) is a distinct member of the cysteine peptidases. The name legumain was first used by Kembhavi and colleagues (1993). It was given to an endopeptidase that was initially isolated and characterized from leguminous seeds and *Vigna aconitifolia* (moth beans). Legumain is expressed in a diverse range of organisms such as parasites, plants and mammals (Halfon *et al*. 1998). They are involved in functions such as processing stored vacuole proteins in plants, digestion of haemoglobin in blood flukes, processing precursor polypeptides site-specifically during protein maturation and also protein degradation (Halfon *et al*. 1998). Legumain belong to a very large family of proteolytic enzymes known as cysteine proteinases (CPRs) that can be further categorized into 30 individual sub-families based on their variations in molecular structure. Initially, only three families of CPRs, were assumed to be represented in mammals-consisting of:

- C1- Papain family, including cathepsins-B, H, L, S that are predominantly lysosomal in nature, responsible for lysosomal/endosomal system proteolysis but can also be processed to act extracellularly.
- C2- Calpain
- C14- Caspase; C2 and C14 were secreted in the cytosolic fraction of cells and were involved in restricted proteolysis of cytosolic substrates.

However, Chen and colleagues (1997) were later able to demonstrate the presence of legumain (C13) along with other mammalian CPRs. According to *MEROPS* database, legumain was further grouped along with four other sub-families of CPRs comprising of clostripain (C11), caspases (C14), gingipain (C25) and separin (C50) to form a distinct cluster characterized as clan CD, (Rawlings and Barrett 1999; Chen *et al*. 1998) mainly based on the amino acid sequence of their catalytic motif. In 1991, Hara-Nishimura and colleagues were able to extract Asn-specific CPRs from sprouting seeds of castor beans and soybean cotyledons (Scott *et al*. 1992). These enzymes were found to transform precursor polypeptides of 2S albumin and 11S globulin into mature proteins by catalysing, restricted proteolysis of Asn residues *in vitro*. The proteinase sequence as in Schistosoma mansoni-an invertebrate published by Klinkert and colleagues as early as in 1989 was found to be similar to the legumain family. It is also being stated that legumain is a vital

enzyme for matrix degradation (Morita *et al.* 2007), antigen processing (Maehr *et al.* 2005), atherosclerosis (Clerin *et al.* 2008), tumorigenesis (Liu *et al.* 2003; Luo *et al.* 2006) and is also associated with several pathological disorders (Lee and Bogyo, 2010). The functions of legumain vary according to the conformational state of the corresponding substrate protein. For example, with access to the regular processing sites in a precursor polypeptide they can only act as VPE, but when the substrate is open for unrestricted proteolysis due to any mutated conformation, then it can function as a degradative enzyme. The sensitivity and quality of these reactions appears to be the effect of co-evolution of enzyme and substrate (Muntz *et al.* 2002).

1.3.1. Structure and role of legumain in mammals

Chen and colleagues (1997) were able to demonstrate that this enzyme was also present in humans by expressed sequence tags (ESTs), which coded the presence of sequence similarity to legumain. The first mammalian legumain was isolated from the kidney of pigs (Chen *et al.* 1997). It was also demonstrated that this enzyme regulated the biosynthetic process of lysosomal enzymes such as cathepsin B, L and also activated proMMP-2 (Shirahama-Noda *et al.* 2003; Smith *et al.* 2012). Legumain is expressed in a number of mammalian tissues (kidney, placenta, spleen, liver and testis); however, their activity was found to be predominant in the kidneys (Morita *et al.* 2007).
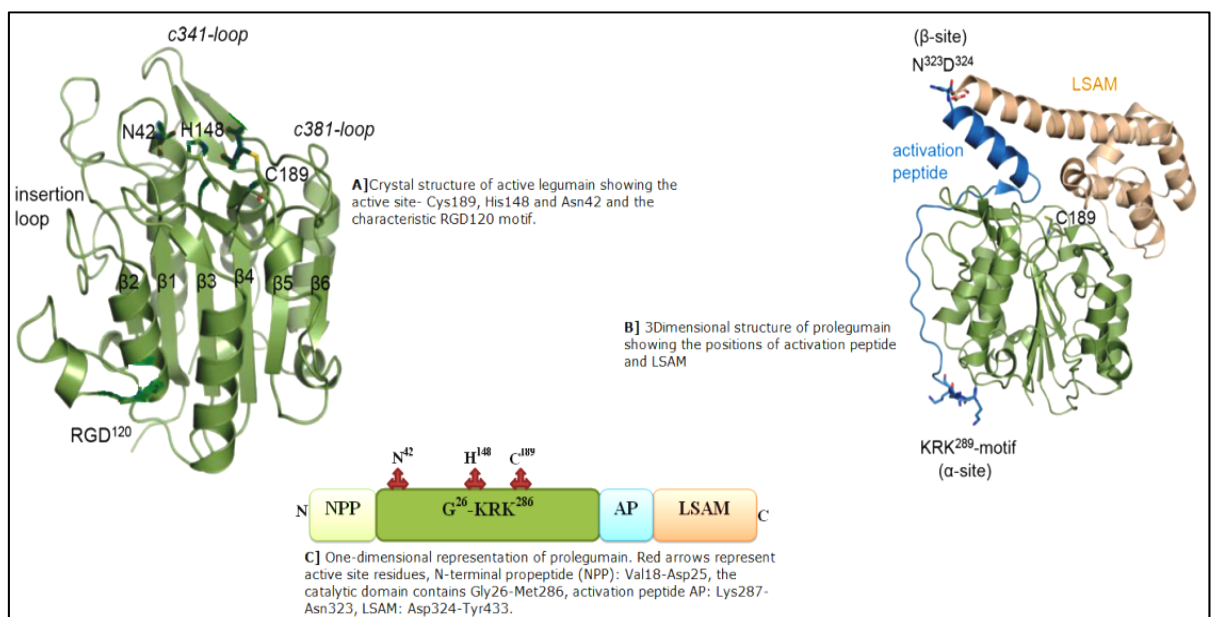
Figure 9: Structure of prolegumain and active legumain (adapted from Dall and Brandstetter, 2013).

Legumain is a well conserved sequence from plant to mammals, including humans and it was found that they were localized intracellularly in endosome/lysosomal structures and is involved in protein degradation intracellularly (Chen *et al*. 1998). However, it was proved that they were also present extracellularly, linked with matrix, cell surfaces and efficient in tumour microenvironments at reduced pH (Wu *et al*. 2006). Legumain is generally produced as inactive forms- prolegumains (50-60kDa) and upon activation at pH 4.5 they are converted to active form (46kDa) by auto cleavage at carboxyl end of asparagine. Further activation at acidic conditions produces the mature enzyme (36kDa) known as active human legumain (Stern *et al*. 2009). Chen and co-workers (2000) have determined 4.5pH to be the optimal pH for prolegumain activation, which was comparable with the pH range of lysosomal units as in macrophages, kidney cells and fibroblasts (Ling *et al*. 1998; Kundra and Kornfeld 1999). Any further increase in pH after the release of the active enzyme would lead to inactivation due to denaturation of the enzyme (Chen *et al*. 2000). The crystal structure of legumain was well studied by Dall and Brandstetter (2013). The structure of active legumain showed the presence of a six-stranded β-sheet located centrally and bordered by 5 major α-helices, as shown in figure 9[A]. They also confirmed the presence of a unique ~30aa inset between β2 and α2 helix which regulates zymogen activation (Fuentes-Prior and Salvesen, 2004). Although the mechanism of action, localization and pH dependent activation and maturation are similar to cathepsins, legumains show no sequence similar to the above making them completely unique from other CPRs (Dall and Brandstetter 2012). The crystal structure of prolegumain indicated the presence of a helical 'pro-domain' placed at the top of the protease domain blocking access to the active site of the enzyme, as shown in figure 9[B]. The presence of this pro-domain regulates the conformational stability and dormant nature of the enzyme prior activation. The pro-domain contains legumain stabilization and activity modulation (LSAM) domain otherwise known as C-terminal death domain (Asp324-Tyr433) and activation peptide (Lys287-Asn323), as shown in figure 9[B] (Dall and Brandstetter 2013).

Any interaction with the protease is mainly electrostatic in nature involving both activation peptide and LSAM domain. The positively charged pro-domain interface balances the highly negatively charged protease surface ultimately clarifying the rationale behind:

1) Stability of prolegumain at neutral pH.
2) Release of pro-domain upon protonation of protease surface.
3) Stability of the activated AEP at acidic pH.

4) Denaturation of active AEP at pH >6.0 (Fuentes-Prior and Salvesen 2004). Dall and Brandstetter (2013) also showed the presence of an electrostatically encoded stability switch near the activation peptide, which can be activated by pH or ligands explaining the reason why legumains are pH-sensitive. Further analyses of legumain have explained its potential role in various disease states including parasitic infections, matrix degradation, antigen processing, atherosclerosis, and tumours (Lee and Bogyo 2012). With mounting evidence illustrating the role of legumain in triggering extracellular matrix turnover in several disease states, it has also been demonstrated that they play a vital role in many tumour development and metastasis (Liu *et al*. 2003) and in growth of atherosclerotic plaques (Papaspyridonos *et al*. 2006). Morita and co-workers (2007) were able to clearly present the degradation of fibronectin- a vital element of extracellular matrix protein in renal proximal tubes to be caused by legumain. Moreover many peptidases such as matrix metalloproteinase (MMP-2/9), serine proteases and cathepsins which are generally involved in matrix degradation and remodelling are synthesized by tubulointerstitial cells of the renal system (Jernigan and Eddy, 2000), thus increasing the possibility of legumain, (which is said to be over expressed in the renal proximal tubules) in controlling matrix remodelling by activating these peptidases (Morita *et al*. 2007).

In their findings Mattock and co-workers (2010) demonstrated that active legumain was found to be over expressed in unstable regions of plaques. This is one of the key factors for atherosclerotic plaque progression, thus confirming the association between legumain and atherosclerosis. It was also shown that the enzyme was found to be co-localized in the endothelial cells of mature plaques in coronary arteries of humans but no presence of the enzyme was noticed in normal arteries or during the early stage of plaque development (Clerin *et al*. 2008). Legumain was also found to be expressed in fibrin rich regions of unstable plaques presenting a probability that it is involved in extracellular matrix remodelling (Morita *et al*. 2007). These results also demonstrate the fact that the enzyme requires an acidic surrounding for activation similar to the lysosome, macrophage microenvironment and plaque microenvironment, where areas of acidic pH are correlated with regions which are rich in lipids, inflammatory cells and metabolically active (Naghavi *et al*. 2002).

1.3.2. Legumain expression in atherosclerosis

Amongst the various proteases involved in atherosclerosis, MMPs and serine proteases have been shown to be more active in this pathological condition. Recent

studies have shed light on the potential role of lysosomal CPRs in atherogenesis by demonstrating legumain expression in stable and unstable human atherosclerotic plaques (Clerin *et al*. 2008; Papaspyridonos *et al*. 2006). Liu and co-workers (2003) were able to suggest that the role of extracellular legumain may actively be associated with the metastatic behaviour of tumour cells. Wu and colleagues (2006) were also able to show the presence of legumain in the tumour microenvironment linked with the tumour cell surface. Moreover, Choi and colleagues (1999) in their findings were able to demonstrate that legumain could also function in a protease independent fashion inhibiting osteoclast formation and bone resorption. In experiments conducted by Clerin and colleagues (2008) in mouse models of atherosclerosis, they were able to characterize the gene and protein expression of legumain as well as increased legumain concentration in human samples of atherosclerotic tissues. They also proposed that legumain expressed by macrophages could lead to atherogenesis by protease regulated as well as protease independent pathways. Foam cells, macrophages and their precursors (monocytes) are the main cell types that expressed legumain in both mouse and human atherosclerotic arterial tissue (Clerin *et al*. 2008).

Studies have shown that differentiated macrophages presented high concentration of legumain secreted extracellularly (Clerin *et al*. 2008). However, the pro-form of legumain secreted extracellularly can be converted to mature forms by pH driven autocatalytic activation in the acidic microenvironment of the cell. These active forms in turn could activate other proteases such as MMPs, cathepsins-L and S that have already been shown to be involved in atherosclerosis and are present in the cell microenvironment (Liu *et al*. 2006). Plaque instability and rupture is highly increased by continuous ECM destabilization and tissue remodelling which is said to be enhanced by legumain which activates and regulates elastolytic or collagenolytic enzymes such as MMP-2 (Chen *et al*. 2001).

Besides this protease dependent activity of legumain in atherosclerosis, they can also induce chemotaxis of monocytes. They perform this by initially acting as a chemo attractant of monocytes and then conduct a chemotactic effect on the same (Clerin *et al*. 2008). Choi and colleagues (2001) have also shown a similar behaviour of legumain in inhibiting the differentiation of monocytes to osteoclast by its C-terminal peptide (17kDa). Thus this characteristic feature of legumain could cause ECM destabilization by a protease dependent mechanism and at the same time can also function as a chemo attractant recruiting monocytes into atherosclerotic plaques and retaining macrophages in the lesions (Clerin *et al*. 2008). Therefore, they may play a dual role in atherosclerosis by acting both as a

protease and a chemo attractant. Moreover, legumain are also found to be expressed by endothelial cells of mouse and human atherosclerotic tissues. These endothelial legumain are expressed only by advanced stage plaques of both human and mouse atherosclerotic models (Clerin *et al*. 2008). The presence of legumain expressed by inflammatory cells in areas of plaque neovascularization of coronary arteries in humans proposes a crucial role of legumain in angiogenesis (Clerin *et al*. 2008). On the basis of all these findings, this research study also focused on the synthesis of a library of peptide substrates that could be site specifically cleaved by legumain and hence could act as a diagnostic tool for detecting atherosclerotic plaques. In summary, legumain produced by macrophages overexpressed in atherosclerotic lesions could contribute to neovessel development by initiation and migration of endothelial cells, invasion and propagation (Clerin *et al*. 2008). All this evidence supports the fact that legumain can be exploited in various ways as a vital diagnostic tool for identifying, characterizing and treating atherosclerosis.

1.3.3. Legumain–substrate specificity

Inspite of the numerous findings in exploiting legumain as an important therapeutic target, the available procedures to characterize the role of legumain in disease states still depends on genetic alterations and antibodies. These elements make it a challenge to study the enzyme in its natural state (Lee and Bogyo 2012). Hence, there is a need for small chemical probes with high sensitivity and selectivity for the enzyme, which can monitor its function and parameters in a wide range of biological systems. When compared to other cysteine proteases, which generally exhibit broader specificity, proteases belonging to clan CD have a specific requirement at the P_1 position of the amino acid sequence with legumain having strict specificity for asparagine at the P_1 amino acid position (Mathieu *et al*. 2002). Mathieu and colleagues (2002) proved the specificity of legumain for asparagine at P_1 by performing a variable library screen test. They constructed a library with 19 wells corresponding to 19 different amino acids at P_1 of a tripeptide sequence and results confirmed that asparagine was the only amino acid tolerated by legumain at the P_1 site. Knowledge about substrate specificity for legumain apart from the P_1 position is not available; hence determining specific amino acids ideal for binding would also serve the same purpose of studying the enzyme (Backes *et al*. 2000). Mathieu and colleagues (2002) using a positional scanning synthetic combinatorial library (PS-SCL) having asparagine constant at P_1 and the respective P_2 and P_3 positions randomized using 19 amino acids were able to find

the most suitable substrates for legumain. The result of these studies demonstrated that Pro-Thr-Asn was the ideal sequence for human legumain. Another study by Manoury and colleagues (1998) stated that tetanus toxin antigen of microbial origin when processed by mammalian legumain, cleaved the sequence at the carboxyl side of asparagine. They further confirmed that human legumain could not tolerate His or Tyr at P_2 or P_3 positions and has a very low reactivity towards substrates containing Asn, Glu, Gln, Phe or Trp at these positions. Chen and colleagues (1997) conducted experiments by treating legumain with a range of oligopeptides and concluded that Asn was the hot spot for substrate recognition and cleavage (figure 10). They also stated that there was no data available for any other protease apart from legumain and other variants species of legumain with such high specificity for asparagine. In another similar study by Dando and colleagues (1999) on substrate specificity for legumain isolated from pig kidney, the role of amino acids on other positions apart from P_1 was evaluated.

Neurotensin
pGlu-Leu-Tyr-Glu-Asn---↓--Lys-Pro-Arg-Arg-Pro-Tyr-Ile-Leu

Human VIP 1-2
His-Ser-Asp-Ala-Val-Phe-Thr-Asp-Asn---↓--Tyr-Thr-Arg

Chicken VIP 16-28
pGlu-Met-Ala-Val-Lys-Lys-Tyr-Val-Asn---↓--Ser-Val-Leu-Thr-NH₂
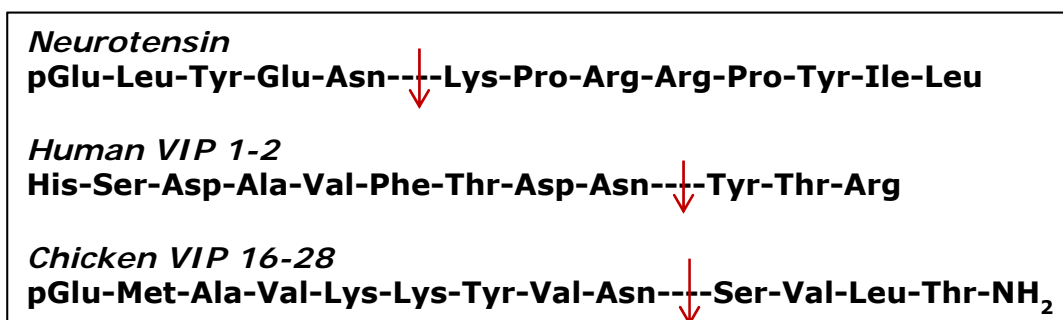
Figure 10: Asn as the required hot spot for legumain cleavage of peptide substrates
(Chen *et al*. 1997).

They concluded that there was no distinctive necessity for a particular amino acid in any subsite apart from P_1. However they also presented evidence of Pro and Ala at P_3 and P_2 subsites in most of the substrates cleaved by the enzyme proposing the idea that this might be an important detail while designing a substrate for legumain. To date there are two prodrugs designed based on legumain activity for tumor studies. One of which has a peptide substrate: N-Suc-Ala-Ala-Asn-Leu-doxorubicin (LEG-3) cleavable by legumain (figure 11), which releases Leu-Doxorubicin at the tumour microenvironment (Wu *et al*. 2006).
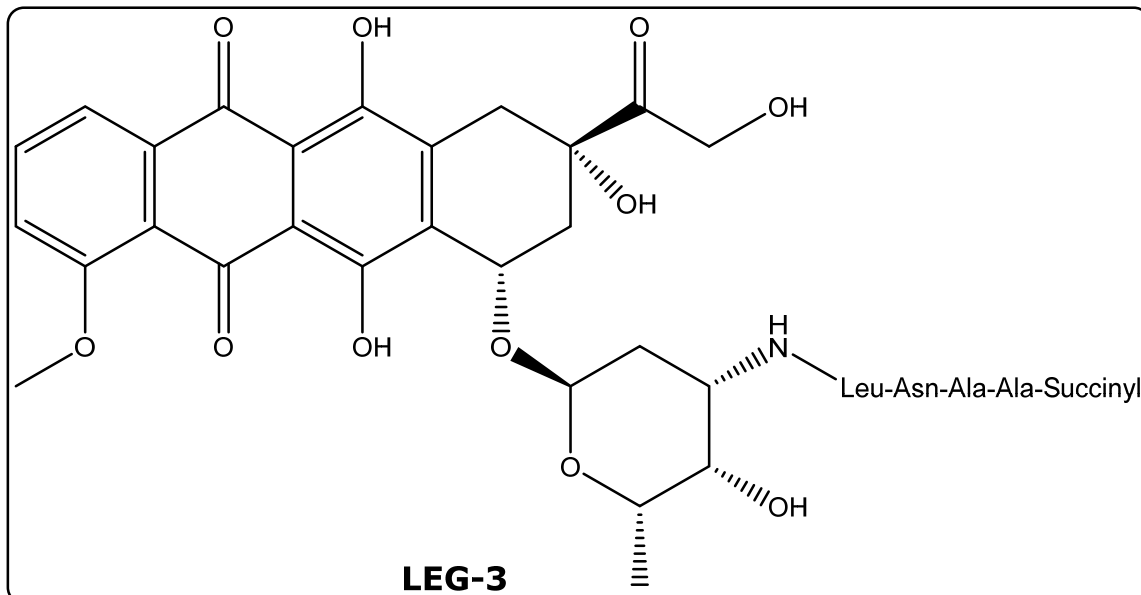
Figure 11: Structure of LEG-3, Legubicin (Liu *et al*. 2003).

Table 3a				
P3	P2	P1	K_m (µM)	V_{max} (µM s^{-1})
Ala	Ala	Asn	80 ± 6	4.4 ± 0.5
Thr	Ala	Asn	128± 10	5.7 ± 0.6
Pro	Thr	Asn	122 ± 13	6.9 ± 0.6
Phe	Tyr	Asn	n/a	n/a

Table 3b			
P3	P2	P1	P1´
Hydrophobic pocket	Hydrophobic, non-polar amino acids with varying side chains	Strict specificity for an asparagine residue	Positive, hydrophilic amino acids
Proline, β-Alanine	Alanine, Glycine and Leucine	Asparagine	Lysine, Histidine and Arginine

Table 3a and 3b: Most preferred substrate sequences for legumain (Mathieu *et al*. 2002).

The second one is a legumain based DNA vaccine constructed to down regulate tumour development and metastasis (Luo *et al*. 2006; Lewen *et al*. 2008). Thus based on the available literature (Table 3a and b), peptide substrates suitable for legumain have been synthesized using both solid phase and solution phase peptide synthesis in this research study. Research and advancements have enhanced the knowledge on disease etiology resulting in effective therapeutic schemes. Despite these improvements heart failure causes 300,000 deaths per year (Leuschner and Nahrendorf 2011). Hence the need for new diagnostic approaches that identify individuals with atherosclerotic plaques and enable personalized treatment before irreversible damage occurs is in demand. For instance, molecular imaging can be used to locate an inflamed plaque at the event of rupture and initiate significant action in order to prevent any further complication (Fuster *et al*. 2005). Moreover molecular imaging could also help in prioritising the lesions that need treatment from those that can be left alone hence avoiding unwanted problems and re-interventions (Leuschner and Nahrendorf 2011). Hence, there is a significant need in combining experimental evaluations with molecular imaging to improve the process of developing therapeutic and diagnostic tools for atherosclerotic plaques.

1.4. Molecular imaging

The *in vivo* characterization and evaluation of biological events at cellular and molecular level without any surgical biopsies is known as molecular imaging (MI) (Weissleder and Mahmood 2001). One of the objectives of molecular imaging is to detect specific targets for various molecular pathways, in order to study fundamental biological events and also activities of cells in their natural living environments (Weissleder and Mahmood 2001). This multi-disciplinary field is of extreme importance to researchers of molecular biology and chemistry with its significant applications rooted in chemical, biological and imaging sciences (Jaffer *et al*. 2007). Imaging modalities require probes, which are highly specific and sensitive in their purpose of detection incorporating two important features. A high-resolution output source, capable of interpreting even the weak signals sent out from the target compound and a target specific imaging moiety that distinguishes the required molecular or cellular region. The major challenge for synthesizing these imaging agents will arise when the target compound is of low-abundance or is remote (Jaffer *et al*. 2007). Molecular imaging of subjects can be enhanced by, using appropriate probes with:

- High specificity and affinity

- Strength to withstand biological barriers
- Increased selectivity
- Quick response to stimulus and clarity in imaging (Jurasz 2003).

Some examples of probes for detecting signals include fluorophores used in near infrared fluorescence imaging, paramagnetic (gadolinium) and super paramagnetic (iron oxides) ligands used in MRI, radioisotopes used in PET and single-photon-emission computed tomography (SPECT) and micro bubbles used in ultrasound imaging (Jurasz 2003). These probes can be further categorized into 3 classes:

- **Compartmental probes**: These probes are used for evaluating physiological functions, for example flow and perfusion. In this type of imaging the target is not exactly a molecular process but a surrogate as in diffusion MRI (dMRI).
- **Targeted probes**: These probes are designed for a very specific target such a particular molecule, receptor or enzyme.
- **Smart probes**: These probes are tailored to become active only in the presence of the target molecule. In this type of imaging the presence of background noise is significantly reduced enhancing the resolution when compared to targeted probe imaging (Perez *et al*. 2003).

1.4.1. Atherosclerosis and molecular imaging

Various advancements in this field of molecular imaging have reformed the quality of medical science by offering reasonable and adequate anatomical details that improve diagnostic proficiency and also help monitor therapeutic interventions (Quillard *et al*. 2011). The biological and anatomical origination and development of atherosclerosis has been studied applying various advanced techniques in molecular biology, from genetically modified mice and also from meticulous observations of human pathological samples. All these findings have provided a complex picture of the disease and the diverse targets responsible for the condition as shown in figure 12. It is hoped that eventually, this will lead to the development of specific probes. These probes coupled with various imaging platforms have widened ventures for medical imaging and preclinical research (Sadeghi *et al*. 2010).

Both cells and the endothelial surface within the arterial vessel wall in atherosclerosis release a number of molecular targets. Notably there is differential accumulation or release of these targets right from early plaque formation up to advanced stages such as susceptible plaques and thrombotic complications

(Choudhury and Fisher 2009). Some examples of molecular targets for recognizing premature or early artherogenesis consists of endothelial cell adhesion molecules such as vascular cell adhesion molecule-1 (VCAM-1), P and E selectin and intercellular adhesion molecule-1 (ICAM-1).
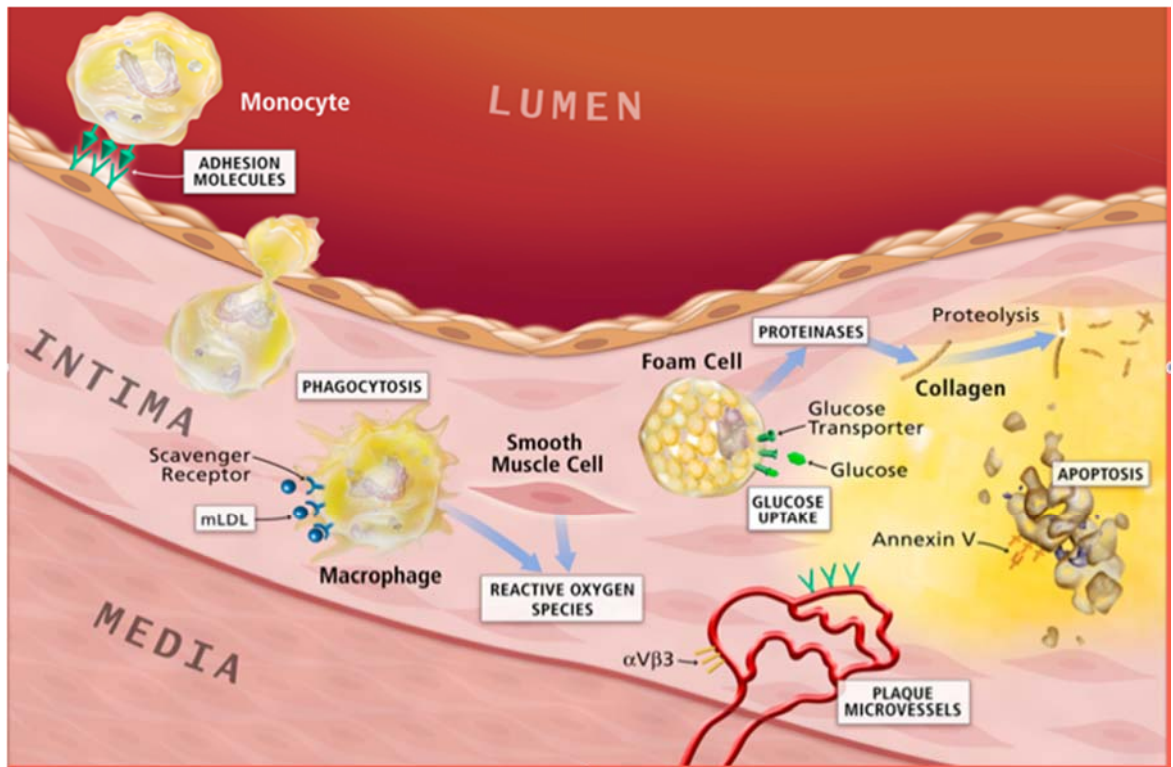
Figure 12: A pictorial representation of an artery presenting potential molecular targets for imaging and hence detecting atherosclerotic development (Libby *et al*. 2010).

These molecular targets are responsible for mononuclear leukocyte concentration to endothelial activation and subsequent migration towards the subendothelial space (Galkina and Ley 2007; Cybulsky MI and Gimbrone Jr 1991). These endothelial adhesion molecules have always been on the forefront of molecular imaging mainly because they are overexpressed in a wide range of diseases such as atherosclerosis, cancer, ischemic stroke and ischemia-reperfusion (McAteer *et al*. 2010). Vascular imaging is less complicated compared to neurological imaging, as targets are available in the circulating blood stream. However, it has its own limitations such as, the location of blood vessels deep inside the body; impeding access for low penetrance techniques similar to fluorescent and ultrasound imaging (Choudhury and Fisher 2009). Moreover, the small size of these plaques and movement due to cardiac and respiratory function, along with challenges to concentrate contrast agents due to the shear force on blood stream applied from the blood vessel walls, impose further restrictions on imaging (Choudhury and

Fisher 2009). The main reason for imaging such targets will help to serve two purposes:

- In understanding the course of treatment to the patient and will also aid in monitoring the progress of the treatment.
- In synthesizing novel drugs for the disease (Overall and Kleifeld 2006).

Techniques	Advantages	Disadvantages
Optical	Use of non-ionizing radiations, high resolution, for imaging soft tissues, can be applied for longitudinal studies, aid in imaging structures over a diverse range of size and category, comparatively cheap and less time consuming.	Low penetrance, invasive in procedure.
Ultrasound	Real-time, no harmful radiation, less expensive, readily available, non-invasive and quick.	Skilled operator required, chances for false-positive results are quite high leading to further evaluation, less quantitative.
Computer Tomography	More detailed compared to ultrasound, quick, 3-dimensional resolution, painless and quick.	Use of strong radiation, low resolution for soft tissues and on inflammations, allergic reactions.
Nuclear	Functioning of body parts can be imaged, painless, highly sensitive technique, quantitative.	Exposure to ionizing radiation may cause allergic reactions, quite high-priced.
Magnetic resonance imaging	Non-invasive, painless, no harmful radiation used, can be applied for a diverse range of disorders, highly sensitive, both spatial and temporal resolution obtained.	Expensive, quite a lengthy procedure, not recommended for people with metal implants, advice required for people with renal problems.

Table 4: Advantages and disadvantages of popular imaging procedures for vascular disorders (Leuschner and Nahrendorf 2011).

A wide range of imaging modalities are available to image vulnerable plaques, including invasive techniques such as, virtual histology, IVUS, OCT, thermal heterogeneity and integrated backscatter, are being evaluated *ex-vivo* and are now under clinical evaluation. A few non-invasive imaging techniques such as, MRI, PET and CT are applied to analyse metabolic activity and also unstable plaques (Fayad and Fuster 2001). Some advantages and disadvantages of these imaging techniques are listed and compared in Table 4.

Of these various imaging procedures only nuclear imaging has been used for medical purposes. FDG (2-[^{18}F]fluoro-2-deoxy-d-glucose) a ^{18}Fluorine agent has been used in PET scans for imaging inflammatory regions in atherosclerotic conditions by exogenous administration. Moreover, it has been demonstrated that FDG-PET can successfully detect inflamed and metabolically active plaques (McAteer *et al*. 2010). The major advantages of PET scans are that they are highly sensitive and have quantitative outputs. By synthesizing novel radio-ligands which can then detect molecular targets such as MMPs (Wagner *at al*. 2006), serine proteases (Li *et al*. 2008) and apoptosis (Korngold *et al*. 2008) PET can be used beyond macrophage imaging and also to visualize the response to treatments in high risk patients (McAteer *et al*. 2010). Another means of enhancing molecular imaging of plaques will be the use of multimodality imaging. This involves evaluation of several vascular regions at the same time. To achieve this, the required molecular probe is administered into the body intravenously and allowed to aggregate at the specific region (Sadeghi *et al*. 2010).
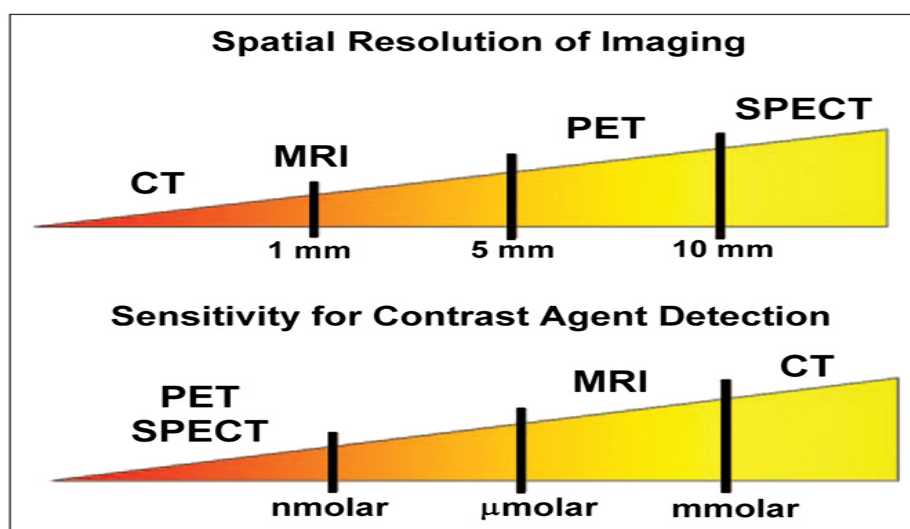
Figure 13: Comparison of spatial resolution and sensitivity of imaging techniques. (Sadeghi *et al*. 2010).

Upon optimizing the conditions for a positive target to background signal, the highly sensitive nuclear procedure (SPECT, PET) which, exhibits sensitivity up to picomolar scale along with superior spatial and temporal resolution of CT and MRI (figure 13) will be the source behind hybrid imaging platforms. Hybrid procedures such as PET/CT and PET/MRI are becoming popular (Sadeghi *et al.* 2010). Compared to other imaging modalities, MRI provides well-defined information on both anatomy and plaque configuration. Furthermore, coupling the imaging agents with specific molecular probes that target the biomarkers of atherosclerotic plaques (Sadeghi *et al.* 2010) will further enhance the resolution of the image captured via a MRI scanner. These imaging techniques provide very promising application in studying the various causes and events in atherosclerosis. But clinical manifestation will require a safe, low-molecular weight, non-toxic and cost effective probe and imaging technique. This research study is focused on the use of chemical agents for MRI, which is one of the efficient techniques for vascular imaging. MR imaging produces no harmful radiation and hence can be applied for longitudinal studies and can also provide high spatial and temporal resolution required for studying vascular injury by atherosclerosis (Fayad and Fuster 2001). Moreover, its proficiencies are further improved by the application of increasingly refined contrast agents that can specifically target molecular markers, cells and biological progressions (McAteer *et al.* 2010).

1.4.2. Magnetic resonance imaging (MRI)

Magnetic resonance imaging is a diagnostic tool that penetrates deep into soft tissues of the body and can produce images of resolution up to submillimeter scale (Hasegawa and Zaidi 2006). It has been developed as a sensitive technique for non-invasively imaging thrombosis within the carotid artery (Flacke *et al.* 2001). With its property of non-invasive imaging MRI produces precise three-dimensional images of the structures that are scanned. A very brief comparison between MRI and other non-invasive imaging techniques (table 5) rationalizes that MRI produces images with enhanced contrast and resolution. Because it gives clear analysis on anatomy, function and metabolic activity of tissues it is readily applied in musculoskeletal, neurological, cardiovascular and cancer studies (Strijkers *et al.* 2007). As previously discussed MRI requires no ionizing radiation and moreover the contrast agents (CA) used for improving the clarity of imaging are far less toxic when compared to CT contrast agents (Sand and Levitin 2004). MRI with high resolution helps to detect and characterize atherosclerotic plaques and also mature lesions. It also helps in distinguishing between thin and broken fibrous cap (figure

15) with an intact and thick fibrous cap *in vivo*. Along with the available CAs the required molecular specificity could be easily attained by MR imaging (Caravan 2009). In a study on MRI by Corti (2006) it was stated that magnetic resonance imaging is a promising technique for studying plaques. The procedure for this imaging technique is mainly based on the signal emitted from the experimental subject under a strong magnetic field in accordance to the concentration of water and also the relaxation times (T1 and T2).

Imaging procedure	Energy used	Spatial resolution (mm)	Acquisition time per frame (s)	Required probe quantity (ng)	Depth of tissue penetration (mm)
PET	Annihilation photons	1-4(A), 4-5(C)	1-300	1-100	>300
SPECT	Gamma rays	0.5-5(A), 7-15 (C)	60-2000	1-100	>300
Fluorescence imaging	Visible to infrared light	2-10	10-2000	103-106	1-20
MRI	**Radio-frequency waves**	**0.025-0.1(A), 0.2 (C)**	**60-3000**	**103-106**	**>300**
Ultrasound	High-frequency sound waves	0.05-0.5(A), 0.1-1 (C)	0.1-100	103-106	1-200
CT	X-rays	0.03-0.4(A), 0.3-1(C)	1-300	N/A	>300

Table 5: Comparison between various non-invasive imaging tools (Camici 2012).

The rapid spinning of the positively charged protons around the nucleus creates a magnetic field. Therefore, when they are placed inside a MRI scanner emitting strong magnetic field strength, these protons align themselves with the external magnetic field. When specific radio waves are passed through them some of these protons are raised to a higher or excited energy level and while returning to their original frequency in the absence of the radio wave, termed as relaxation, these protons emit radiofrequency signals. These signals are then detected by a receiver or antenna in the MRI device producing clinical images (Ibanez *et al*. 2009; Singer 2012) as shown in figure 14.

The time taken by the protons to return to its original longitudinal energy state after the radio wave is switched off is marked as T1 and technically denoted as spin-lattice relaxation whereas the time taken to return to its transverse state to

release energy is marked as T2 and technically denoted as spin-spin relaxation (Singer 2012). The T1 and T2 values are said to vary according to various tissues; for example water has a longer relaxation time when compared to fat and hence the characteristic features can be studied.

Random movement of protons in the absence of a magnetic field

B_O

Alignment of protons in the presence of a strong magnetic field

Rf wave

B_O

Realignment of protons in the presence of radio-frequency wave

B_O

Release of energy

When the radio frequency is switched off, the protons return to their original position and release energy
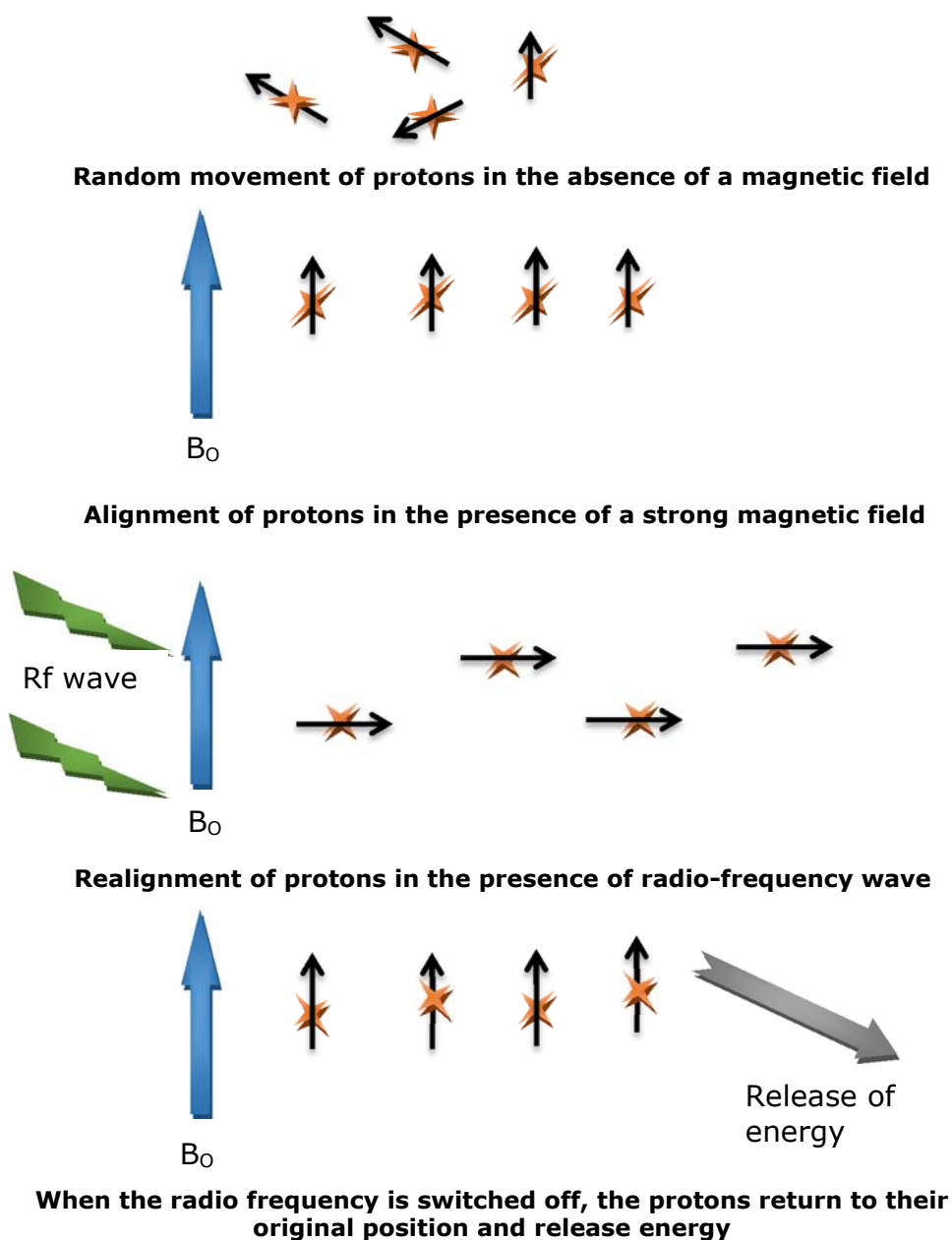
Figure 14: A pictorial representation of the behaviour of protons when exposed to a strong magnetic field to radio frequency signal (adapted from, Singer 2012).

Since the human body contains a lot of hydrogen atoms, which is the underlying basis for MR imaging, it helps in producing a clinically significant image. The mode of MR imaging is based on the biophysical and biochemical aspects, which include molecular change, chemical composition, water content and physical conditions. The first human atherosclerotic plaque sample to be viewed *in vivo* was that of the coronary lesions by Fayad and colleagues (2000).
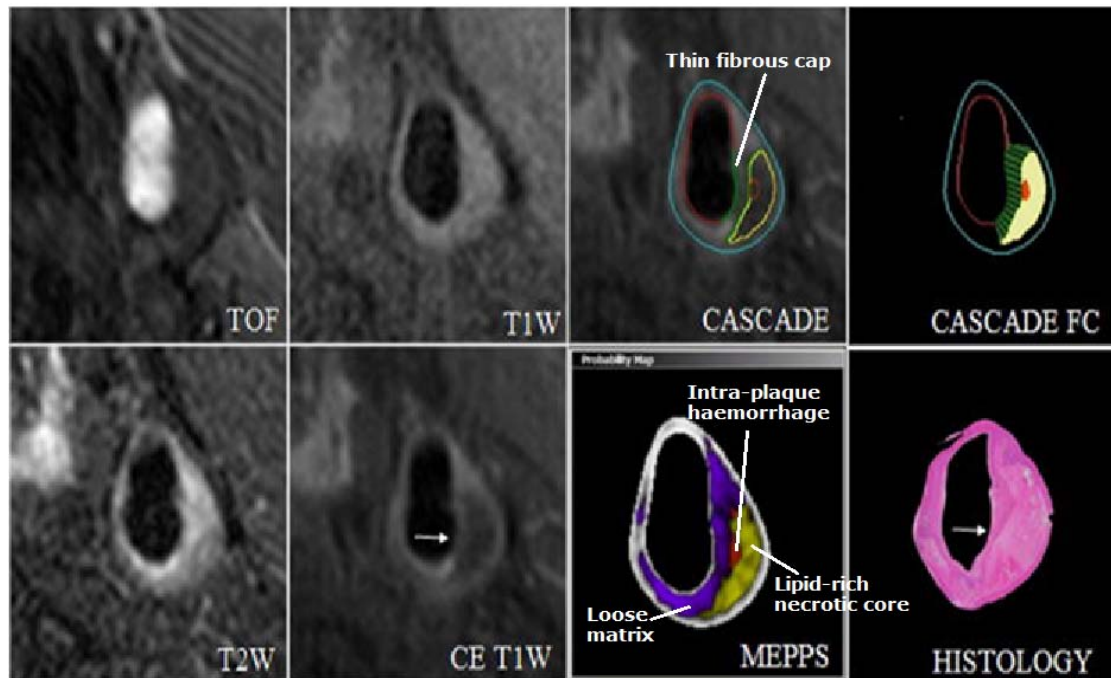
Figure 15: Automated segmentation of high-spatial resolution, multi contrast, bright and black blood *in vivo* magnetic resonance imaging. (Chu *et al*. 2010).

This procedure was further improved by Botnar and colleagues (2000) to give a better image of the coronary plaque during free breathing conditions. Since then the clinical applications of MRI have been exploited in studying plaques in human carotids (Yuan *et al*. 2001), aortic (Corti *et al*. 2001), peripheral (Coulden *et al*. 2000) and diseases of the coronary artery (Fayad *et al*. 2000). Thus many researchers have concluded that MRI is a sensitive diagnostic technique for studying the progress of the plaque (Corti 2006) and also has the potential to explain the thickness of the lesion, the growth rate along with the composition of the atherosclerotic plaques (figure 15) (Fayad *et al*. 2000). In most imaging situations the generally evaluated tissue-proton spin density and T1/T2 relaxation times are normally sufficient to provide the required images (figure 15). However under certain pathological conditions when the changes in local relaxation times are not enough to distinguish between the normal and abnormal regions of the tissue, the need for a better contrast between them becomes necessary (Singer 2012). This can be brought about by using CAs, which work by altering the normal relaxation times and thus, enhance the resolution of the target location (Strijkers *et al*. 2007). This combination between MRI and contrast agents will enhance understanding about various diseases including arthritis, tumour related conditions and atherosclerosis.

Further work on MRI contrast agents, for example, small iron oxide particles that can be easily ingested by macrophages accumulated in plaques producing better images and quickening the process of diagnosing high risk plaques (Kooi *et al*. 2003) are in progress. Advancements in the clinical setting of MRI with the release of 3-T and 7-T scanner have further increased the opportunities for distinguishing and studying atherosclerotic disease (Yuan *et al*. 2012). High spatial resolution black-blood techniques (figure 15) might enhance non-invasive imaging of human coronary and carotid arteries thus helping in early evaluation of atherosclerotic conditions (Fayad and Fuster 2000). These progressions in MRI and contrast agents will help in developing an accurate and early diagnostic tool for susceptible patients of atherosclerosis (Ostergaard *et al*. 1999).

1.4.2.1. Contrast agents (CAs)

For molecular imaging, the selection of an appropriate contrast agent (CA) is important to produce the required contrast between the abnormal and normal regions. CA's constructed for any particular target generally consist of two main elements:

- A unit, which is highly specific to the overexpressed molecular marker similar to proteins, peptides, carbohydrates, small molecules, antibodies and fragments of antibodies.
- An imaging tag that can send significant signals to an external imaging device of choice such as fluorophores (NIRF), radio-isotopes (SPECT/PET), sonic enhancers (US) or magnetic groups (MRI) (Jaffer *et al*. 2006).

However, the choice of a CA mainly depends on the location, concentration and nature of the corresponding molecular target more than the imaging modality. For instance, ECM proteins such as collagen, elastin and fibrins are normally expressed in high concentrations when compared to other molecular targets and hence low molecular weight CA's equivalent to a few nanometres can be used to image them. Imaging such molecular targets enables greater accumulation of CAs leading to the use of imaging devices with comparatively low intrinsic sensitivity. On the other hand molecular targets of very low concentrations such as 10^{-9} to 10^{-13} mol/g require different targeting procedures (Nunn *et al*. 1997).

As stated earlier, image quality of the MRI scan can be improved by the use of appropriate contrast agents, which work by altering the relaxation time of the tissue water. In a recent evaluation 35% of MRI examinations are said to use contrast agents to provide high-resolution imaging (Li *et al*. 2008). In a review by

Sosnovik and Caravan (2009) it has been stated that the CAs for MRI have been broadly classified into paramagnetic gadolinium (Gd)-based CAs and super paramagnetic (iron-oxide) nanoparticles. CAs in general alter both T1 and T2 but it is always important to distinguish them into transverse relaxation agents (1/T2) or longitudinal relaxation agents (1/T1). The 1/T2 agents function by increasing transverse rate (1/T2) to the same extent as the longitudinal rate (1/T1) increase and they are denoted as T1 emitting a positive scan with an increase in signal intensities. When there is a large increase in the transverse rate (1/T2) alone then they are denoted as T2 agents causing a reduction in signal intensity and are denoted as negative contrast agents. Ferromagnetic iron oxide CAs are denoted as T2 agents and paramagnetic Gd CAs are examples of T1 agents. There are several other contrast agents available but the gadolinium based paramagnetic CAs are more widely used in several clinical analyses (Caravan 2006).

1.4.2.2. Paramagnetic gadolinium (Gd)-based contrast agents

Gadolinium complexes are used as extracellular contrast agents for MRI in recognizing and treating a wide range of vascular dysfunction. Since they are classified as T1 contrast agents they produce positive images with good signal intensities such as Gd-DTPA (Mohs *et al*. 2004; McAteer *et al*.2010) as shown in figure 16. Most of the Gd based analyses are pre-clinical with the exception of EP-2104R, which is a fibrin targeted probe (Spuentrup *et al*. 2008). Contrast agents targeted to MMPs (Lancelot *et al*. 2008), high-density lipoproteins (Frias *et al*. 2004) and also non-targeted complexes such as gadofluorine-M (Meding *et al*. 2007) and long chain fatty acids coupled to Gd (Lipinski *et al*. 2006) have imaged plaque growth in rabbit and mouse models for atherosclerosis.
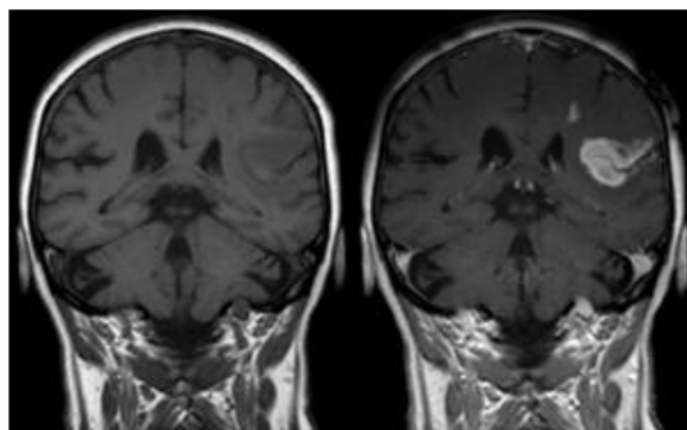
Figure 16: MRI image (T_1) of blood-brain barrier defect due to stroke. The image on the left has no CAs but the one on right has CAs.

Some of the commercially available Gd-based contrast agents are Gd-DTPA (figure 17-A), Gd-DOTA or Gd-BOPTA (figure 18-A) which are known for their low toxicity but at the same time posses low molecular weight leading to immediate clearance from the blood stream producing poor signals for imaging (Mohs *et al*. 2004). However, this problem can be solved by coupling them to a polymeric carrier for example Gd-DTPA and Gd-DOTA coupled to polyamidoamine (PAMAM) dendrimers (Nwe *et al*. 2010), Gd-DTPA coupled to dextran (Li *et al*. 1992), Gd-DTPA coupled to albumins (Spanoghe *et al*. 1992), Gd-DTPA coupled to polyethylene glycol. These CAs reduce the toxicity of the metal ion, increases their retention time in the blood stream, act as carriers of therapeutic compounds and also can increase the specificity of the CA by carrying a probe specific to the target of interest (Mulder *et al*. 2005). Thus these macromolecular agents act as blood pool contrast molecules delivering more information on vascular anatomy, capillary permeability, and blood volume. Moreover, pathological information pertaining to myocardial infarction, inflammation, plaque deposits, break-down of blood brain barrier and tumour can also be evaluated.

Free gadolinium metal ion is toxic and therefore it is always bound to chelating agents such as DTPA or DOTA. Many factors are assured to impact the stability of these complexes such as charge, number, basicity of co-ordinating groups, conformational effects and ligand field including intrinsic energy and entropy (Frullano and Meade 2007). The gadolinium (III) complex of diethylenetriamine pentaacetic acid (Gd-DTPA/Gadopentetate dimeglumine/Magnevist) is one popularly used MRI contrast agent. This water-soluble complex is stable and nontoxic with a very high demetallation (K_{Gd}) stability constant (breaking of a bond between the substrate and metal atom) of logK_{Gd} = 22.4 and is approved for both experimental and medical research. Gd-DTPA aggregates in tissue by perfusion-controlled routes and has been permitted for use in adult patients since June 1988 (Frullano and Meade 2007). Revealing abnormal vascularity and imaging disruptions of blood-brain barrier triggered by lesions are well documented by using such image enhancing contrast agents (Hayes *et al*. 2002). Laboratory analyses of animals and humans involving a Gd(III) complex of DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetra acetic acid) a macrocyclic ligand has proved its conformational rigidity and demetallation stability constant to be 25.3 and also has a very low dissociation constant at physiological pH. This CA has also been approved for medical examinations (Frullano and Meade 2007). Further research in this field of MRI contrast agents have categorised CAs into four major groups:

- Targeting CAs: Tailored to spot specific receptors or organs and hence aggregate at the target site.
- Activatable CAs: Designed to undergo modification in relaxivity corresponding to changes in specific physiological factors
- Blood pool agents: In contrast to other CAs they have longer residence time and also find application in magnetic resonance angiography.
- Chemical exchange saturation transfer agents (CEST): These novel agents work using an irradiated group of protons ultimately leading to their activation. Their reaction is controlled by few variables such as chemical shift and exchange rate between their proton group and water protons (Zhang *et al*. 2003; Frullano and Meade 2007).

Generic name	Trade name
Gadopentetate dimeglumine	Magnevist (Gd-DTPA)
Gadoterate meglumine	Dotarem
Gadodiamide	Omniscan
Gadoteridol	ProHance
Gadobutrol	Gadovist
Gadoversetamide	Optimark
Gadobenate disodium	MultiHance
Gadoxetate disodium	Primovist/Eovist
Gadofosveset trisodium	Vasovist/Ablavar

Table 6: List of clinically approved contrast agents (Singer 2012).

Some of the approved CAs, in use, intravenously in the United States and European Union are shown in table 6 (Li *et al*. 2008). Based on the chemical structure of the chelating agent they can be classified as macrocyclic and linear contrast agents (figure 17 and 18).
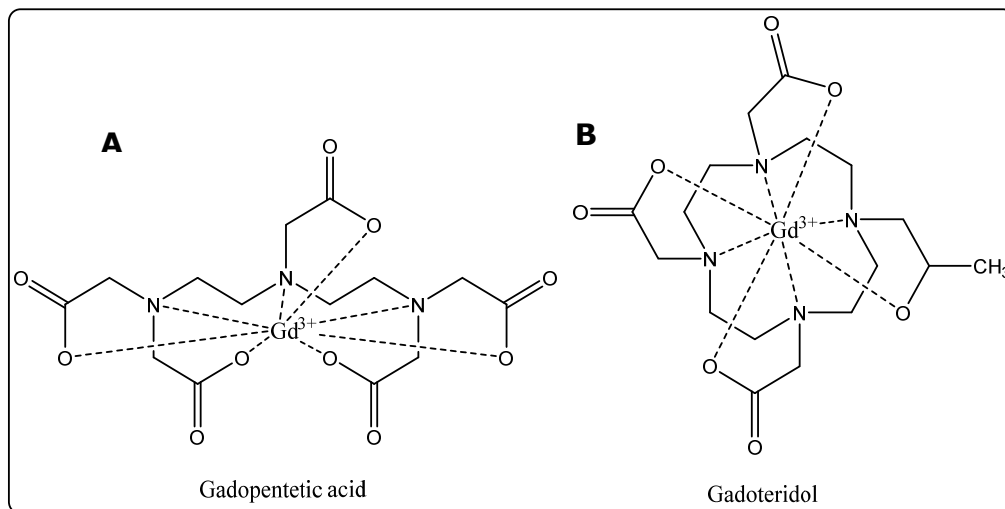
Figure 17: Structure of (A) Magnevist (Gadopentetic acid; Gd-DTPA) and (B) ProHance (Gd-HP-DO3A; Gadoteridol) (Kunnemeyer *at al.* 2009).
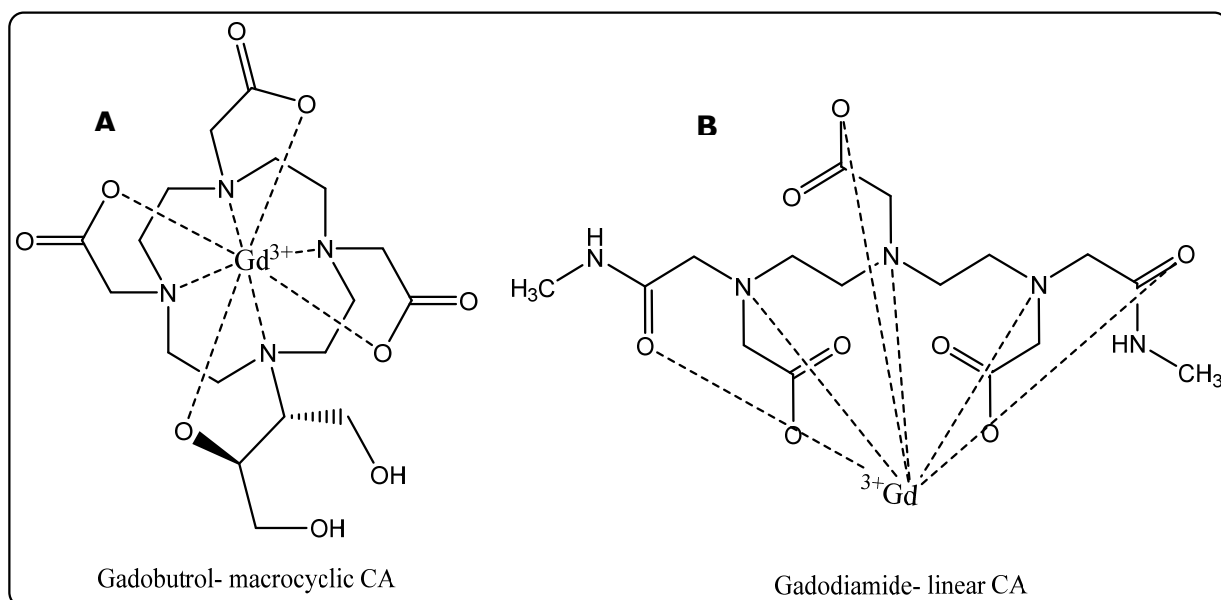
Figure 18: structure of (A) macrocyclic and (B) linear contrast agents (Singer 2012).

Theoretically macrocyclic chelates are more stable than the linear counterpart as they bind securely to the metal ion and hence the chance for disassociation of the metal is low (Kei and Chan 2008). Some of these CAs are inexpensive in production and have a very good safety profile allowing use even in children (Caravan 2006). However, the use of these agents in people with renal disorder needs special consideration as they might cause severe renal failure leading to death, also pregnant women will require the advice of their doctors as CAs are prone to cross the blood-placenta barrier (Webb *et al.* 2009).

1.5. Fluorescence resonance energy transfer (FRET)

The enormous growth in magnetic resonance imaging has already widened its application from medical diagnosis to consistent cellular analysis and imaging

(Frullano and Meade 2007). However, along with MRI techniques, optical imaging procedures are also being suggested as a potent tool for both *in vivo* and *in vitro* molecular imaging because of their excellent sensitivity and selectivity (Chen *et al.* 2014). Optical imaging effectively helps in the study of interactions and structural modifications along with microscopic visualization and subcellular localization of biochemical reactions (Schmid and Sitte 2003). FRET is one of the most widely suggested techniques for studying live interactions between molecules without causing any damage to the native biological evolutions in living cells (Kalab and Soderholm 2010).
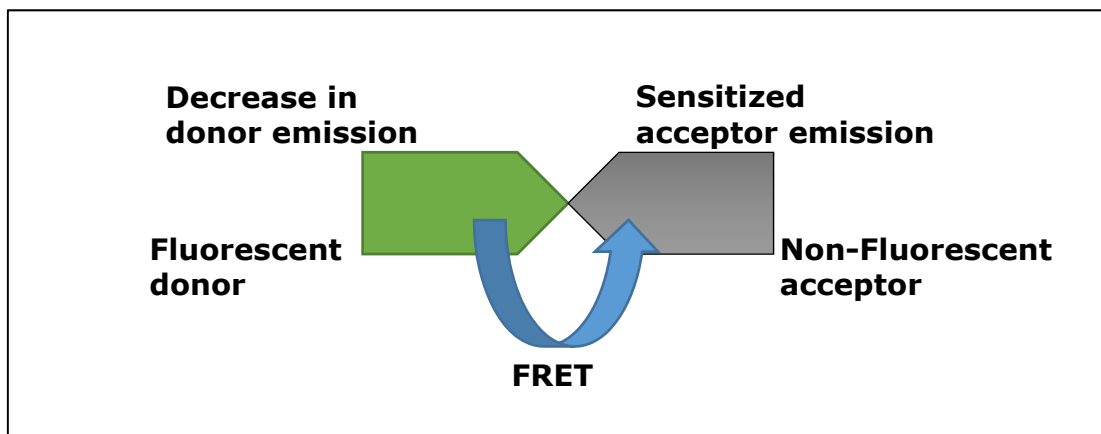
Figure 19: Pictorial representation of fluorescent energy transfer.

FRET is a technique used to measure distance between molecules even at the nanometre range, thus helping to predict any changes between the molecules (Selvin 2000). This procedure can be defined as the transfer of energy from an excited fluorophore, which is the donor molecule to another fluorophore denoted as the acceptor molecule as in figure 19 (Schmid and Sitte 2003). It is commonly known as Forster resonance energy transfer due to the fact that the expression 'fluorescence resonance energy transfer' evokes the idea of resonance energy transfer or transfer of energy due to photon effect which would completely mislead any lay person (Kalab and Soderholm, 2010). This technique helps to image localized proteins and other surrounding molecules with 3-dimensional resolution compared to conventional optical techniques.

The efficiency of FRET is based on optimising a few parameters such as:

- 10-100nm has been proposed as the required distance between donor and acceptor pair (Hayward *et al.* 2010).
- The excitation spectrum of the donor and acceptor pair should produce an overlap (Zaccolo 2004).

- The direction of the fluorophore otherwise denoted as the dipole moments (Selvin 2000).

In FRET, many different combinations of fluorescent dyes have been examined and used for both transfer of energy and labelling macromolecules. Some of the most prevalently used dyes as FRET pair are fluorescein/rhodamine, Cy3/Cy5 (cyanine dyes) and also combinations of Alexa Fluor® dyes are used (Schmid and Sitte 2003).
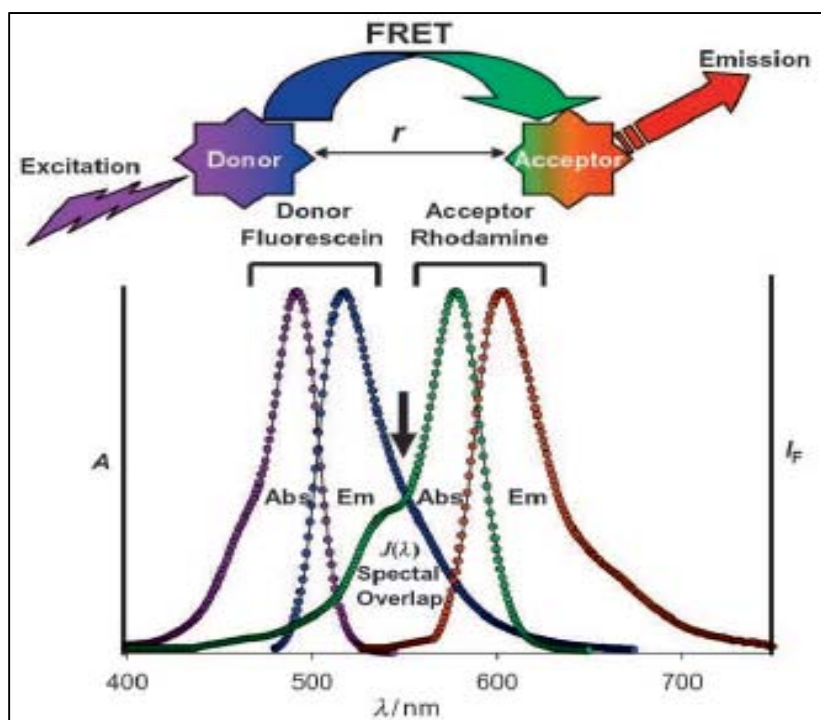
Figure 20: A pictorial expression of FRET between FITC and rhodamine (Sapsford *et al.* 2006).

The spectra in figure 20, shows the relative absorption (Abs) and emission (Em) of a commonly used FRET pair, where fluorescein is the donor and rhodamine is the acceptor unit. The energy transfer rate k_t between the donor/acceptor pair depends on the distance 'r' and is generally denoted as 'R_0'. This is defined as the distance between the donor and acceptor unit during which 50% of the excited molecules at the donor side deteriorate and cause an increase in the acceptor molecule emission (Sapsford *et al.* 2006; Selvin 2000). The efficacy of a FRET process can be evaluated using the equation: $E = 1/[1+r/R_0)^6]$ (Buntru *et al.* 2009).

1.5.1. FRET construct

FRET is a technique widely used for various bio analyses mainly due to its intrinsic sensitivity towards nanoscale alterations in donor-acceptor pairs. This can be achieved only by designing a suitable FRET construct based on experimental

needs. The components for making a FRET design can be categorised into three divisions: organic materials such as traditional fluorophore dyes, dark quenchers and polymers (Widengren 2010); inorganic substances such as metals and semi-conductors (Medintz *et al*. 2003) and finally biological origin fluorophore similar to fluorescent proteins (FP), amino acids and gene-based compounds that would fluoresce by particular enzyme catalytic reactions (Touryan *et al*. 2009). These components may either be used as a donor or as an acceptor or can be both depending upon the experimental set up (Sapsford *et al*. 2006).

1.5.1.1. Anthraquinone as quencher

The acceptor molecule used in this present study was 1-((2-(2-(2-amino ethoxy)ethoxy)ethyl)amino)anthracene-9,10-dione (figure 21). The acceptor unit is otherwise denoted as the quencher, which has the same meaning as an acceptor.
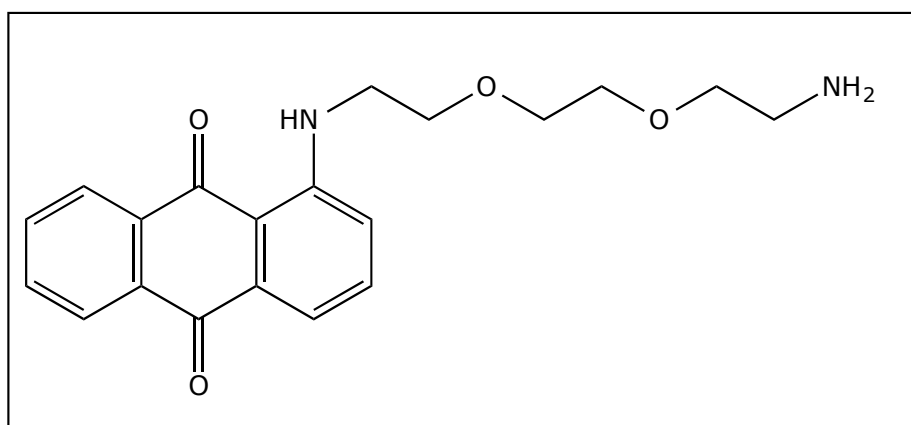
Figure 21: Structure of 1-((2-(2-(2-aminoethoxy)ethoxy)ethyl)amino)anthracene-9,10-dione.

Labelling of oligonucleotides with fluorescent elements has become popular and also a large number of non-fluorescent quenchers varying from gold nanoparticles to nucleotides are being used as probes (Marras *et al*. 2002). Substituted anthraquinones are suggested to be efficient quenchers because of their non-fluorescent properties, which make them effective dark quenchers. As they do not fluoresce like many other quenchers they can effectively absorb heavy fluorescent emissions. They have a broad absorption wavelength, which closely matches many fluorescein emissions with wavelengths between 490nm to 520nm (Nakayama *et al*. 2003). They can be easily purified and also in certain cases, isolation using reverse phase HPLC results in a highly purified compound (Laikhter *et al*. 2004). These quenchers are popular because of the ease in handling the compound and also because of the cost when compared to other radioactive labelling procedures. One such example is 15mer linear oligo DNA probe with fluorescein as donor and

anthraquinone as quencher (Kodama *et al*. 2006); where anthraquinone was found to efficiently quench the fluorescent moiety.

Anthraquinone has favourable characteristics as an effective quencher, as found by May and colleagues in their research work (2005), studying most FRET pairs. Anthraquinones have been used efficiently as quenchers in many experimental analyses (Laikhter *et al*. 2004) mainly due to their broad wavelength capacity. The use of efficient quenchers in a FRET construct has become vital because of the reduction in background fluorescence (Sapsford *et al*. 2006). Quenchers are used to study the binding and separating process with ease and also save much time, for processes that would otherwise be very difficult to follow (Walter and Burke 1997). Marras and colleagues (2002) have demonstrated that quenching of fluorophore can be made more efficient with effective binding between the quencher and fluorophore (relating to the dipole moment). Also in certain DNA related studies, quenchers are used as probes in combination with organic dyes (Zuo *et al*. 2010). Van Valckenborgh and colleagues (2005) have also proven the effective quenching of a fluorophore (FITC) by anthraquinone in their research studies. Thus taking advantage of all these properties 1-((2-(2-(2-aminoethoxy)ethoxy)ethyl)amino)anthracene-9,10-dione was used as the quencher for the FRET construct in this research work.

1.5.1.2. 5(6)-Carboxyfluorescein (5(6)-CBF) as donor

5(6)-CBF is a mixture of 5 and 6-carboxy substituted fluoresceins. Grignon and colleagues (1989) have used 5(6)-CBF to study cell:cell fusion and intercellular transport in plants. They have also demonstrated that 5(6)-CBF was successfully used as a tracer element to monitor phloem sap in real time, in short and long term experiments. 5(6)-CBF is one of the most popular organic dyes used in many fluorescent based analyses mainly because of its favourable characteristics such as availability of instrumentation for interpreting results, comparatively low price and compatibility with living cells (Fischer *et al*. 2003). Carboxyfluorescein (CBF) has been the most preferred compound of choice for synthesising fluorescent-labelled peptides requiring hydrolytic stability and for protein conjugates (Weber *et al*. 1998).
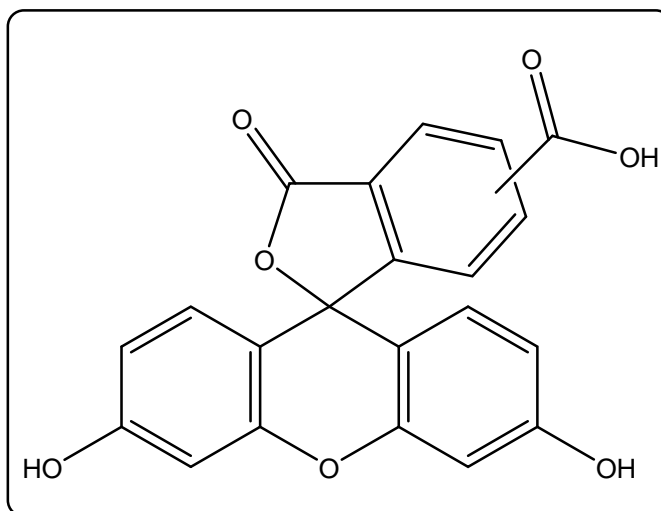
Figure 22: Chemical structure of 5(6)-carboxyfluorescein.

Mordon and colleagues (1994) in their studies on imaging tumor tissue using a fluorescent probe have stated that 5(6)-CBF is widely used in pharmacological and cellular studies without significant toxicity and its biocompatibility has been well documented. Taking advantage of 5(6)-CBF's sensitivity towards variations in pH Mordon and colleagues were able to differentiate between normal and cancerous tissues. 5(6)-CBF is less vulnerable to photo-bleaching and is more stable, moreover it has also been demonstrated that its fluorescence emission (max: 515nm) is directly proportional to increase in pH stating that it is active in acidic to near neutral pH (5.0-7.8) (Bidmanova *et al*. 2012) and also has two wavelengths of maximum absorbance (465 and 490nm) (Mordon *et al*. 1992). In this research study 5(6)-CBF (figure 22) has been used as the donor unit by coupling to the required peptide sequences. It reacts with the primary amine forming an amide bond and this chemistry is more stable and better yielding (Fischer *et al*. 2003). In a study to image the movement of pepducin in living cells using the principle of FRET, Tsuji and colleagues (2013) have used 5,6-CBF as the donor reagent for the FRET construct and have shown positive results. Bidmanova and colleagues (2012) have studied the photo stability and temperature response and demonstrated that CBF's showed less or no significant changes to these conditions. Thus considering these advantages 5(6)-carboxyfluorescein was preferred as the donor compound in the FRET construct for this research study.

1.6 Aim

The aim of this research project was the design, synthesis and evaluation of selective, non-invasive imaging agents for atherosclerotic plaques. To achieve the aim the study was split into three objectives.

The first objective was the design, synthesis, purification and characterization of a small library of FRET peptide substrates specific to legumain (JC- 1st and 2nd series). The sensitivity of the FRET compounds was evaluated by incubation with the protease (rhLegumain) to observe the release in fluorescence. The results were measured fluorometrically using a multi-mode microplate reader. The successful compounds were then used as templates for the synthesis of a peptide based contrast agent with the potential to be used for MR imaging (JS series).

The second objective was the synthesis of a library of compounds carrying the diagnostic contrast agent (Gd-DOTA). The design of these compounds (JS linear and JS lysine series) followed a similar pattern as in the JC library. Upon synthesis, purification and characterization of the JS library of compounds, they were incubated with the Recombinant Human Legumain/Asparaginyl Endopeptidase (rhLegumain). The end products of the enzyme treated compounds were purified and characterized by ESI mass spectrometry. The cytotoxicity of the JS series of compounds was also evaluated.

The final objective was the synthesis, purification and characterization of a small discrete library of peptide motifs, which were designed to be peptide substrate specific to MMP-2 and MMP-9 (JJ series). The substrates were then conjugated to the contrast agent *via* *N*-terminal conjugation (compound: JD) or by coupling to the side chain of a lysine residue (compound: JL-87). This synthesis involved both solid and solution phase peptide synthesis. Upon successful synthesis of these compounds they were incubated with the protease to evaluate the release of the contrast agent *in vitro*.

1.6.1. Rationale

Atherosclerosis causes the death of more people than any other disease in the Western world (American heart association). Atherosclerotic plaques contain significantly higher quantities of activated MMP-2 than normal tissue extracts

(Galis *et al*. 1994) and subjects with acute coronary syndrome have increased plasma levels of MMP-1,2 and 9 (Alberto *et al*. 2006). Their expression is not only a part of atherosclerotic lesions but also predictive of plaque instability (Galiz 2004). McGeehan and colleagues initially characterized peptide substrates specific to MMP-1 and MMP-9 in 1994. Since then its function has been exploited to not only deliver site specific pharmacological products in a range of disease states both *in vitro* and *in vivo* (Van Valckenborgh *et al*. 2005) but also as a device to localise imaging agents (Lancelot *et al*. 2008). An experiment by Ishizaki and colleagues (2010) has demonstrated the over expression of both MMPs and legumain in stroke caused by atherosclerosis. Legumain is a distinctive late cysteine endopeptidase found in many mammalian tissues (Chen *et al*. 1998). Legumain is found to be active under acidic environments in an autocatalytic manner (Li *et al*. 2003) expressing specificity to an asparagine residue at the P1 position of a substrate sequence (Chen *et al*. 1997; Ishizaki *et al*. 2010). Legumain is also involved in a number of pathological conditions such as tumour, atherosclerosis and inflammation, though their biological function in these conditions is not clearly understood (Chen *et al*. 2014). It has also been demonstrated that legumain may assist in plaque development by activating other proteases such as MMP-2, cathepsin-L, whose functioning has already been strongly linked with atherosclerotic plaque progression (Mattock *et al*. 2010).

More recently a peptide model of MRI contrast agent (Gd-NBCB-TTDA-Leg (L)) and a near infrared fluorescent probe (CyTE777-Leg (L)-CyTE807) were successfully synthesized to specifically target the over expressed legumain in the tumour microenvironment (Chen *et al*. 2014). A more or less similar approach has been undertaken in this research project to specifically and efficiently target legumain activity *in vitro* (table 7, 8 and 9). Thus synthesising highly specific and potent substrates for legumain will not only be effective for demonstrating the expression of legumain in atherosclerosis, but will also be valuable for therapeutic interventions of atherosclerosis in the future (figure 23-A).

Prodrugs that are synthesized to be specifically activated in the presence of over expressed or exclusively expressed enzymes in conditions such as cancer or atherosclerosis were not significantly successfully in clinical research. This was mainly due to the lack of a clear understanding about the enzyme activity and expression *in vivo* (Bagshawe 1993; Connors and Knox 1995). Hence in order to specifically image these enzymes (overexpressed in abnormal regions) it will be effective, to have a relatively non-invasive procedure which will enable real time evaluation of enzyme activity. Therefore, prodrugs attached to molecular imaging

agents may be advantageous in the improvement of enzyme-activated prodrugs (Gabriel *et al*. 2011). As discussed in previous sections, there are a number of imaging techniques to detect enzymatic activity, among which MRI and optical imaging (FRET) were the main focus of study in this research project. It has already been demonstrated that MRI could be used to monitor vulnerable plaques using molecular probes (Sirol *et al*. 2004). Molecular imaging may substantially improve the accuracy of MRI by selectively revealing markers of instability. As previously discussed MMPs and legumain can be exploited to site specifically deposit these contrast agents/therapeutic drugs specifically *via* enzyme cleavage in vulnerable plaques.

To achieve this, the synthesis of a paramagnetic gadolinium chelate covalently linked to a small peptide (typically 7-9 amino acids in length for MMPs and a tetrapeptide for legumain), where the peptide is a tailored substrate for specific MMP/legumain which are found to be over expressed in vulnerable plaques will be attempted which is the main aim of this research project (figure 23-B). Lancelot and colleagues have already reported the use of MMP inhibitor, rather than a substrate for non-invasive imaging in 2008. However, a reduction in biological clearance time and also an increase in MRI contrast can be enhanced by the use of a MMP-substrate/gadolinium chelate conjugate in a method analogous to that of Jastrzebska and colleagues (2009). The authors in their work have developed a novel solubility-switchable contrast agent for the *in vivo* imaging of tumours.

The aim of this research program was the synthesis and characterization of two main series of compounds reported as the Legumain series (table 7, 8 and 9) and the MMP series (table 10). Each library of compounds was synthesized *via* combination of both solid and solution phase peptide synthesis. The synthesis of these compounds could lead to site-specific deposition of the paramagnetic contrast agent by successful cleavage of the peptide substrates *via* the corresponding proteases (MMP/Legumain) for imaging atherosclerotic plaques.
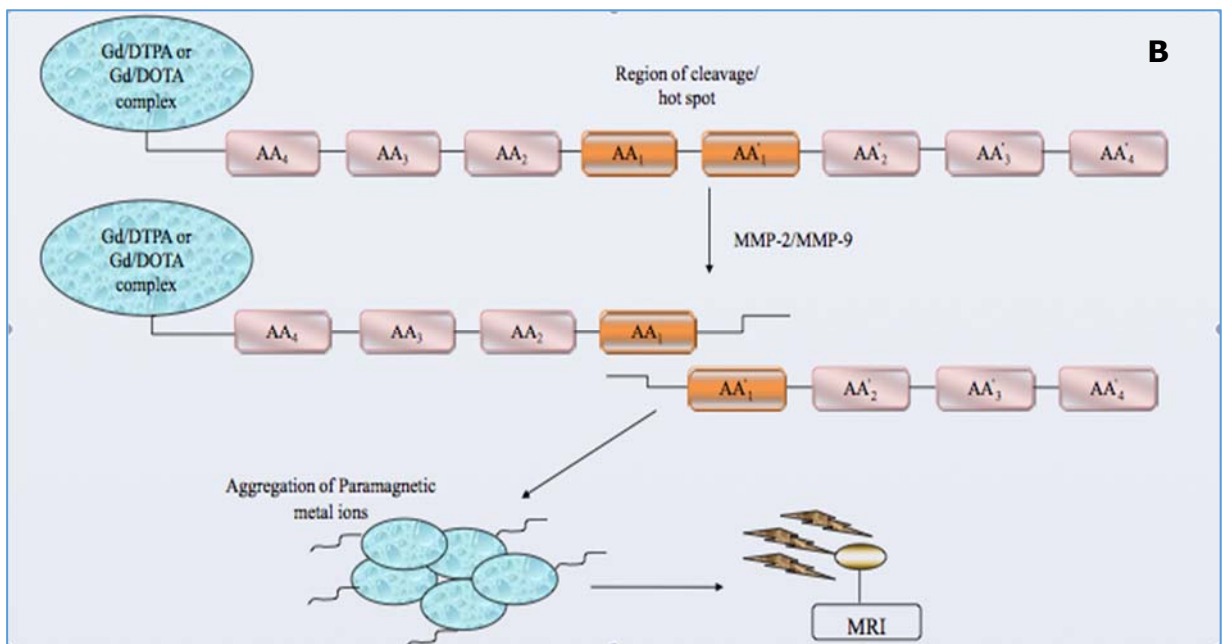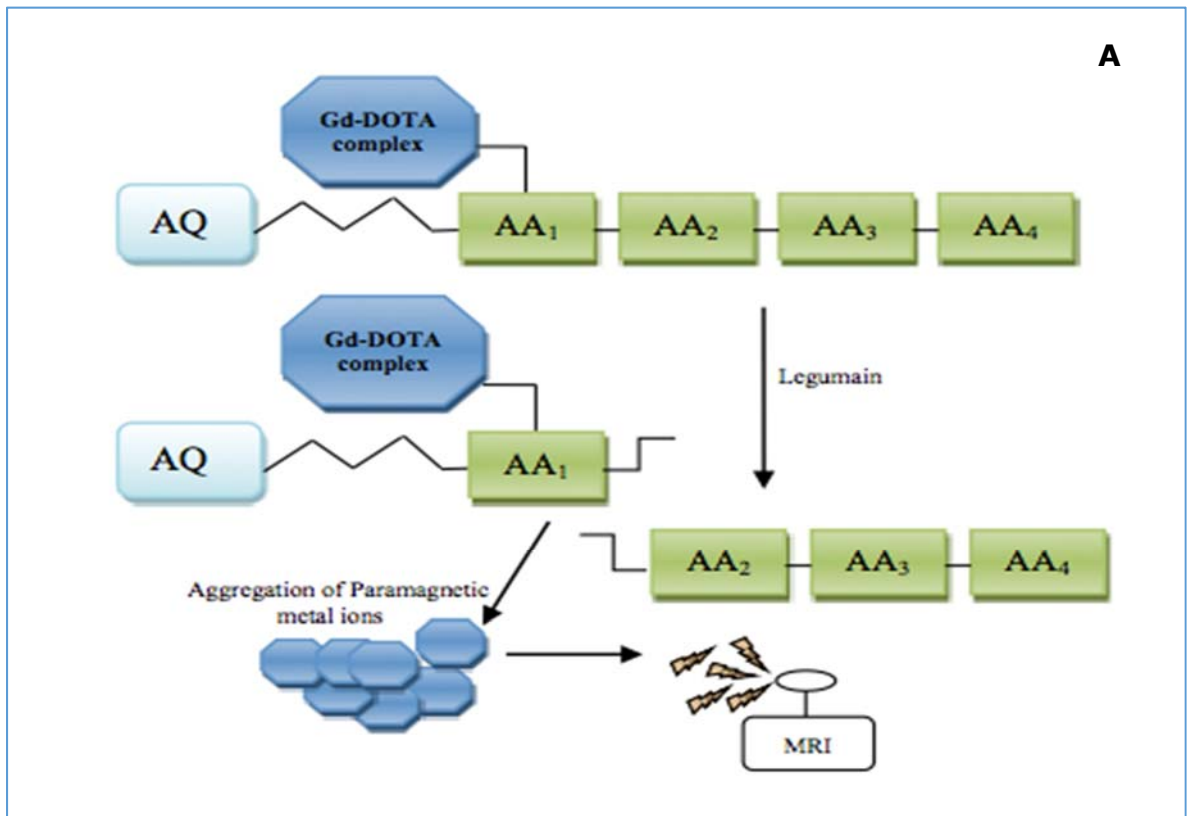
Figure 23: A pictorial representation of peptide substrate cleaved by legumain (A)/MMPs (B) to aggregate metal ions for imaging.

1.6.1.1. Peptide synthesis

The amino acids were coupled sequentially to one another through amide bonds to synthesize the peptide substrates. These substrates were designed to be specific to the enzyme of interest, which in this research study was MMP-2/9 and legumain. The arrangement of amino acids along the sequence was done to mimic the corresponding enzyme subsites or pockets so that the positioned amino acid projects into the relevant spaces. The peptide sequence containing the active site is normally labelled as P3-P2-P1-P1′-P2′-P3′ and the corresponding enzyme cleaves the sequence around the region between P1-P1′ (Lim *et al*. 2010) denoted as the hot spot of the sequence.
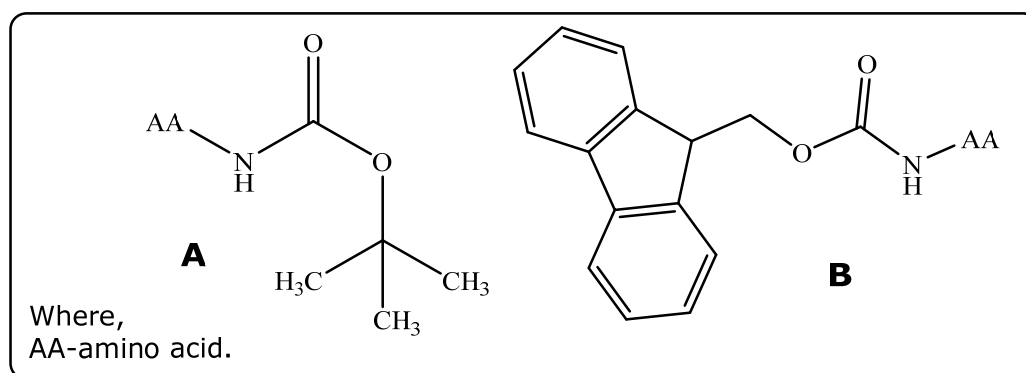
Figure 24: A- tertiarybutyloxycarbonyl (tBoc) group, B- 9-fluorenylmethyloxycarbonyl (Fmoc) group.

The amino acids chosen for the peptide synthesis included a wide range of aliphatic, aromatic, hydrophobic and neutral functional groups. In order to prevent any unwanted side chain reactions or polymerization all amino acids were *N*-protected either by the tertiarybutyloxycarbonyl (tBoc) group or the 9-fluorenylmethyloxycarbonyl (Fmoc) group (figure 24). Solid phase peptide synthesis is a process by which a desired peptide sequence can be synthesised by the sequential addition of amino acids on to a solid support. This method was first developed by Bruce Merrifield in 1984 for the synthesis of polypeptides. Some of the major advantages of this procedure when compared to conventional synthesis are easy work-up, stable, quick and simple isolation procedures (Merrifield 1963; Erickson and Merrifield 1976; Leznoff 1978). In synthesizing longer sequences, solution phase peptide synthesis may become labour-intensive due to the repetition of coupling and deprotection cycles and may also require isolation of peptide intermediates. Peptide substrates for MMPs were synthesized manually applying the SPPS procedure and a 2-chlorotrityl resin (figure 25A) preloaded with glycine was used as the solid support.
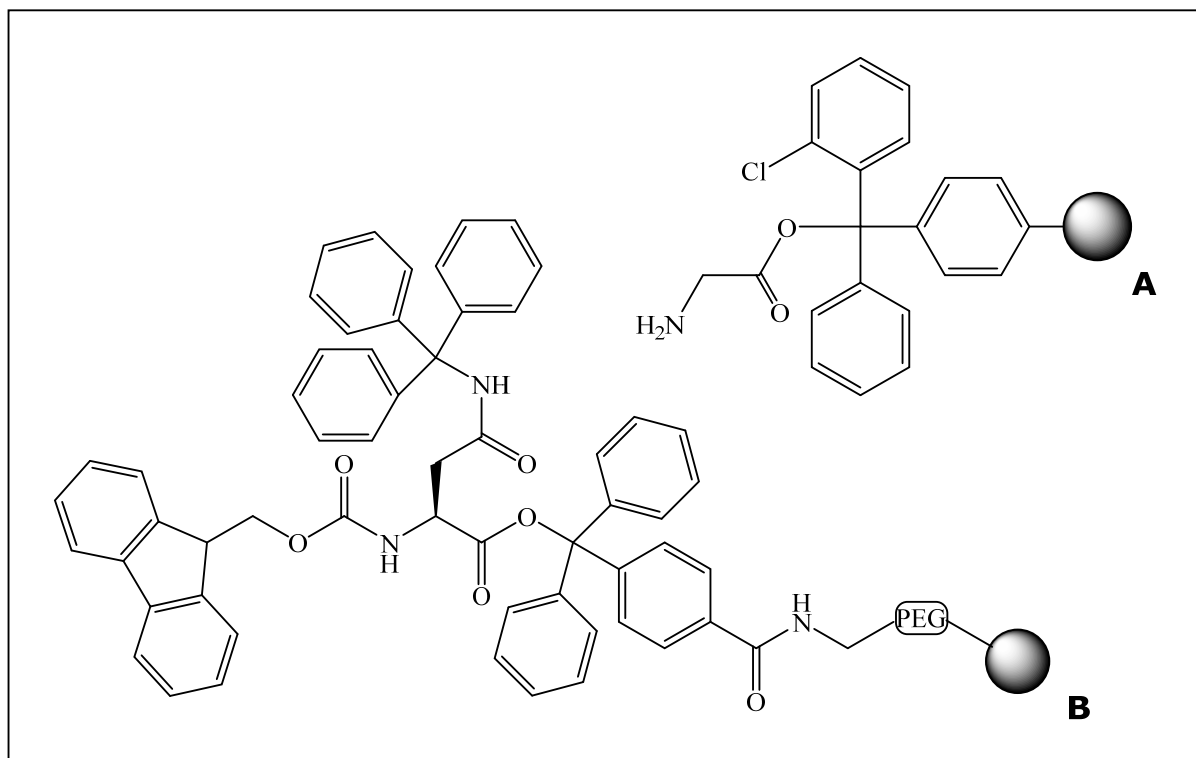
Figure 25: A: 2-chlorotrityl resin preloaded with glycine (H-Gly-2-ClTrt); B: Novasyn TGT resin preloaded with N-Fmoc protected asparagine (Fmoc-Asn(Trt)-Novasyn TGT).

The 2-chlorotrityl resin was chosen as polymeric support because it is acid-labile and the rate of racemisation is highly reduced when using Fmoc protected amino acids (Barlos *et al*. 1989). The stability of the resin is good under storage conditions of 2-6ºC in DCM and final cleavage of the peptide sequence from the resin was done using trifluroacetic acid (TFA) in dichloromethane (DCM). The tetrapeptide substrates for legumains were synthesized manually applying standard fluorenylmethoxycarbonyl (Fmoc) SPPS procedure on preloaded Fmoc-Asn(trt)/Leu-Novasyn TGT resin (figure 25B). The chemical properties of this resin are similar to that of 2-chlorotrityl resin but also have other advantages such as highly polar poly ethylene glycol (PEG) and polystyrenes (PEG-PS) matrix and have good solvation properties (Kates *et al*. 1998).

1.6.1.2. Coupling, protection and deprotection procedure

Coupling agents were used during both solid and solution phase peptide synthesis. This was done to activate the carboxyl group of the corresponding amino acid prior to coupling to the sequence. The formation of a peptide bond is the vital step in peptide synthesis, which involves amide bond formation (figure 26). For this reaction to take place, coupling reagents are required for activating the carboxylic

unit of the amino acid. The formation of a peptide bond is a nucleophilic substitution reaction of an amino group (nucleophile) at a carboxyl group *via* a tetrahedral intermediate. Carboxylic acids react with ammonia or amines at rt to produce an ammonium salt instead of a carboxamide. Hence the carboxyl component of an amino acid has to be activated prior to a peptide bond formation by introducing electron-accepting moieties. The following reaction between the activated intermediate and amine unit is defined as the coupling reaction and the reagents used for activation are termed as coupling reagents (Han and Kim 2004).
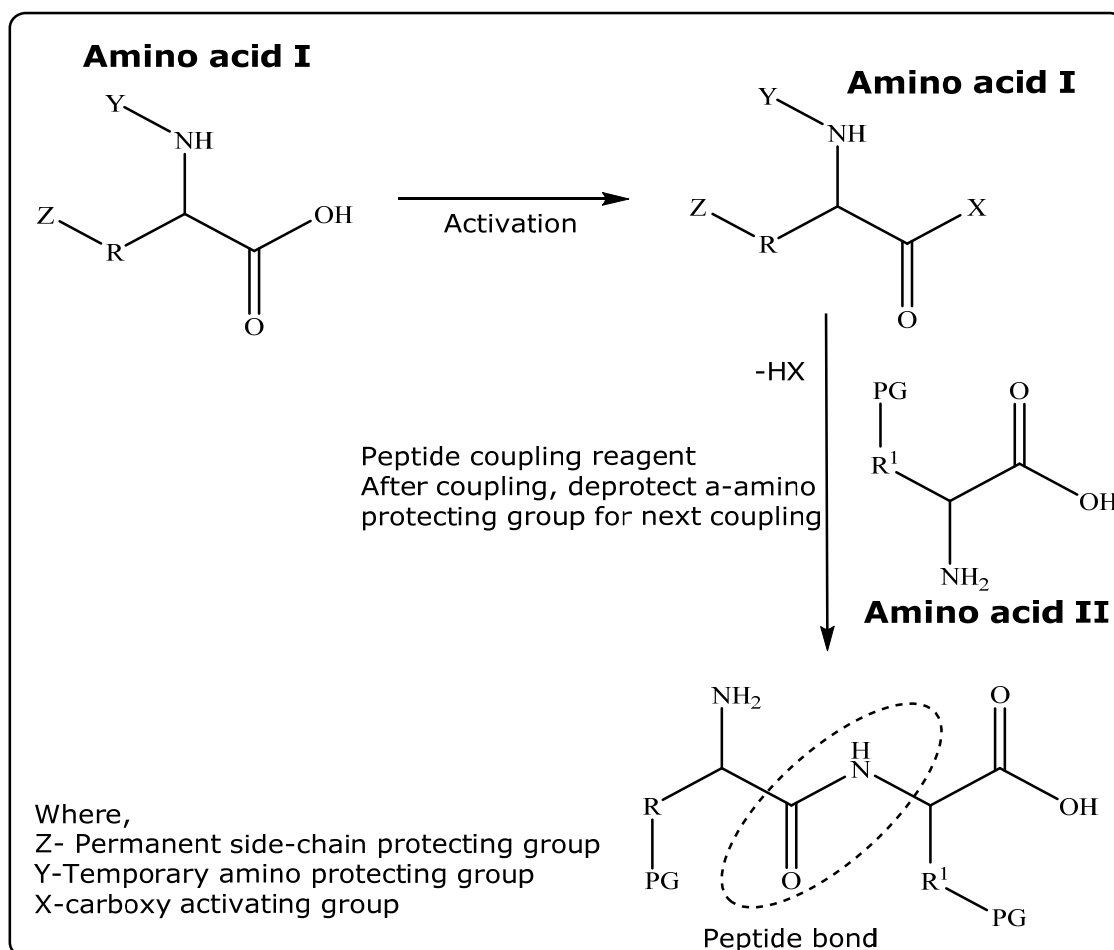
Figure 26: Schematic representation of a coupling reaction.

There are a number of coupling agents available of which the most commonly used are uronium and phosphonium salts and hydroxybenzotriazole (HOBt) (Miranda and Alewood 2000). The coupling agents used during this research project consisted of TBTU/HOBT and DIPEA or PyBOP/DIPEA, which were used interchangeably. The synthesis of the intact peptide sequences for conjugation to the aminoanthraquinone conjugate/Gadolinium conjugate was done following the procedure for solid phase peptide synthesis (figure 27). While sequentially arranging amino acids onto the solid support it is always important to protect any reactive side chain groups together with the primary amine. The temporary

protecting groups on the primary amine were removed at each stage after the coupling and the side chain protecting groups were removed along with the final cleavage of the sequence from the resin (Kent 1988). The most commonly used protecting groups are ^{t}Boc/Bzl and Fmoc/^{t}Bu (Chan and White 2000) mainly because they can be removed easily using an acid, base or by hydrogenolysis respectively.
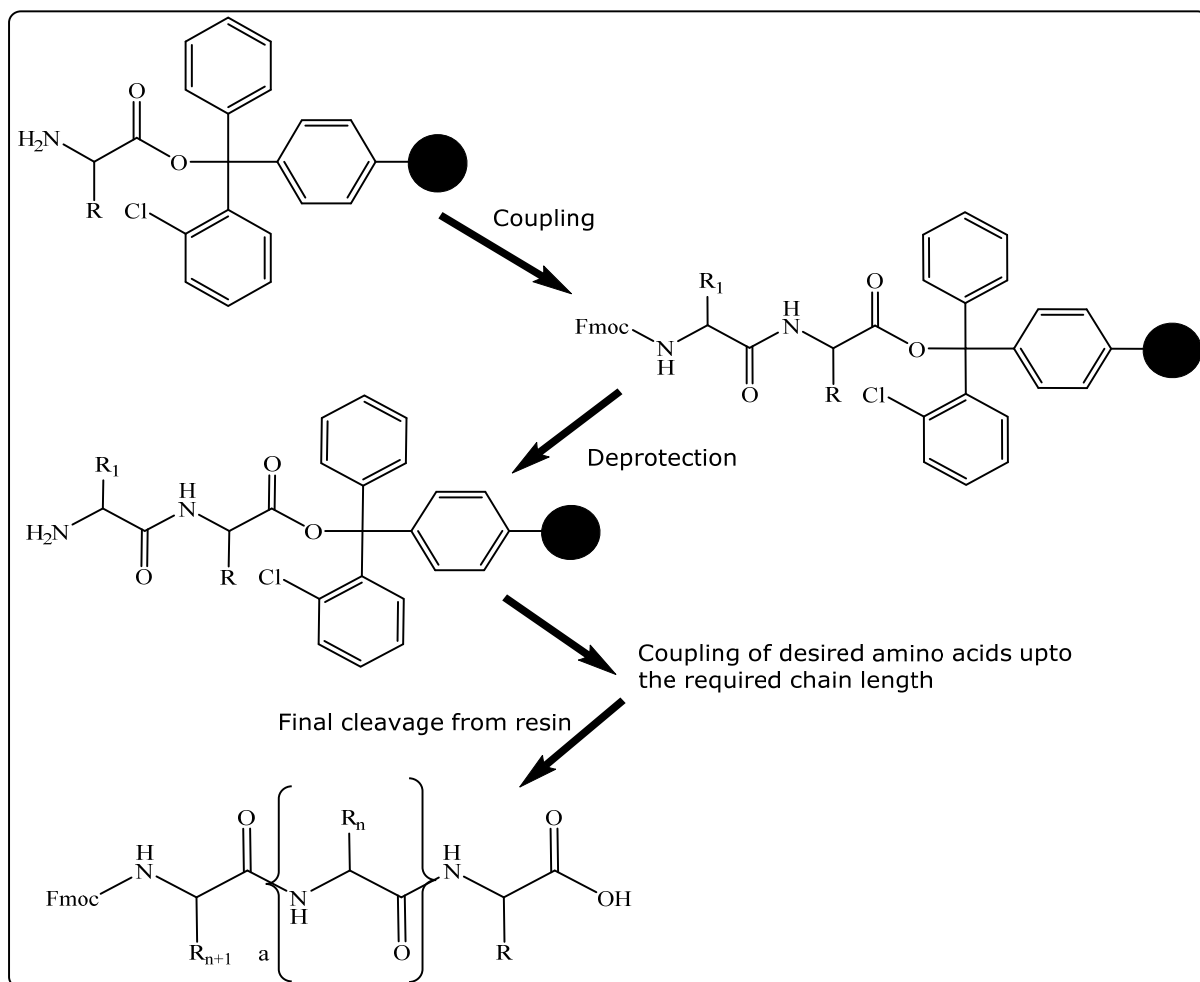
Figure 27: Representation of steps involved in SPPS.

Moreover these groups provide better orthogonality, are cost effective and many amines are available commercially with these protecting groups (Sheppeck *et al*. 2000). In the ^{t}Boc/Bzl system, the primary amine is protected by the ^{t}Boc group and the benzyl or cyclohexyl group protects the side chains. Similarly for Fmoc/^{t}Bu, the primary amine is protected by the Fmoc group and the side chains are protected by the tert-butyl group (Albericio 2000).
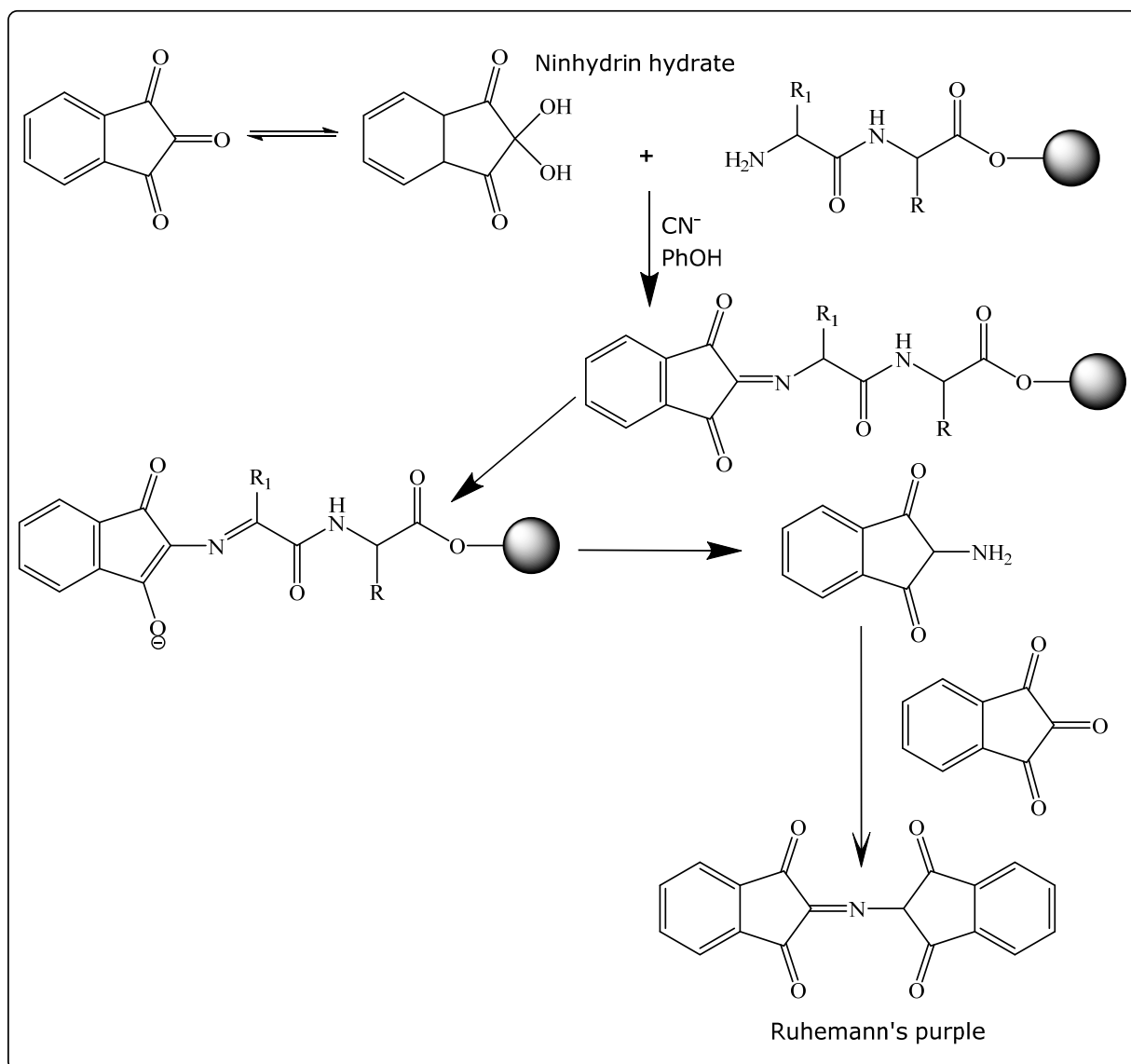
Figure 28: Schematic representation of Kaiser test (Friedman, 2004).

In order to monitor the successful completion of coupling and deprotection of amino acids in a SPPS system, the Kaiser test was undertaken to analyse the colour changes. This test is based on the reaction of ninhydrin with the primary amine (Kaiser *et al.* 1970) giving rise to an intense blue colour. The intensity of the colour developed depends on the terminal amino acid (figure 28). After the final coupling, the peptide sequence was removed from the resin and characterized using mass spectrometry. Some of the peptide substrates for the legumain series of compounds were synthesized following the procedure for solution phase peptide synthesis.
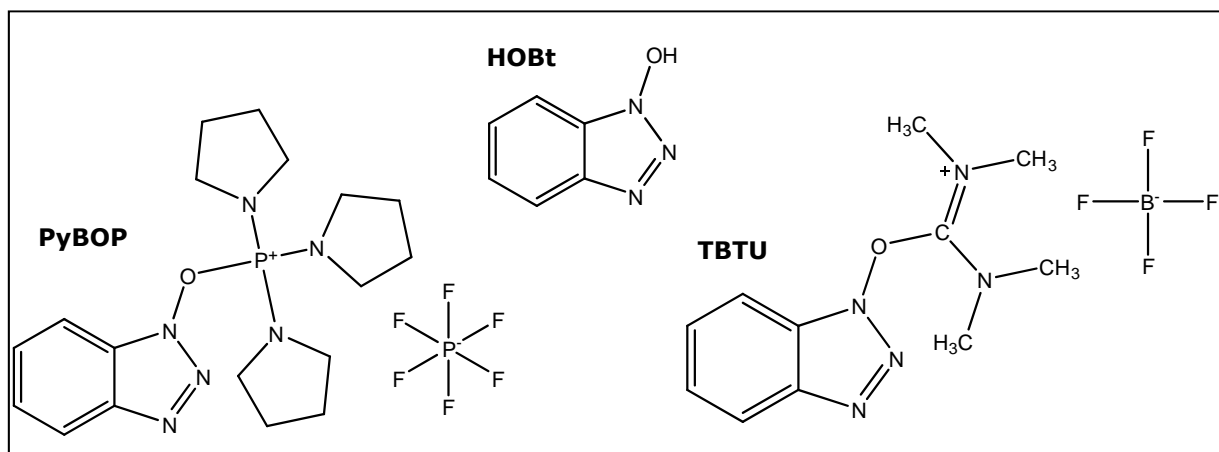
Figure 29: Structure of coupling agents used in this project.

An aminoanthraquinone unit conjugated to a PEG like spacer acted as a soluble support for immobilization of the amino acids. Coupling of the desired *N*-protected amino acids was performed in DMF using activating/coupling reagents such as HOBt/TBTU and DIPEA or the PyBOP and DIPEA system (figure 29). Both TBTU and HOBt were used to accelerate the process of coupling but the presence of HOBt was also found to suppress racemisation by forming less reactive HOBt esters. This additive also inhibited dehydration of the carboxamide side chain of Asn residues (Stawikowski and Fields, 2002). An excess of the activated amino acid in the range of 1.1 to 2 folds compared to the anchoring support (anthraquinone moiety) was used to drive the coupling reaction to completion.

1.6.1.3. Choice of chelating agents and contrast agent

In order to produce a sharp contrast between the normal and abnormal tissues MR imaging depends on the use of contrast agents (CAs). This clarity in images can be produced by increasing the electronic retention time of the imaging agent and the magnetic moment of the complex, which can be achieved by the use of chelated gadolinium complexes (Gd-DOTA/Gd-DTPA). As previously discussed the use of macrocyclic (DOTA) and linear (DTPA) chelating agents have long been recognized as efficient conjugates for MRI contrast agents (Nwe *et al*. 2011). Gadolinium (Gd), a rare lanthanide known for its unique magnetic property has been one of the most prominently used contrast agents for MR imaging (figure 30). However, there are a few other contrast agents such as iron particle CA and manganese (II) chelate, but Gd is preferred due to the nature of the metal ion and also the requirement of MR imaging (Caravan *et al*. 1999).
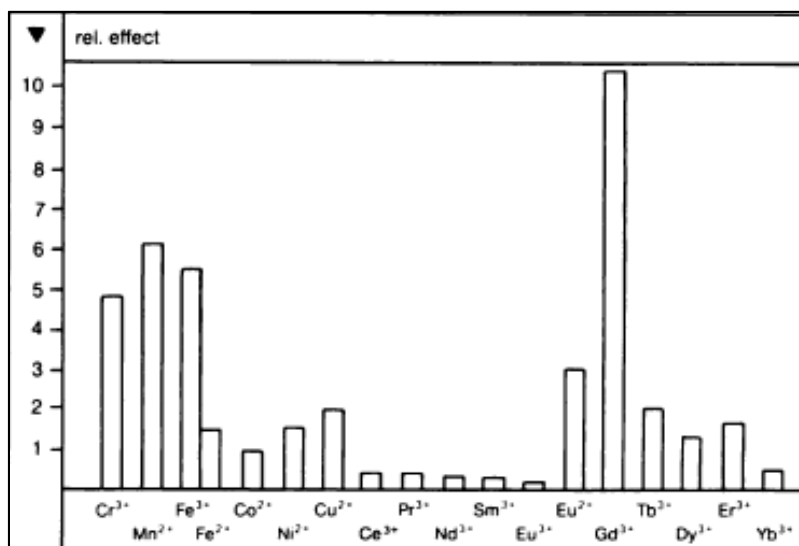
Figure 30: Various paramagnetic ions and their influence on proton relaxation time (Pople *et al*. 1959).

However, due to the toxic nature of the free metal ion, it has to be chelated to ligands such that they remain intact inside the body and can also be excreted without dissociation from the complex. For example, as in Magnevist $[Gd(DTPA)(H_2O)]^{2-}$ a clinically approved and commercially available CA is found not to enter cells and does not have any intercalating groups to damage DNA (Caravan *et al*. 1999). Of the available ligands, DOTA is considered to possess good thermodynamic stability and kinetic resistance when compared to DTPA although both DOTA and DTPA are commercially used as chelating agents for MRI contrast agents (Duncan *et al*. 1994; Loncin *et al*. 1986). Gd-DOTA has been successfully used as a contrast agent for evaluating intracranial lesions of the central nervous system and results show that they have been well tolerated *in-vivo*. (Parizel *et al*. 1989).

A ligand should be designed in such a way that it would increase the number of inner-sphere water molecules (q) in the complex, causing a very slow tumbling rate ($1/\zeta r$) and finally less water residence time (ζm) as shown in figure 31 (Raymond and Pierre 2005). In order to increase the sensitivity of contrast agents, they are coupled to proteins or receptor molecules or small peptide sequences which in turn increases the specificity by binding to the corresponding target molecule. This coupling also increases the concentration and retention of the entire complex in the target region leading to increased relativity and also reduced background noise (Caravan *et al*. 1999), for example, Gd-DOTA-Gal80 specific peptide is unique for binding to galactose regulatory protein Gal80 (Li *et al*. 2002).
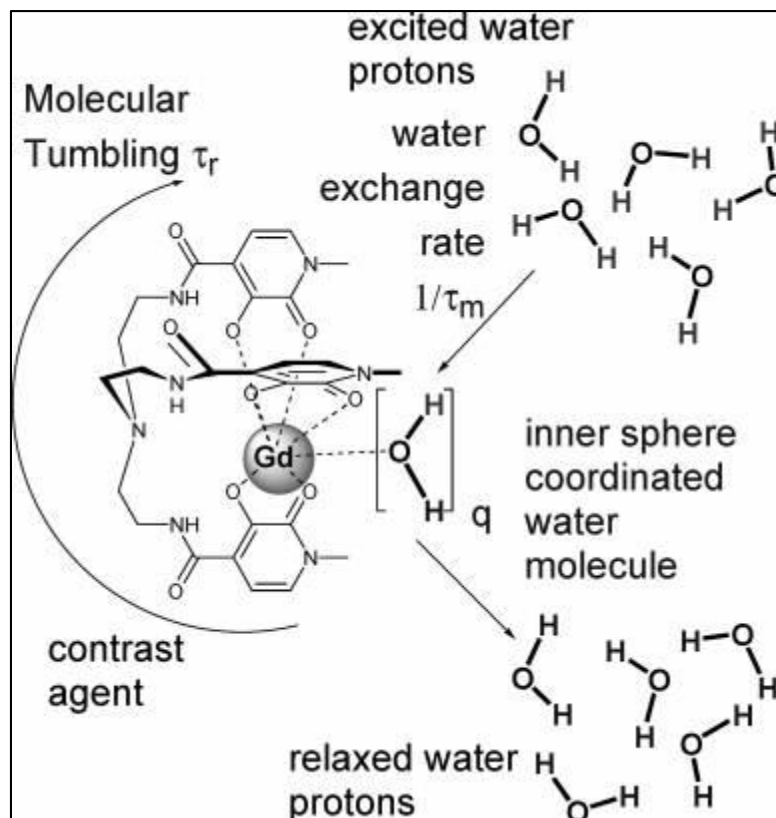
Figure 31: Features influencing relativity (Raymond and Pierre 2005).

Based on the same idea, a small library of Gd-DOTA-peptide conjugates was synthesized with the peptide sequence having the favourable cleavage site for the corresponding enzyme of interest, which was MMP-9/2 and legumain during this research program. Upon cleavage by the target enzyme the contrast agent (Gd-DOTA) is anticipated to quicken water exchange and increase relativity leading to enhanced contrast between the target and the surrounding region (Meade and Allen 2004).

2. Results and discussion

2.1. Library of FRET peptide substrates designed to be specific to legumain

This section explains the design, synthesis, characterization and preliminary biological evaluation of a library of peptide fluorogenic substrates. Each peptide sequence was designed to be a specific substrate for legumain and consisted of four amino acids with the anticipated cleavage site (*) being on the carboxyl end of an asparagine residue.

- **JC-series**: consisted of a fluorescent probe (5(6)-CBF) conjugated to a peptide substrate: 5(6)-CBF-Z-Y-Asn-(*)-Lys/Leu-Spacer-AQ; [where Y and Z denoted the amino acids used in the sequence and the spacer consisted of a PEG like diamino chain] (Figure 32).
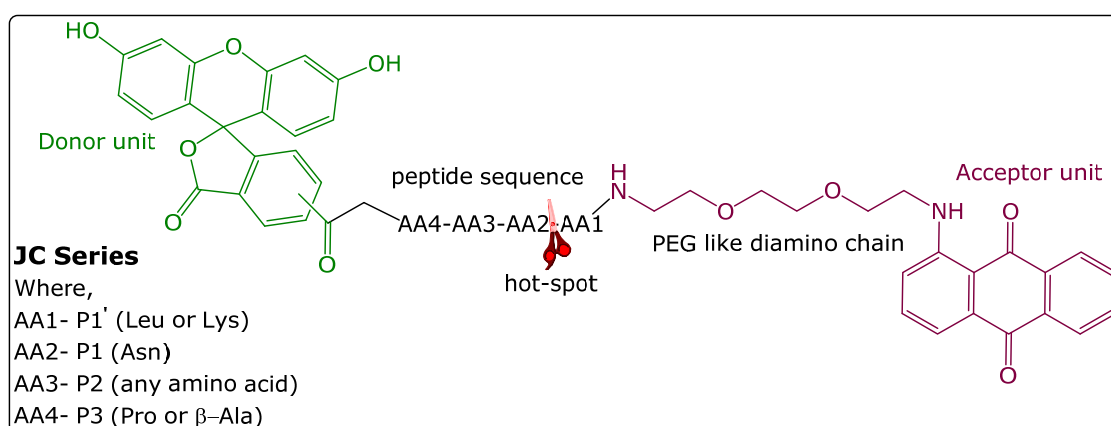
Figure 32: Structure of JC (1st and 2nd) series of peptide substrates.

Two batches of FRET peptide substrates designed to be specific to legumain were synthesised (**JC series 1 and 2**). Fluorescently tagged amino acids have been exploited as powerful tools for the evaluation of biological events such as receptor-ligand binding, enzyme activities and protein structures (Ozawa and Umezawa 2002). Fluorescent labeling is rapidly replacing radiolabelling for production of kits for binding and enzyme assays with the application of FRET techniques and fluorescent polarization. Use of fluorescent labeling involving the FRET principle has become increasingly attractive due to advantages such as high stability, quick detection and ease of instrumentation. Also the disadvantages, involved with the use of radioactive isotopes can be avoided (Tung 2004). The aim of this research study was the synthesis of imaging agents which could specifically identify

atherosclerotic plaques with the aid of biomarkers such as legumain that are overexpressed in the microenvironment of the plaques. In order to achieve this, a FRET peptide library was designed in such a way that the positioning of the amino acids along the sequence was done to mimic the subsites of the enzyme of interest, which in this study was legumain. Hence, the purpose for the synthesis of the FRET compounds was to identify those peptide substrates that could be readily cleaved by legumain which was confirmed by the release of fluorescence *via* a legumain-based assay. In this manner, the successful peptide substrates can be used to carry the contrast agent that could then be deposited site specifically on atherosclerotic plaques, which could possibly improve the quality of image obtained from a MRI scanner. The design outline for this library of peptide substrates will follow the template as shown in figure 32.

The fluorophore was 5(6)-carboxyfluorescein whilst the acceptor moiety was an aminoanthraquinone derivative with a hydrophilic diamino chain (PEG like structure). Some of the advantages of the 5(6)-CBF and PEG-like diamino chain was the improved solubility during synthesis; the carboxylic group on the fluorophore can form an amide bond with the adjacent amine, such amides are resistant to hydrolysis, which coupled to the non-toxic nature of the fluorophore makes it biocompatible for *in-vitro* studies. Peptides and nucleotides labelled with FITC tend to deteriorate more quickly when compared to 5(6)-CBF conjugates (Banks and Paquette. 1995). Another advantage of 5(6)-CBF was its sensitivity at acidic and near-neutral pH making it suitable for the rh-legumain assay where the release in fluorescence is studied in similar conditions (Bidmanova *et al*. 2012). Moreover, 5(6)-CBF can be used for both solid and solution phase peptide synthesis and is cost effective when compared to other fluorescein derivatives such as fluorescein isothiocyanate and carboxyfluorescein-*N*-succinimidyl ester (Fisher *et al*. 2003; Fernandez-Carneado and Giralt 2004). The design of the JC series had the fluorophore at the *N*-terminus of the tetrapeptide and the aminoanthraquinone conjugate at the carboxyl terminus. From the available information on legumain substrate specificity it was decided to place the asparagine residue at the P1 position and the P3 position was occupied by either proline or β-alanine (Mathieu *et al*. 2002; Sexton *et al*. 2007; Liu 2009). The FRET properties of the peptide fluorogenic substrates were evaluated using a spectrophotometer prior to demonstrating the release of fluorescence when incubated with the rh-legumain.

2.1.1. Synthesis of FRET peptide substrates (JC series-1st batch)

From previous discussions (section-1.3.3) it can be demonstrated that legumain is one promising biomarker for the identification of atherosclerotic plaques. Therefore, synthesizing compounds that could be recognized by legumain and subsequently release the diagnostic/therapeutic agent could be the next step towards early detection of atherosclerotic plaques (Papaspyridonos *et al*. 2006). A library of peptide substrates designed to be specific to legumain were synthesized in order to study the efficiency in fluorescence release upon incubation with the enzyme. The successful peptide substrates were then used as a template for the synthesis of peptide coupled contrast agents which could perhaps be used for the imaging of atherosclerotic plaques *via* MRI.
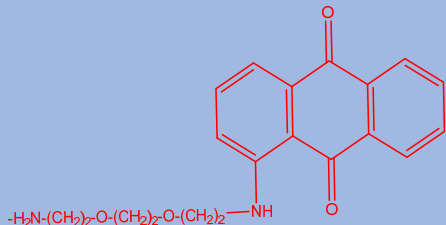
Comp. no.	Fluorophore	P3	P2	P1	P1′	Donor unit
8		Pro	Ala	Asn	Leu	
11		Pro	Gly	Asn	Leu	
14		Pro	Leu	Asn	Leu	
17	5(6)-CBF	β-Ala	Ala	Asn	Leu	
20		β-Ala	Leu	Asn	Leu	
25		β-Ala	Gly	Asn	Leu	

Table 7: List of FRET peptide sequence for JC series-1.

This section details the design, synthesis and characterization of the 1st batch of JC series of FRET peptide substrates. An overview of the 1st batch of FRET compounds is shown in table 7. The complete synthesis of the JC series (1st) of peptide substrates will involve the intermediate and target compounds as detailed in figure 33.
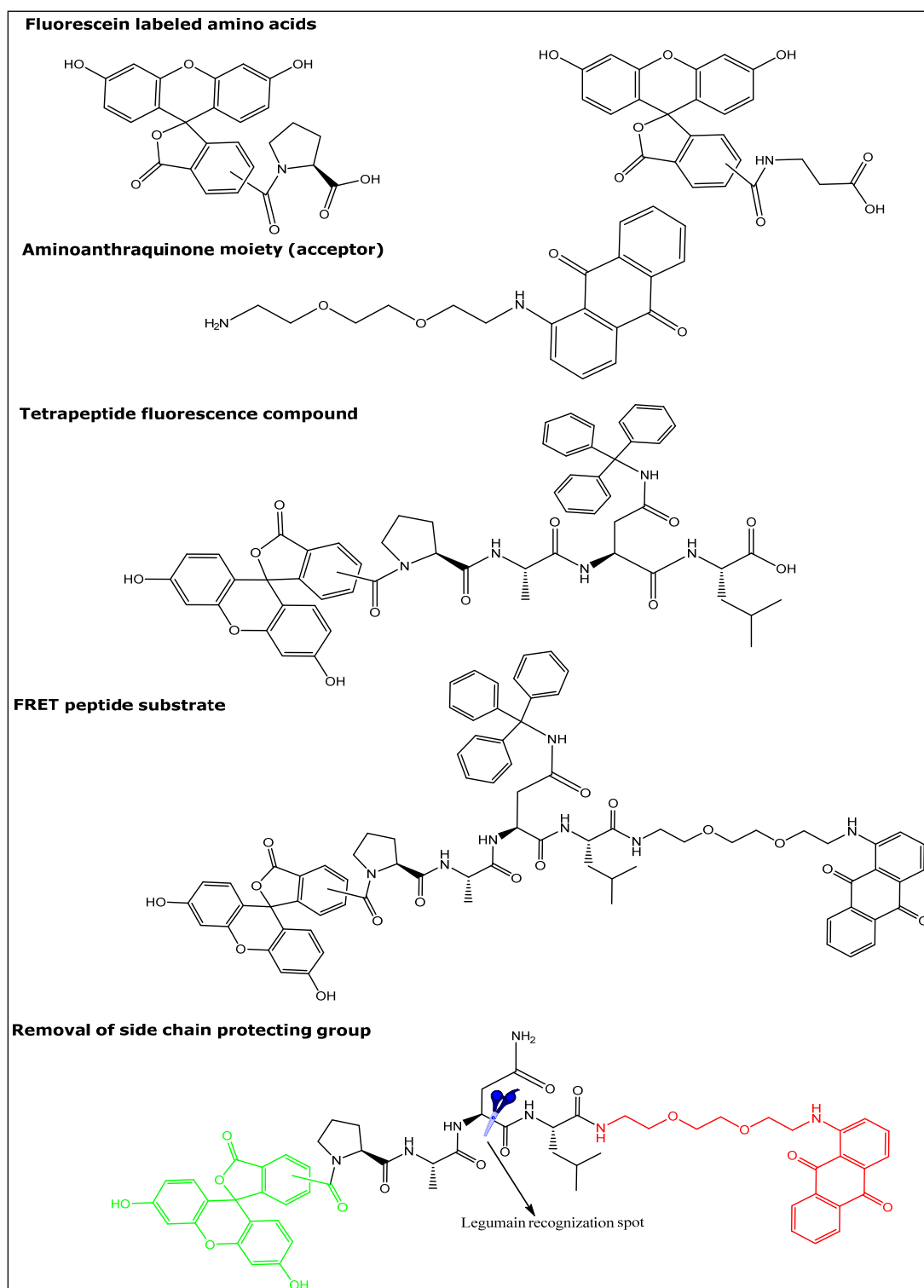
Figure 33: An overview of the intermediate and target compounds for the JC series-1 of FRET peptide substrates.

2.1.1.1. Step 1: Synthesis of proline/β-alanine labeled amino acids (donor unit) (Compound 2, 4)

In this study fluorescein labeled amino acids were initially synthesized as part of the tetra peptide sequence. The fluorescein labeled compound was then coupled

to the amino terminus of the tripeptide sequence containing the donor unit (AQ) of the FRET pair.
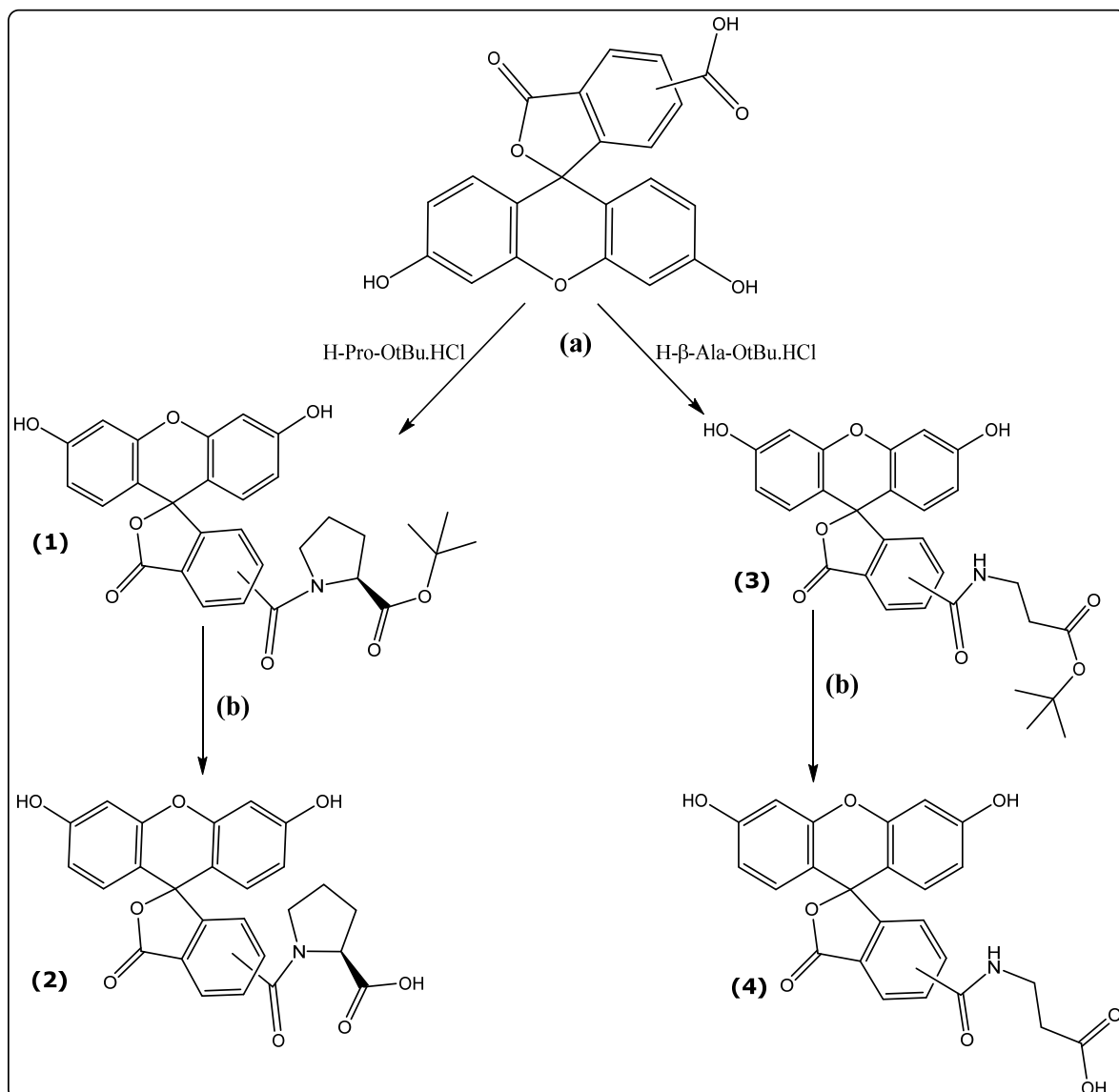
Scheme 1: Synthesis of fluorescein labeled amino acids. Conditions and reagents: (a) DMF, DIPEA, PyBOP (Proline/β-alanine), rt. (b) TFA, rt.

The *tert*-Butyl protected fluorescein conjugates (**1** and **3**) were synthesized by the reaction of 5(6)-carboxyfluorescein with H-Pro-O^{t}Bu hydrochloride or H-β-Ala-O^{t}Bu hydrochloride in DMF as shown in scheme 1(a). The mixed isomer 5(6)-carboxyfluorescein was activated using PyBOP in DMF prior to coupling with proline (**1**) or β-alanine (**3**) (scheme 1a). Reaction progress was monitored by TLC analysis, using a chloroform extract of the reaction mixture, which was referred against the starting material (5(6)-CBF). The TLC plate showed two components with a bright yellow fluorescein band corresponding to the product (with a R$_f$ value of 0.9 demonstrating the less polar nature of the compound due to the protected proline/β-alanine residue) and a pale yellow band corresponding to the unreacted 5(6)-CBF. The crude product was then partitioned between chloroform and an

aqueous solution of citric acid in order to facilitate the extraction of the product into the organic solution. The organic solution was then washed with water several times until the aqueous solution was colorless or pale yellow. The organic solution was then purified by column chromatography, using 10% MeOH in chloroform to elute the first pale fraction followed by a gradient elution using MeOH in 5% increments (up to 20%) to elute the major band.
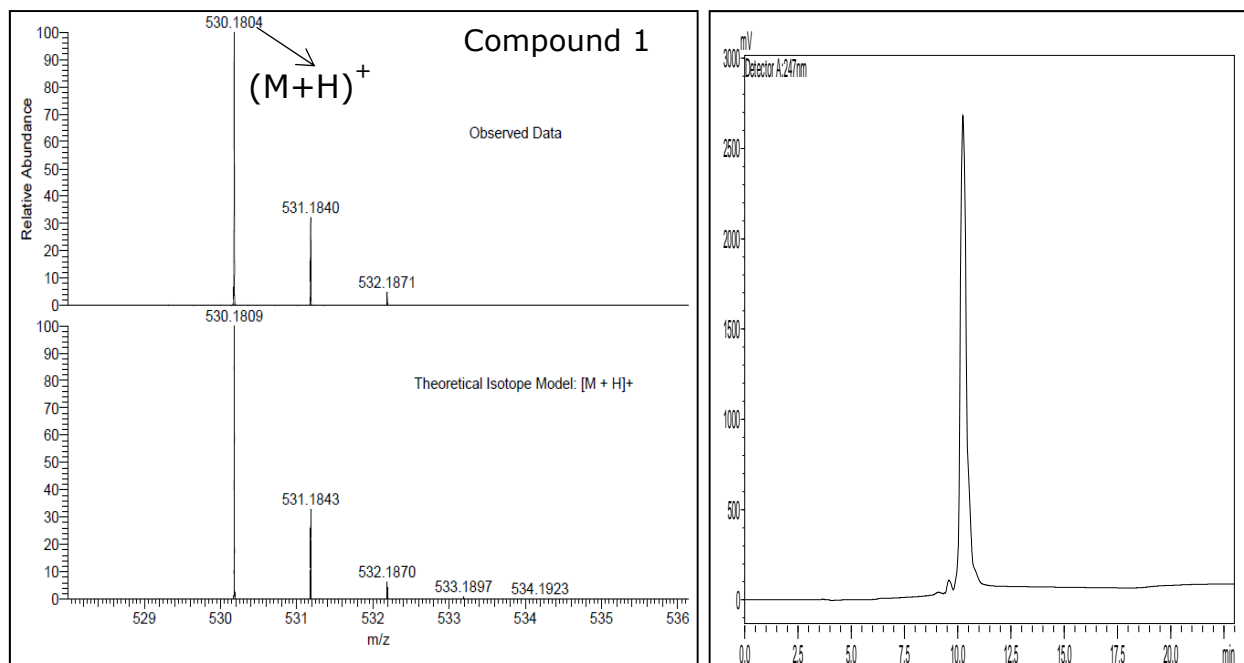
Figure 34: Mass spectrometry result for compound **1**. RP-HPLC elution profile of compound **1** (0.5mg/mL in methanol); Conditions: Flow rate, 0.5mL/min; Detection, 247nm; stationary phase: Eclipse XDB-C18 RP column; mobile phase A= 0.1%TFA/water; mobile phase B=0.1%TFA/methanol; gradient: low pressure gradient; 21min.

The purity of compounds **1** (figure 34) and **3** were found to be 92% and 94% using a Shimadzu Prominence HPLC system equipped with an eclipse XDB-C18 RP column. Additionally the compounds were also characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 530.1804 (M+H)$^+$ corresponding to a molecular mass of 529 relating to the proline conjugate (**1**-figure 34). Similarly compound **3** containing the β-alanine conjugate was also characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 504.1642 (M+H)$^+$ corresponding to a molecular mass of 503 (figure 35).

In order to couple the labeled amino acid to the amino terminus of the tri-peptide sequence, the *tert*-butyl protected compounds **1** and **3** were deprotected using trifluoroacetic acid (scheme 1b). Reaction progress was monitored by TLC, which showed the presence of a single new component, with a low R$_f$ value due to the polar nature of the deprotected product. The purity of the compounds **2** and **4** were found to be 81% and 51% using a Shimadzu Prominence HPLC system

equipped with an eclipse XDB-C18 RP column. The products (**2** and **4**) were then characterized by high resolution ESI (+) mass spectrometry, which displayed a signal at m/z 474.1179 (M+H)⁺ corresponding to a molecular mass of 473 relating to the proline conjugate (**2**). Similarly the β-alanine conjugate (**4**) was also characterized by high resolution ESI (+) mass spectrometry, which displayed a signal at m/z 448.1025 (M+H)⁺ corresponding to a molecular mass of 447 (figure 35).
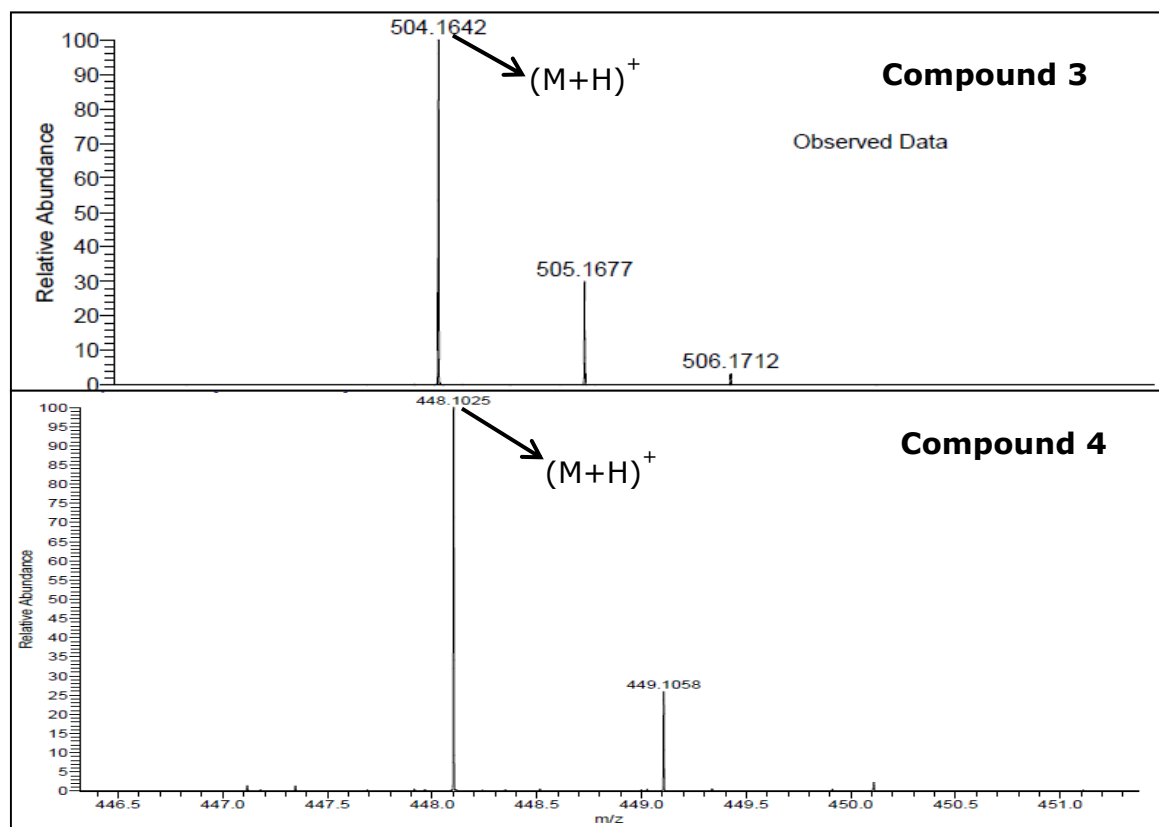
Figure 35: Mass spectrometry result for compounds **3** and **4** was confirmed using a high resolution-electrospray ionization mass spectrometry (ESI-MS).

2.1.1.2. Step 2: Synthesis of anthraquinone-spacer conjugate (acceptor unit) (Compound 5)

The synthesis of the aminoanthraquinone spacer conjugate (**5**) was attempted, following an adapted procedure by Turnbull (2003). 1-Chloroanthraquinone was reacted with 2,2'-(ethylenedioxy)bis(ethylamine) in DMSO (scheme 2). The mixture was heated over a water bath for 3 hours. A gradual color change from a pale yellow to an intense pink, characteristic of the modified anthraquinone chromophore was noticed.
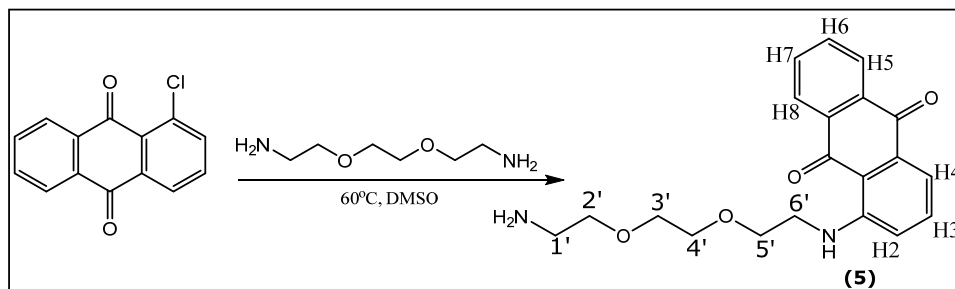
Scheme 2: Synthesis of anthraquinone spacer conjugate.

The reaction mixture was then cooled to room temperature and a large excess of water was added to facilitate precipitation of the aminoanthraquinone product. The product was loaded onto a chromatographic column followed by elution of a front running pale yellow impurity using DCM. The polarity of the solvent system was increased gradually by adding methanol in increments of 5% (up to 20%). Since the elution of the major red band was slow, triethylamine (TEA) was added to quicken the elution of the polar compound. Adding TEA prevents ionization of the free amine in the compound due to the slightly acidic nature of silica and thus reduces the retention time of the compounds on the silica gel column. The purity of the sample was found to be 83% using a Shimadzu Prominence HPLC system equipped with an eclipse XDB-C18 RP column. The product was then characterized using a high resolution ESI (+) mass spectrometry, which gave a signal at m/z 355 $(M+H)^+$ corresponding to a molecular mass of 354. Furthermore, the ^{1}H NMR spectrum (d-DMSO) (figure 36) showed a 2-proton triplet signal at 2.65ppm and another 2-proton triplet signal at 3.39ppm which was assigned to 1′ and 2′ protons. Signals for 4-proton multiplet between 3.51 and 3.58ppm were assigned to 6′ and 3′ protons. A two proton multiplet between 3.62ppm and 3.64ppm was assigned to 4′ protons. Signal for a 2-proton triplet at 3.72ppm was assigned to 5′ protons. The protons attached to the aromatic group were assigned as follows: H-2 and H-4 gave a single proton doublet at 7.28ppm and a single proton double doublet at 7.44ppm respectively and a 1-proton multiplet between 7.62ppm and 7.66ppm was assigned to H-3. Another two proton multiplet between 7.81ppm and 7.92ppm was assigned to H6 and H7 protons. The H-5 and H-8 gave a single proton double doublet signals each at 8.12ppm and 8.20ppm respectively. A triplet for a single proton at 9.78ppm was assigned to the anthraquinone amino group proton. To further confirm the assignment in the ^{1}H NMR, a 2D-COSY spectrum of compound **5** was also undertaken (appendix 1).
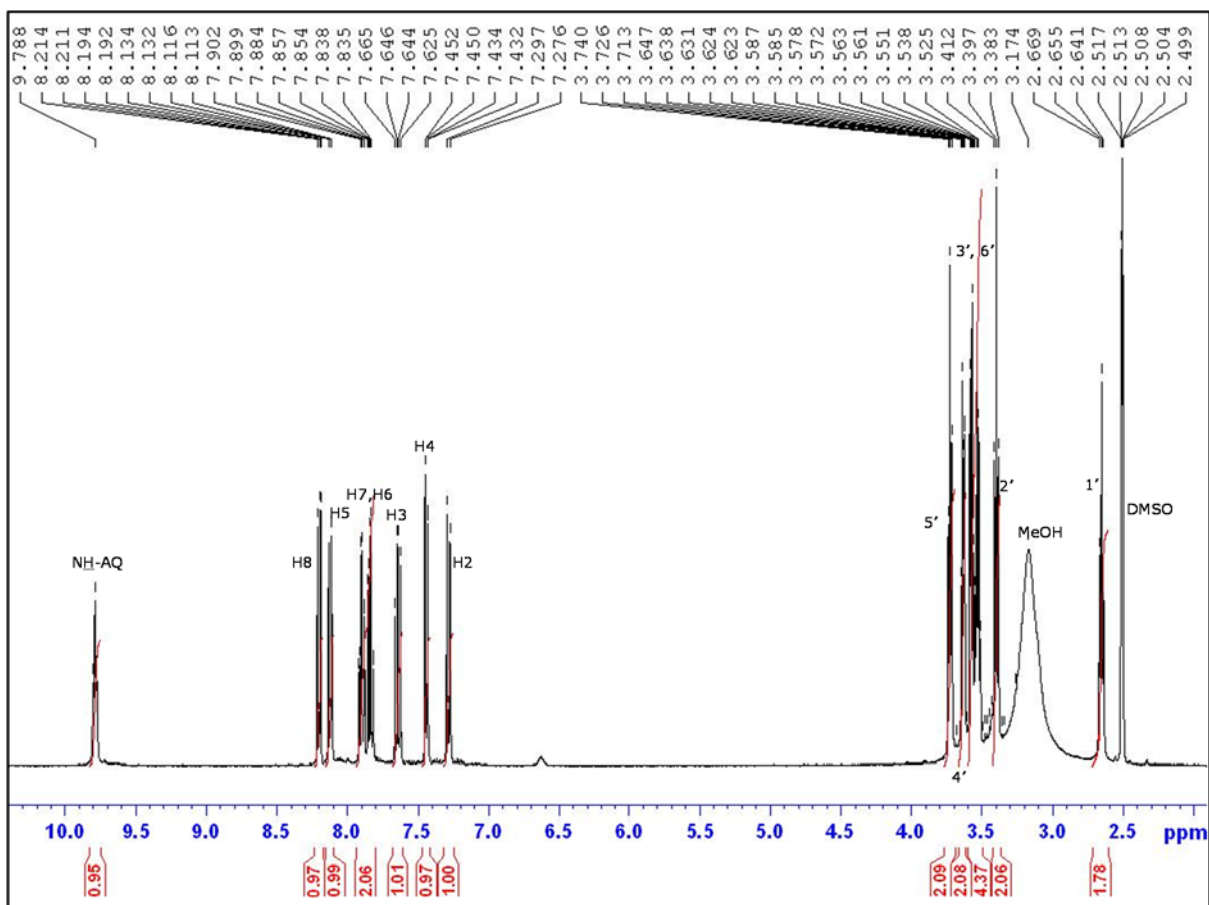
Figure: 36 ¹H-NMR in d-DMSO was recorded for compound **5** on a Bruker NMR spectrometer operating at ¹H frequency of 400MHz.

2.1.1.3. Step 3, 4 and 5: Synthesis of FRET tetra peptide substrates (1ˢᵗ batch-JC series)

Upon successful synthesis, isolation and characterization of the donor and the acceptor molecule, the next step was the chemical synthesis of the peptide substrate *via* SPPS (scheme 3). The rationale for the selection of amino acids for the synthesis of the peptide substrates for legumain was based on the information available from the literature (Mathieu *et al*. 2002; Dall and Brandstetter, 2015). From experiments conducted by Mathieu and colleagues (2002) on substrate specificity for legumain it was demonstrated that the enzyme showed strict specificity towards asparagine at P1 position of the peptide chain. In the same study the authors demonstrated that amino acids such as Thr, Ala, Pro, Val and Gly were desirable at P2 and P3 positions and on the prime position amino acids with small side chains (Ala, Gly, Ser, Cys) were well tolerated, however proline at this position was found to be unfavorable. Moreover it was also demonstrated that except for the most conserved feature of Asn at P1 position, the rest of the amino acid positions varied between species. This was proven by the ideal cleavage

sequence for *schistosoma* legumain to be Thr-Ala-Asn (P3-P2-P1) whereas the human legumain preferred Pro-Thr-Asn (Dall and Brandstetter 2015).
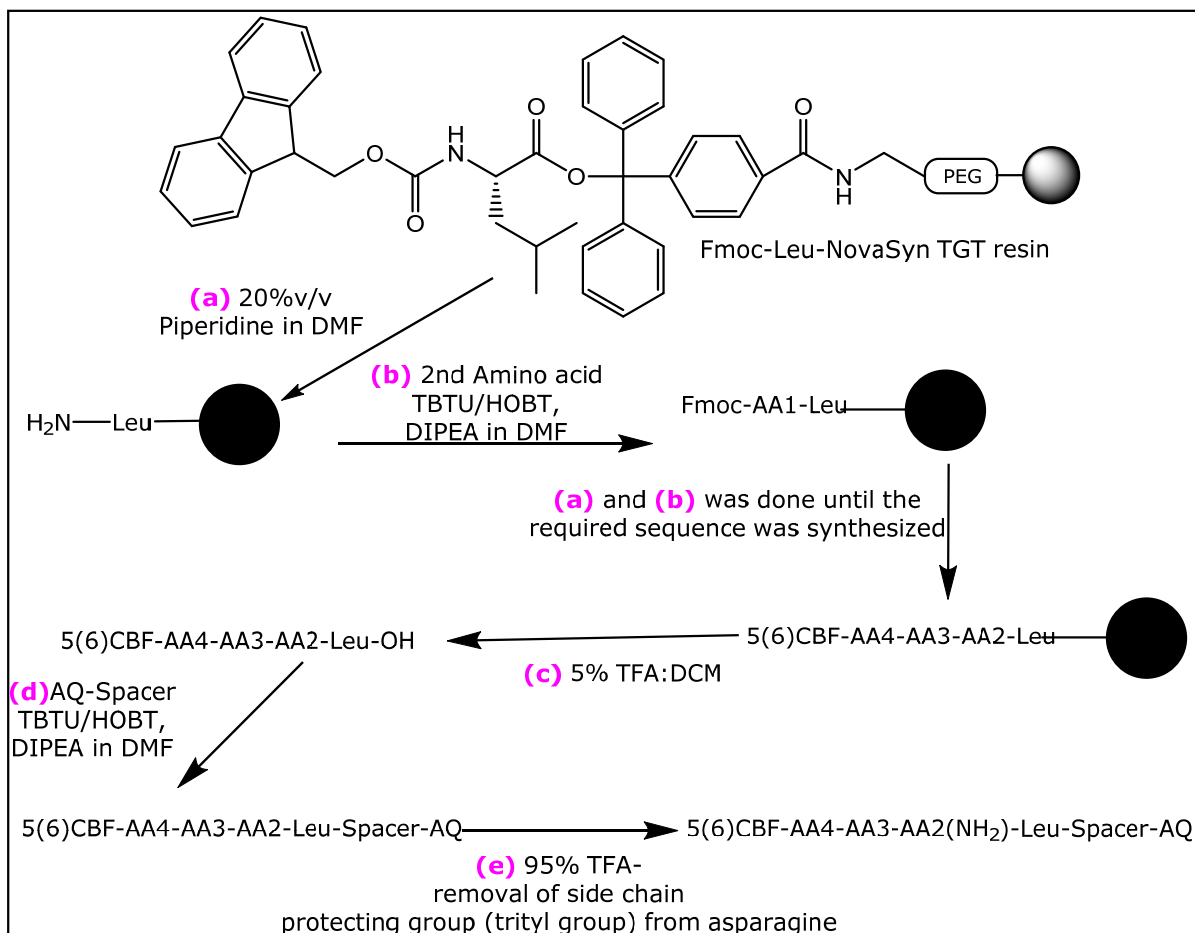
Scheme 3: Steps involved in synthesis of JC series-1.

Information from the *MEROPS* database on the available legumain activity probes suggest that P2 position was well tolerated by amino acids such as threonine, alanine, glycine, leucine and serine (Rawlings *et al*. 2012). Considering these findings, the FRET peptide substrates for legumain were structured spanning the P3 to P1′ positions (table 7, pg. 62).

The FRET tetra peptide substrates were synthesized according to a stepwise process, which included:

I. Synthesis of the tetra-peptide chain was done by sequential addition of amino acids onto a leucine pre-loaded resin for solid phase peptide synthesis. The pre-loaded resin was reacted in cycles of coupling and deprotection with commercially available amino acids (scheme 3a,b). The *N*-Hydroxybenzotriazole (HOBt) esters are commonly used during peptide synthesis mainly because of their excellent racemization suppressing effect and coupling acceleration properties (Gutte 1995). For the final coupling a fluorescein labelled amino acid (5(6)-CBF-Pro-OH (**2**)/ 5(6)-CBF-βAla-OH

(**4**)) was conjugated to the *N*-terminal of the tri-peptide sequence. The fluorescein tagged tetra-peptide sequence was then removed from the resin using TFA in DCM (scheme 3c). The product was then characterized by high-resolution mass spectrometry and the purity was determined using HPLC.

II. The next step involved the coupling of the fluorescein labelled tetra-peptide sequence to the aminoanthraquinone conjugate using the coupling agents TBTU/HOBT and DIPEA (scheme 3d). The aminoanthraquinone conjugate is expected to quench the fluorescence emission from the 5(6)-carboxyfluorescein in the absence of the target enzyme (legumain). The crude product was purified by column chromatography and then characterized by high-resolution mass spectrometry (ESI-HR).

III. The final step involved removal of the γ-trityl protecting group from the asparagine residue. The rational for the side chain protection of the primary amide in the asparagine residue was to avoid any unwanted side reactions during the process of coupling and deprotection. Moreover unprotected asparagine had poor solubility which could delay reaction progress. The *N*-trityl protection of the side chain in asparagine could prevent the dehydration of the amide structure to a nitrile group, thus keeping the asparagine residue intact and recognizable by legumain. This is important because of the strict specificity of legumain towards asparagine at the P1 position of the peptide substrate. The trityl protecting group was used because of its stability during the cycles of coupling and deprotection and finally it can be removed by 95%TFA in DCM (Ding, 2014). This should allow the peptide substrate to be recognized by legumain leading to hydrolysis on the carboxyl end of asparagine. This hydrolysis of the peptide substrate leads to an increase in distance between the FRET pair (amino anthraquinone/5(6)-carboxyfluorescein) and the release in fluorescence can be measured fluorimetrically, which was the objective for synthesizing the FRET peptide substrate (scheme 3e).

I. Synthesis of 5(6)-CBF-Pro-Ala-Asn-Leu-Spacer-AQ (Compound 8)
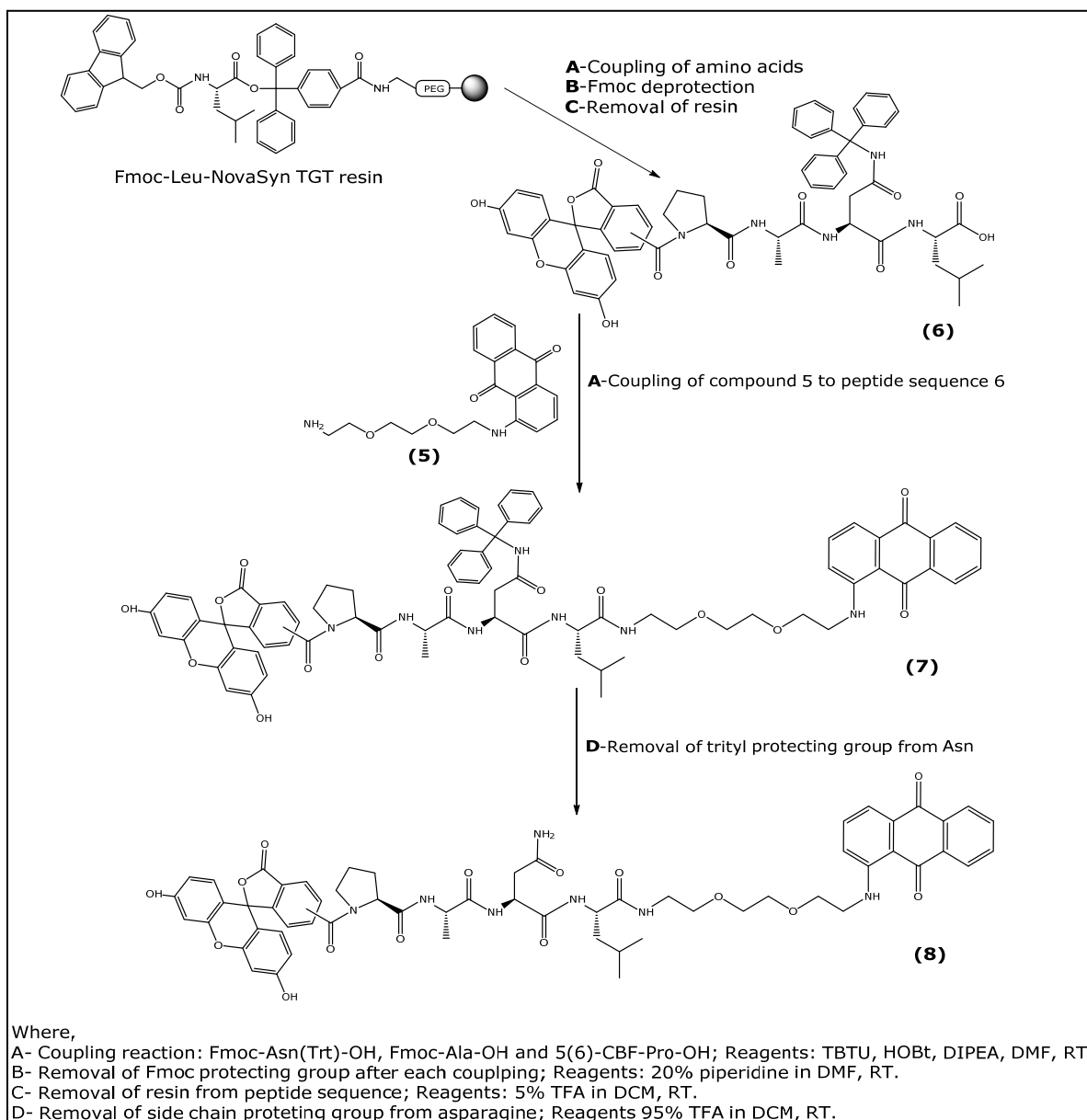
Scheme 4: Steps involved in the synthesis of compound **8**.

A leucine pre-loaded resin was reacted in cycles of coupling and deprotection with Fmoc protected amino acids in the following order: *N*-Fmoc-Asn(Trt)-OH, *N*-Fmoc-Ala-OH followed by coupling of the fluorescein labeled amino acid (**2**). Reagents and reaction conditions are as mentioned in scheme 4. Leucine (leu) at P1′ position of the peptide sequence was chosen based on the fact that most of the tumor related prodrugs targeting legumain expression (Liu 2009) were found to be tolerated by the enzyme. After the final coupling, the peptide sequence was removed from the resin with trifluoroacetic acid (5%) in DCM.
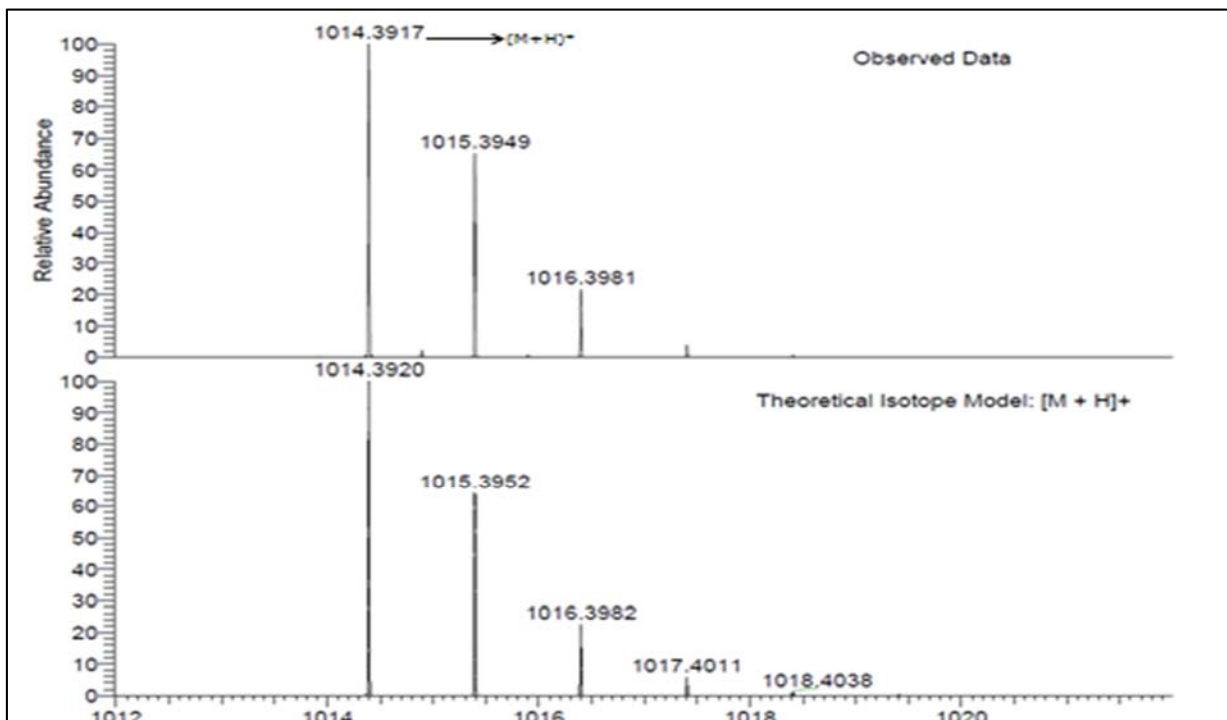
Figure 37: Expansion of the ESI (+) mass spectrum of compound **6** that shows the isotopic pattern of the molecular ion.

HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 44%. The rest of the signals were accounted for as belonging to reagents and other unidentified mass. Product **6** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1014.3917 $(M+H)^+$ corresponding to a molecular mass of 1013.38 (figure 37).

The FRET peptide sequence **8** was synthesized by reacting compound **6** with compound **5** in DMF (scheme 4A) followed by liquid/liquid extraction and chromatographic purification as detailed in the previous section 2.1.1.3 (II). The yield of the product obtained after column chromatography was 53%. The resulting compound (**7**) was characterized by ESMS (+), which gave a signal at m/z 1350.4 $(M+H)^+$ corresponding to a molecular mass of 1349 and also a signal for the doubly charged ion at m/z 675.8 $[(M+2H)/2]^{2+}$. Removal of the γ-trityl protecting group from the asparagine residue was achieved by treating compound **7** with TFA as detailed in section 2.1.1.3 (III) (scheme 4D). Deprotection progress was monitored by TLC analysis, which showed the presence of a new single component, which gave a low R_f value indicating the polar nature of the deprotected compound (**8**).

II. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Leu-Spacer-AQ (Compound 11)

(11)
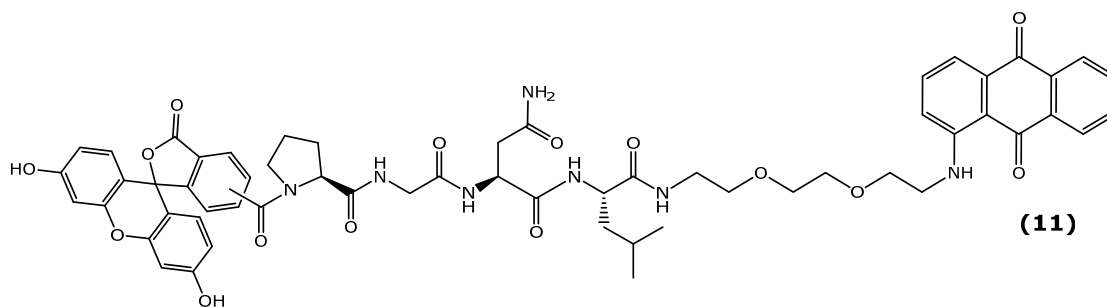
The FRET peptide sequence **11** was synthesized consisting of 5(6)-CBF at the *N*-terminus of the tetra peptide and an aminoanthraquinone-spacer acceptor unit at the C-terminus of the sequence. The position of asparagine and proline were preserved from the previous sequence whereas alanine at P2 was replaced with glycine. This decision was based on the fact that glycine was also tolerated at the P2 position of the tetra peptide sequence (Mathieu *et al*. 2002). Moreover the current database on legumain active probes demonstrates that five of the probes had a glycine residue at the P2 position of the peptide sequence (Rawlings *et al*. 2012). The synthesise of the target FRET peptide substrate **11** required the removal of the trityl protecting group from the side chain of Asn in order to be recognized by the activated legumain. The protecting group was removed by treating the compound with TFA (scheme 4b) to yield the deprotected compound **11** as detailed in section 2.1.1.3 (III).

III. Synthesis of 5(6)-CBF-Pro-Leu-Asn-Leu-Spacer-AQ (Compound 14)

(14)
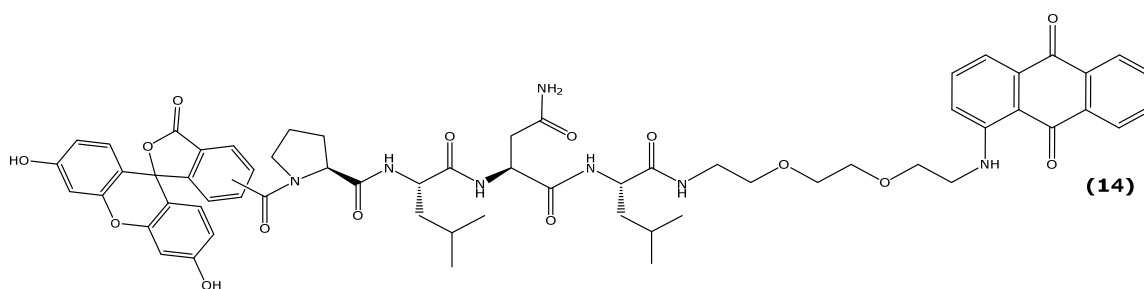
Synthesis of the FRET peptide substrate **14** followed the same protocol as compound **8**. The positing of the amino acids from P3-P1′ including the FRET pair remained the same except for the amino acid at the P2 position of the sequence. Leucine (leu) was placed at P2 rather than ala or gly as in compounds **8** and **11**. The choice of leu at that position was based on the information from the *MEROPS* database which demonstrates that three of the legumain peptide substrates had leu at this position. However, the coupling and deprotection methods remained the

same as detailed in scheme 4. Though TLC analysis indicated that the reaction had gone to completion by the presence of a new component close to the base line indicating the polar nature of the compound **14**, it also showed the presence of a fluorescein band lying close to the red spot and hence it was decided to purify the sample by silica gel column chromatography. Various solvent systems were tried (ethyl acetate: hexane: MeOH-3:1:1; ethyl acetate: hexane- 4:1; ethyl acetate: DCM: isopropanol- 8:1.5:0.5; varying gradients of MeOH-3% up to 10% in DCM) to obtain a separation between the two bands and it was found that 10% MeOH in DCM was able to produce a separation between the front running red band presumed to be the deprotected product (due to the presence of the anthraquinone conjugation) and the fluorescein band. The sample was dissolved in 10% MeOH in DCM and loaded onto the silica gel column prepared using the same solvent mixture. However, the fluorescein band was found to overlap with the deprotected sample (red band) during the course of elution. Two fractions were eluted from this column with the first fraction being eluted with 10% MeOH in DCM and the other fraction was eluted with a higher percentage of MeOH in 5% increments (up to 30%). Upon TLC analysis of the first eluted fraction it still appeared to be a binary mixture of two close running spots (a reddish band and a fluorescein band), which could not be resolved by column chromatography. Mass spectrometric analysis on the first eluted fraction showed the presence of many unresolved signals. Hence it was decided to repeat the peptide synthesis applying solution phase peptide synthesis, which involved purification after each coupling and deprotection cycle in the hope to isolate the desired product eliminating any undesired reagents after each coupling and deprotection cycles. Moreover, optimizing the equivalence of fluorophore to be coupled to the peptide sequence may also reduce the presence of excess/unreacted dye.

IV. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Leu-Spacer-AQ (Compund 17)
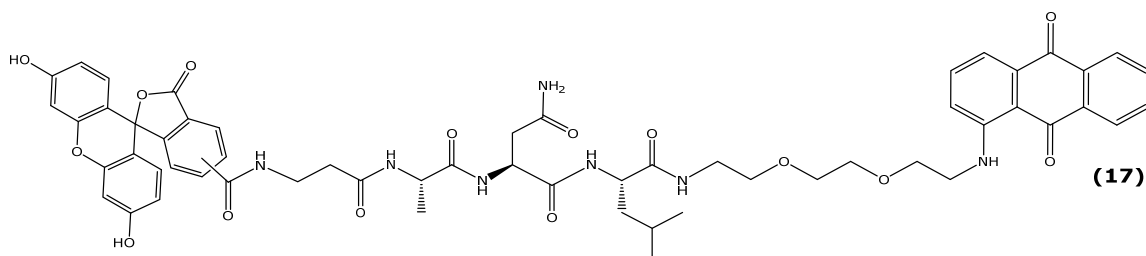
The FRET tetra peptide sequence (**17**) was synthesized following a procedure analogous to compound **8**. The amino acids were coupled in the subsequent order: *N*-Fmoc-Asn(Trt)-OH, *N*-Fmoc-Ala-OH followed by coupling of the fluorescein

labeled amino acid (**4**). The position of asparagine residue was maintained at P1 and the positon of alanine at P2 was also similar to sequence **8**; however proline at P3 positon was replaced by β-alanine based on the information on legumain substrate specificity as discussed in section 2.1.1.3. TLC analysis of compound **17** showed a binary mixture of two close running components with one corresponding to the deprotected sample perceived by the reddish brown stain due to the presence of the anthraquinone moiety and a bright yellow fluorescein spot which may be due to the presence of unreacted fluorescein residue, which upon column chromatography separation was not possible to be isolated.

V. 5(6)-CBF-β-Ala-Leu-Asn-Leu-Spacer-AQ (Compound 20)
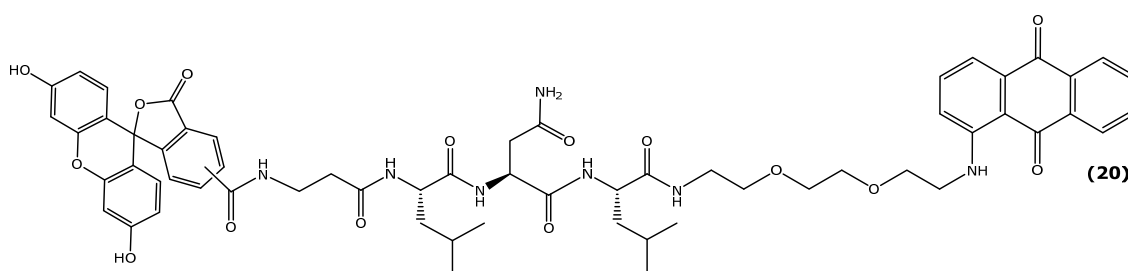
Compound **20** was synthesized following a procedure analogous to compound **8**. The amino acids were coupled in the subsequent order: *N*-Fmoc-Asn(Trt)-OH, *N*-Fmoc-Leu-OH followed by coupling of the fluorescein labeled amino acid (**4**). However, the same problem as encountered with compound **15** and **17** was observed after deprotecting the sample with TFA. Hence it was decided to synthesize the next FRET peptide substrate *via* solution peptide synthesis in an attempt to investigate the purity, yield and homogeneity of the intermediates and target compound.

VI. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Leu-Spacer-AQ (Compound 25)

The aminoanthraquinone conjugate was used as a support for growing the amino acid sequence. The coupling of subsequent *N*-protected amino acids was performed in DMF along with activating/coupling agents such as TBTU and HOBT (scheme 5) in the presence of a base (DIPEA). The amino acids used for the peptide synthesis were *N*-protected by the Fmoc group except for asparagine, which also had a γ-trityl protecting group along with the Fmoc group at the *N*-α terminus. The Fmoc group was removed using 20%v/v piperidine in DMF and the general mechanism for Fmoc removal is detailed in figure 38.
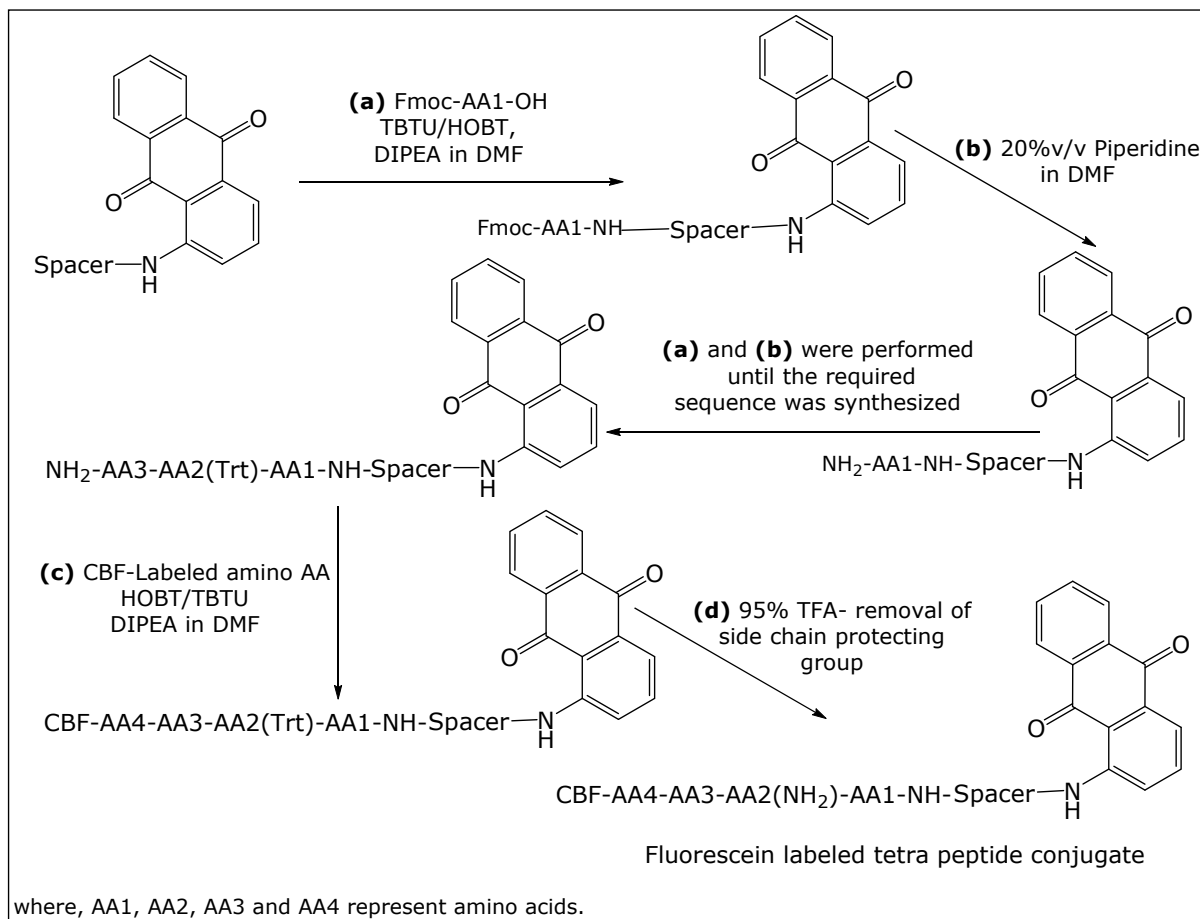
Scheme 5: Synthesis of fluorescein labelled tetra-peptide compound *via* solution phase peptide synthesis.

Following each coupling and deprotection reaction, the intermediates were purified by column chromatography and the isolated samples were characterized by low-resolution and/or high-resolution mass spectrometry and some of the intermediates were also characterized *via* NMR. Upon successful synthesis of the FRET peptide sequence the γ-trityl protecting group from asparagine residue was removed using TFA (scheme 5). Some of the advantages of solution peptide synthesis are the isolation and characterization of the intermediates obtained after each coupling and deprotection reaction and hence every reaction cycle can be easily controlled (Chandrudu *et al*. 2013). Purification at each step is done by chromatography using a silica gel column which provided the intermediates in reasonable yield and purity (Dunn 2015). This is important because no matter how carefully the SPPS was done to avoid deletion compounds due to incomplete coupling, there are still possibilities for the presence of such molecules (Doonan 2002). Moreover this technique is well suited for the synthesis of small peptides composed of few amino acid residues and is highly flexible with respect to the chemistry of coupling and the combination of amino acids (Guzman *et al*. 2007).
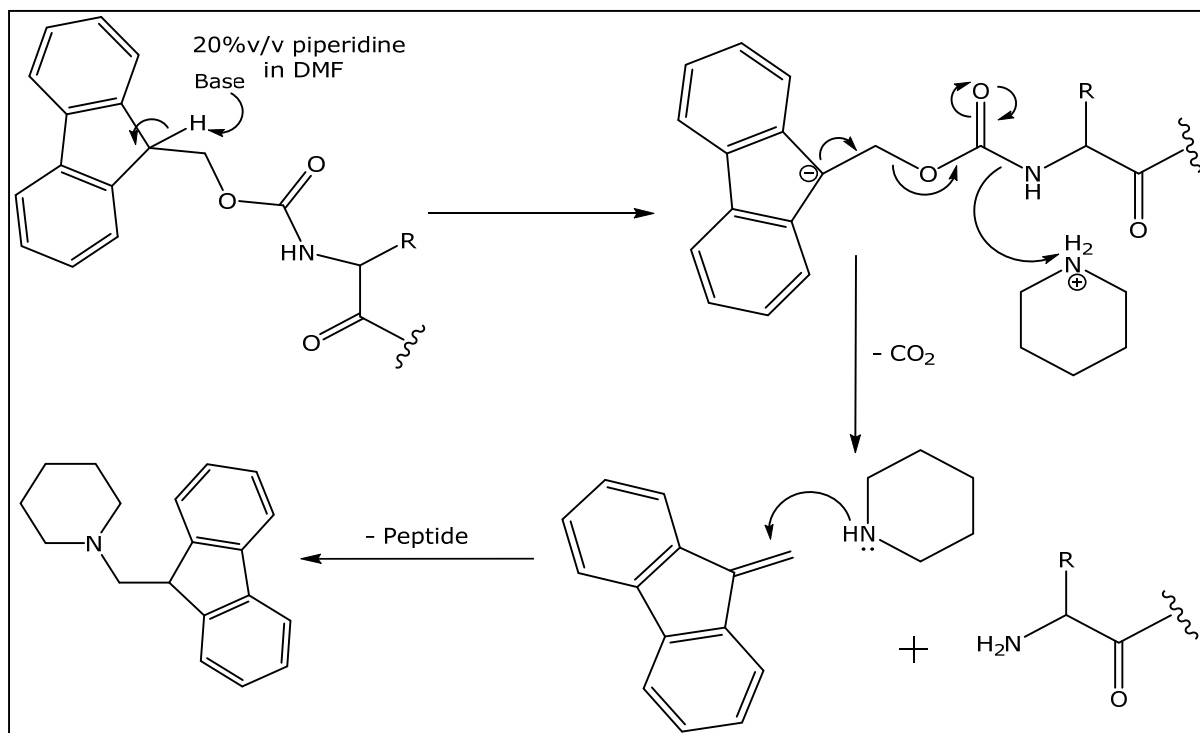
Figure 38: General mechanism for Fmoc group deprotection.

Another advantage of this technique over SPPS is the reduced cost for synthesizing peptides. In this method amino acids, reagents, derivatives and coupling agents are generally used in a ratio of 1:1 and not in excess as in SPPS. Additionally other limitations in SPPS such as incomplete coupling and deprotection reactions, buildup of byproducts and finally because the peptide chain remains attached to the resin, it meant that if any coupling was incomplete then the final compound will have a population of molecules with missing amino acid residues. And the process of removing the deletion product is quite difficult (Chandrudu *et al.* 2013). The following sub-sections demonstrate in detail the synthesis of FRET peptide substrate **25**.

A. Synthesis of intermediate compound H₂N-Leu-Spacer-AQ (Compound 21)

(21)
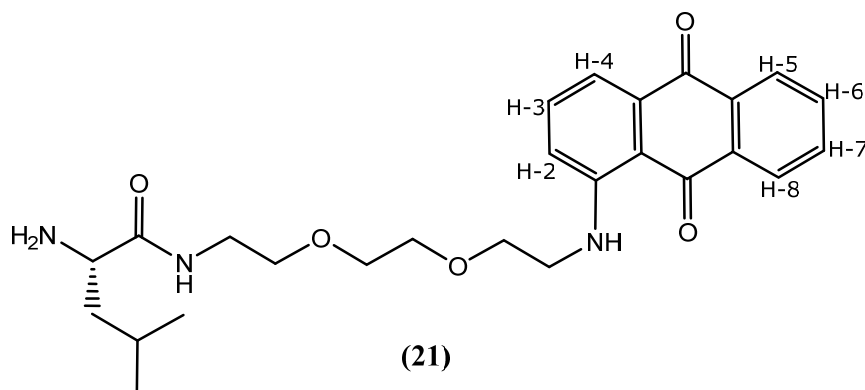
The Fmoc-protected intermediate (Fmoc-Leu-Spacer-AQ) was synthesized by reacting *N*-Fmoc-L-Leucine with compound **5** (Spacer-AQ) in DMF. The reaction conditions were as mentioned in scheme 6. The carboxyl group of the *N*-protected amino acid was activated *via* an esterification reaction with TBTU/HOBt in the presence of the base (DIPEA) prior to addition to the aminoanthraquinone conjugate (**5**). Reaction progress was monitored by TLC, which showed the presence of a single new component different from the starting material. The Fmoc protecting group was then removed using piperidine in DMF followed by purification by column chromatography to produce compound **21**. The yield of the compound after column chromatography was 81%. The purity of the sample was found to be 98% *via* analytical HPLC performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column. The product **21** was then characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 468.2488 $(M+H)^+$ corresponding to a molecular mass of 467. Furthermore, the ^{1}H NMR spectrum (d-DMSO) showed, for example, a multiplet signal for the 6 protons at the δ carbon of the leucine side-chain between 0.80 and 0.85ppm. Multiplet signal for the 2 protons at the β carbon was found between 1.14 and 1.21ppm and 1.33ppm and 1.40ppm. Another multiplet for the 1 proton at the γ carbon was found between 1.63 and 1.70ppm. A quartet signal at 3.11ppm was assigned to the signal proton at the α carbon of leucine. A two proton multiplet signal between 3.18 and 3.30ppm and another two proton triplet at 3.46ppm was assigned to C$\underline{H}_2$-NH-COO and to C$\underline{H}_2$-CH$_2$-NH-COO protons. A two proton quartet at 3.53ppm was assigned to C$\underline{H}_2$-NH-AQ protons. A multiplet signal between 3.57ppm and 3.64ppm was assigned to O-C$\underline{H}_2$-C$\underline{H}_2$-O protons of the spacer arm. A two proton triplet at 3.73ppm was assigned to C$\underline{H}_2$-CH$_2$-NH-AQ protons. The protons attached to the aromatic group were assigned as follows: H-2 and H-4 gave a single proton double doublet at 7.29ppm and 7.45ppm and a single proton multiplet between 7.63ppm and 7.67ppm was assigned to H-3. A multiplet between 7.82ppm and 7.94ppm for three protons was assigned to H-6 and H-7 and the amide proton of leucine. The H-5 and H-8 protons gave a double doublet signal at 8.12ppm and 8.20ppm respectively. The anthraquinone amino group proton gave a triplet at 9.79ppm. To further confirm the assignment in the ^{1}H NMR, a 2D-COSY spectrum of compound **21** was also done (appendix 2).

B. Synthesis of intermediate compound H₂N-Asn(Trt)-Leu-Spacer-AQ (Compound 22)

(22)
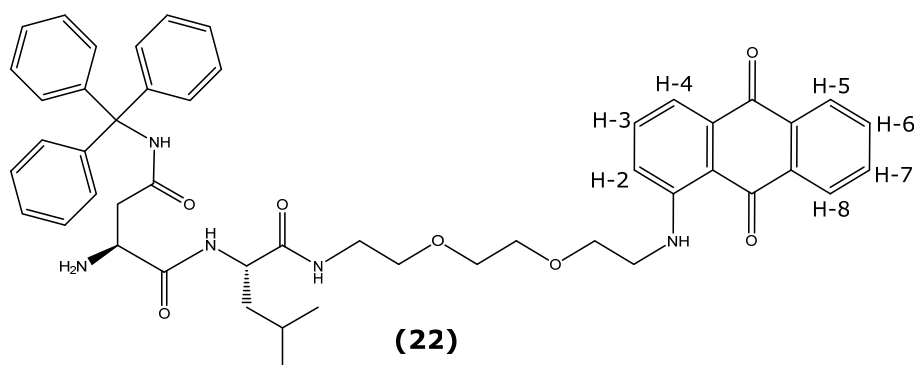
Compound **22** was synthesized following a procedure analogous to compound **21**. The yield of the sample after column chromatography was 94% and the purity of the sample was determined using analytical HPLC performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column which was found to be 91% (figure 42). The product **22** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 824.4013 (M+H)⁺ corresponding to a molecular mass of 823.4013 (figure 39); the ESMS (-) mass spectrometry displayed a signal at m/z 822 corresponding to the proton abstraction product (M-H)⁻ .
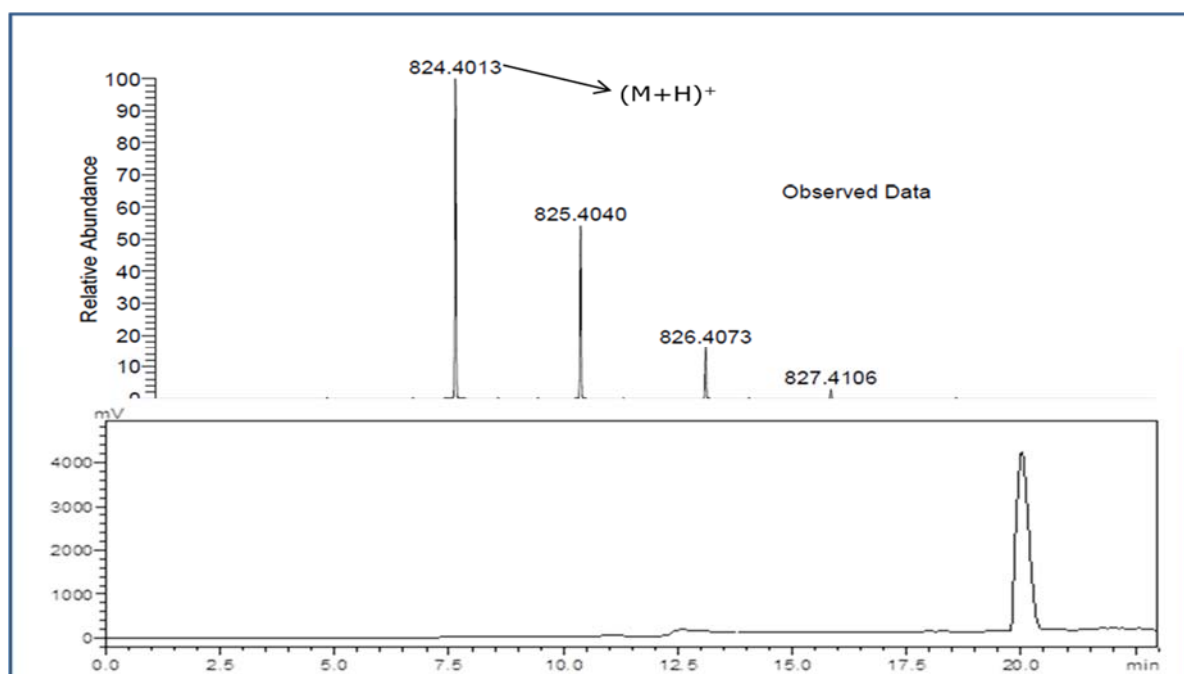
Figure 39: Mass spectrometry result for compound **22**. RP-HPLC elution profile of compound **22** (1mg/mL in methanol); Conditions: Flow rate, 1mL/min; Detection, 247nm; stationary phase: Eclipse XDB-C18 RP column; mobile phase A= 0.1%TFA/water; mobile phase B=0.1%TFA/methanol; gradient: low pressure gradient; 23min.

Furthermore, the structure of the compound **22** was confirmed by ^{1}H NMR spectroscopy (d-DMSO), which showed, for example, a 6 proton multiplet between 0.80ppm and 0.85ppm was assigned to the two methyl group protons in leucine. A two proton multiplet signal between 1.37ppm and 1.48ppm was assigned to the protons at the β carbon of the leucine side chain. Another multiplet signal for the single proton at the γ carbon was found between 1.51ppm and 1.61ppm. A 2 proton multiplet between 2.38 and 2.46ppm was assigned to the methylene group of the asparagine side chain. A two proton multiplet signal between 3.07 and 3.18ppm was assigned to C$\underline{H}_2$-NH-Leu protons. Another signal for two protons between 3.34 and 3.41ppm was assigned to C$\underline{H}_2$-CH$_2$-NH-Leu protons was seen as an overlap with the water signal which was further confirmed using 2D-analysis (see appendix 3). A multiplet for five protons between 3.45 and 3.53ppm was assigned to protons of C$\underline{H}_2$-NH-AQ, C$\underline{H}$-CH2-CONH-trityl protecting group and C$\underline{H}_2$-O-CH$_2$-CH$_2$-NH-AQ respectively. Another two proton multiplet between 3.58 and 3.61ppm was assigned to Leu-NH-CH$_2$-CH$_2$-O-C$\underline{H}_2$ protons. A two proton triplet signal at 3.71ppm was assigned to C$\underline{H}_2$-CH$_2$-NH-AQ. A quartet at 4.26ppm was assigned to the single proton at the α carbon of leucine. Signals for a 1 proton and a 15 proton multiplet between 7.17 and 7.30ppm were assigned to the H-2 proton of the aromatic ring along with the protons of the phenyl rings of the trityl protecting group. The protons attached to the aromatic group were assigned as follows H-4 gave a single proton double doublet at 7.45ppm and a multiplet between 7.62ppm and 7.66ppm was assigned to H-3. A single proton triplet of doublets at 7.83ppm was assigned to H6 followed by another single proton triplet of doublets at 7.89ppm was assigned to H7 proton. A multiplet for two protons between 7.95ppm and 8.02ppm was assigned to N$\underline{H}$-spacer arm proton and the amide proton of leucine. The H-5 and H-8 protons gave a double doublet signal at 8.13ppm and 8.20ppm respectively. A singlet at 9.23ppm was assigned to N$\underline{H}$-Trityl protecting group proton. Signal for a one proton triplet at 9.80ppm was assigned to the anthraquinone amino group proton. To further confirm the assignment in the ^{1}H NMR, a 2D-COSY spectrum of compound **22** was also done (appendix 3).

C. Synthesis of intermediate compound H₂N-Gly-Asn(Trt)-Leu-Spacer-AQ (Compound 23)

(23)
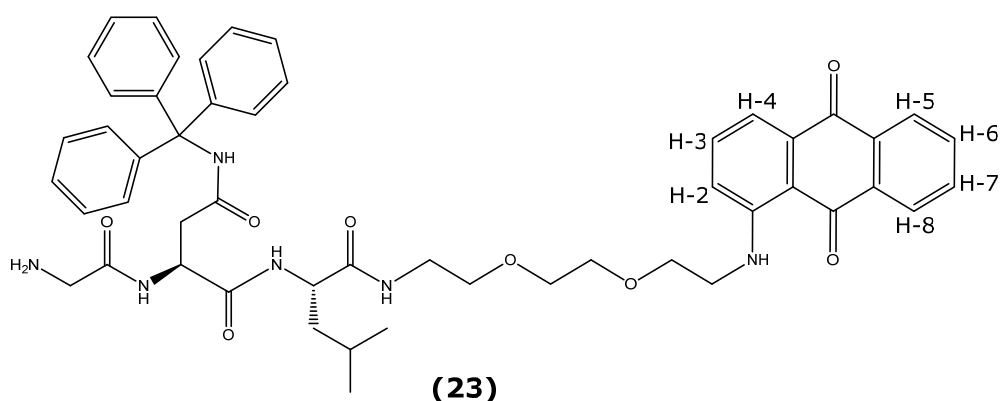
Compound **23** was synthesized following a procedure analogous to compound **21**. The yield of the sample was found to be 82% post column chromatography. The purity of the sample was determined using HPLC performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column which was found to be 88%. The product **23** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 881.4232 (M+H)⁺ corresponding to a molecular mass of 880; the ESMS (-) low resolution mass spectrometry gave a signal at m/z 879 corresponding to the proton abstraction product (M-H)⁻ and m/z 915 corresponding to the chloride adduct (M+Cl)⁻. Furthermore, the structure of the compound **23** was confirmed by ¹H NMR spectroscopy (d-DMSO), which showed, for example, a 6 proton double doublet at 0.82ppm which was assigned to the two methyl group protons in the side chain of leucine. A multiplet signal for the 3 protons at the β and γ carbons in leucine was found between 1.38 and 1.60ppm. A 2 proton multiplet between 2.65 and 2.76ppm was assigned to the methylene group protons of the asparagine side chain. A 2 proton multiplet signal between 2.97 and 3.10ppm was assigned to C$\underline{H}_2$-NH-Leu and a singlet for two protons at 3.13 was assigned to C$\underline{H}_2$-NH₂. A two proton triplet at 3.26ppm was assigned to C$\underline{H}_2$-CH₂-NH-Leu protons. A multiplet signal for six protons between 3.47ppm and 3.59ppm was assigned to O-C$\underline{H}_2$-C$\underline{H}_2$–O and C$\underline{H}_2$-NH-AQ protons. A triplet at 3.71ppm was assigned to the two protons of C$\underline{H}_2$-CH₂-NH-AQ protons. A multiplet between 4.17 and 4.23ppm was assigned for a single proton at the α carbon of leucine. A broad signal at 4.56ppm for a single proton was assigned the methine group proton of asparagine. Signals for a 1 proton and a 15 proton multiplet between 7.18ppm and 7.29ppm was assigned to the H-2 proton of the aromatic ring along with the protons of the phenyl rings of the trityl protecting group. The protons attached to the aromatic group were assigned as

follows H-4 gave a single proton double doublet at 7.45ppm and a single proton multiplet between 7.62ppm and 7.66ppm was assigned to H-3. A multiplet for four protons between 7.81ppm and 8.03ppm was assigned to H-6 and H-7 protons along with the protons for N<u>H</u>-spacer-AQ and the amide proton of leucine. The H-5 and H-8 protons gave double doublet signals at 8.13ppm and 8.21ppm respectively. Signal for a one proton singlet at 9.22ppm was assigned to N<u>H</u>-Trityl protecting group proton followed by another one proton triplet at 9.79ppm which was assigned to the anthraquinone amino group proton.

D. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Leu-Spacer-AQ (Compound 25)
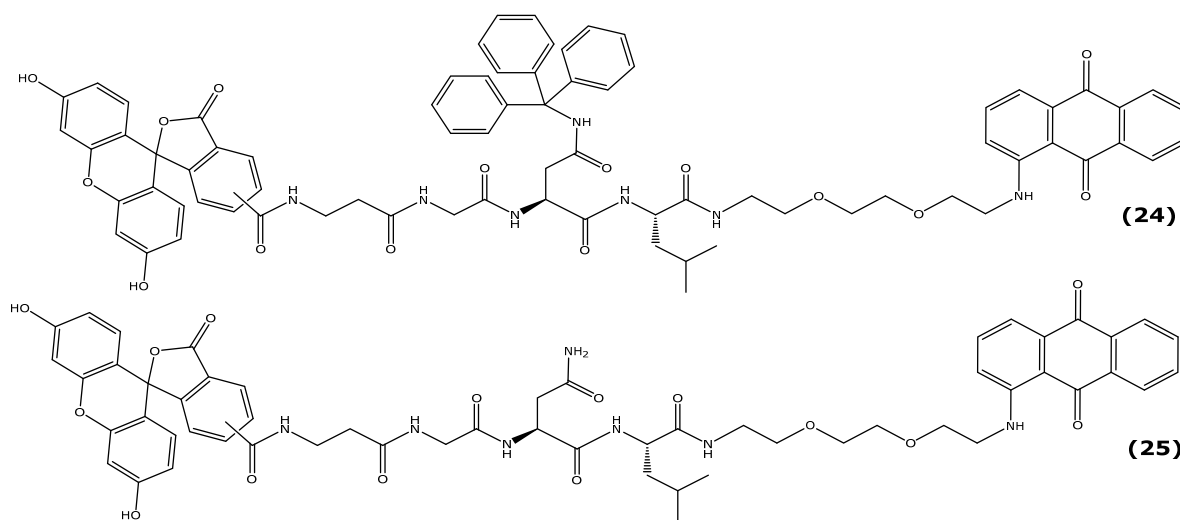
The fluorescein tagged peptide sequence **24** was synthesized by reacting compound **23** with the fluorescein labeled amino acid (**4**) in DMF. The carboxyl group of the fluorescein labeled amino acid was activated *via* an esterification reaction with TBTU/HOBT in the presence of DIPEA as base.
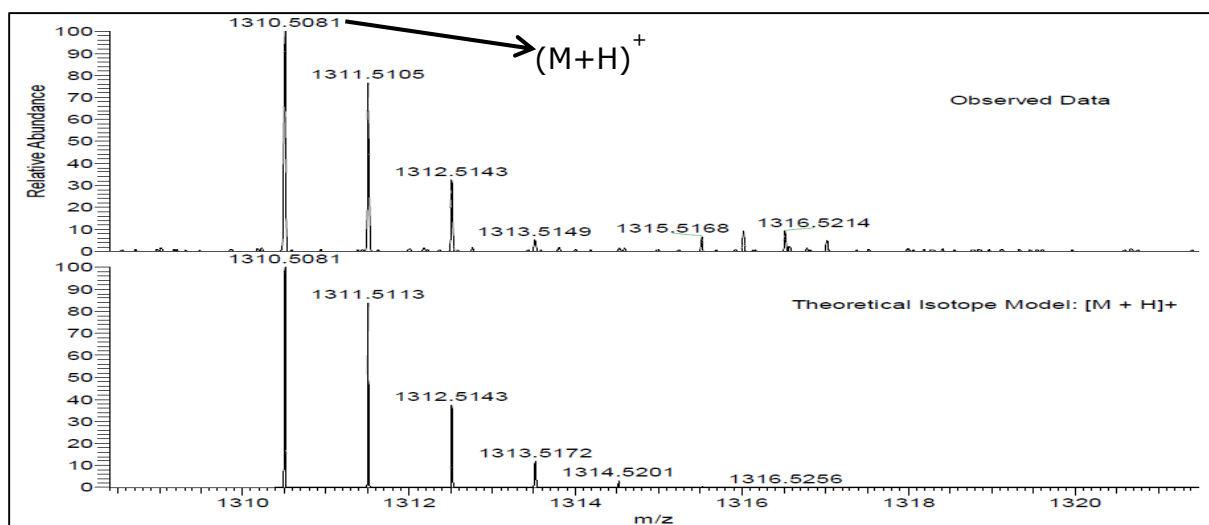
Figure 40: Mass spectrometry result for compound **24**.

The compound was purified by column chromatography using MeOH: DCM by gradually increasing the polarity of the solvent system (up to 10% MeOH) to elute the product. The yield of the sample after column chromatography was found to be 50%. The compound (**24**) was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1310.5081 (M+H)$^+$ corresponding to a molecular mass of 1309 (figure 40) and also a signal at m/z 1333 corresponding to the sodium adduct (M+Na)$^+$; the ESMS (-) mass spectrometry gave a signal at 1308 (M-H)$^-$ corresponding to the proton abstraction product and also a signal m/z 653 (M/2-2H)$^{2-}$. Subsequent γ-trityl deprotection of the asparagine residue was achieved by treating the compound **24** with TFA to yield the corresponding deprotected compound **25** as detailed in section 2.1.1.3 (step:3-III). Upon TLC analysis, the γ-deprotected tetra-peptide sequence appeared to have a close running fluorescein band. The same problem as encountered with the previous compounds (**14, 17** and **20**) which were synthesized *via* SPPS was noticed which upon further column chromatography was not possible to resolve. It was decided to repeat the peptide synthesis by coupling the fluorescent dye directly to the tetra peptide sequence in future reactions. This idea was further supported by the work done by Fernandez-Carneado and Giralt (2004) on synthesizing fluorescently labelled peptides. The authors suggest that the method of fluorescent labeling could potentially be improved if the fluorophore was directly coupled on to an assembled peptide sequence. In the same study it was also mentioned that 5(6)-CBF was demonstrated to react in a reproducible manner with amino groups producing highly fluorescent 5(6)-carboxy fluoresceinamido derivatives.

2.1.1.4. Fluorescence measurements

In this study, a library of FRET peptide substrates consisting, of an aminoanthraquinone/5(6)-CBF pair tagged to the C and N termini of a peptide sequence, designed to be recognizable by legumain have been synthesized (section 2.1.1.3 pg. 68). Legumain is one of the proteases which is found to be up regulated during atherosclerotic plaque progression. The role of this enzyme has been linked to atherosclerosis both directly by proteolytic breakdown of the ECM and indirectly by activation of other proteases that are associated with degradation of the ECM. Hence, the synthesis of suitable probes, which are designed to be specific to legumain that are found to be overexpressed in atherosclerotic plaques has been undertaken in this research project.
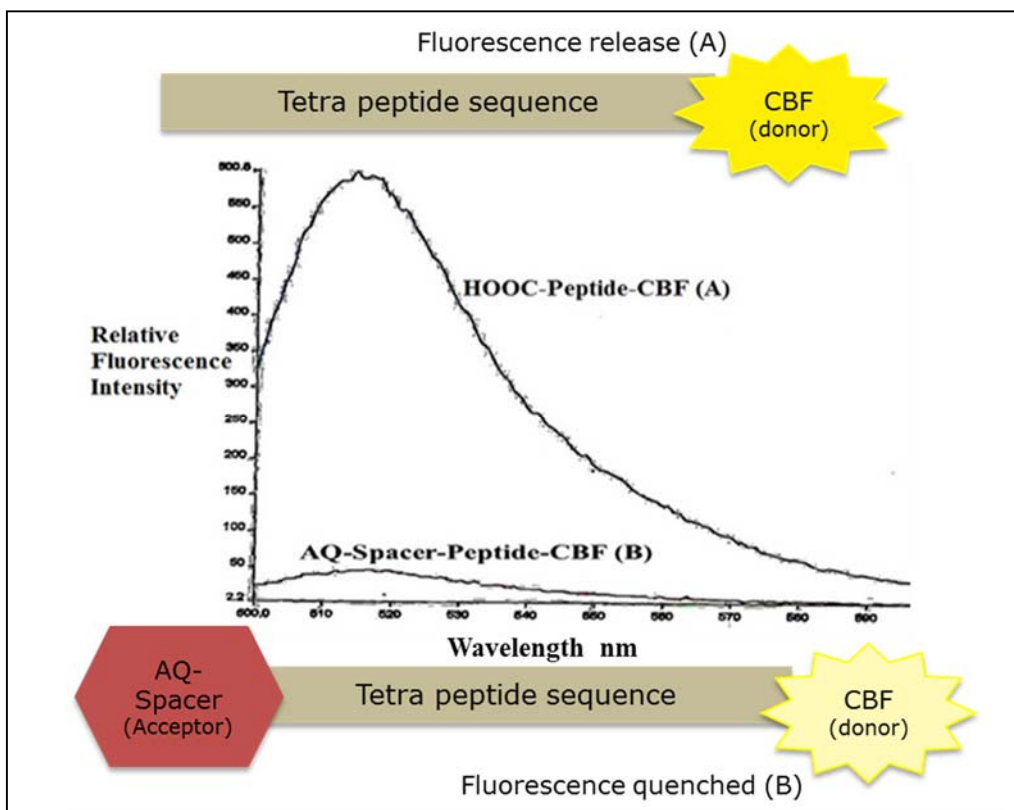
Figure 41: Pictorial explanation of FRET mediated quenching effect.

In order to evaluate the efficiency of the FRET peptide library (JC series-1) of compounds, fluorescence spectra was recorded for compound **11**. The fluorescence measurement was performed to demonstrate that the absorption spectra of the FRET compounds (with the acceptor/donor pair) would completely overlap with the emission spectrum of the fluorescein counterpart (donor) as described in figure 41. Hence the experiment was done by analyzing the fluorescence spectra of the peptide-fluorophore conjugate before and after coupling to the acceptor (quencher) moiety. The UV-Vis absorption spectrum of compound **11** (5(6)-CBF-Pro-Gly-Asn-Leu-Sp-AQ) had a maximum absorbance at 490nm and the fluorescence emission spectrum of compound **9** (5(6)-CBF-Pro-Gly-Asn-Leu-OH) showed a maximum fluorescence intensity wavelength at 515nm, emission slit width: 10nm. From the data as seen in figure 42, it can be clearly understood that the relative fluorescence intensity of the compound **11** was almost completely quenched by the acceptor unit (AQ) which can be seen by a low lying curve (4). This result suggests that there is very little or no fluorescence detected prior to incubation with the activated enzyme.

1] Water
2] 5(6)-carboxyfluorescein
3] 5(6)-carboxyfluorescein-Pro-Gly-Asn-Leu-OH (**9**).
4] 5(6)-carboxyfluorescein-Pro-Gly-Asn-Leu-OH-Spacer-AQ (**11**).
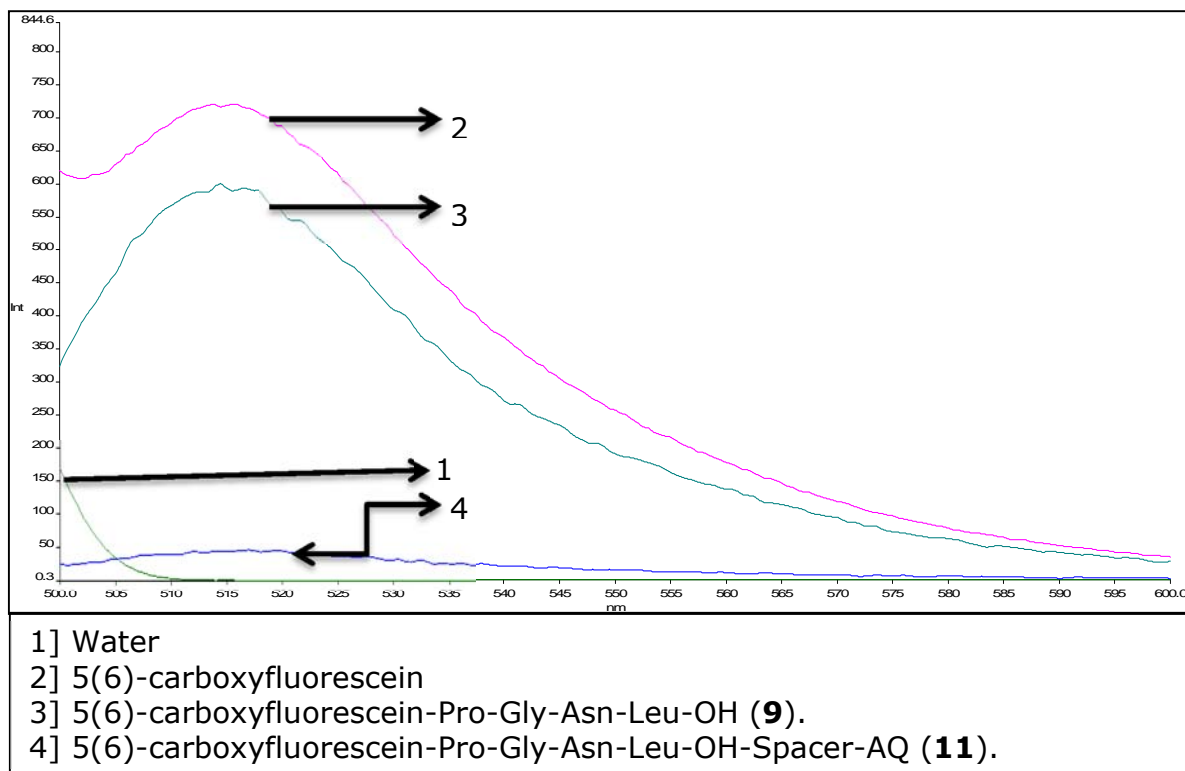
Figure 42: Comparison of relative fluorescence intensities between 5(6)-CBF-Pro-Gly-Asn-Leu-OH **9** (1μM) and the FRET substrate-compound **11** (1μM).

2.1.1.5. Legumain activity assay

The human genome comprises of 2% of the total genetic information related to members of the protease family which sum up to a total of 560 members (Edgington *et al*. 2011). One of the functions of this diverse family of proteases (serine, aspartate, metalloproteases, and cysteines) is to cleave peptide bonds that are specific to each family. Although this biological mechanism is vital to maintain normal cellular activities, it is also a serious regulatory process for several pathologies. Any unregulated proteolysis would be extremely detrimental to the stable cellular environment and hence the family of proteases is subjected to strict regulatory pathways. In order to have a clear understanding about both the normal and pathological functions of proteases, direct evaluation of the parameters for their enzymatic activities is essential (Edgington *et al*. 2011).

Initially proteases were thought to degrade proteins completely to preserve the homeostasis of proteins in the molecular system but they also perform restricted proteolysis of substrates at only certain cleavage sites. This ability to cleave a substrate at a specific site is regulated by a number of factors such as concentration of target and protease; and tertiary structure however, in numerous cases this is controlled by specific amino acid chains surrounding the scissile bond

(Edgington *et al*. 2011). Therefore, on developing this idea it may be possible to synthesize peptide sequences consisting of a fluorescein dye and a quencher tagged at opposite ends of the sequence (figure 41). These substrates on incubation with the corresponding enzyme can then be cleaved to release fluorescence. This method of enzymatic detection is known as substrate based probes. Another method involves the application of activity based probes which are designed to carry a reactive functional group (warhead) coupled to a protease recognizable sequence and a fluorogenic/radiolabelled tag for imagining purposes (Edgington *et al*. 2011). These probes function by covalently modifying the target enzyme leading to direct identification and quantification of the labeled enzyme. Moreover, contrasting the substrate based probes, these probes continue to be linked to the target enzyme enabling live cell imaging ultimately leading to dynamic analyses on enzyme activation and localization (Blum *et al*. 2005; Blum *et al*. 2007). This research study focused on the synthesis of FRET peptide substrate-based probes that could be activated by legumain.

The rationale for the design of the JC library of compounds was to observe the increase in fluorescence intensity when incubated with legumain. Excitation at 492nm and emission at 520nm based on the fluorescence properties of free 5(6)-CBF was initially subdued by the aminoanthraquinone (absorbance 200nm-650nm) conjugate in the absence of the enzyme as demonstrated in section 2.1.1.4 and this model could be predicted to generate a means for efficient substrate cleavage detection.

8	Time (min)	0	20	40	60	80	100	120
	Relative Fluorescence Intensity	298	1954	3159	3811	4177	4312	4382
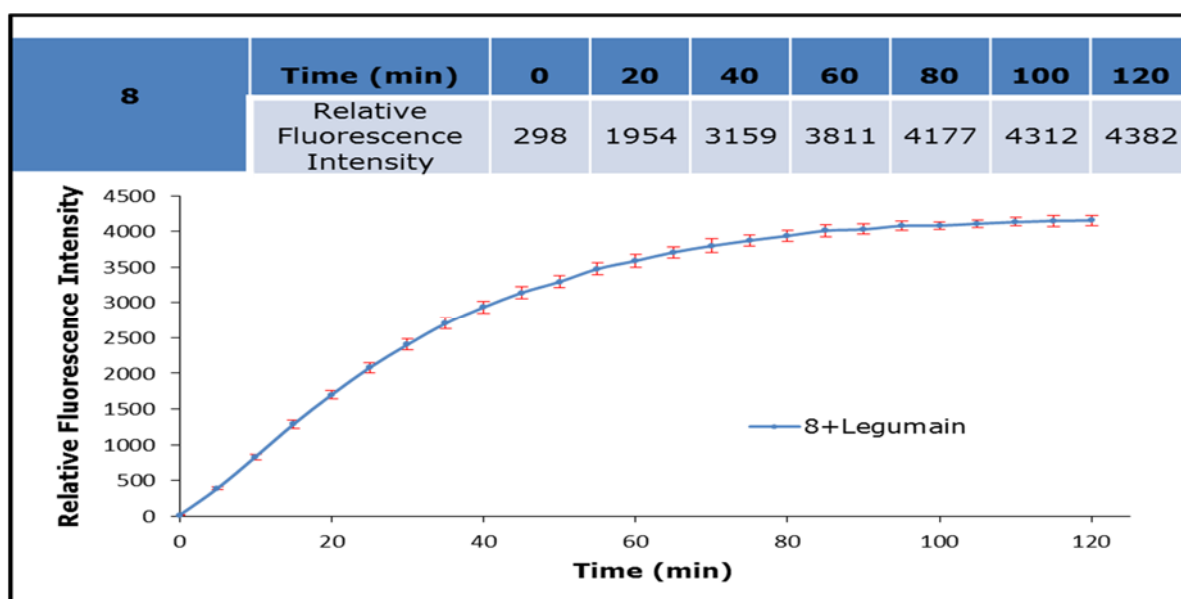
Figure 43: Relative fluorescence intensity upon incubation of probe **8** (10μM) with legumain (40ng in a final assay volume of 100μL per well) in legumain assay buffer, pH 5.0 λ_{ex} 492nm, λ_{em} 520nm. Data represents mean values ± SD from triplicates from one experiment recorded individually.

Time (min)	0	20	40	60	80	100	120
11 Relative Fluorescence Intensity	11	448	678	776	818	833	839
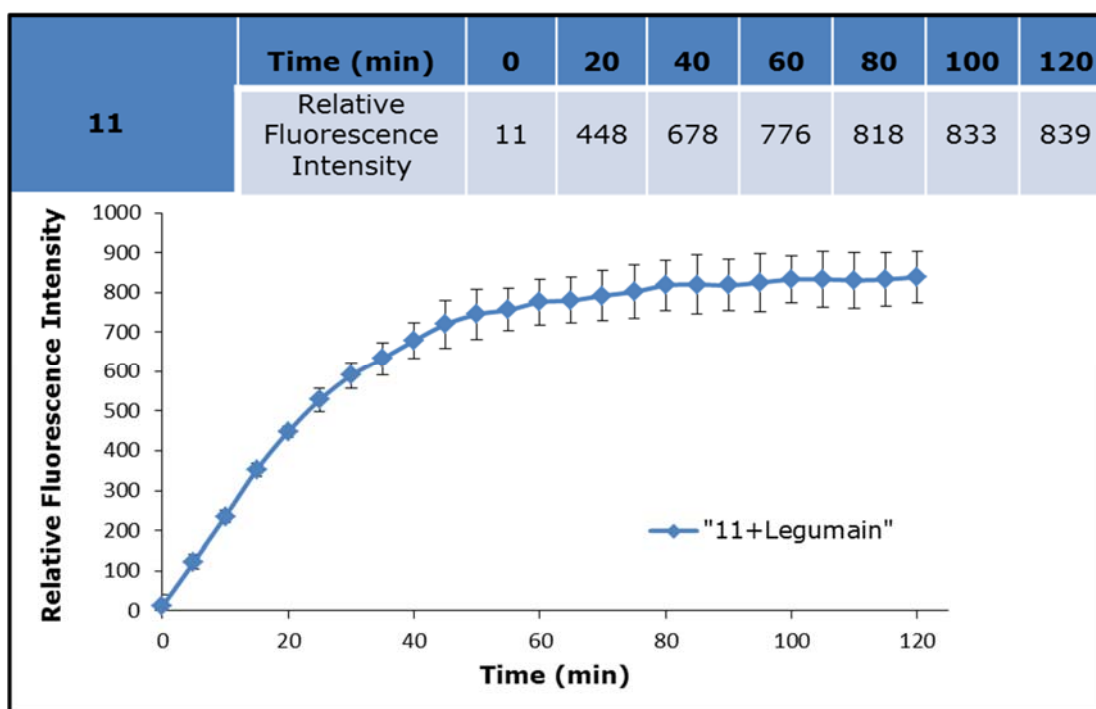
Figure 44: Relative fluorescence intensity upon incubation of probe **11** (10µM) with legumain (40ng in a final assay volume of 100µL per well) in legumain assay buffer, pH 5.0 λ_{ex} 492nm, λ_{em} 520nm. Data represents mean values ± SD from triplicates from one experiment recorded individually.
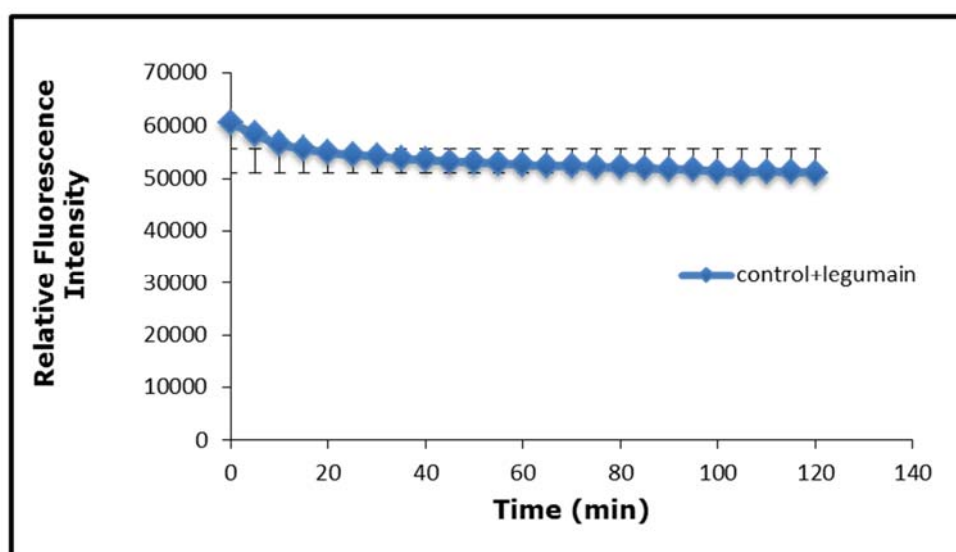
Figure 45: Relative fluorescence intensity upon incubation of control (10µM) with legumain (40ng in a final assay volume of 100µL per well) in legumain assay buffer, pH 5.0 λ_{ex} 492nm, λ_{em} 520nm. Data represents mean values ± SD from triplicates from one experiment recorded individually.

Six FRET peptide substrates were synthesized (**8**, **11**, **14**, **17** and **25** section: 2.1.1.3) and these compounds were incubated individually with activated human recombinant legumain at 37°C for two hours. Increase in fluorescence due to increase in the distance between the FRET pair (AQ/5(6)-CBF) by hydrolysis of the

substrate in the presence of activated legumain was observed only in substrates **8** and **11**. Compounds **8** and **11** showed an increase in fluorescence emission when compared to the control (figure 43, 44 and 45). The control was the corresponding peptide sequences without the anthraquinone unit (acceptor) for example compounds **6** and **9** respectively. The other four compounds (**14**, **17, 20** and **25**) did not produce the gradual increase in fluorescence as observed in compounds **8** and **11**. During the synthesis of compounds **14**, **17**, **20** and **25** it was observed that there was a close running fluorescence band near to the presumed product, which was not possible to separate using column chromatography. This could have also been one of the reasons for the background fluorescence noise even before incubating with the activated enzyme. Another observation from the legumain assay was that two of the six compounds (**8** and **11**) having proline at P3 positon were found to be readily cleaved by the enzyme which was consistent with the data that the preference for proline at this position was high compared to other amino acid residues (Sexton *et al*. 2007; Mathieu *et al.* 2002). Based on the results obtained from the legumain assay, the successful FRET peptide substrates (**8** and **11**) were used as a model for the synthesis of the 2nd batch of JC FRET peptide substrates. The 2nd batch of FRET peptide substrates were synthesized in order to eliminate the problems encountered during the synthesis of the 1st batch of JC compounds.

2.1.1.6. Synthesis of FRET peptide substrates (JC series-2)

Schwarz and colleagues (2002) in their experiment to determine the substrate preferences of legumain (isolated from pig, *schistosoma* and humans) synthesized many peptide libraries. The results from these experiments indicate that Lys/Arg/His were well tolerated at P1$^{'}$ position. The authors also demonstrated that the presence of amino acids such as pro, leu, asp and glu at P1$^{'}$ was less effective when compared to other amino acid residues at that position. Based on this study the following peptide substrates in the JC- 2nd series had lysine at P1$^{'}$ rather than leu as seen in the 1st batch of compounds. The FRET peptide substrates for legumain were structured spanning the P3 to P1$^{'}$ positions except for compound 46 which has an extra amino acid at P4 position. The most conserved feature of the FRET peptide library was the FRET pair (AQ/5(6)-CBF) which was similar to the previous series (JC series-1) along with the positioning of Asn at the P1 of the peptide sequence. The design for the 2nd batch of JC series followed the protocol

as detailed in section 2.1.1.3. This series of compounds included the selection of amino acids as shown in table 8.
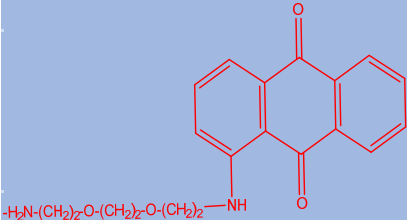
Comp. no.	Acceptor unit	P4	P3	P2	P1	Hot spot	P1'	Donor unit
31			Pro	Gly	Asn	❖	Lys	
34	5(6)-CBF		β-Ala	Gly	Asn	❖	Lys	
38			Pro	Ala	Asn	❖	Lys	
41			β-Ala	Ala	Asn	❖	Lys	
46		β-Ala	Pro	Ala	Asn	❖	Lys	

Table 8: List of tetra peptide sequence for JC series (II) of FRET peptide substrates.

The adaptations made during this peptide synthesis were that the terminal amino acids (proline/β-alanine) were not initially labeled with the fluorescein dye. 5(6)-CBF was directly coupled to the aminoanthraquinone-peptide conjugate in the hope that the close running fluorescein impurity as noticed in JC series-1 could potentially be avoided.
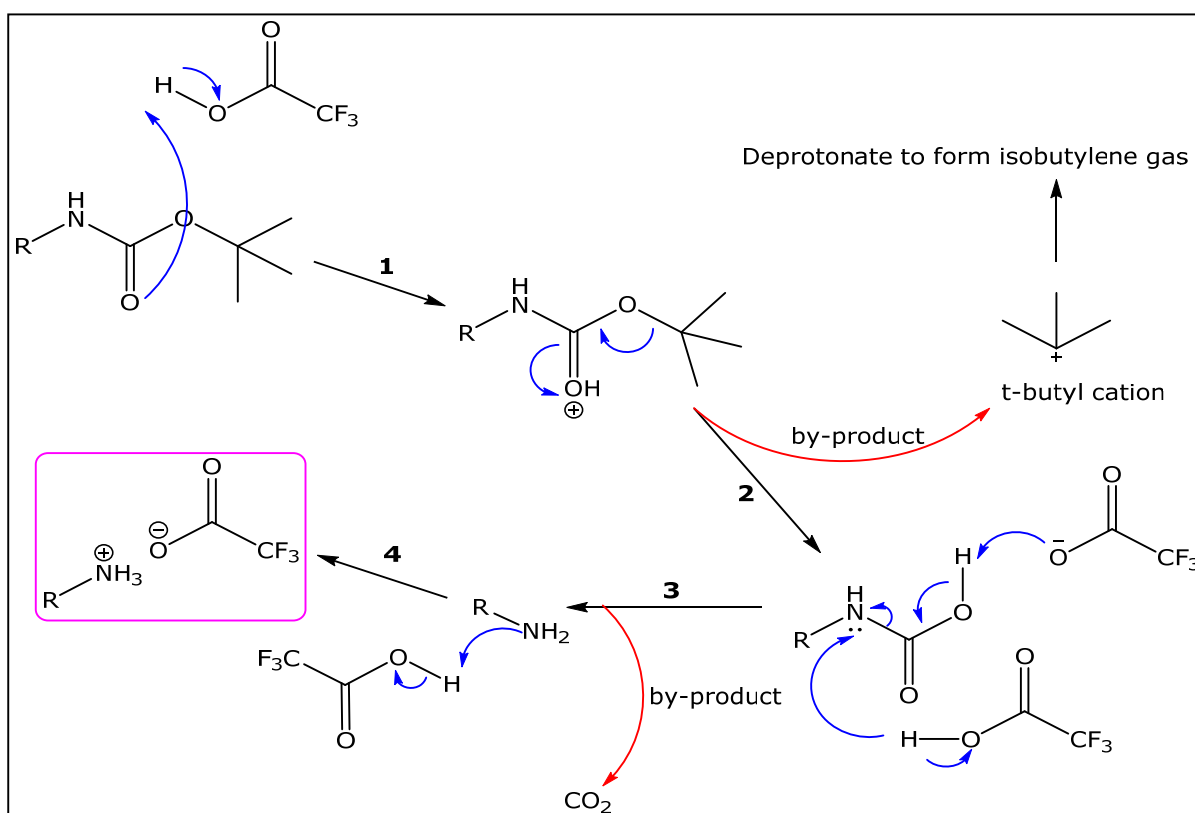
Figure 46: General mechanism for Boc group deprotection.

Moreover, it was decided to reduce the equivalence of the fluorophore while coupling to the aminoanthraquinone-peptide conjugate as a means to reduce the presence of any unreacted starting material. The final step involved the removal of the side chain protecting group from ε-lysine and γ-asparagine residues as explained in section 2.1.1.3 (III). The general mechanism of the Boc deprotection reaction is shown in figure 46. During the removal of the Boc protecting group using TFA, the *tert*-butyl carbamate becomes protonated leading to the loss of the *tert*-butyl cation resulting in a carbamic acid. Decarboxylation of the carbamic acid results in the release of the free amine and the final product, as the TFA salt results by the protonation of amine under acidic conditions as shown in figure 46. The rationale for the selection of amino acids for JC-2nd series was based on the information as discussed previously in section (2.1.1.3).

The synthesis of the 2nd batch of FRET tetra peptide substrates is shown in an abbreviated form in scheme 6.
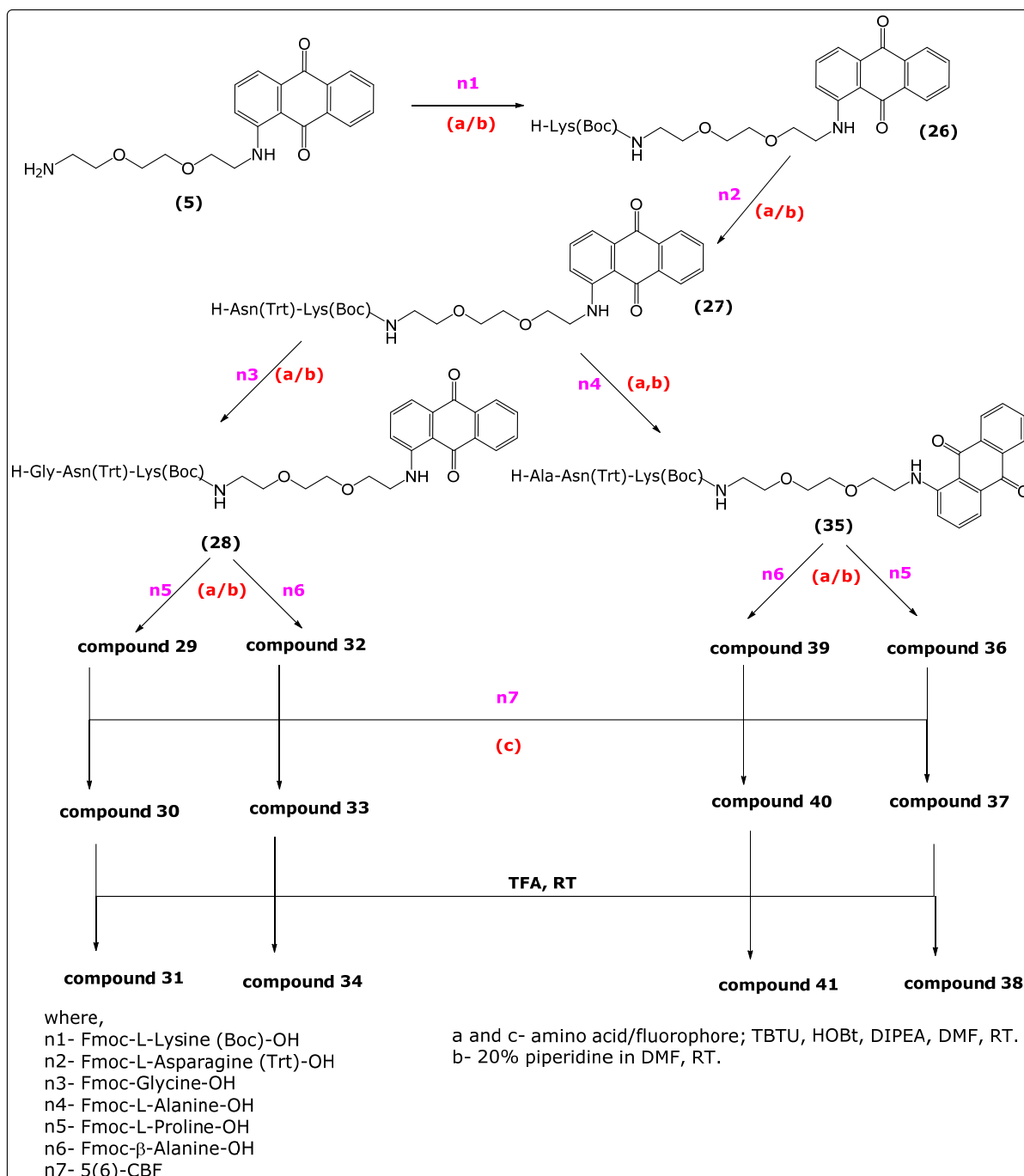
Scheme 6: A schematic representation of the synthesis of the compounds detailed in table 8.

where,
n1- Fmoc-L-Lysine (Boc)-OH
n2- Fmoc-L-Asparagine (Trt)-OH
n3- Fmoc-Glycine-OH
n4- Fmoc-L-Alanine-OH
n5- Fmoc-L-Proline-OH
n6- Fmoc-β-Alanine-OH
n7- 5(6)-CBF

a and c- amino acid/fluorophore; TBTU, HOBt, DIPEA, DMF, RT.
b- 20% piperidine in DMF, RT.

A. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Lys-Spacer-AQ (Compound 31)

I. Synthesis of intermediate compound- H$_2$N-Lys(Boc)-Spacer-AQ (Compoud 26)

(26)
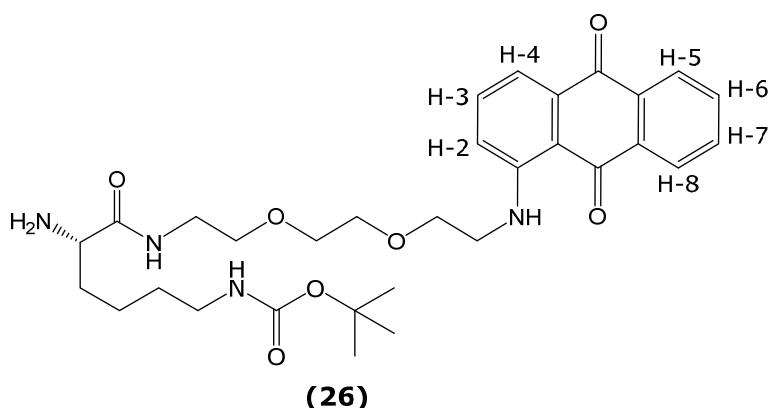
The Fmoc-protected intermediate (AQ-Spacer-Lys(Boc)-Fmoc) was synthesized by reacting *N*-Fmoc-Lys(Boc)-OH with compound **5** (scheme 6/n1) in DMF. Subsequent liquid/liquid extraction followed by *N*-Fmoc deprotection was performed using piperidine in DMF (scheme 6/n1). The deprotected sample was purified by column chromatography using 5% MeOH in DCM as the initial solvent to elute the front running impurities and then gradually increasing the polarity of the solvent system by using methanol up to 10% to elute the required product. The yield of the sample after column chromatography was 90%. The deprotected product **26** was then characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 583 (M+H)$^+$ corresponding to a molecular mass of 582 and also a signal at m/z 604 corresponding to the sodium adduct (M+Na)$^+$. The ESMS (-) mass spectrometry gave a signal at m/z 581 (M-H)$^-$ corresponding to the proton abstraction product and also a signal at m/z 616 corresponding to the chloride adduct (M+Cl)$^-$. Furthermore, the ^{1}H NMR spectrum (d-DMSO) showed, for example, a multiplet signal for the 15 protons between 1.20ppm and 1.52ppm which was assigned to β, γ, δ and the Boc group protons. A two proton quartet signal at 2.86ppm was assigned to the ε-methylene group protons of the lysine side chain. A quartet at 3.06ppm was assigned to the single α-proton of the lysine side chain. A two proton multiplet signal between 3.21 and 3.26ppm and another two proton triplet at 3.46ppm was assigned to C$\underline{H}_2$-NH-COO and to C$\underline{H}_2$-CH$_2$-NH-COO protons. A six proton multiplet between 3.52 and 3.64ppm was assigned to C$\underline{H}_2$-NH-AQ and O-C$\underline{H}_2$-C$\underline{H}_2$-O protons of the spacer arm. A two proton triplet at 3.73ppm was assigned to C$\underline{H}_2$-CH$_2$-NH-AQ. A single proton triplet at 6.75ppm was assigned to NH-COO-Boc group. The protons attached to the aromatic group were

assigned as follows: H-2 and H-4 gave a single proton doublet and a double doublet at 7.30ppm and 7.46ppm and a single proton triplet at 7.66ppm was assigned to H-3. A multiplet between 7.82ppm and 7.92ppm for three protons was assigned to H-6 and H-7 and NH-spacer arm protons. The H-5 and H-8 protons gave a double doublet signal at 8.13ppm and 8.21ppm respectively. The anthraquinone amino group proton gave a triplet at 9.80ppm. To further confirm the assignment in the ^{1}H NMR, a 2D-COSY spectrum of compound **26** was also performed (appendix 4).

II. Synthesis of intermediate compound- H₂N-Asn(Trt)-Lys(Boc)-Spacer-AQ (Compound 27)

(27)
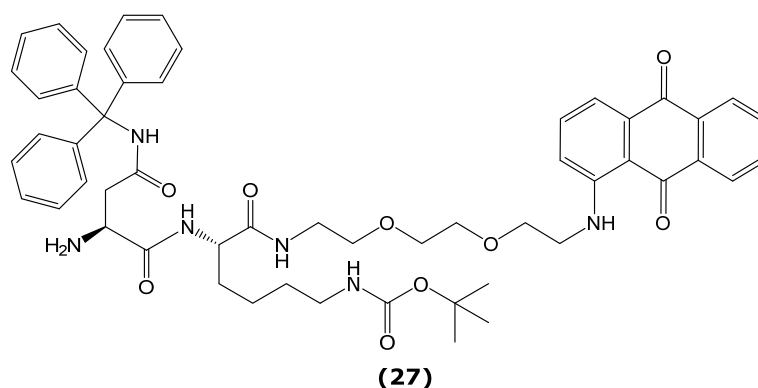
Compound **27** (scheme 6/n2) was synthesized following a procedure analogous to compound **26** and the yield for the post chromatographic sample was found to be 76%. The product **27** was then characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 939 (M+H)$^+$ corresponding to a molecular mass of 938; and the ESMS (-) mass spectrometry gave a signal at 937 (M-H)$^-$ corresponding to the proton abstraction product.

III. Synthesis of intermediate compound- H₂N-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (Compound 28)

(28)
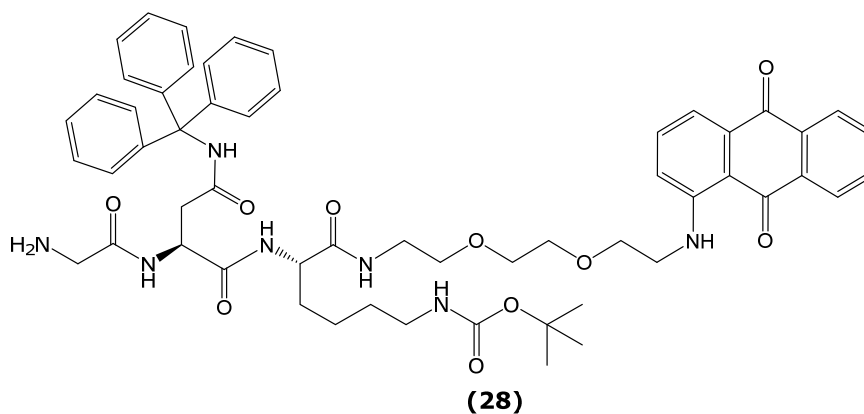
Compound **28** (scheme 6/n3) was synthesized following a procedure analogous to compound **26** and the yield for the sample after column chromatography was found to be 67%. The product **28** was characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 996 (M+H)$^+$ corresponding to a molecular mass of 995 and also a signal at m/z 1017 corresponding to the sodium adduct (M+Na)$^+$; the ESMS (-) mass spectrometry gave a signal at 993 (M-H)$^-$ corresponding to the proton abstraction product and m/z 1029 corresponding to the chloride adduct (M+Cl)$^-$.

IV. Synthesis of intermediate compound HN-Pro-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (Compound 29)

(29)
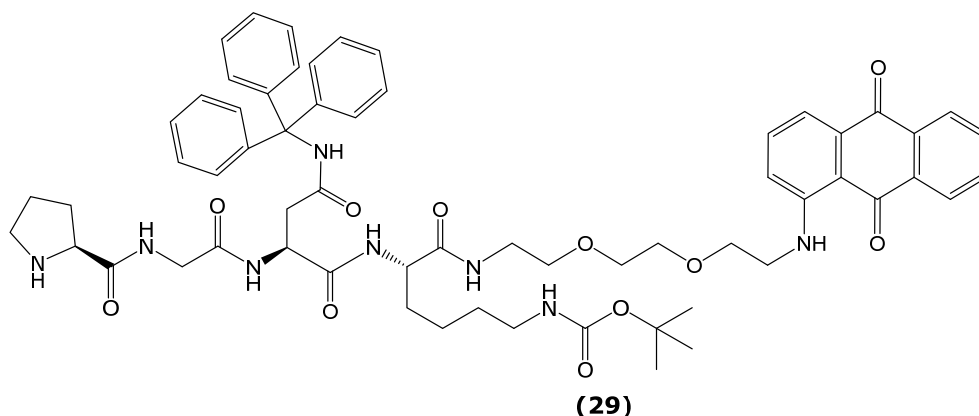
Compound **29** (scheme 6/n5) was synthesized following a procedure analogous to compound **26** and the yield for the sample was found to be 86%. The product **29** was then characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1093 (M+H)$^+$ corresponding to a molecular mass of 1092 and also a signal at m/z 1115 corresponding to the sodium adduct (M+Na)$^+$; the ESMS (-) mass spectrometry gave a signal at 1091 (M-H)$^-$ corresponding to the proton abstraction product and m/z 1127 corresponding to the chloride adduct (M+Cl)$^-$.

V. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Lys-Spacer-AQ (Compound 31)

| 5(6)-CBF | Pro | Gly | Asn | Lys | Spacer | AQ |
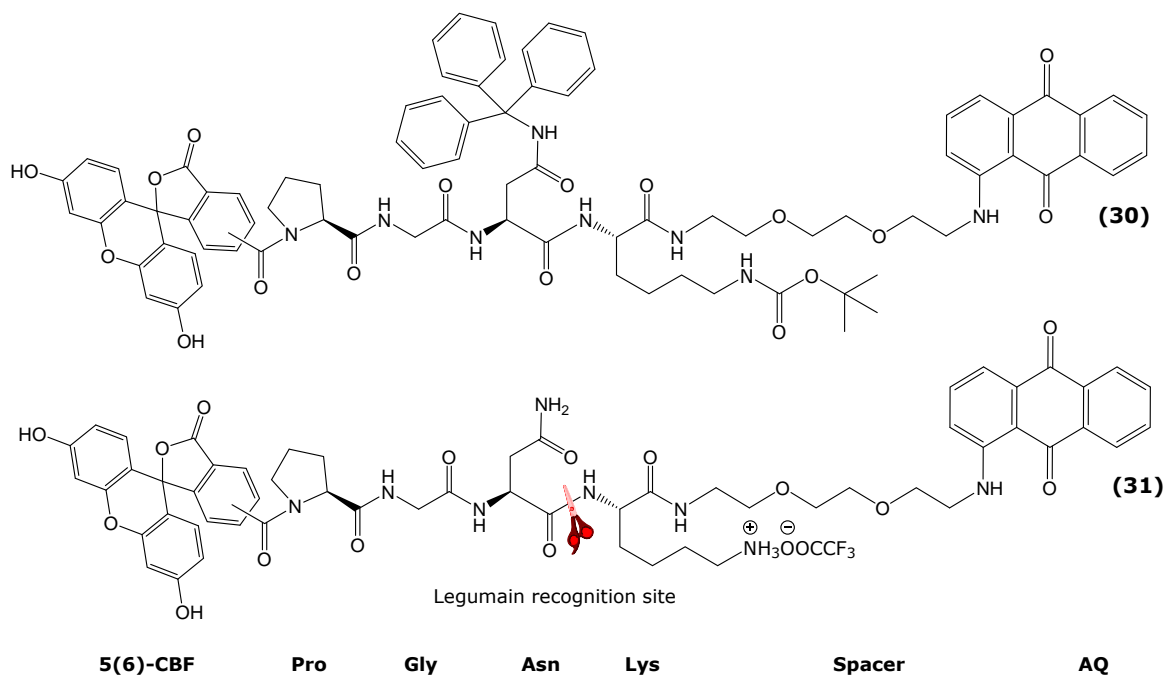
The fluorescein labeled tetra peptide sequence (**30**) was synthesized by reacting 5(6)-CBF with compound **29** in DMF (scheme 6/n7). The crude product was purified by column chromatography. The crude sample was dissolved in a polar solvent system (chloroform: MeOH: pyridine:35% ammonia- 20:55:15:10) and loaded onto a chromatographic column. The same solvent system was used to elute three fractions with the first and last fraction corresponding to impurities and unreacted fluorescein dye. The yield for the sample after column chromatography was found to be 82%. The second band consisting of a reddish brown solution was then filtered to remove silica and was characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1451 (M+H)$^+$ corresponding to a molecular mass of 1450 and also a signal at m/z 1473 corresponding to the sodium adduct (M+Na)$^+$; the ESMS (-) mass spectrometry gave a signal at 1449 (M-H)$^-$ corresponding to the proton abstraction product and m/z 1485 corresponding to the chloride adduct (M+Cl)$^-$.

Deprotection of the γ-trityl (asparagine) and ε-Boc (lysine) protected compound (**30**) with TFA gave the deprotected compound **31** after 3 hours reaction at rt (scheme 6/TFA). TLC showed the presence of a single new component close to the base line, due to the polar nature of the deprotected compound. Analytical HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 93% (figure 47) and the yield of the compound (**31**) was 97%. Additionally the

target compound **31** was also characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1109.4252 $(M+H)^+$ (figure 47) corresponding to a molecular mass of 1108 and also a signal for the doubly charged ion at m/z 555 $[(M+2H)/2]^{2+}$. The observed data was consistent with the theoretical isotope model confirming the structure of compound **31**. It was noticed that the yield and purity of the target compound was high when compared to the 1st batch of JC compounds.
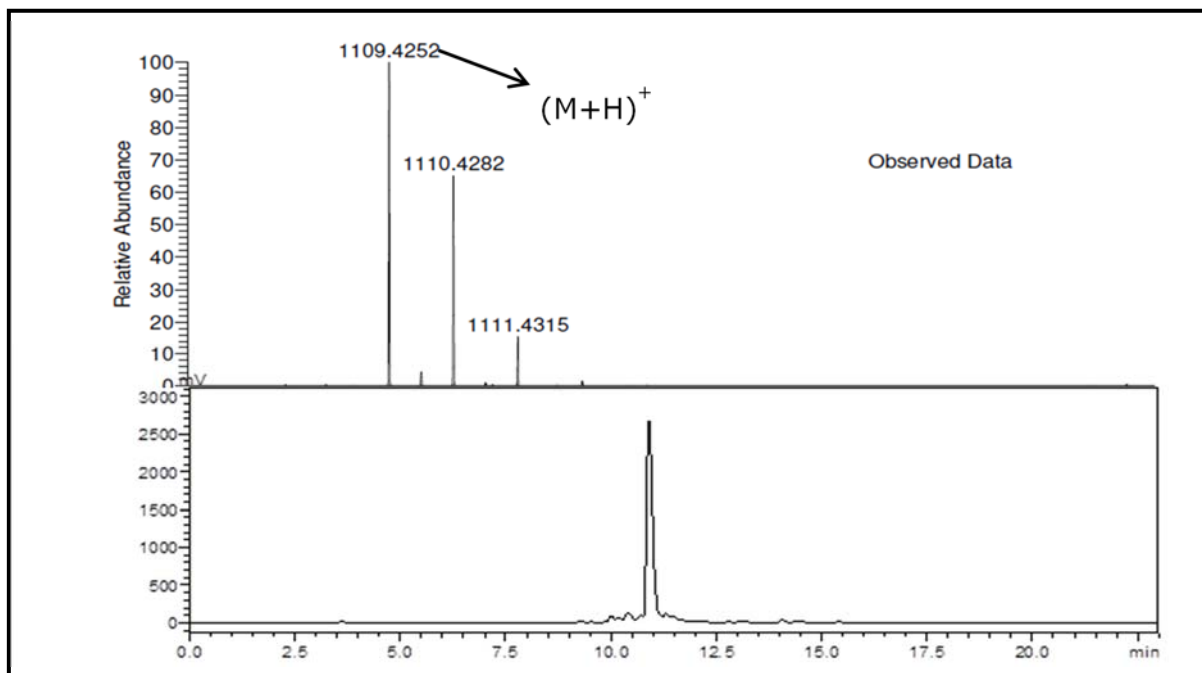
Figure 47: Mass spectrometry result for compound **31**. RP‑HPLC elution profile of compound **31** (1mg/mL in methanol); Conditions: Flow rate, 1mL/min; Detection, 275nm; stationary phase: Eclipse XDB-C18 RP column; mobile phase A= 0.1%TFA/water; mobile phase B=0.1%TFA/methanol; gradient: low pressure gradient; 23min.

This could have been possible because of the switch to solution peptide synthesis which normally is the favourable technique for synthesis of peptide sequences less than 8 amino acids long. Moreover the isolation of intermediates after each coupling and deprotection cycle by column chromatography could have possibly removed any undesirable products. Furthermore Fernandez-Carneado and Giralt (2004) have demonstrated in their studies on fluorescent labeled peptides that the coupling of 5(6)-CBF into a solid support-attached peptide sequence was found to be difficult.

B. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Lys-Spacer-AQ (Compound 34)

5(6)-CBF   β–Ala   Gly   Asn   Lys   Spacer   AQ
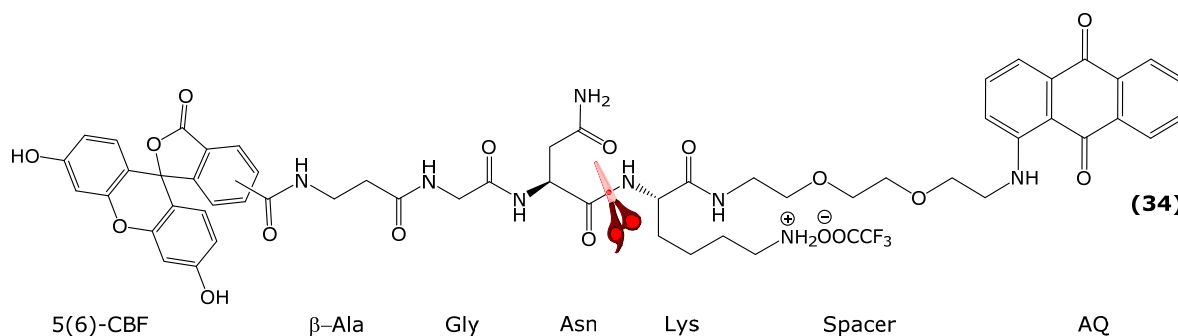
The entire chemical synthesis of the FRET peptide substrate (**34**) was similar to compound **31**. The only difference was the replacement of proline with β-alanine at P_3 position of the peptide sequence. The coupling and deprotection methods and the reaction conditions were the same as in compound **31** (scheme 6). HPLC for the FRET tetra peptide **34** was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 77% and the yield of the compound was 83%. The product **34** was also characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1083.4097 (M+H)$^+$ corresponding to a molecular mass of 1082 and also a doubly-charged signal at m/z 542 [(M+2H)/2]$^{2+}$.

C. Synthesis of 5(6)-CBF-Pro-Ala-Asn-Lys-Spacer-AQ (Compound 38)

5(6)-CBF   Pro   Ala   Asn   Lys   Spacer   AQ
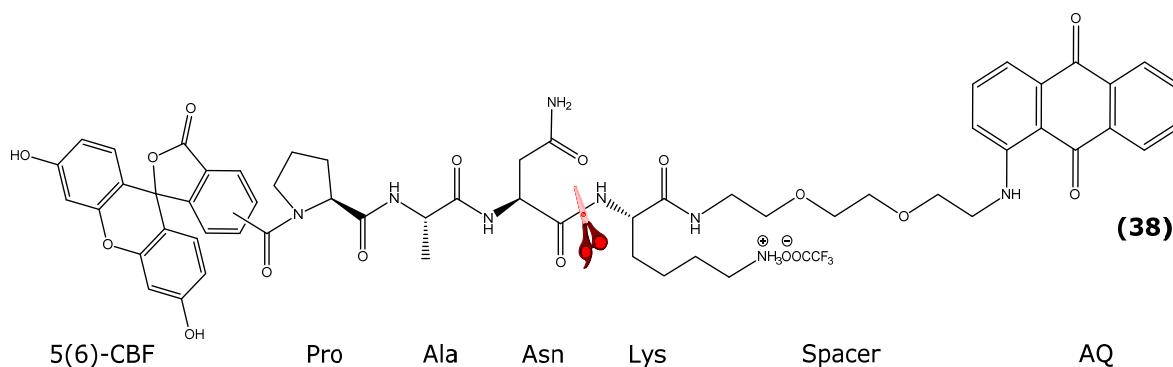
The chemical synthesis of the FRET peptide substrate (**38**) was similar to compound **31**. The only difference was the replacement of glycine with alanine at P_2 position of the peptide sequence. The coupling and deprotection methods were similar to compound **31** as mentioned in scheme 6. HPLC for the target compound **38** was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 82% and the yield of the compound was 92%. The deprotected product **38** was characterized by high resolution ESI (+) mass spectrometry, which gave

a signal at m/z 1123.4411 (M+H)$^+$ corresponding to a molecular mass of 1122 and also a doubly-charged signal at m/z 562 [(M+2H)/2]$^{2+}$.

D. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Lys-Spacer-AQ (Compound 41)

5(6)-CBF	β−Ala	Ala	Asn	Lys	Spacer	AQ
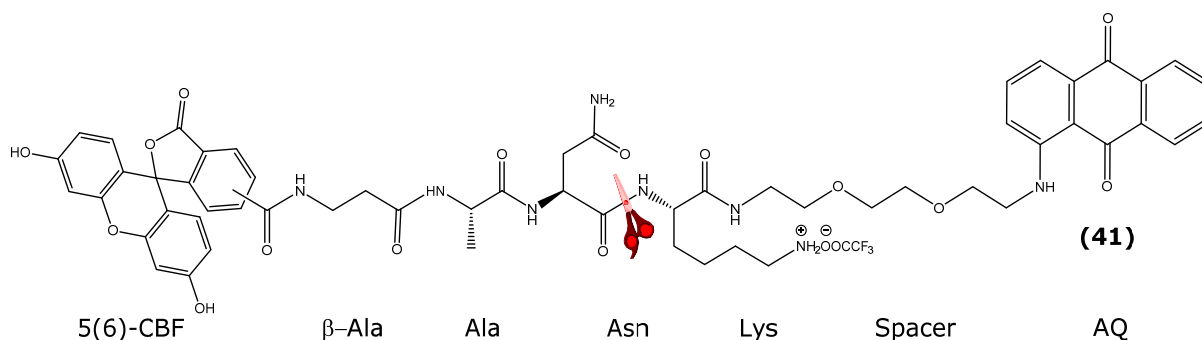
The synthesis of the FRET peptide substrate followed the same procedure as compound **34**. The major difference was the replacement of gly with ala at the P2 positon of the sequence. The reaction conditions remained the same as mentioned in scheme 6. The purity of the target compound **41** was determined to be 84% *via* HPLC analysis performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the yield of the compound was 91%. The deprotected product **41** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1097.4250 (M+H)$^+$ corresponding to a molecular mass of 1096 and also a doubly-charged signal at m/z 549 [(M+2H)/2]$^{2+}$.

E. Synthesis of 5(6)-CBF-β-Ala-Pro-Ala-Asn-Leu-Spacer-AQ (Compound 46)

5(6)-CBF	β−Ala	Pro	Ala	Asn	Leu	Spacer	AQ
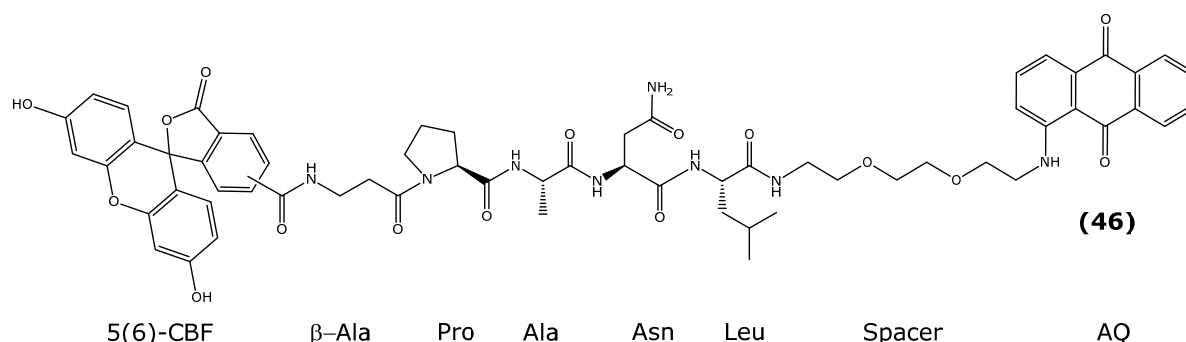
The chemical synthesis of the FRET peptide substrate (**46**) was similar to compound **41**. The major difference was the addition of another amino acid to the peptide sequence making it a penta-peptide substrate. This peptide sequence was synthesized in order to evaluate the influence of additional amino acids on the recognition of the peptide substrate by the enzyme. The coupling and deprotection

methods were similar to compound **31** as mentioned in scheme 6. The γ-trityl (asparagine) protected compound (45) was deprotected using trifluoroacetic acid (scheme 6/TFA). Reaction progress was monitored by TLC analysis, which showed the presence of a single new component with a low R_f value indicating the polar nature of the deprotected sample. HPLC for the target compound **46** was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be only 64% and the yield of the compound was 75%. The product **46** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1179.4674 $(M+H)^+$ corresponding to a molecular mass of 1178 and also a doubly-charged signal at m/z 590 $(M/2+2H)^{2+}$.

2.1.1.7. Fluorescence measurements for the 2nd series of JC library

Earlier in this section (2.1.1) it was discussed that fluorescent-labeled peptide substrates were effectively used as sensitive tools for monitoring various biological events. Chen and colleagues (2014) in their work on fluorescent probes have demonstrated that the principle of FRET may be an effective technique to image legumain-expressing abnormal tissues for both diagnosis and targeted treatments. Hence, by applying the principle of FRET, a small library of peptide substrates designed to be specific to legumain was synthesized. It was hoped that these self-quenched FRET substrates may release detectable fluorescence signals upon cleavage by legumain. Moreover the presence of a non-fluorescent quencher moiety (AQ) with a broad absorption wavelength of 550nm-700nm was found to be efficient in quenching 5(6)-CBF by eliminating much of the background fluorescence. Based on this idea the JC series of compounds were synthesized consisting of a tetra peptide substrate between the FRET pair. In order to evaluate the ability of the anthraquinone chromophore to quench the fluorescence emission of 5(6)-CBF an initial fluorescence measurement study was perform in a similar manner to that of compound **8**.

Compound **31** along with its corresponding peptide-fluorophore sequence was dissolved in water and the fluorescence spectrum was recorded using a luminescence spectrophotometer. From figure 48, the data for fluorophore quenching studies demonstrated that the relative fluorescence intensity of the compound **31** was almost completely quenched by the acceptor unit (AQ) which can be seen by a low lying curve (4) when compared to its fluorogenic-peptide conjugate (3). Thus the fluorescence measurement studies demonstrated that the relative fluorescence intensity of the fluorophore (5(6)-CBF) has been completely

quenched by the anthraquinone chromophore. Therefore any release in fluorescence that would be observed during the *in vitro* studies with rh legumain was not because of background fluorescence from the FRET compounds but only because of the release in fluorescence due to the increase in distance between the acceptor and donor pair as a result of proteolytic activity by the enzyme.

1] Water
2] 5(6)-carboxyfluorescein
3] 5(6)-carboxyfluorescein-Pro-Gly-Asn-OH (A)
4] 5(6)-carboxyfluorescein-Pro-Gly-Asn-OH-Spacer-AQ (**31**).
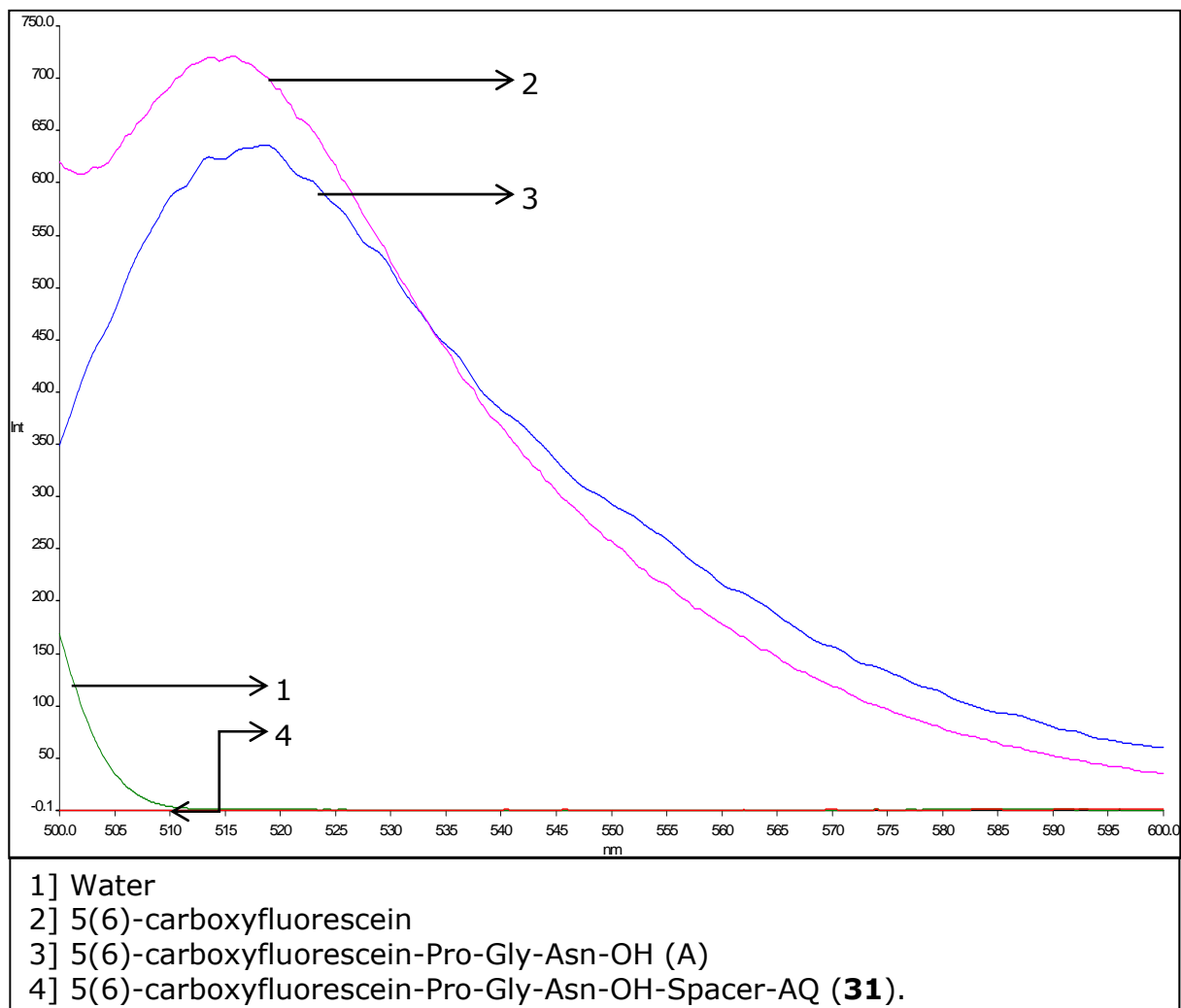
Figure 48: Comparison of relative fluorescence intensities between 5(6)-CBF-Pro-Gly-Asn-OH (1µM) and the FRET substrate-compound **31** (1µM) in water.

2.1.1.8. Legumain activity assay for the 2nd series of JC library

The 2nd series of FRET peptide substrates contained five compounds labeled with 5(6)-CBF. The compounds **31,34,38,41** and **46** had a design similar to 5(6)-CBF-AA2–AA1-Asn-Lys-Spacer-AQ, where AA1 and AA2 corresponded to amino acids. The five compounds were incubated with activated legumain individually following a protocol similar to the previous assay conditions as in section 2.1.1.6. The FRET pair in each of the compounds was expected to quench the emission by the fluorophore. However, on cleavage of the peptide substrate by the enzyme, an

increase in the distance between the donor/acceptor pair would lead to fluorescence release (Chen *et al*. 2014). Thus release in fluorescence was noticed as soon as the compounds were incubated with the activated legumain. The relative fluorescence intensity reached a maximum after 120 minutes incubation (figure 49).
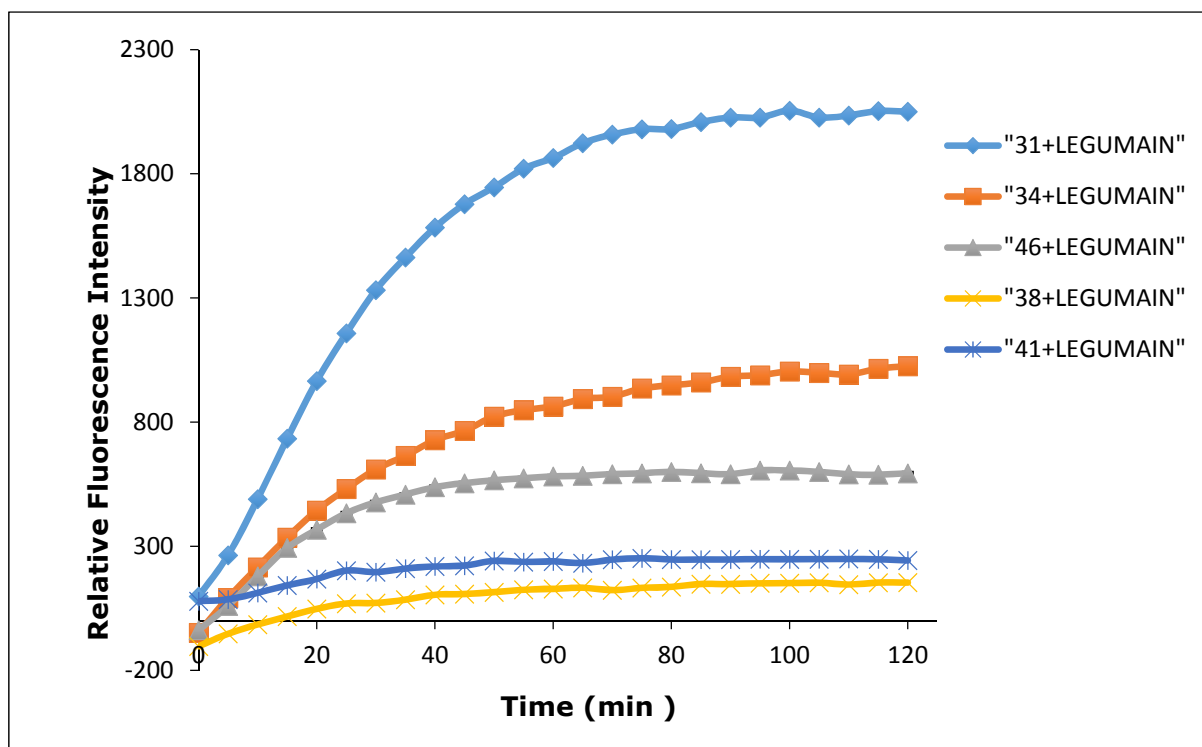
Figure 49: Relative fluorescence intensity upon incubation of **31**, **34**, **46**, **38** and **41** (10µM) with legumain (40ng in a final assay volume of 100µL per well) in legumain assay buffer, pH 5.0 λ_{ex} 492nm, λ_{em} 520nm. Data represents mean values ± SD from triplicates from one experiment done individually.

The results indicated that the peptide substrates could be recognized by legumain and hence these compounds could potentially be used to synthesize diagnostic tools for targeting legumain, which is overexpressed in atherosclerotic plaques. From the legumain enzyme assay it was observed that compounds **31**, **34** and **46** produced a gradual increase in fluorescence intensity with time when incubated with the enzyme, whereas compounds **46** and **49** did not produce the gradual increase in fluorescence. Moreover, compound **34** having β-Ala at P3 position was found to release fluorescence which was in contrast to the results obtained from the 1st batch of JC compounds. This may suggest that compounds **17**, **20** and **25** did not produce the required result because of background fluorescence as well as the position of leucine at P1 (Schwarz *et al*. 2002) and not necessarily because of β-Ala at P3 position. More information that can be gained from the legumain assay was that, from the five successful compounds (**8**, **11**, **31**, **34** and **46**) including the 1st-batch compounds, three of the compounds (**11**, **31** and **34**) had Gly at P2

position and the other two compounds (**8** and **46**) had Ala at P2 position. This result may perhaps suggest that the preference for glycine could be higher than alanine, even though both the amino acids are thought to be favored at P2 position as suggested in the experiments conducted by Mathieu *et al*. 2002.

Although compounds **14**, **17**, **20** and **25** suffered severe background fluorescence, this problem was rectified during the synthesis of compounds **31**, **34**, **38**, **41** and **46**. The latter compounds, denoted as the 2nd batch of FRET peptides, were synthesized *via* solution peptide synthesis, which may possibly have influenced the purity and yield of the target compounds when compared to the 1st batch of FRET peptide substrates. As previously discussed legumain is a distinct enzyme, which has strict specificity towards an asparagine residue at the P1 position of the peptide sequence (Mathieu *et al*. 2002). It was also mentioned that the amino acids on positions apart from P1 did not influence enzyme specificity however, proline and alanine residues at the P3 position were found to be the most desired amino acids at that position (Mathieu *et al*. 2002). Moreover in order to analyze the influence of the length of the peptide sequence on enzyme recognition, a penta-peptide substrate (**46**) was also synthesized as part of the 2nd batch of FRET peptide substrates. The compounds **8**, **11**, **31**, **34**, **38** and **41** were adapted as the design template for synthesizing the library of JS series in the following section 3.1.2. The JS library of peptide substrates were complexed to a paramagnetic metal ion-gadolinium linking to the aim of this research study, which was the synthesis of a contrast agent that would ultimately improve MR imaging, facilitating a non-invasive means for detecting atherosclerotic plaques.

2.1.2. MRI based contrast agents for imaging atherosclerotic plaques (JS series)

Magnetic resonance imaging is a non-invasive technique that can provide information on the anatomy, function and metabolism of tissues *in vivo*. Apart from the basic contrast in MR images produced by the difference in the local water content, the signals from regional differences in the relaxation time T1 and T2 also dominate image contrast. However, the intrinsic contrast produced by the signals provided by water T1 and T2 about tissue pathology are often too limited for a sensitive and specific diagnosis (Strijkers *et al*. 2007). For this reason, MRI contrast agents such as gadolinium complexes of either DOTA or DTPA are increasingly used to alter and improve image contrast. Gadolinium is a positive-contrast paramagnetic agent, which works by shortening the relaxation time (T_1) of water protons in the tissue. Liu and colleagues (2012) in their studies have

demonstrated that the MR images acquired with the application of gadolinium complexes were able to expose tissue lesions, atheroma and metastases more clearly than the images without the contrast agent.

A biocompatible molecule that can bind efficiently to a paramagnetic ion to produce a stable metal chelate is defined as a ligand. Many such gadolinium chelates (as discussed in the introduction chapter) are commercially available for research and clinical applications. Novel MRI contrast agents have been synthesized by covalently attaching them to suitable frameworks using dendrimers, nanoparticles, liposomes, proteins and bioactive peptides (Huang *et al*. 2013). In a study by Nwe and colleagues (2010) on contrast agents, they have synthesized gadolinium complexes either using DTPA or DOTA conjugated to G4 PAMAM dendrimers in order to evaluate MRI quality and contrast. Their findings suggested that gadolinium-DOTA complexes produced better quality images and also a faster blood clearance compared to Gd-DTPA.

Another major limitation for the use of gadolinium contrast agents for clinical application is their low sensitivity. Hence the need for targeted contrast agents to specifically image regions of interest is desired. Many studies have suggested the use of folic acids, lactoferrin, nucleic acids, peptides and monoclonal antibodies as some of the targeting components of MRI contrast agents (Chen *et al*. 2014). Such targeted MRI contrast agents have proven useful in research and clinical diagnosis for highlighting specific metabolites and biomarkers (Liu *et al*. 2014). Chen and colleagues (2014) in their studies on MRI contrast agents have successfully synthesized a novel contrast agent and a near infrared fluorescent probe for both *in vivo* and *in vitro* detection of legumain, which is generally over expressed in unstable atherosclerotic plaques. The application of contrast agents conjugated to legumain activatable peptide substrates to detect legumain accumulation in atherosclerotic plaques using MR imaging was the basis for this current research study.

The aim of this research study was to synthesize a small library of substrate based MRI contrast agents designed to specifically image atherosclerotic plaques by targeting some of the biomarkers (enzymes/proteases-legumain) that are over expressed in the microenvironment of the disease. Moreover, it can be anticipated that these MRI-images can give a better understanding about the type of therapeutic agent that could be effective for treating the plaques or the imaging agent can alternatively be used to deliver the therapeutic drug to the atherosclerotic plaques (Nelson 2012). This is one possibility for the future development of these substrate based contrast agents for therapeutic purposes.

In this way, the contrast agents can be used for both research and clinical evaluations as a novel means to diagnose and monitor drug administration and treatment. A targeted contrast agent, in context to this research study, can be defined as an agent, which is expected to produce little or zero signal until activated by the target enzyme. In this way the activity of the enzyme of interest can be monitored thus creating a platform for using them as diagnostic tools to locate plaques. The atherosclerotic plaques could potentially be imaged using a peptide substrate conjugated to MRI contrast agents analogous to the JS series of compounds. This library of compounds consisted of a paramagnetic gadolinium chelate covalently coupled to a peptide sequence designed to be specific to legumain. Hence, legumain driven cleavage of the peptide substrate can lead to accumulation of the contrast agent at the site of the plaque ultimately leading to better contrast MR images as described in figure 50 (Leuschner and Nahrendorf 2011). Following the successful synthesis and characterization of JC series of FRET peptide substrates, the same method of synthesis was applied to construct a small library of MRI contrast agents conjugated to peptide substrates designed to be specific to legumain. The sequence design for the JS series was based on the results from the legumain assay on the JC FRET peptide library. Compounds **8** and **11**, which were readily hydrolysed in the presence of legumain, were chosen as the model for the JS linear series (figure 50). The linear sequences were synthesized *via* solution phase peptide synthesis. Similarly, based on the results from the second batch of JC FRET peptide library, the JS lysine series was designed and synthesized. The lysine sequences were synthesized *via* a combination of both solution and solid phase peptide synthesis (figure 50).
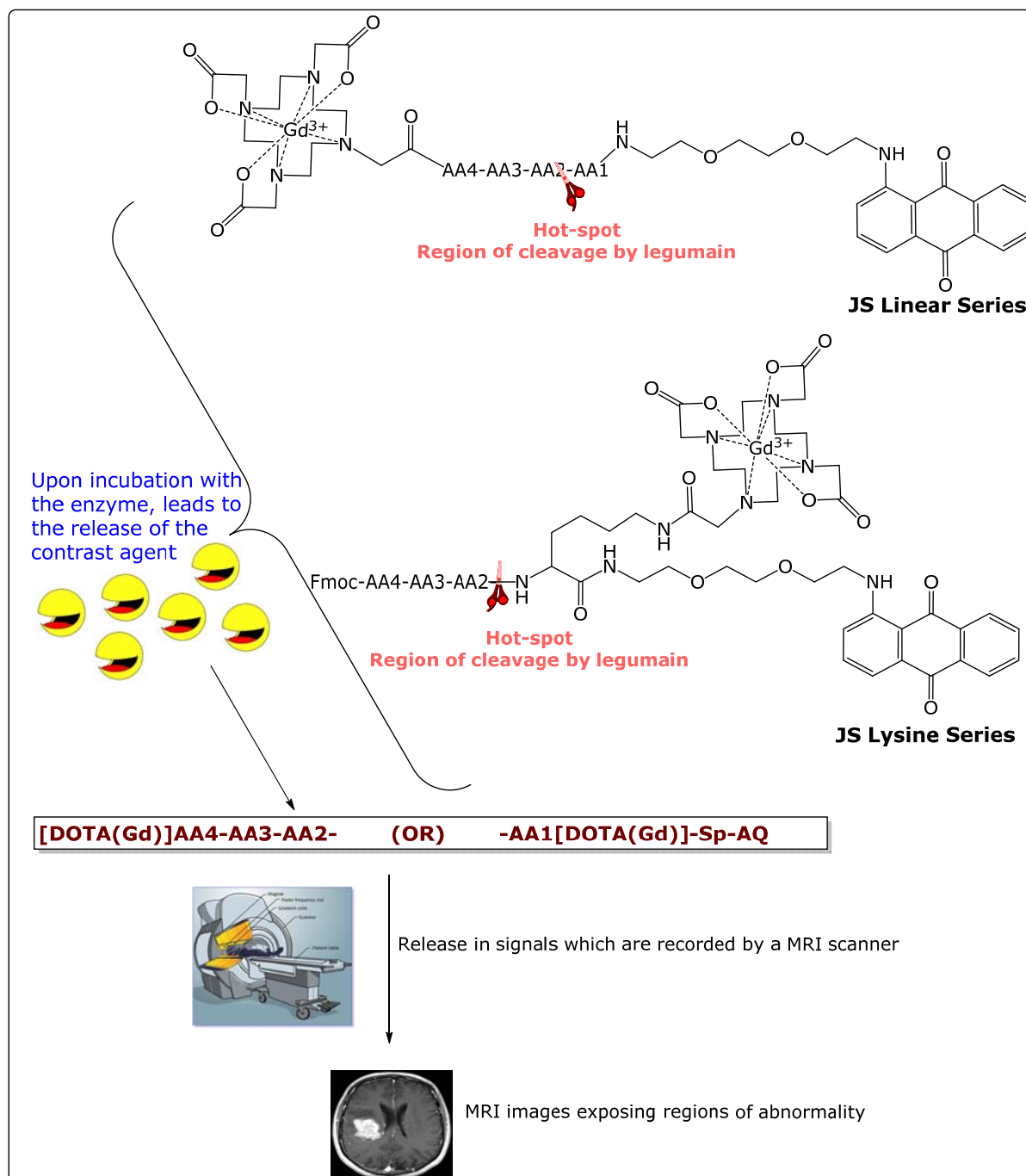
Figure 50: Template describing the JS series of compounds and the anticipated image enhancement on release of the CA when the MRI peptide CA is activated by legumain.

2.1.2.1. Synthesis of Gadolinium-DOTA-Peptide substrates-JS Linear sequences

The rationale for the synthesis of the gadolinium-DOTA-peptide substrates was to image atherosclerotic plaques by specifically identifying the biomarkers (legumain) over expression in these plaques. Hence two series of compounds were synthesized and this section explains in detail the chemical synthesis of the JS-linear series. The linear series of JS compounds had a design: Paramagnetic gadolinium complex-AA4-AA3-AA2-AA1-Spacer-AQ, where AA represents amino

acids and the complex consisted of Gd-DOTA chelate. The linear sequences were synthesized following the procedure for solution peptide synthesis. The JS-linear library of compounds consisted of a hydrophobic group comprising of an aminoanthraquinone moiety, which was conjugated to the C-terminus of a tetra peptide sequence in order to, potentially, increase the retention time of the contrast agent (DOTA-Gd) within the atherosclerotic plaque environment. This increase in retention time of the contrast agent could perhaps enhance the quality/resolution of the image from a MRI scanner. The presence of a short diamino chain between the carboxyl end of the peptide sequence and anthraquinone moiety was incorporated to produce an amphiphilic compound in order to improve the solubility of the compound. Furthermore, the intense reddish-pink colour of the aminoanthraquinone chromophore helped to identify the required compound on a silica gel chromatographic column; thus making the process of identifying the required fraction easier. The chelating agent-DOTA was either covalently attached to the side chain of a lysine residue (**JS-lysine series**) or to the terminal proline (**JS-linear series**). The rationale for such a design was based on the concept that the aminoanthraquinone-peptide-contrast agent conjugate can accumulate and bind to the extracellular protein moieties in the ECM and when cleaved by the target enzyme, the intact contrast agent may be released locally. The released contrast agent (DOTA-Gd), along with the hydrophobic site (proline, β-alanine, and AQ-spacer-lysine) attached to them, could increase the retention time of the CA. This may enhance the contrast between the normal and abnormal regions of investigation. As discussed in previous sections, DOTA is a favourable bifunctional chelating agent (BFCA) when compared to DTPA mainly because of its thermodynamic stability and kinetic inertness preventing *in vivo* dissociation of free metal ions into the blood stream. Furthermore, most of the clinically relevant MRI contrast agents are based on DOTA-gadolinium complexes such as Gadovist®, ProHance®, Dotarem® and Gadomer-17®. Several derivatives of DOTA-based BFCA's are available commercially such as active DOTA esters, DOTA with a coupling unit present either at the α- site of one carboxylate arm (McMurry *et al*. 1992; James et al. 2012) or into the macrocycle (Kruper *et al*. 1993) and protected DOTA species (Jamous *et al*. 2013). However, the protected DOTA derivative (DOTA-tris-^{t}Bu ester) has been used in this research project as it can be readily coupled to the peptide sequence following a standard coupling procedure during both solution and solid phase peptide synthesis.

A. Synthesis of Gd(III)DOTA-Pro-Ala-Asn-Leu-Spacer-AQ (Compound 49)

I. Synthesis of intermediate compound DOTA(*t*Bu)-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ (Compound 47)

(47)
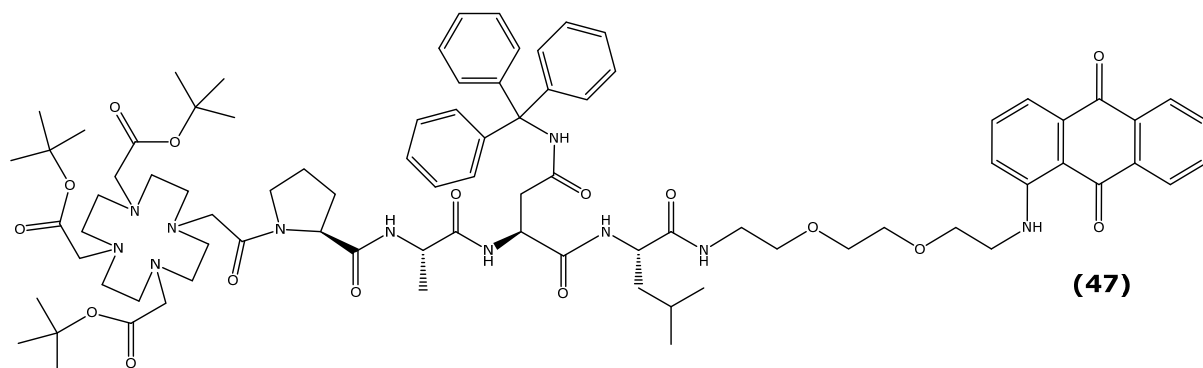
The *tert*-butyl protected compound (**47**) was synthesized by reacting DOTA-tris-*t*Bu-ester with compound **43** in DMF along with the in situ coupling agents: TBTU/HOBT and DIPEA. The crude sample was purified by column chromatography and the yield of the product (**47**) post column chromatography was 43%. After purification the compound **47** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1568.8402 (M+Na)$^+$ corresponding to the molecular mass of 1545 and also a signal at 784 [(M+Na)/2]$^{2+}$.

II. Synthesis of intermediate compound DOTA-Pro-Ala-Asn-Leu-Spacer-AQ (Compound 48)

(48)
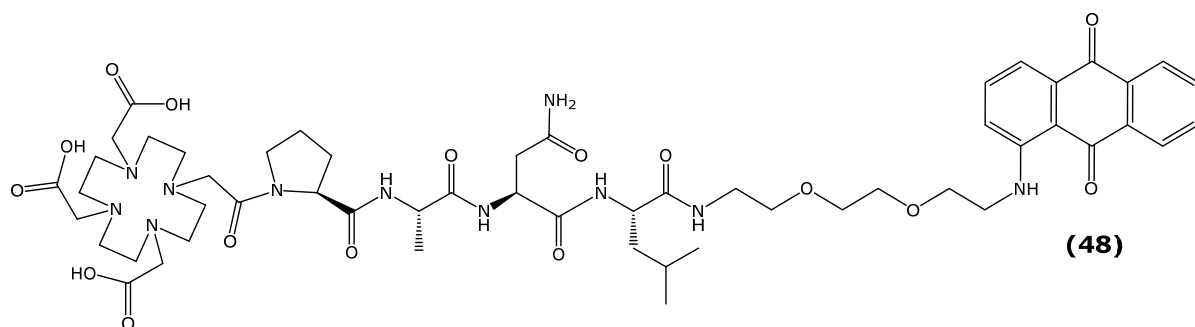
The γ-trityl and *tert*-butyl protected aminoanthraquinone-tetra peptide sequence (**47**) was deprotected using trifluoroacetic acid. The reaction conditions and reagents were the same as detailed in section (2.1.1.3). Reaction progress was monitored by TLC analysis, which showed the presence of a single new component close to the base line confirming the polar nature of the deprotected sample. The residual TFA was evaporated under reduced pressure and any remaining solvent was re-evaporated with ethanol. The condensate was triturated with diethyl ether

followed by cooling on ice to facilitate further precipitation. The yield of the crude sample **48** was found to be 87%. The product (**48**) was characterized by high resolution ESI (+) mass spectrometry, which gave a doubly-charged signal at m/z 568.7844 $[(M+2H)/2]^{2+}$ corresponding to a molecular mass of 1135 and also a signal at m/z 587 corresponding to the potassium adduct $[(M+H+K)/2]^{2+}$.

III. Synthesis Gd(III)DOTA-Pro-Ala-Asn-Leu-Spacer-AQ (Compound 49)
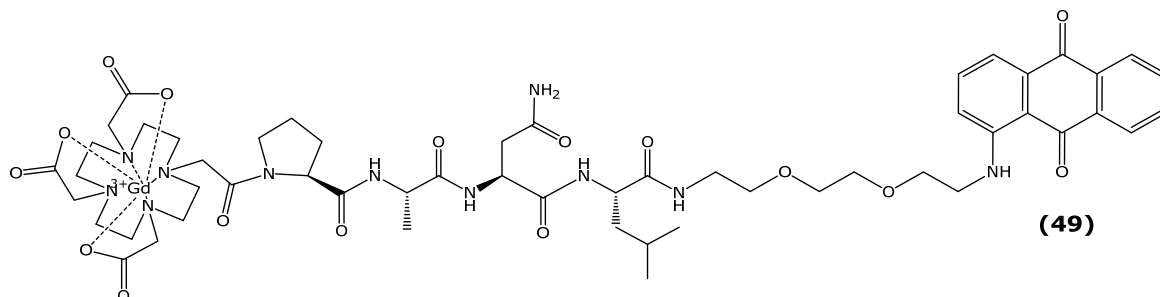
(49)

Synthesis of the target gadolinium complex of DOTA-tetrapeptide-amino anthraquinone spacer conjugate (**49**) was performed in accordance to the reaction conditions proposed by Digilio and co-workers (2010). The Gd complex was prepared in a neutral aqueous solution by reacting GdCl$_3$(6H$_2$O) with compound **48**.
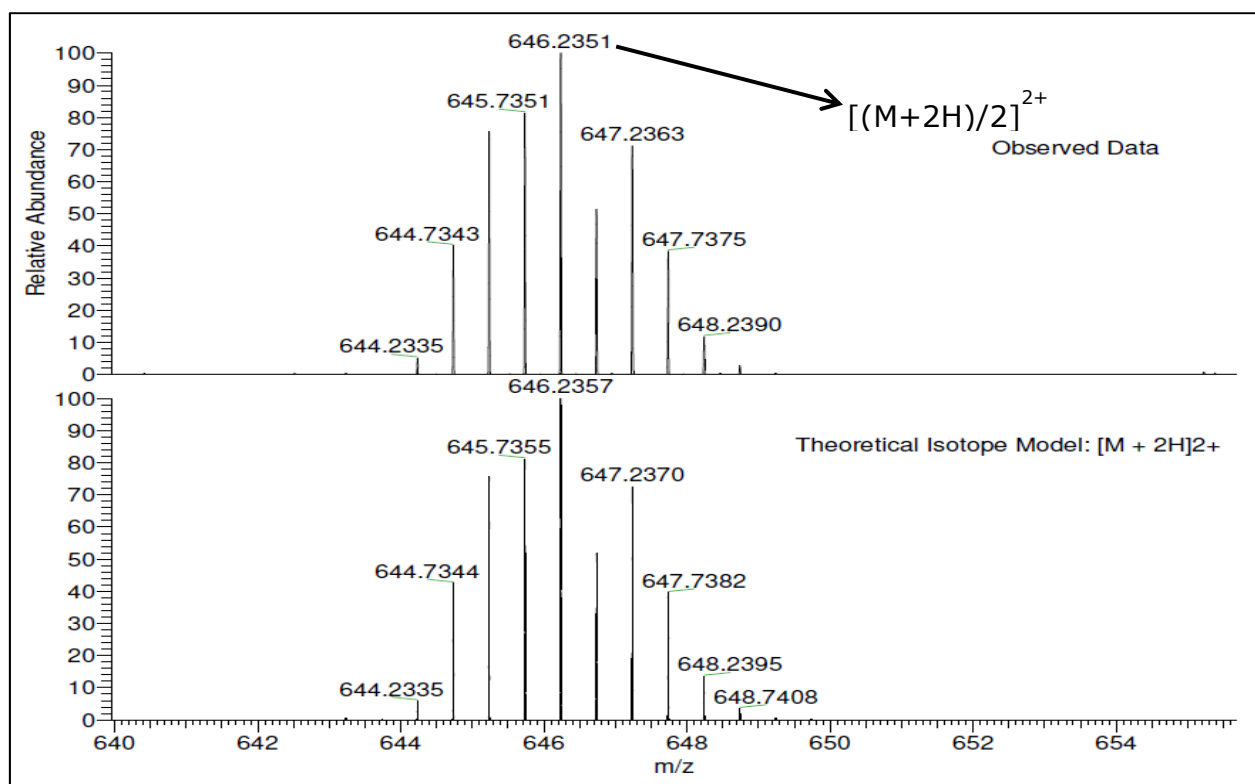
Figure 51: Expansion of the ESI (+) mass spectrum of the Gd-DOTA-peptide sequence (**49**) that shows the isotopic pattern of the molecular ion (doubly charged at half mass).

The reaction pH was adjusted to 7 by adding a solution of ammonia 0.35% drop wise. After two hours the reaction progress was checked by TLC analysis using chloroform: MeOH: pyridine:35%ammonia (20:55:15:10), which showed the presence of two components with one corresponding to the starting material. After 13 hours the reaction mixture was reduced to low volume by evaporation at room temperature and then purified by column chromatography using the same solvent system as for the TLC analysis. Analytical HPLC for the target compound **49** was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 72%. The product (**49**) was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1291 (M+H)$^+$ corresponding to a molecular mass of 1290 and also a signal at m/z 1313 corresponding to the sodium adduct (M+Na)$^+$. A doubly-charged signal at m/z 646.2351 [(M+2H)/2]$^{2+}$ was also observed (figure 51).

B. Synthesis of Gd(III)DOTA-Pro-Gly-Asn-Leu-Spacer-AQ (Compound 53)

| DOTA-Gd^{3+} | Pro | Gly | Asn | Leu | Spacer | AQ |
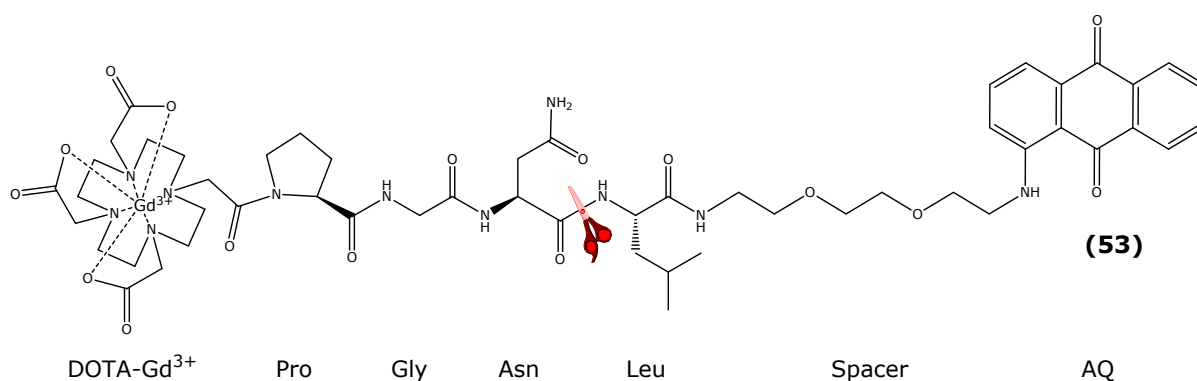
Compound **53** was synthesized following a procedure analogous to compound **49**. The only difference was the swapping of alanine with glycine at the P$_2$ position of the peptide sequence. Analytical HPLC for the target compound (**53**) was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 77%. The product **53** was characterized by high resolution ESI (+) mass spectrometry, which gave a doubly-charged signal at m/z 661.2134 [(M+2Na)/2]$^{2+}$ (figure 52) corresponding to a molecular mass of 1276 and also a signal at m/z 1299 corresponding to the sodium adduct (M+Na)$^+$.
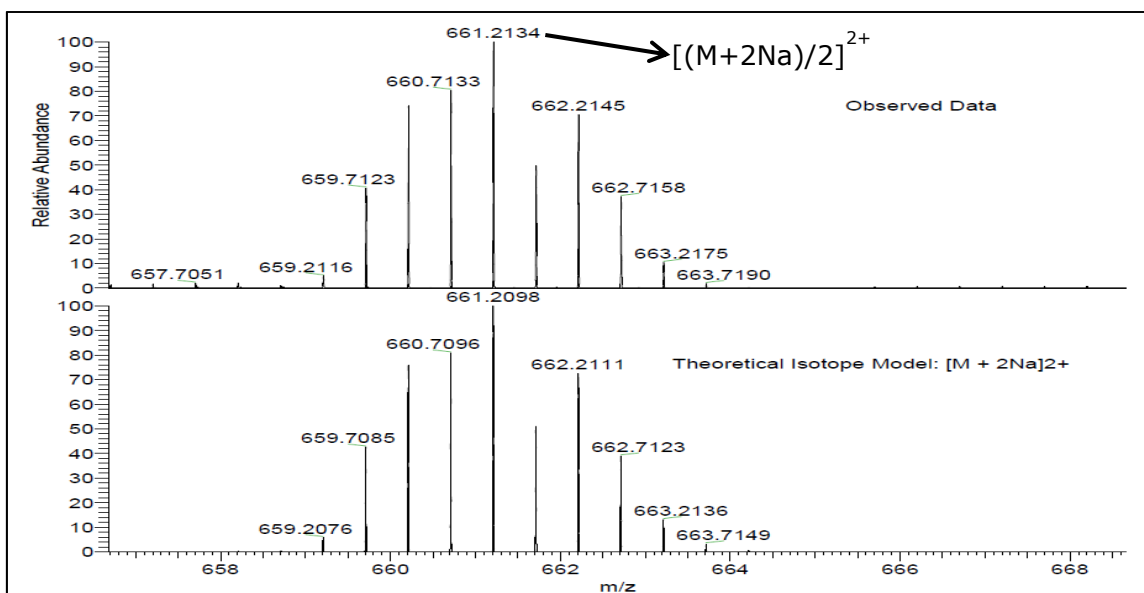
Figure 52: Mass spectrometry result for compound **53**.

2.1.2.2. Synthesis of Gadolinium-DOTA-Peptide substrates-JS Lysine sequence

Yang and colleagues (2014) in their study on MRI contrast agents have demonstrated that the relaxivity of the contrast agent- Gd-DO3A-Lys was found to be significantly higher than the commercially available macro cyclic MRI contrast agent-Dotarem and Gd-DTPA based CA's. They also demonstrated that the CA conjugated to lysine was found to be sensitive enough for clinical evaluation on a 3T human scanner. In the same study the authors have also demonstrated that when DOTA was conjugated to the side chain of a lysine residue, it produced a major difference in bio-distribution by lengthening the half blood-life (68.1 min) of the compound. Signal enhancement of the GD-DO3A-Lys contrast agent remained high even at 52 min post injection (Yang and Chuang 2012). The authors also suggested that the mode of excretion was not only through the renal system but also *via* the reticuloendothelial system (RES- liver and spleen) and visceral organs such as the gastrointestinal (GI) tract (Yang *et al.* 2014; Shiraishi *et al.* 2010). Thus considering these findings, leucine was replaced by lysine in the following series of compounds.

The JS lysine series of peptide substrates were synthesized *via* a combination of both solid and solution phase peptide synthesis; the gadolinium complex was covalently coupled to the side chain of a lysine residue rather than the terminal amino acid as in the JS-Linear series. In order to increase the retention time of the contrast agent within the atherosclerotic plaques, a hydrophobic unit consisting of an aminoanthraquinone moiety coupled to a small diamino chain was conjugated

to the C-terminal of the lysine residue. Upon deprotection of the ε-Boc group from the lysine residue, the tert-butyl protected DOTA unit was coupled followed by removal of the capping unit from DOTA. The paramagnetic metal ion-Gadolinium (III) was then complexed to the ligand. It was anticipated that upon peptide hydrolysis by legumain the contrast agent should retain the hydrophobic unit, which may then increase the retention time of the gadolinium complex at the site of activation.

Fmoc-Asn(Trt)-NovaSyn TGT resin **a/b** → Fmoc-AA2-Asn(Trt)-NovaSyn TGT resin

(a/c) ↓

Fmoc-AA3-AA2-Asn(Trt)-OH ← **(d)** Fmoc-AA3-AA2-Asn(Trt)-NovaSyn TGT resin

Coupling to compound **56**

Fmoc-AA3-AA2-Asn(Trt)-Lys(DOTA-*t*Bu)-Spacer-AQ **(e)** → Fmoc-AA3-AA2-Asn-Lys(DOTA-OH)-Spacer-AQ

(f) ↓

Fmoc-AA3-AA2-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ

where,
(a)- Fmoc deprotection *via* 20% Piperidine in DMF, RT
(b)- Coupling to second amino acid (Gly/Ala); TBTU, HOBt, DIPEA in DMF, RT
(c)- Coupling of third amino acid (Pro/b-Ala); TBTU, HOBt, DIPEA in DMF, RT
(d)- Removal of resin; 5% TFA in DCM, RT
(e)- Removal of *tert*-Butyl and trityl protecting groups from DOTA and Asparagine; TFA, RT
(f)- Adding the gadolinium metal ion to the peptide substrate; water, neutral pH, RT.

Scheme 7: Chemical synthesis of the MRI based peptide contrast agent-JS lysine series.

The protocol for the synthesis of the following compounds involved (scheme 7):

(i) The synthesis of the (DOTA)-Lys-Spacer-AQ conjugate.

(ii) The synthesis of a tri-peptide sequence *via* solid phase peptide synthesis.

(iii) The next step was the coupling of the tri-peptide sequence to the (DOTA)-Lys-Spacer-AQ conjugate.

(iv) The target gadolinium-peptide substrate-AQ conjugate was synthesized by coupling Gd(III)Cl$_3$(6H$_2$O) to the peptide substrate.

2.1.2.2.1. Synthesis of Gadolinium complex using 1,4,7,10-Tetraaza cyclododecane-1,4,7,10-tetraacetic acid (Compound 58)

I. Synthesis of Fmoc-Lys(Boc)-Spacer-AQ (Compound 54)

(54)
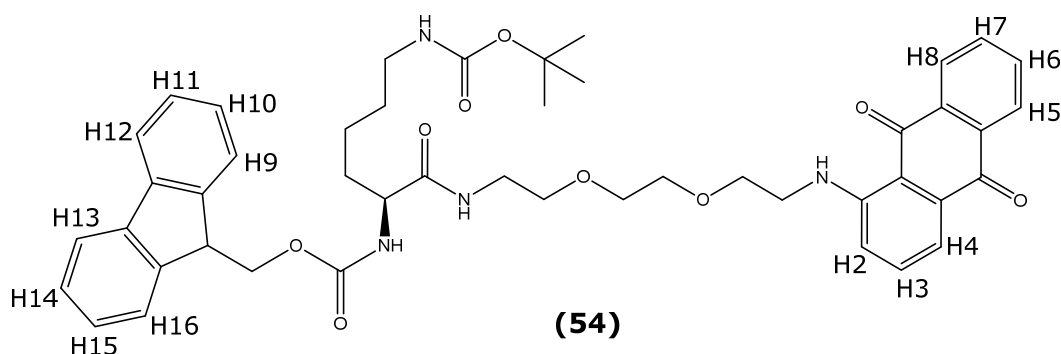
The Boc-protected conjugate (**54**) was synthesized by reacting *N*-Fmoc-Lys(Boc)-OH with compound **5** in DMF along with in situ coupling agents TBTU, HOBt and DIPEA. Reaction progress was monitored by TLC analysis; the crude product was then partitioned between DCM and water (x3), and washed with saturated sodium hydrogen carbonate (x3) and dried with anhydrous magnesium sulphate. The yield for the sample was found to be 81%.
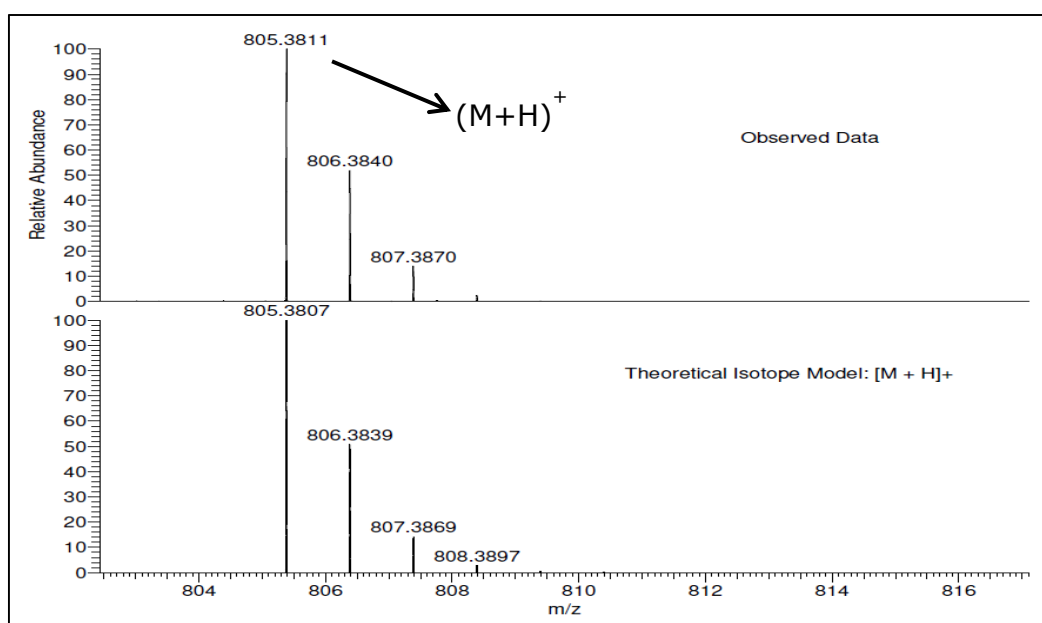
Figure 53: Mass spectrometry result for compound **54**.

The product **54** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 805.3811 (M+H)$^+$ corresponding to a molecular mass of 804 (figure 53). Furthermore, the structure of the compound was confirmed by ^{1}H NMR (d$_6$-DMSO) spectrum, which showed, for example, the presence of the ^{t}Boc-protecting group along with the γ-methylene group and the δ-methylene

group of the lysine side chain as a thirteen-proton multiplet between 1.17ppm and 1.40ppm. Also a two-proton multiplet between 1.46 ppm and 1.62ppm was assigned to the β-methylene group of the lysine side chain. Signal for a two proton multiplet between 2.82ppm and 2.89ppm was assigned to the ε-methylene group protons of the lysine side chain. A two proton multiplet between 3.17ppm and 3.29ppm was assigned to C$\underline{H}_2$-NH-Lys protons. A triplet at 3.45ppm was assigned the two protons of C$\underline{H}_2$-CH$_2$-NH-Lys. A multiplet signal for six protons between 3.49ppm and 3.61ppm were assigned to C$\underline{H}_2$-NH-AQ, O-C$\underline{H}_2$-C$\underline{H}_2$-O of the spacer arm. A two proton triplet at 3.70ppm was assigned to C$\underline{H}_2$-CH$_2$-NH-AQ. A single proton multiplet between 3.90ppm and 3.95ppm was assigned to the α-H methine proton of the lysine side chain. A three proton multiplet between 4.17ppm and 4.29ppm was assigned to C$\underline{H}_2$-C$\underline{H}$-Fmoc group protons. A signal proton triplet at 6.76ppm was assigned to N$\underline{H}$-^{t}Boc proton. A multiplet for three protons between 7.26ppm and 7.33ppm was assigned to H2, H10 and H15 protons. Another multiplet between 7.36ppm and 7.45ppm for three protons was assigned to H4, H11 and H14. A multiplet for a single proton between 7.62ppm-7.66ppm was assigned to H3 proton of the anthraquinone group. A two proton multiplet signal between 7.70ppm and 7.73ppm was assigned to H9 and H16 protons. A five proton multiplet between 7.81ppm and 7.95ppm was assigned to H6, H7, H12 and H13 protons along with N$\underline{H}$-spacer arm protons. Two double doublet signals at 8.12ppm and 8.20ppm for one proton each was assigned to the H5 and H8 protons. Finally a single proton triplet at 9.78ppm was assigned to the anthraquinone amino group proton.

II. Synthesis of Fmoc-Lys-Spacer-AQ trifluoroacetate salt (Compound 55)

(55)
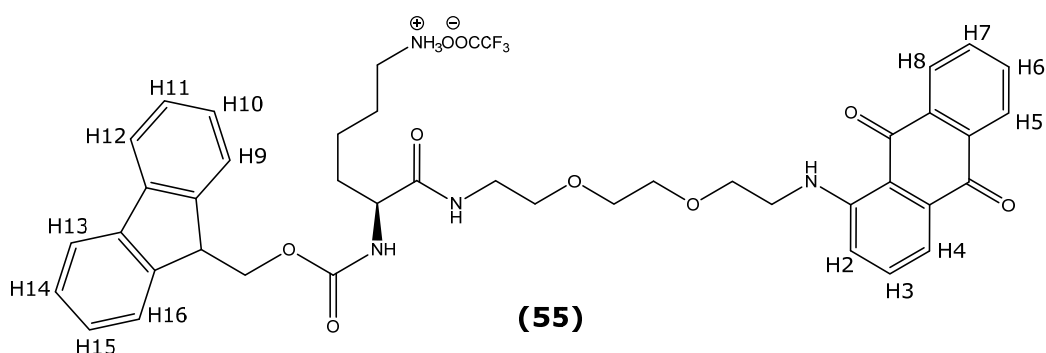
The Boc protected compound (**54**) was deprotected using trifluoroacetic acid. Reaction progress was monitored by TLC analysis, which showed the presence of a single new component, close to the base line of the TLC sheet suggesting the polar nature of the deprotected salt. The yield of the crude product was 54%.

Analytical HPLC for the compound (**55**) was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 95% (figure 54). The deprotected product **55** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 705.3280 (M+H)$^+$ corresponding to a molecular mass of 704 (figure 54). Furthermore, the removal of the ^{t}Boc-protecting group was confirmed by ^{1}H NMR (d$_6$-DMSO) spectrum, which showed the absence of the signal for the ^{t}Boc protecting group which was seen in compound **54**.
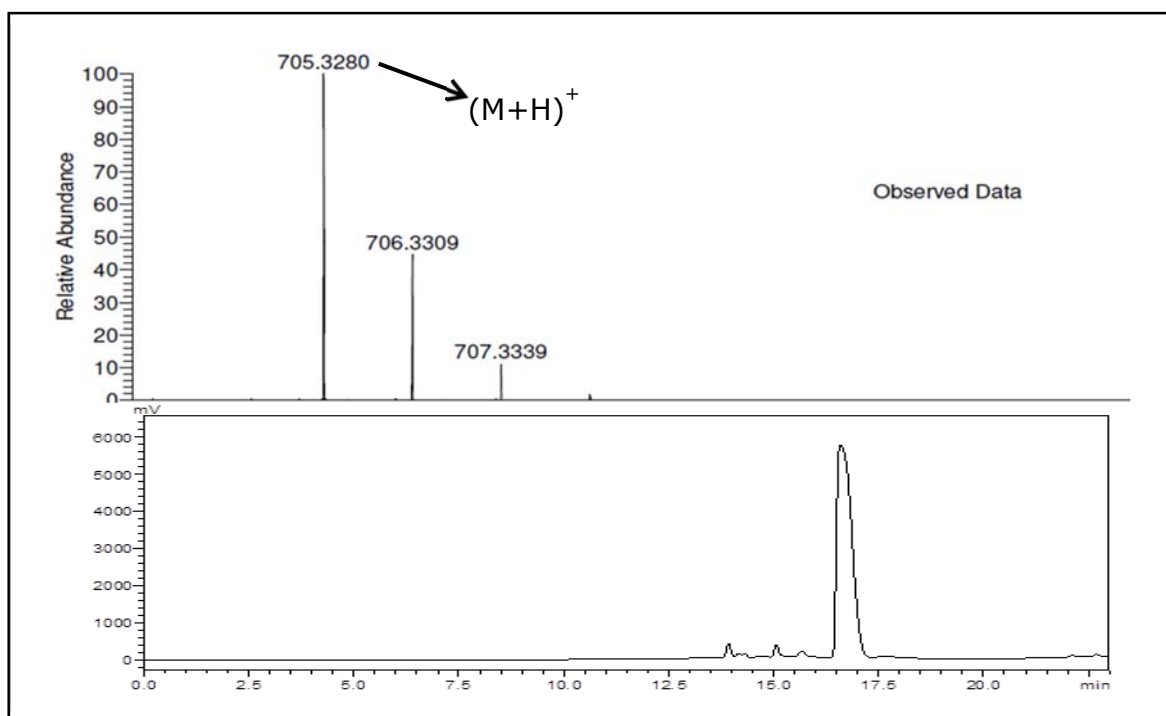
Figure 54: Mass spectrometry result for compound **55**. RP-HPLC elution profile of compound **55** (1mg/mL in methanol); Conditions: Flow rate, 1mL/min; Detection, 254nm; stationary phase: Eclipse XDB-C18 RP column; mobile phase A= 0.1%TFA/water; mobile phase B=0.1%TFA/methanol; gradient: low pressure gradient; 23min.

A multiplet for two protons between 1.27ppm and 1.32ppm was assigned to the γ-methylene group protons of the lysine side chain. Another multiplet for four protons between 1.51ppm and 1.61ppm was assigned to the β and δ methylene group protons of the lysine side chain. A two proton multiplet between 2.74 and 2.76ppm was assigned to the ε- methylene group protons. A two proton quartet at 3.24ppm was assigned to C$\underline{H}_2$-NH-COO protons. A triplet at 3.46ppm was assigned the two protons of C$\underline{H}_2$-CH$_2$-NH-COO. A multiplet signal for six protons between 3.51ppm and 3.61ppm were assigned to protons of C$\underline{H}_2$-NH-AQ, O-C$\underline{H}_2$-C$\underline{H}_2$-O of the spacer arm. A two proton triplet at 3.70ppm was assigned to C$\underline{H}_2$-CH$_2$-NH-AQ. A single proton multiplet between 3.92 ppm and 3.98ppm was assigned to the α-H methine proton of the lysine side chain. A single proton triplet

at 4.19ppm was assigned to C<u>H</u>-Fmoc proton and a two proton doublet at 4.26ppm was assigned to C<u>H</u>$_2$-CH-Fmoc protons. A multiplet for three protons between 7.26ppm and 7.33ppm was assigned to H2, H10 and H15. Another multiplet between 7.38ppm and 7.48ppm for three protons was assigned to H4, H11 and H14 protons. A multiplet between 7.62ppm and 7.73ppm was assigned to the H3 proton of the anthraquinone group and H9 and H16 protons of the Fmoc group. A four proton multiplet between 7.81ppm and 7.91ppm was assigned to H6, H7, H12 and H13 protons followed by a triplet at 7.94ppm for a single proton was assigned to N<u>H</u>-spacer arm-AQ. A double doublet at 8.13ppm and a doublet at 8.19ppm for one proton each were assigned to H5 and H8 protons. A single proton triplet at 9.77ppm was assigned to the anthraquinone amino group proton.

III. Synthesis of H$_2$N-Lys(DOTA-^{t}Bu)-Spacer-AQ (Compound 56)

(56)
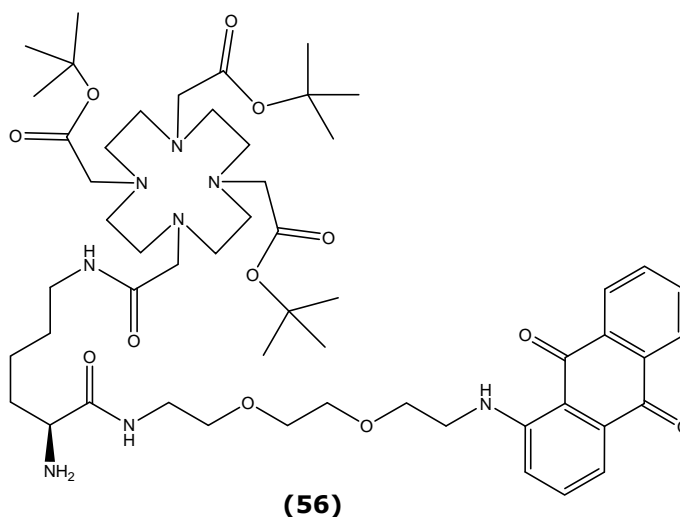
The *tert*-butyl protected intermediate (Fmoc-Lys(DOTA-*t*-Bu)-Spacer-AQ) was synthesized by reacting compound **55** with DOTA-tri-^{t}Bu-ester in DMF along with the in situ coupling agents: TBTU/HOBT and DIPEA. After purification by column chromatography the Fmoc-protecting group was removed by treating the compound with piperidine in DMF at rt. Reaction progress was monitored by TLC analysis and on reaction completion the crude product was purified by column chromatography. The dried sample was re-dissolved in a slightly polar solvent system containing 5% MeOH in DCM. The same solvent system was used to elute the front running impurities. The major product was then eluted by gradually increasing the polarity of the solvent system (up to 10% MeOH in DCM). The yield of the sample was found to be 28%. The purity of the product **56** was found to be 90% determined *via* analytical HPLC performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column. The product **56** was also characterized by high resolution ESI (+) mass spectrometry, which

gave a signal at m/z 1059.6071 (M+Na)$^+$ corresponding to a molecular mass of 1036 (figure 55) and the low resolution ESMS(+) gave a signal at m/z 519 [(M+2H)/2]$^{2+}$; ESMS (-) mass spectrometry gave a signal at m/z 1035 (M-H)$^-$ corresponding to the proton abstraction product.
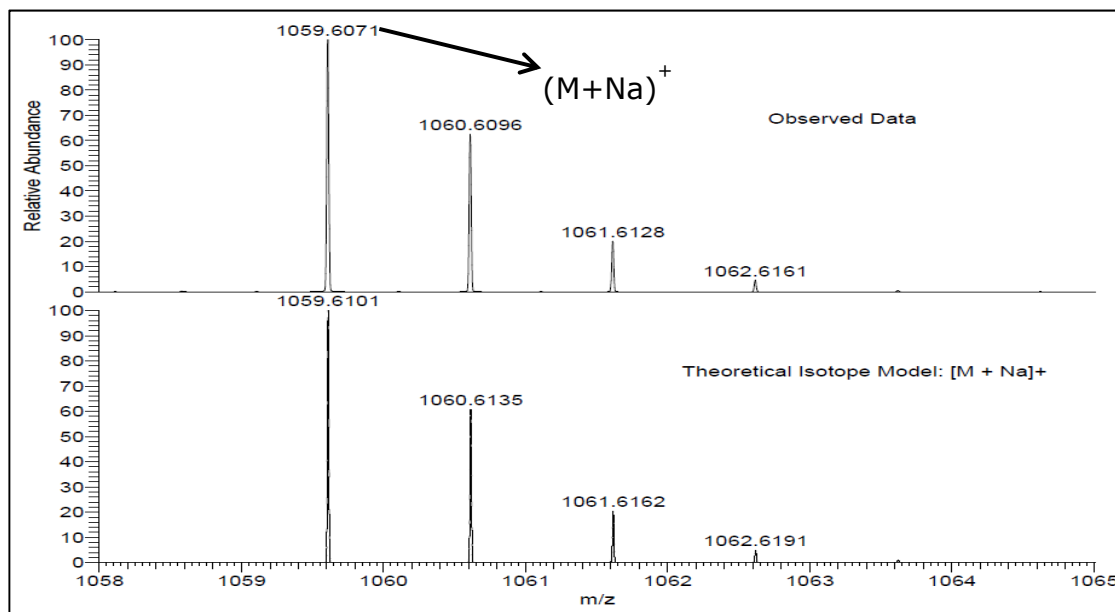
Figure 55: Mass spectrometry result for compound **56**.

IV. Synthesis of H$_2$N-Lys(DOTA)-Spacer–AQ (Compound 57)

(57)
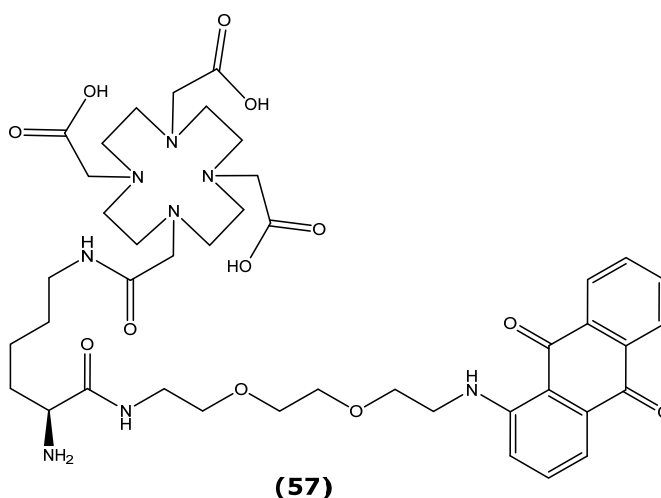
The *tert*-butyl protected compound (**56**) was deprotected using trifluoroacetic acid. Reaction progress was monitored by TLC analysis, and on completion the residual TFA was evaporated under reduced pressure. The purity of the product **57** was found to be 95% determined *via* analytical HPLC performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column. The deprotected product **57** was characterized by high resolution ESI (-) mass spectrometry, which gave a signal at m/z 867.4259 (M-H)$^-$ corresponding to a molecular of 868 (figure 56).

- 115 -
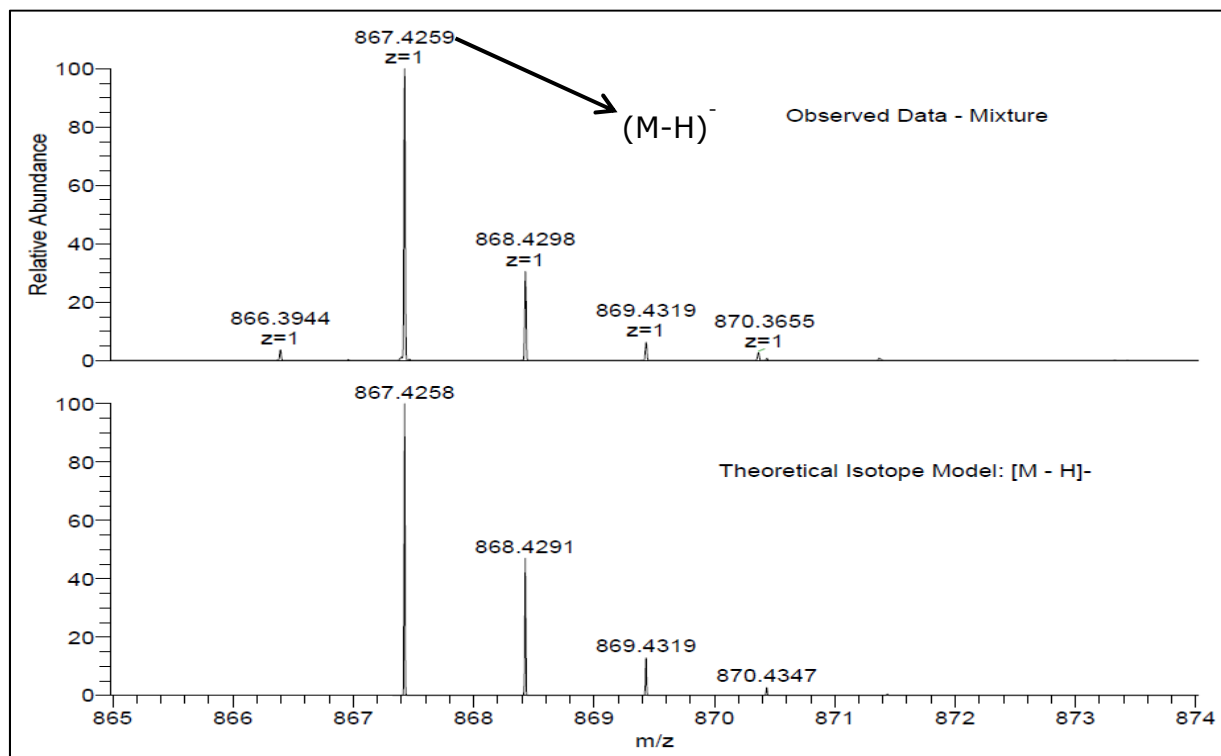
The low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 869 $(M+H)^+$ corresponding to a molecular mass of 868 and also a signal at m/z 435 $[(M+2H)/2]^{2+}$.

V. Synthesis of H$_2$N-Lys(DOTA-Gd^{3+})-Spacer-AQ (Compound 58)

(58)
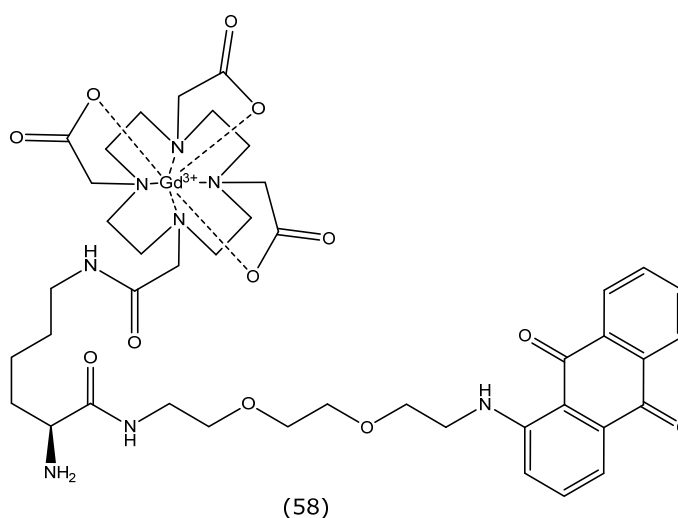
Synthesis of the target gadolinium complex of DOTA-Lysine-aminoanthraquinone spacer conjugate (**58**) was performed in accordance to the reaction conditions mentioned in scheme 7. The Gd complex was prepared in a neutral aqueous solution by reacting GdCl$_3$(6H$_2$O) with compound **57**. The reaction pH was adjusted to 7 by adding a solution of ammonia 0.35% drop wise. Reaction progress was monitored by TLC analysis and the crude product was purified by column

chromatography. During purification it was observed that the compound was found to have very high affinity to the silica gel in the column and even by increasing the polarity of the eluting solvent system it failed to move through the silica gel. The silica slurry from the glass column was collected and washed several times with 10% methanol in DCM to collect the sample. The filtered solution was dried at rt and then characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1024 (M+H)$^+$ corresponding to a molecular mass of 1023; ESMS (-) mass spectrometry gave a signal at 1022 (M-H)$^-$ corresponding to the proton abstraction product and also a signal at m/z 1057 for the chloride adduct (M+Cl)$^-$

2.1.2.2.2. Synthesis of JS-Lysine series of tetra peptide substrates

The design template for the JS lysine series was Fmoc-AA3-AA2-AA1-Lys(DOTA-Gd)-Spacer-AQ, where AA represents amino acid. The synthesis of the JS-Lysine series of tetra peptide substrates is shown in an abbreviated form in scheme 7. The tripeptide sequence corresponding to positions P1, P2 and P3 were synthesized *via* SPPS. A NovoSyn®TGT resin preloaded with asparagine was used for SPPS because the peptide chain can be cleaved from the resin by treatment with 5% TFA in DCM. This is vital because higher concentration of TFA could cause the removal of the side chain protecting group from Asn before the synthesis of the target compound. Moreover, unprotected Asn could undergo a dehydration reaction on the free primary amide causing it to form a nitrile group.

Comp. no.	P3	P2	P1	P1′	
62	Fmoc-Pro	Gly	Asn	Lys-(DOTA-Gd)	
66	Fmoc-β-Ala	Gly	Asn	Lys-(DOTA-Gd)	
70	Fmoc-Pro	Ala	Asn	Lys-(DOTA-Gd)	
74	Fmoc-β-Ala	Ala	Asn	Lys-(DOTA-Gd)	
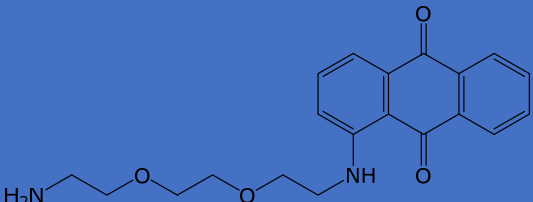
Table 9: List of tetra peptide sequence for JS-lysine series.

This dehydration reaction could modify the structure of Asn in the final compound thus making it unproductive as the enzyme can only recognize and cleave at the

carboxyl end of asparagine. This JS-lysine series of compounds included the selection as shown in table 9.

A. Synthesis of Fmoc-Pro-Gly-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (Compound 62)

I. Synthesis of Fmoc-Pro-Gly-Asn(Trt)-OH (Compound 59)
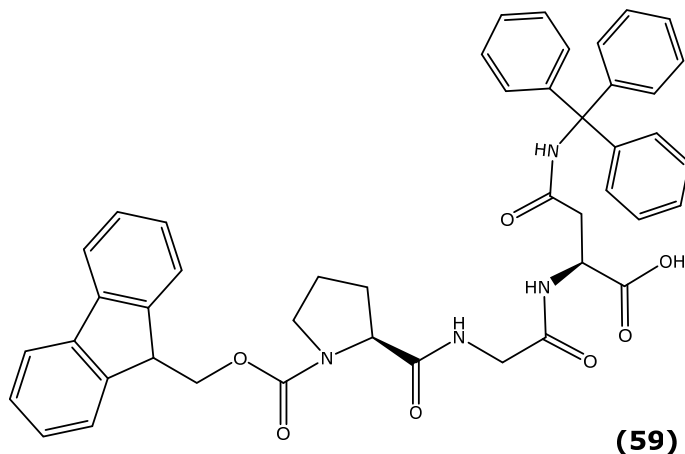
(59)

An asparagine pre-loaded resin was reacted in cycles of coupling and deprotection with Fmoc protected amino acids in the following order: *N*-Fmoc-Glycine and *N*-Fmoc-L-Proline. After the final coupling, the peptide sequence was removed from the resin using trifluoroacetic acid (5%) in DCM. Analytical HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 70% and the yield was 75%. Product **59** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 751.3120 (M+H)$^+$ corresponding to a molecular mass of 750.

II. Synthesis of Fmoc-Pro-Gly-Asn(Trt)-Lys(DOTA-^{t}Bu)-Spacer-AQ (Compound 60)

The *tert*-butyl protected aminoanthraquinone-tetra peptide sequence (**60**) was synthesized by treating the tri-peptide sequence (**59**) with compound **56** in DMF along with the coupling agents TBTU, HOBt and DIPEA (scheme 7). The crude product was purified by column chromatography. The yield after column chromatography was found to be 43%.

(60)
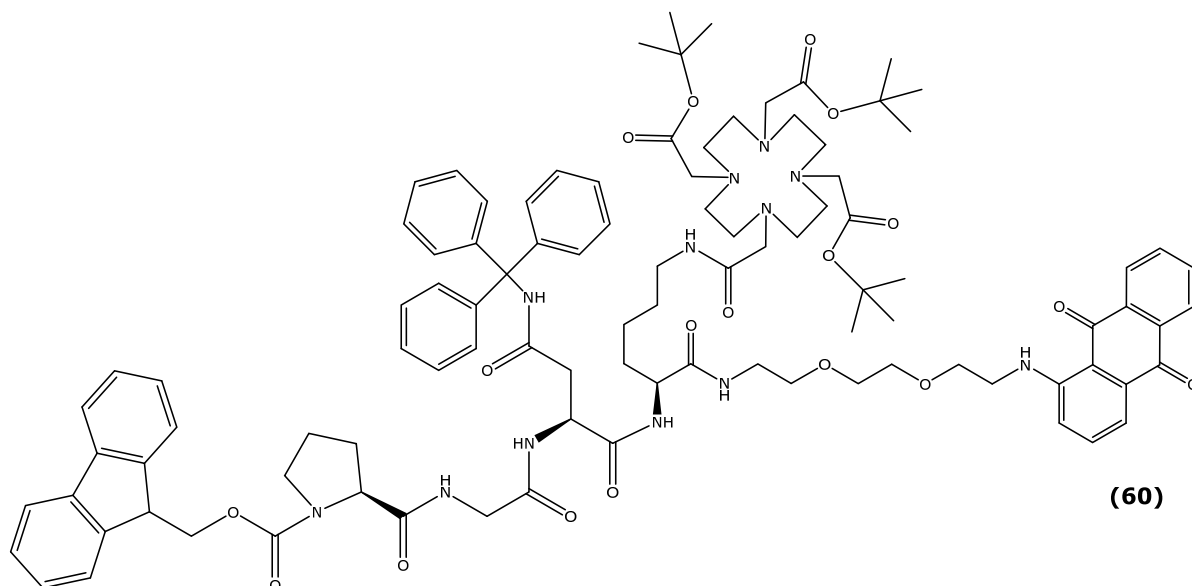
The compound **60** was characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1769 (M+H)$^+$ corresponding to a molecular mass of 1770 and also a signal at m/z 885 [(M+2H)/2]$^{2+}$; ESMS (-) mass spectrometry gave a signal at m/z 1803 corresponding to the chloride adduct (M+Cl)$^-$.

III.   Synthesis   of   Fmoc-Pro-Gly-Asn-Lys(DOTA-OH)-Spacer-AQ (Compound 61)

(61)
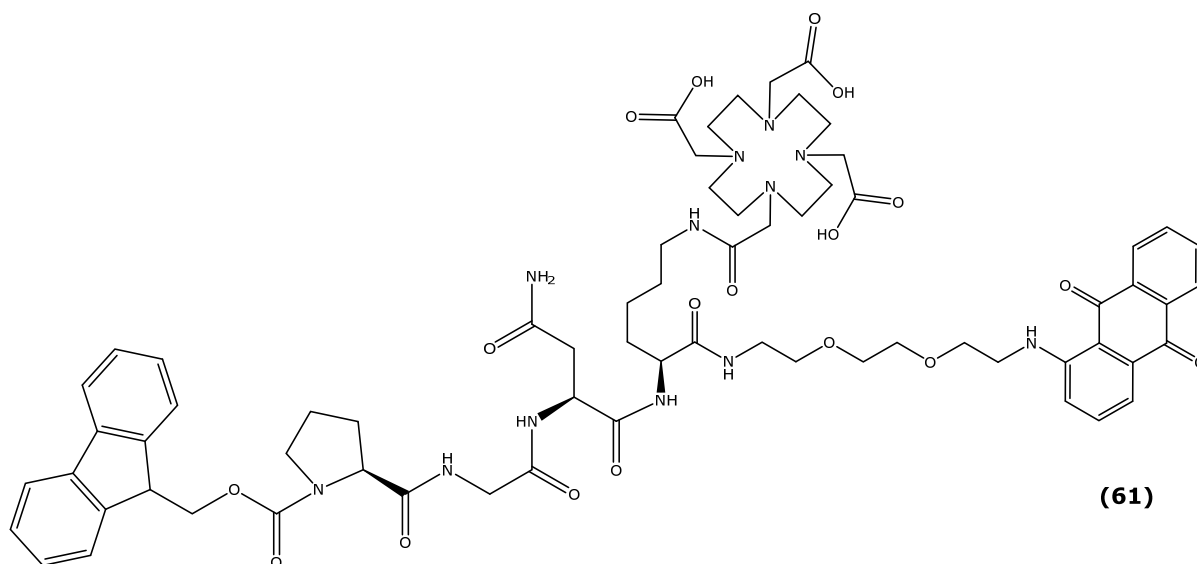
The γ-trityl and *tert*-butyl protected aminoanthraquinone-DOTA-tetra peptide conjugate (**60**) was deprotected using trifluoroacetic acid (scheme 7). Reaction progress was monitored by TLC analysis, which showed the presence of a single new component lying close to the base line indicating the polar nature of the

deprotected product. The deprotected product was deemed sufficiently pure (by TLC) to be used without further purification and the yield of the crude sample was 80%. The deprotected product **61** was characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1359 (M+H)$^+$ corresponding to a molecular mass of 1358 and also a signal at 680 [(M+2H)/2]$^{2+}$ and another signal at 1381 corresponding to the sodium adduct (M+Na)$^+$; ESMS (-) mass spectrometry gave a signal at 1357 (M-H)$^-$ corresponding to the proton abstraction product.

IV. Synthesis of Fmoc-Pro-Gly-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (Compound 62)

(62)
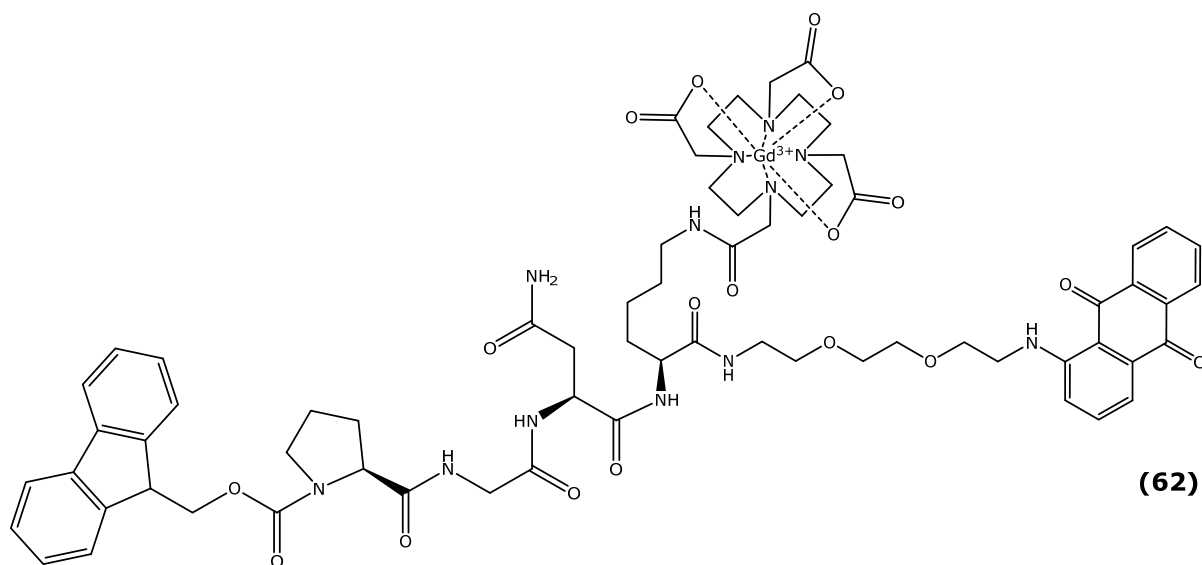
Synthesis of the target gadolinium complex of DOTA-tetrapeptide-amino anthraquinone spacer conjugate (**62**) was performed in accordance to the reaction conditions mentioned in scheme 7. The Gd complex was prepared in a neutral aqueous solution by reacting GdCl$_3$(6H$_2$O) with compound **61**. The reaction pH was adjusted to 7 by adding a solution of ammonia 0.35% drop wise. Reaction progress was monitored by TLC analysis. On completion the reaction mixture was reduced to low volume and then purified by column chromatography. Analytical HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 45%. The product **62** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1514.5264 (M+H)$^+$ corresponding to a molecular mass of 1513 and also a signal at m/z 1536.5082 corresponding to the sodium adduct (M+Na)$^+$ (figure 57). A signal at m/z 779 [(M+2Na)/2]$^{2+}$ was observed on the low resolution ESMS(+).
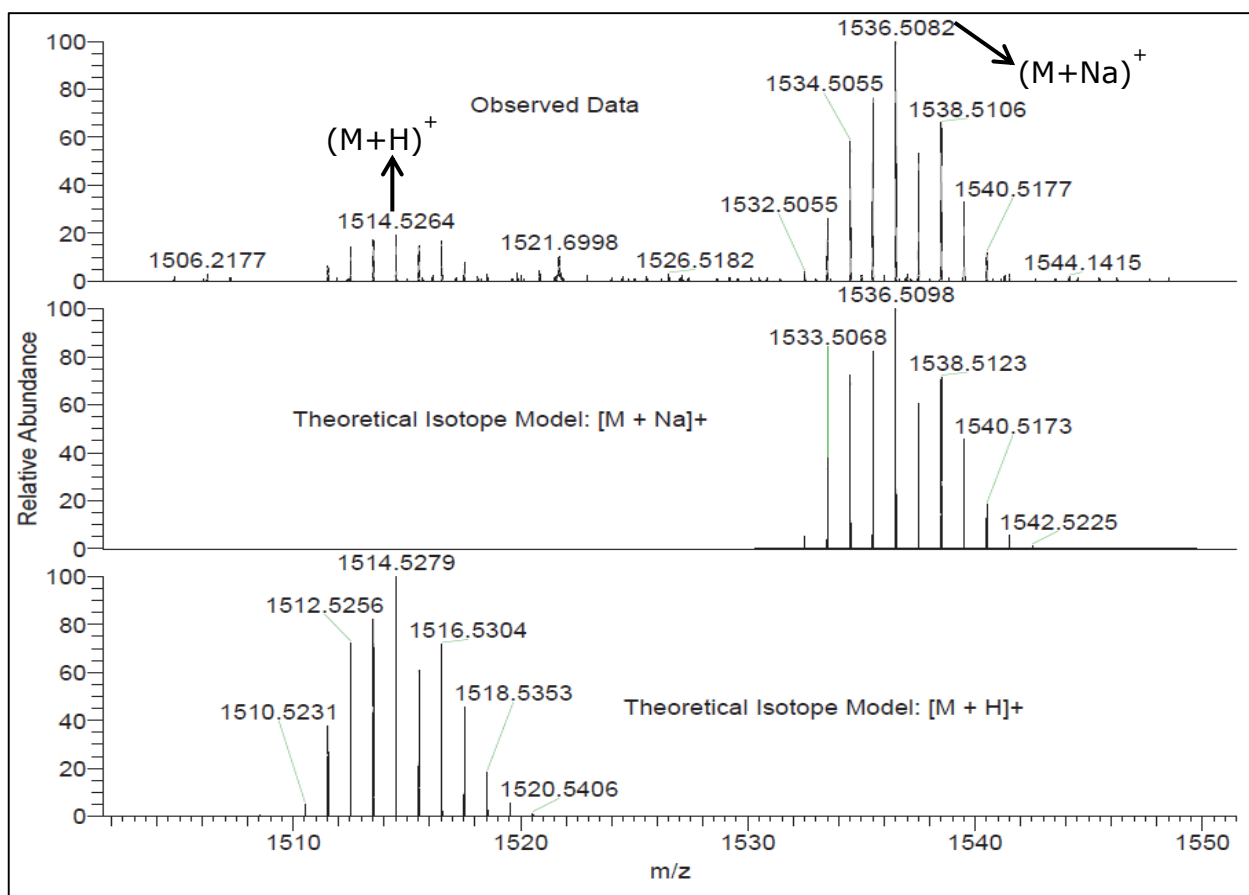
Figure 57: Mass spectrometry result for compound **62**.

B. Synthesis of Fmoc-β-Ala-Gly-Asn-Lys(DOTA-Gd³⁺)-Spacer-AQ (Compound 66)

(66)
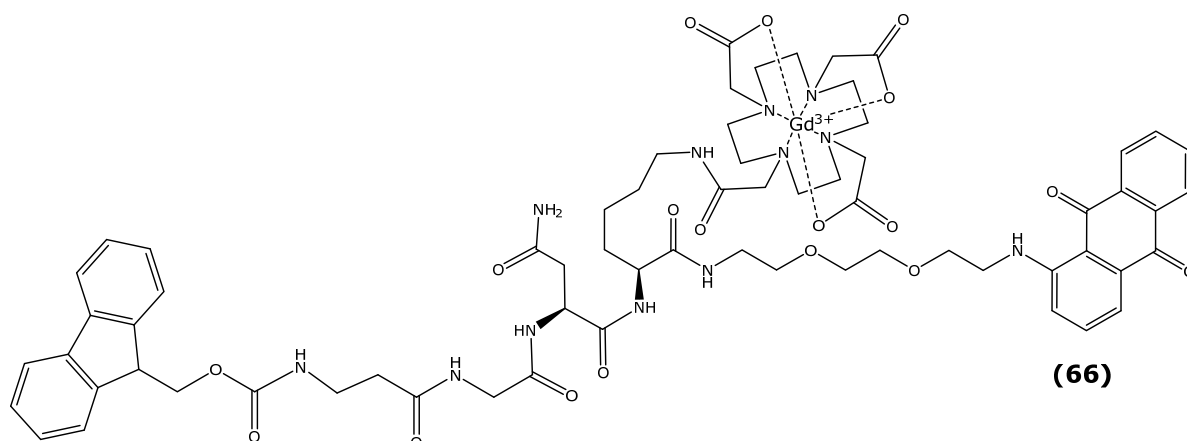
The chemical synthesis of compound **66** was similar to compound **62**. The only difference was that proline at P3 was swapped for β-alanine and the rationale for this change has already been discussed in section 2.1.1.3. The coupling and deprotection methods and the reagents were the same as summarized in scheme

7. The Gd complex was prepared in a neutral aqueous solution by reacting GdCl$_3$(6H$_2$O) with the intermediate compound 65. The reaction pH was amended to 7 by adding a solution of ammonia 0.35% drop wise. Reaction progress was monitored by TLC analysis. On reaction completion the residual solvent was dried at rt. A white precipitate was found along with the dry sample and hence water was added to remove the salt like deposits, which may be due to the use of ammonia in the reaction mixture. The solution was found to contain a few un-dissolvable particles and hence on filtering the mixture it was decided to characterize both separately. The yield was found to be 33%. Low resolution ESMS mass spectroscopy on the red solution gave required peaks and hence it was decided to evaporate the filtered solution. It was also observed that the precipitates collected during filtration were not dissolvable in most solvents and hence it was left aside for further analysis. The dried solution was then characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1533 (M+2Na-H)$^+$ corresponding to a molecular mass of 1487 and also a signal at m/z 1520 (M+CH$_3$OH+H)$^+$.

C. Synthesis of Fmoc-Pro-Ala-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (Compound 70)

(70)
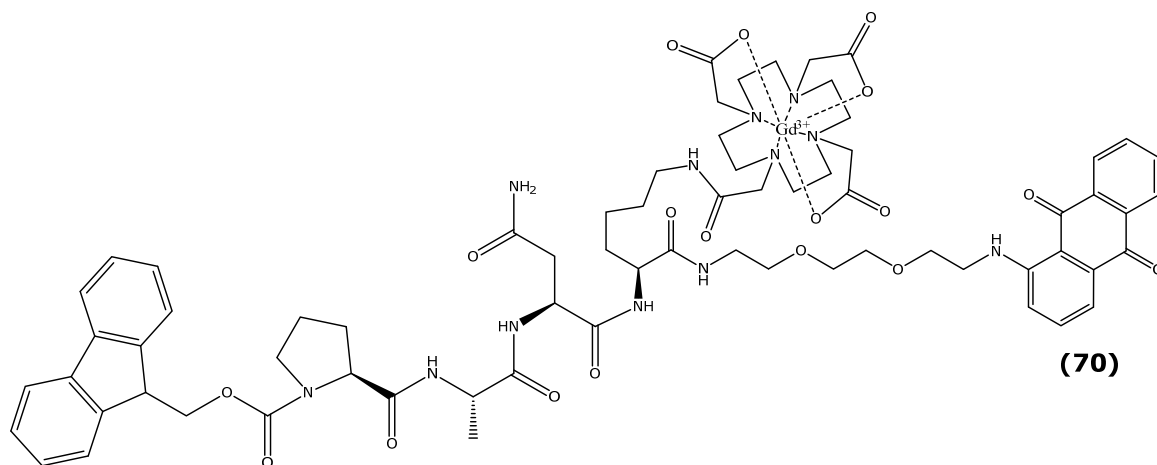
The chemical synthesis of compound **70** was similar to compound **62**. The only difference was that glycine at P2 was swapped for alanine and the rationale for this change has already been discussed in section 2.1.1.3. The coupling and deprotection methods and the reagents were the same as summarized in scheme 7. The target compound **70** was characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1528 (M+H)$^+$ corresponding to a molecular mass of 1527 and also a signal at m/z 1550 corresponding to the sodium

adduct (M+Na)$^+$. Signal at m/z 764.7751 [(M+2H)/2]$^{2+}$ (figure 58) was observed on as high resolution ESI (+) mass spectrometry. The yield of the post column chromatographic product was 29%.
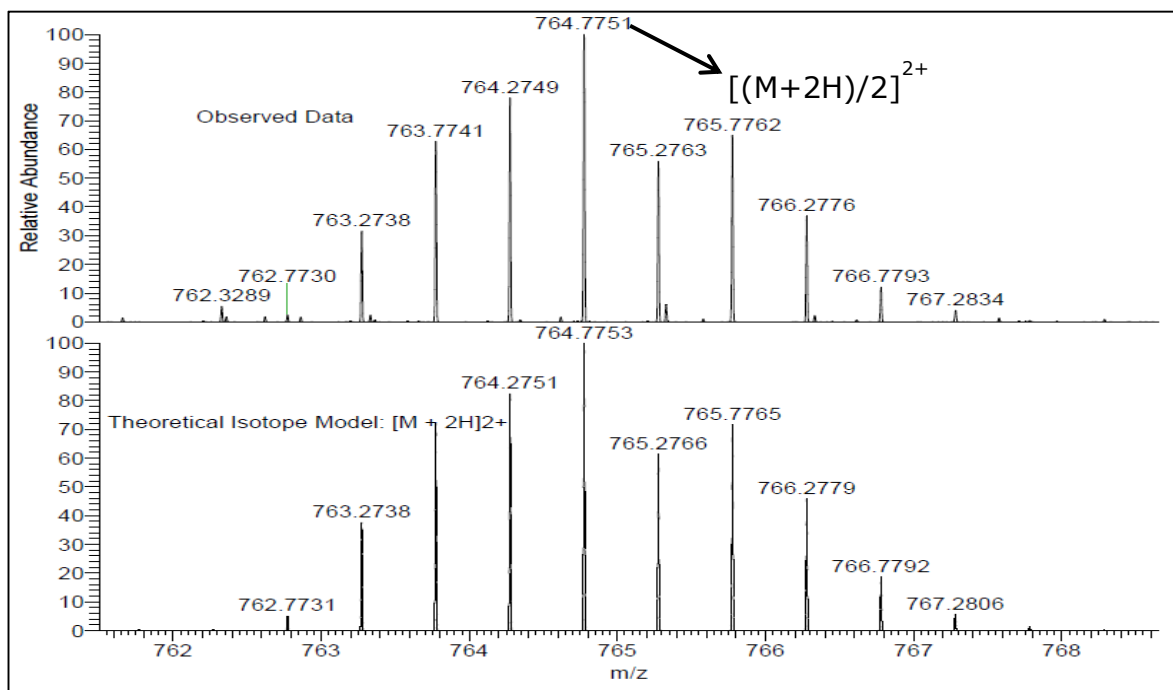
Figure 58: Expansion of ESI (+) spectrum of compound **70** that shows the isotopic pattern of the molecular ion (doubly-charged at half-mass).

D. Synthesis of Fmoc-β-Ala-Ala-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (Compound 74)

(74)
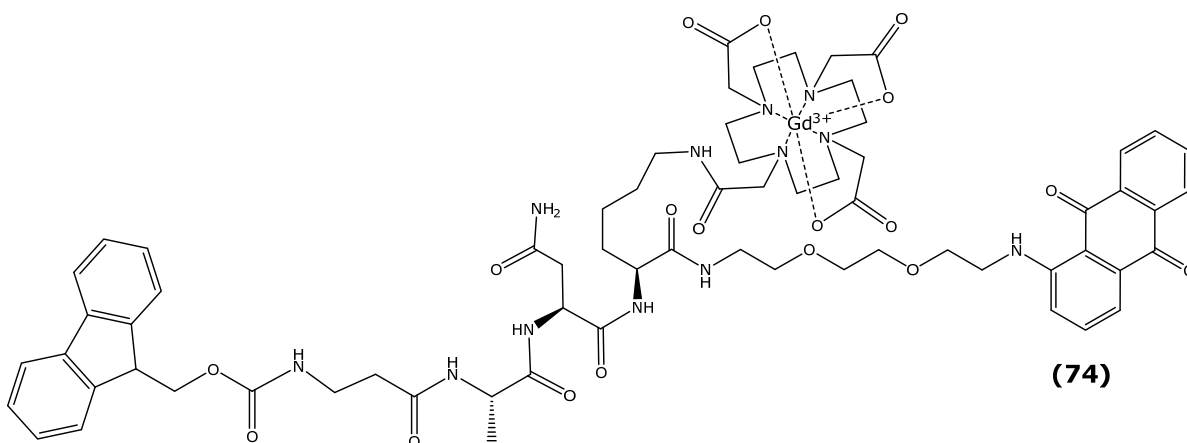
The chemical synthesis of compound **74** was similar to compound **66**. The only difference was that glycine at P2 was swapped for alanine and the rationale for choosing this amino acid has already been discussed in section 2.1.1.3. The coupling and deprotection methods and the reagents were the same as summarized in scheme 7. Analytical HPLC was performed on a Shimadzu

Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column and the purity of the sample was found to be 65%. The target compound **74** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1502.5225 (M+H)$^+$ (figure 59) corresponding to a molecular mass of 1501. A signal at m/z 1524 corresponding to the sodium adduct (M+Na)$^+$ and another signal at m/z 751 [(M+2H)/2]$^{2+}$ was observed on the low resolution ESMS (+) mass spectrometry.
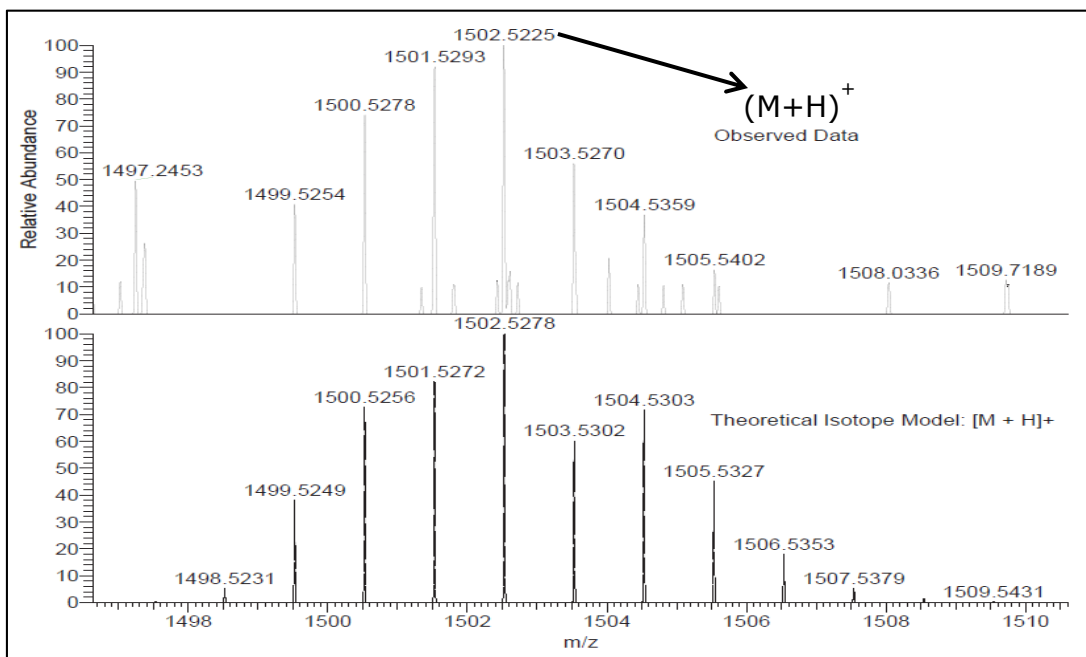
Figure 59: Expansion of the ESI (+) mass spectrum of compound **74** that shows the isotopic pattern of the molecular ion.

2.1.2.3. Legumain activity assay for JS series

Final volume in each eppendorf: 600µL

◆ Assay buffer: 352.3µL
◆ Substrate: 7.7µL
◆ rh Legumain: 240µL
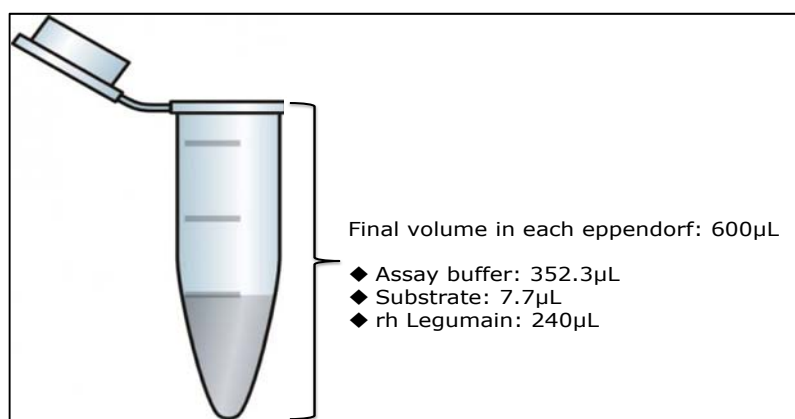
Figure 60: Assay conditions for JS series of compound.

The legumain activity assay was performed to analyse if the compounds of the JS library (linear and lysine series) were recognized by legumain. The contrast agent conjugated tetra peptide compounds (**49, 53, 62, 66, 70** and **74**) were incubated

with activated legumain in sterile eppendorf tubes individually (figure 60). The volume in each tube was set to 600µL to reduce any purification and characterization difficulties due to volume. After two hours of incubation with activated legumain at 37°C, the enzyme digested product was basified using triethylamine and then extracted into chloroform and washed with water. This was done several times and the organic extracts were collected and combined. Both organic/aqueous extracts were dried at room temperature. The dried products of each sequence were then characterized by both low and high-resolution mass spectrometry to identify the molecular mass of the fragmented compound.

2.1.2.3.1. Mass spectrometric results for compound 49 (JS-linear series)
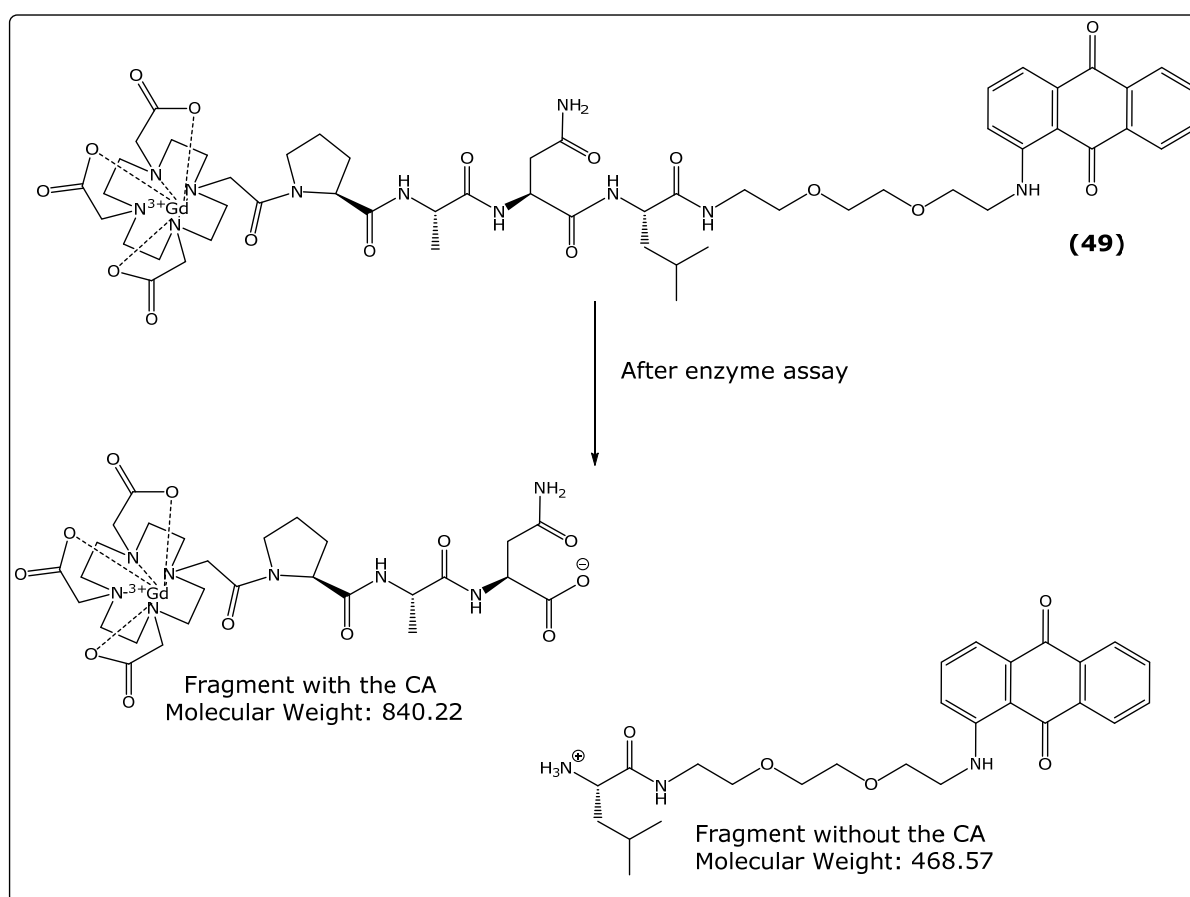
Figure 61: Schematic representation of compound **49** (JS-linear series) before and after enzyme assay.

Compound **49** was incubated with activated legumain and the expected result is shown in figure 61. Mass spectrometric analysis on compound **49** was undertaken to identify the presence of molecular masses 468.25 (fragment without gadolinium complex) and 840.22 (fragment with gadolinium complex). From the low-resolution mass spectrometric analysis (ESMS), it was shown that the compound had been cleaved by legumain producing the required mass 841.4 $(M+H)^+$ corresponding to a molecular mass of 840 as well as its sodium 863.2 $(M+Na)^+$

and potassium 880.6 (M+K)$^+$ adduct. Negative mode ionization (low-resolution MS) for compound **49** also revealed the presence of the fragment with the gadolinium complex with the corresponding signals: 840 (M)$^-$, 875.6 (M+Cl)$^-$, 886.6 (M+FA-H)$^-$, 954.8 (M+TFA-H)$^-$ and also the fragment without the gadolinium complex was observed with the signals for 467 (M-H)$^-$ corresponding to a mass of 468 as well as the chloride adduct 503 (M+Cl)$^-$, and also 511.2 (M+FA-H)$^-$.

However, in contrast to the low-resolution mass spectrometry, the high-resolution mass spectrometry (ESI) analyzed by +ve MALDI using alphacyan hydroxycinnamic acid and DCTB matrices (targeting the peptide and organo-metallic portions for ionization) did not yield the appropriate molecular ions using either matrix. At the same time, the high-resolution ESI-MS results did not show the presence of the intact compound suggesting that the tetra peptide sequence could have been cleaved by legumain. These results suggest the need for a better purification technique to isolate the enzyme-digested product.

2.1.2.3.2. Mass spectrometric results for compound 70 (JS-lysine series)

Similarly, low-resolution mass spectra for compound **70** also gave peaks corresponding to both the fragments. Compound **70** was analyzed for the presence of molecular masses 522.20 (fragment without gadolinium complex) and 1024.34 (fragment with gadolinium complex) (figure 62). The low-resolution mass spectrum suggested that the compound had been cleaved by legumain producing the required signal at 511 [(M+2H)/2]$^{2+}$ corresponding to a molecular mass 1024. Negative mode of ionization (low-resolution MS) for compound **70** also indicated the presence of the fragment with the gadolinium complex with the corresponding signals: 1023 (M-H)$^-$ and 1059 (M+Cl)$^-$ and also the fragment without the gadolinium complex with the peak for 521 (M-H)$^-$ corresponding to a molecular mass 522. However, the high-resolution mass spectrometric (ESI) results conflicted with the low-resolution mass spectrometric (ESMS) results. For ESI the compound (**70**) was dissolved in Methanol and analyzed by nanospray on the Orbitrap, but did not yield the appropriate molecular ions. But the ESI mass spectrometric results showed no signals relating to the molecular mass of the intact compound, suggesting that the tetra peptide sequence was cleaved by legumain. Similar results were observed for compounds **62**, **66**, **74** and **53** confirming the fragmentation of the peptide sequence by the enzyme but lacking confirmation of the presence of required molecular masses.
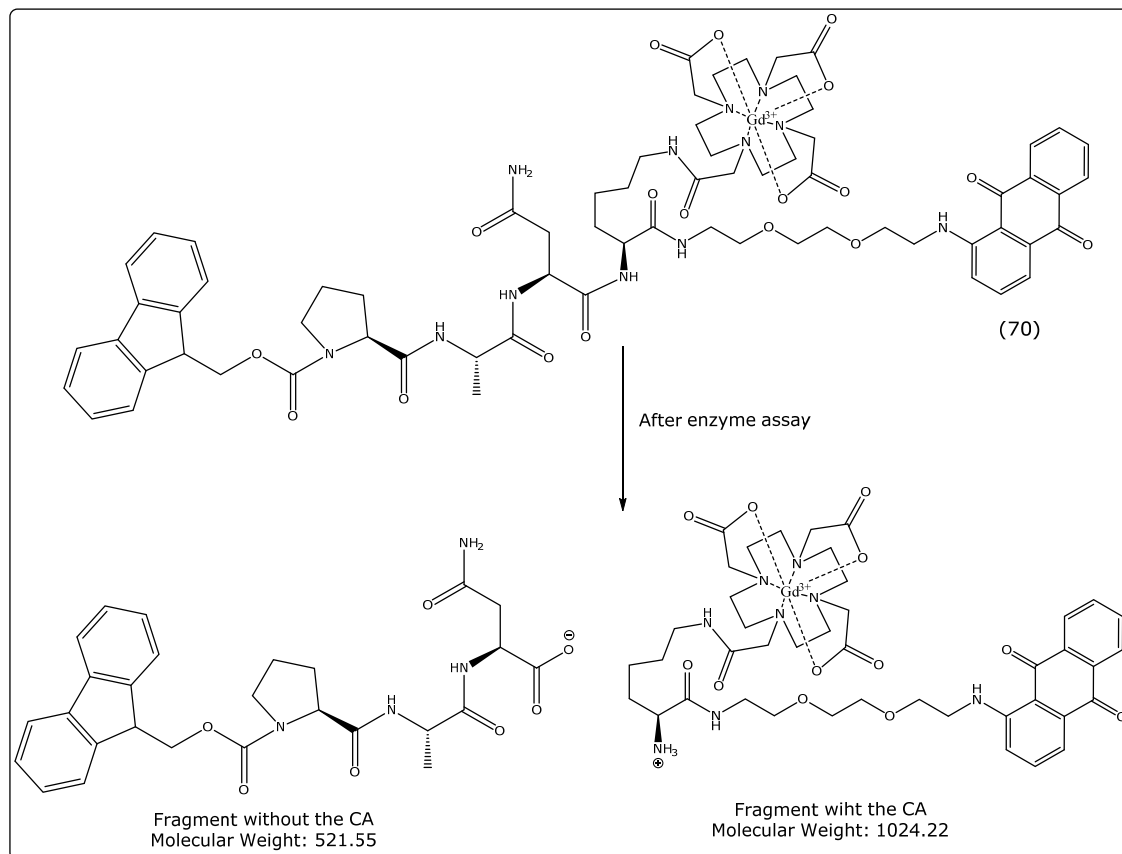
Figure 62: Schematic representation of compound **70** (JS-lysine series) before and after enzyme assay.

From the results obtained it can be suggested that the compounds synthesized had been fragmented by legumain. This could be indicated by the presence of peaks corresponding to the enzyme-digested product, with and without the contrast agent (Gd-DOTA), spotted on the low-resolution mass spectrometric report. However, the high resolution ESI mass spectrometric analysis failed to show the presence of peaks relating to the enzyme digested product. This result may suggest that the enzyme could have hydrolysed the compounds (JS series) but the need for a better purification technique prior to mass spectrometric analysis was necessary. Some of the most commonly used methods for clean up to prepare a sample for analysis post enzyme assay are reversed-phase chromatography in a pipet tip (ziptips), spin column (ultramicrospin columns) or spring column setup (Sep-Pak cartridges) (Gundry *et al.* 2010).

2.1.2.4. Cytotoxicity assay

Studies on the 'drug-like' properties of newly synthesized chemical compounds, during the process of drug development, are important in order to avoid any drug failures during clinical trials (Hamid *et al.* 2004). Evaluation of toxicity of a drug candidate in different human cell lines is vital for its potential use as an imaging agent for biomedical studies (Chen *et al.* 2014). This can be performed using *in*

vitro cytotoxicity assays, which are available commercially and make the process of screening numerous drug libraries easy and quick. Cell viability and proliferation can be evaluated by a number of methods that can measure various cellular responses and effects (Cook and Mitchell 1989). The common assays to study *in vitro* toxicity are lactate dehydrogenase (LDH) release (Decker and Lohmann-Matthes 1988), 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) metabolism (Mossmann 1983); neutral red uptake (Morgan *et al*. 1991) and adenosine triphosphate (ATP) content (Crouch *et al*. 1993).
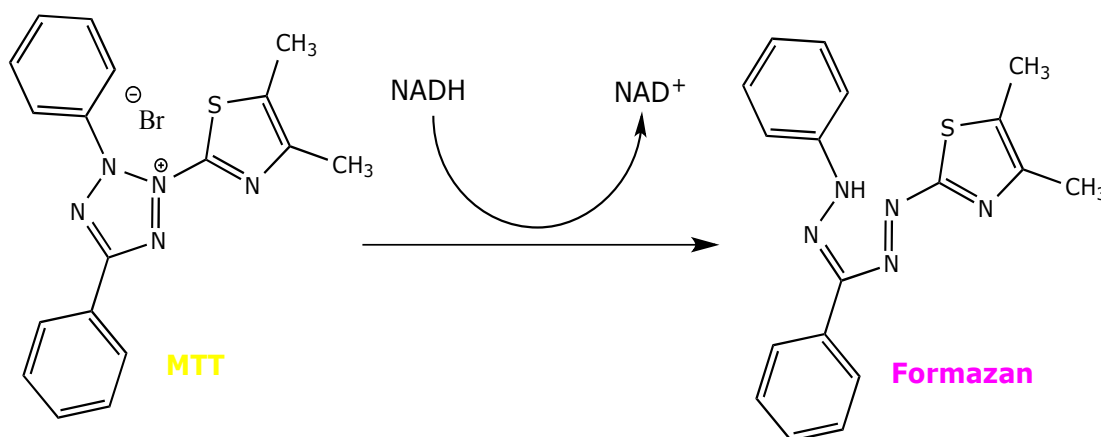
Figure 63: Scheme depicting the conversion of MTT to the metabolized formazan product.

Most modern assays have been adapted for microtiter plates (for example, 96 well plates) and these miniaturizations conveniently allow more than one sample, at a time, to be examined easily and quickly (Weyermann *et al*. 2005). Assays based on traits such as luminescence and colorimetric measurements allow samples to be assayed directly in a microtiter plate using an ELISA plate reader. In this research project, cytotoxicity was determined using the MTT assay. It is one of the most widely used techniques in the investigation of a compound's toxicity mainly due to its simple, quick, sensitive and semi-automated characteristics (Mossmann 1983; Denziot and Lang 1986). MTT is a water-soluble tetrazolium salt, which when added to viable proliferating cells is reduced to insoluble purple formazan crystals (figure 63). These purple crystals are formed by the cleavage of the tetrazolium ring due to the presence of succinate dehydrogenase in the mitochondria of living cells (Fotakis and Timbrell 2005). The MTT metabolite formazan, which is impermeable to the cell membrane, accumulates in viable cells (Berridge and Tan 1992). This accumulation allows the formazan crystals to be solubilized and facilitates their quantification by spectrophotometry (Barron *et al*. 2010). The absorbance value recorded is directly proportional to the number of viable cells (Isobe *et al*. 2001; Pozzolini *et al*. 2003).

2.1.2.4.1. MDA-MB-231 Cells

Human breast epithelial cancer (MDA-MB-231) cells were isolated from a sample of pleural effusion taken from a patient who was undergoing a radical mastectomy in the early 1970s (Cailleau *et al*. 1974). MDA-MB-231 cells demonstrate invasive properties and are commonly used as a model for estrogen receptor negative breast cancer with a function in the evaluation of tumorigenicity, metastasis and cell invasion. These cells are categorized as triple negative/basal-B mammary carcinoma cells (Tate *et al*. 2012). The in-house MDA-MB-231 cells were selected for the MTT assay due to their immortal and proliferative nature and also to determine the effect of the synthesized compound on cell viability (Kong, personal communication).

2.1.2.4.2. Cytotoxicity studies in MDA-MB-231 cells

Figures 64, 65, 66 and 67 show the cell viability (%) of MDA-MB-231 cells when treated with compounds-**62**, **66**, **70** and **58** (0-40µM) after 24 and 72 hours by using the MTT assay.
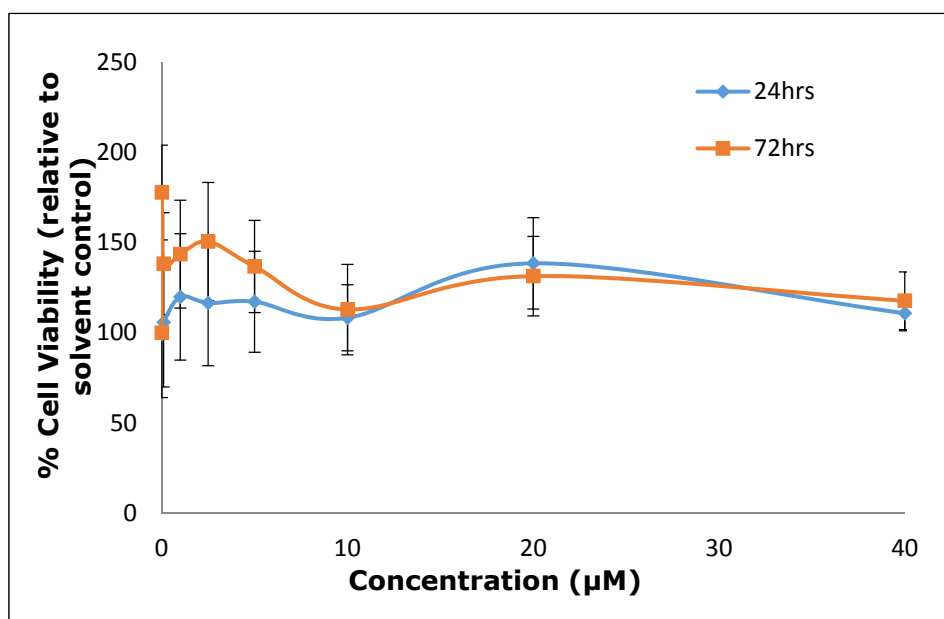
Figure 64: Cell viability of **62** in MDA-MB-231 cells after 24 and 72 hours at 37°C. Data are represented as mean ± SEM of 3 independent experiments (n=3), each experiment consisted of 6 replicates/concentration.
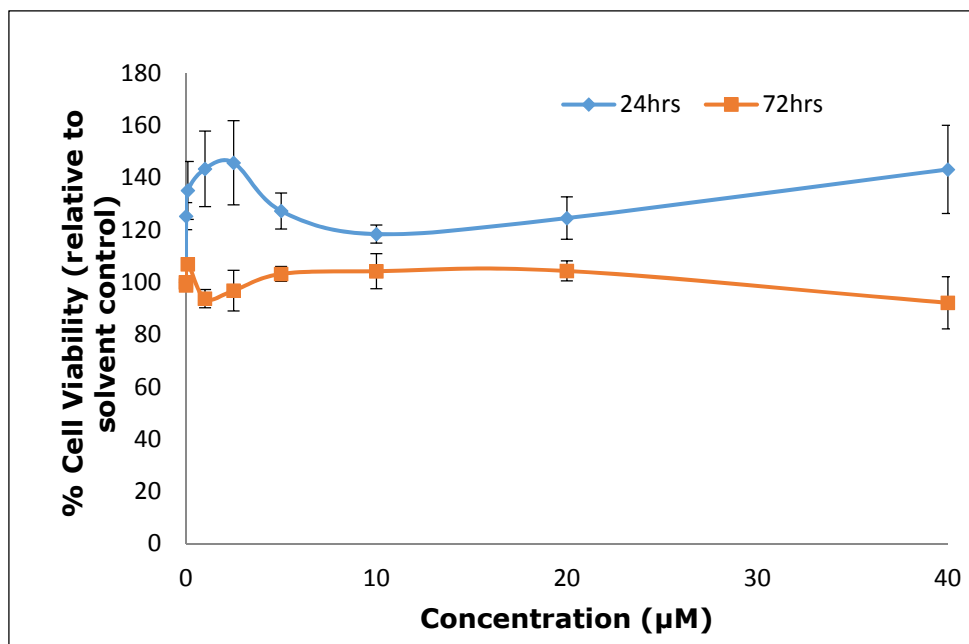
Figure 65: Cell viability of **66** in MDA-MB-231 cells after 24 and 72 hours at 37°C. Data are represented as mean ± SEM of 3 independent experiments (n=3), each experiment consisted of 6 replicates/concentration.
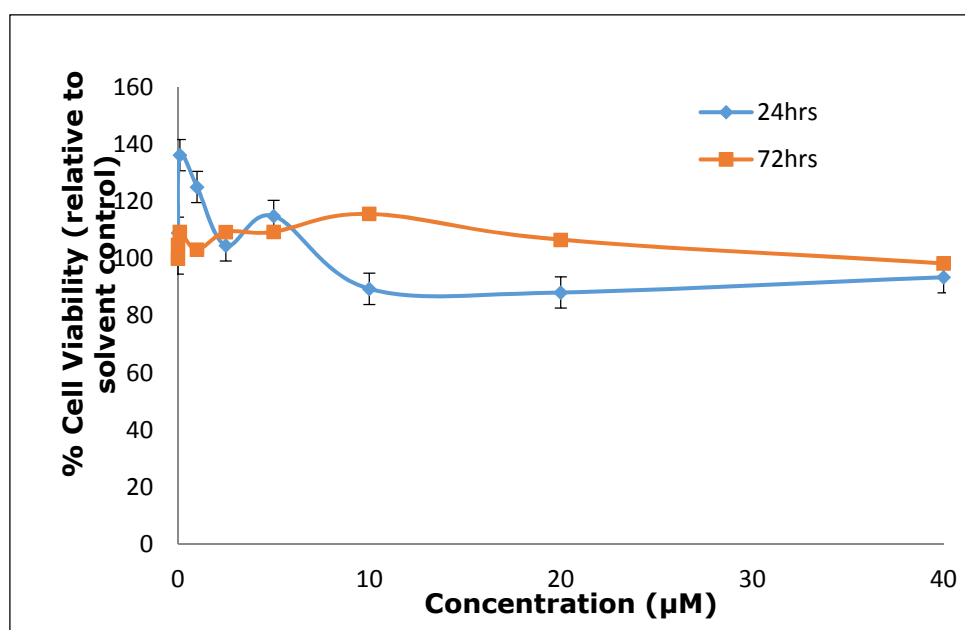
Figure 66: Cell viability of **70** in MDA-MB-231 cells after 24 and 72 hours at 37°C. Data are represented as mean ± SEM of 3 independent experiments (n=3), each experiment consisted of 6 replicates/concentration.
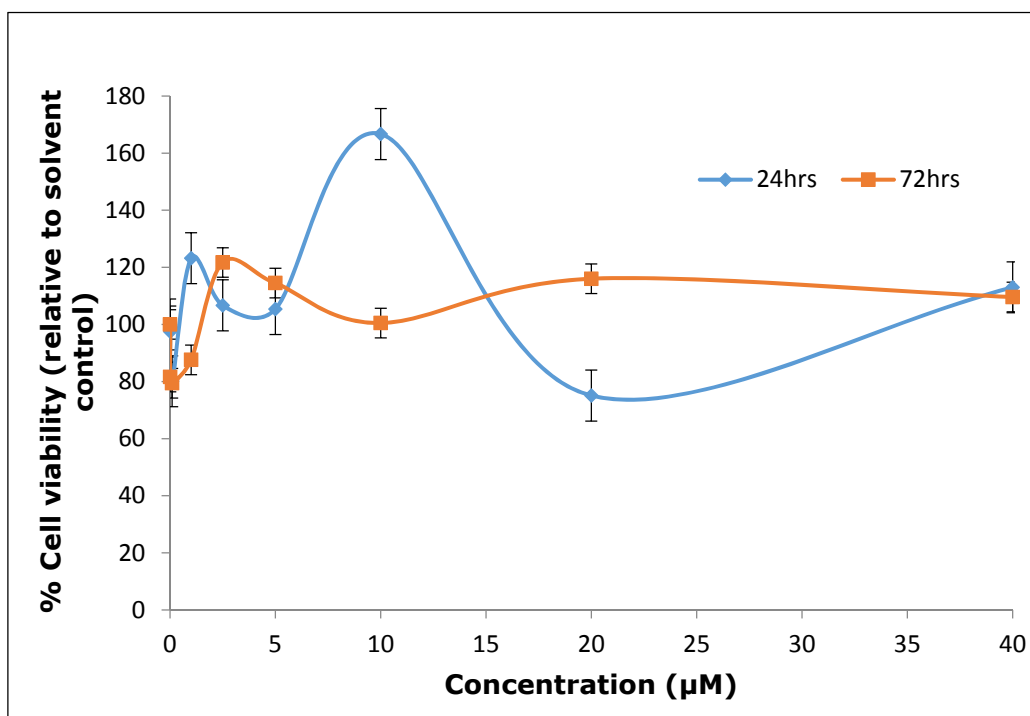
Figure 67: Cell viability of **58** in MDA-MB-231 cells after 24 and 72 hours at 37°C. Data are represented as mean ± SEM of 3 independent experiments (n=3), each experiment consisted of 6 replicates/concentration.

The results showed that **62**, **66**, **70** and **58** did not affect cell viability even at higher concentrations (up to 40μM) in MDA-MB-231 cells. Though compound **62** did not affect cell viability, the values suggest a small variation between the experiments which may be due to experimental errors and hence the MTT assay for compound **62** should be repeated. Data for compound **58** which represents the part of the peptide sequence with the gadolinium chelate (contrast agent) released after the intact compound (for example, **62**) was treated with legumain suggest that it does not affect cell viability. It can be concluded that the compounds synthesized during this research program were found not to affect cell viability of MDA-MB-231 cells. In similar work on synthesizing a peptide based MRI contrast agent and a fluorescent probe to detect legumain activity in tumors, Chen and colleagues (2014) have demonstrated that the MRI contrast agent of the structure Gd-NBCB-TTDA-Leg(L) and a fluorescent probe of structure CyTE777-Leg(L)-CyTE807 were non-toxic to CT-26 and 3T3 cell lines. Similarly, the results of the MTT assay in this study have demonstrated that both the intact compounds (**62**, **66**, **70**) and the paramagnetic gadolinium complex (**58**) which will be released upon cleavage by legumain, were found not to affect the viability of MDA-MB-231 cells and hence can be used for further drug development as a diagnostic tool for atherosclerosis plaques.

2.1.3. Conclusion

The last decade has seen a steady increase in the knowledge and understanding of legumain expression and activity leading to progression of atherosclerotic plaques. Evidence suggested that legumain was directly or indirectly involved in plaque progression either by ECM disruption in an uncontrolled manner or by activation of other proteases such as MMP-2 or cathepsin which also degraded the ECM. Moreover, it was also discussed that legumain showed strict specificity towards asparagine and was active only under acidic to near neutral pH. Hence any diagnostic interventions by exploiting these up-regulated legumains in identifying and treating atherosclerotic plaques would be beneficial.

From the information available on legumain substrate specificity a library of FRET peptide compounds were designed. By utilizing the principle of FRET, which is based on the transfer of energy between a donor (5(6)-CBF) and acceptor (AQ) within close proximities a library of tetra peptide substrates (**JC series**) were synthesized in such a way that they had substrate specificity for legumain. The JC series of compounds was synthesized *via* a combination of both solid and solution phase peptide chemistry. Fluorophore quenching studies on selected compounds (**11** and **31**) demonstrated no release of fluorescence from the compounds in the absence of the enzyme (legumain) suggesting that they could be effective FRET compounds for legumain assay. In another study on FRET probes for detecting legumain activity, it was demonstrated that proline, glycine and asparagine at P3, P2 and P1 position of the peptide sequence produced results similar to the commercial available legumain substrate Z-Ala-Ala-Asn-AMC (Ding, 2014). Hence when the compounds from the JC library were incubated with activated legumain, a gradual increase in fluorescence release was recorded over time. This could have been due to the hydrolysis of the peptide substrate leading to the increase in distance between the FRET pair. The results demonstrated that some of the FRET peptide substrates synthesized during the study were recognizable by legumain. Results from the legumain assay further suggested that the second batch of JC compounds having lysine at P1′ was better substrate than the first batch of compounds which had leucine at that position. This result was also supported by the experiments conducted by Schwars *et al.* (2002) who showed that leucine at P1′ was less effective compared to lysine. Moreover results from the legumain assay also suggested that the preference for glycine at P2 was favourable over alanine. Finally compound **46** which was a penta-peptide was also found to be recognized by legumain by generating a maximum fluorescence intensity released between 800 to 1200 RFI which was comparable with compound **11**. This result

suggested that the length of the peptide sequence has no effect on legumain specificity towards the peptide substrate. Although the results from the legumain assay demonstrate that the FRET peptide substrates were recognized by the enzyme there is a need for improvement in the purity of the compounds. This can be achieved by purifying the compounds by preparative HPLC, which could be an important technique to inform about the purity of the target compounds.

The sequence design of the JC series of compounds was utilized as a model for synthesizing the library of JS compounds. This series of compounds comprised of a tetra peptide sequence (mimicking the JC library) covalently coupled to the MRI contrast agent (Gd-DOTA). This design was undertaken in order to achieve the aim of this research project, which focuses on the synthesis of a non-invasive contrast agent for imaging atherosclerotic plaques. The design of the JS series of compounds was based on the concept that enzyme cleavage of the peptide substrate may lead to the release of the CA and also enhance the hydrophobicity of the CA. Accumulation of the CA at the site of activation could improve the contrast of the surrounding region leading to better image resolution of the atherosclerotic plaques. The JS series of compounds were incubated with the activated enzyme (legumain) and the enzyme-digested product was characterized *via* mass spectrometric analysis. The results of the legumain assay demonstrated that the activated enzyme fragmented the peptide substrates. However, as mentioned in the previous paragraph, it is suggested that the target compounds have to be purified by preparative HPLC. In this manner, the success of the enzyme-substrate assay can be strictly related to the design of the compounds.

Evaluating the cytotoxicity of a chemical drug is an important factor for using them in biomedical studies. Selected compounds (**62**, **66**, **70** and **58**) from the JS series were evaluated for cytotoxicity against MDA-MB-231 cells. The results of the toxicity assay demonstrated that the compounds did not affect the viability of the cells even at higher concentrations of the drug (40µM). However the need for a control to compare the results obtained is desirable. It could be suggested that these compounds may form a new platform of targeted contrast agents for imaging atherosclerotic plaques *via* the application of MRI by specifically targeting the biomarker over expressed in the plaque regions. Thus in this research study MRI contrast agent (Gd-DOTA) conjugated to peptide substrates specific to legumain that is over expressed in unstable atherosclerotic plaques have been successfully synthesized and characterized by spectroscopic analysis.

2.1.4. Limitations

The main limitation of this work was the purity of the peptide sequences synthesized *via* SPPS. The purity of the compounds ranged between 44-80% (synthesized *via* SPPS) and the purity of the target compounds synthesized *via* fragment condensation and solution peptide synthesis ranged from 60-97%. Hence purification of the peptide substrates by preparative HPLC would be desirable. Another problem to explore during the fragment condensation technique is racemization of the amino acid residues. Racemization is a serious issue as it is not easy to detect and more importantly difficulties arise trying to separate residues containing R-amino acids from the S-amino acids. Though racemization suppressants such as HOBt were used during the synthesis process, it must be agreed that there is always a risk for racemization to occur during fragment condensation technique. However racemization cannot occur if the synthesis of the peptide chain proceeds from the C-terminal end, because the carboxyl group is protected by an unreactive group (Doonan 2002). Also, choosing appropriate amino acid residues at the C-terminal position during fragment condensation can also prevent racemization. This can be done by choosing glycine at the C-terminus since it is the only amino acid that lacks a chiral carbon and hence the formation of azlactone which is the main route to racemization by carboxyl activation is prevented (Nyfeler 1994). Proline can also be used at the C-terminus because the structure of proline provides resistance to base-catalysed racemization (Jones 1992).

The other limitation regarding the JS series of compounds was that the purification of the enzyme digest product did not show the presence of the required molecular mass of the product on high resolution mass spectrometry. This highlights the need for another purification technique besides the conventional solvent extraction method followed during the study. The process of purification is important to desalt and concentrate peptide mixtures before any analysis by mass spectrometry because these impurities may hinder peptide ionization and cause chemical noise in the mass spectra (Gundry *et al*. 2009). Some of the most commonly recommended methods are reverse phase chromatography in a pipet tip (Ziptip), spin column or syringe column format (SPE-solid phase extraction columns). For example, RP-SPE columns consist of a non-polar stationary phase and a moderately polar mobile phase can be used. The solution containing the analyte of interest is poured through the column and the interaction between the sample and the hydrophobic functional group on the silica surface causes retention of the

analyte in the column. The sample can then be eluted by using a non-polar solvent which breaks the interaction between the sample and the column. Similarly other separation methods based on charge (ion exchange chromatography), molecular size (gel chromatography), and specific binding group (affinity chromatography) can also be performed to address the problems encountered during analysis of the enzyme digest product in this research study.

2.1.5. Future work

Further work on the determination of the efficacy of the substrates (JS library) towards legumain is desirable. Upon determining the efficacy of the peptide substrate-contrast agent conjugates synthesized during this research study, they may then be utilized as powerful tools for the evaluation of atherosclerotic plaques. Furthermore, analysis on the release of the contrast agent in human atherosclerotic tissue extracts could be the next step in exploiting these compounds for *in vivo* imaging. Moreover, coupling these imaging units to therapeutic drugs could open a new approach for delivering drugs selectively to the diseased tissue and concurrently monitoring treatment progress and side effects.

These small sized molecules (JS series) consisting of a plaque recognizable unit and a gadolinium chelate could benefit from easy and quick endocytosis by the plaques, fast blood clearance leading to excretion of the toxic metal ion and the reproducible nature of these substrate based contrast agents will be one of the key features for transition from lab to clinical analysis (Sosnovik and Caravan 2009). However, one major drawback with such small sized imaging molecules is that the biomarkers (legumain) around the microenvironment of the abnormal tissue are expressed in submicromolar levels. In order to produce signal enhancement the need for a self-assembly approach with multiple (1000s) gadolinium chelates along with a plaque recognizable unit is desirable.

The substrate based probes synthesized during this research study could be used as prototypes to design activity based probes (ABPs) which function by covalently altering the proteases in an activity dependent manner (Edgington et al. 2011). These probes could be used for live cell imaging as the ABPs remains attached to the target biomolecule allowing dynamic analysis on enzyme function and localization. Furthermore the synthesized peptide substrates could be manipulated in such a way as to be used as dual imaging agents via the application of both MRI and optical imaging. The dual probe system could be designed to carry a FRET pair along with a contrast agent covalently coupled to the side chain of an amino acid

residue on a peptide sequence engineered to be specific to the protease/enzyme of interest (Puthenedam 2010). Finally, more biological analysis to understand the turnover kinetics and to evaluate the intercellular contrast agent release properties of the chemical compounds, along with cytotoxic studies (MTT, LDH, ATP content evaluation) on human tissue extracts of atherosclerotic lesions is desirable. This can be achieved by developing optimized assay conditions for treating legumain expressing cell lines such as HEK293 cells and HCT116 colorectal cells known for expressing active legumain (Smith 2014) with the synthesized chemical compounds.

2.2. MMP (2/9) targeted gadolinium based contrast agents for MR imaging of atherosclerotic plaques

The aim of this part of the research study was the synthesis of contrast agents, which exhibited specificity towards enzymes/proteases that are overexpressed in atherosclerotic plaques. The production of excellent MRI contrast between the normal and diseased region depends on endogenous variances in water content and relaxation time of the CA. Several non-specific contrast agents are widely used for examining physiological parameters in clinical applications (Mishra *et al*. 2006; Calcagno *et al*. 2013). However, when it comes to monitoring specific functions/irregularities, the need for a functionalized/targeted CA is essential (Mishra *et al*. 2006) as in the case of atherosclerotic lesions. In order to achieve this (target specific/activatable CA's), biomolecules could be conjugated to the gadolinium chelates (Gd-DTPA/Gd-DOTA). These targeted CAs could potentially be used to image pre-rupture atherosclerotic plaques by detecting the specific enzyme/ protease concentrated in the shoulder region of the plaques *via* a MRI scanner. The targeted CAs are said to localize to the target molecules of interest with great specificity and affinity, ultimately increasing the diversity of MR applications (Mishra *et al*. 2006). The targeted contrast agents can further increase the relaxation time by binding to specific targets (protein/enzyme) and inducing water proton relaxation rate enhancement (Caravan *et al*. 1999).

Gadolinium based contrast agents with MMP recognizable substrates were synthesized potentially to increase sensitivity and delay blood clearance time assisting in accumulation of the CA at the site of activation. Upon activation by the specific MMP (2/9); these CA's produce signal amplification by increasing the rate of relaxation of the surrounding water protons at the site of activation (Lepage *et al*. 2007). Chen and colleagues (2005) in their studies on the detection of MMP activity in atherosclerotic plaques successfully synthesized a fluorogenic probe to monitor protease activity *via* the application of optical imaging. The probe consisted of a peptide substrate cleavable by MMP-9, which was conjugated to a fluorophore (FITC). The probe was injected into mice, which suffered an induced infarction. The images obtained, highlighted the affected areas due to the accumulation of the fluorogenic probe when compared to the control. A similar kind of probe was used to detect gelatinase (MMP-2/9) activity in the aorta of Apo lipoprotein (ApoE-/-) deficient mice with atherosclerotic lesions. This probe produced fluorescence in the presence of the protease confirming MMP activity (Deguchi *et al*. 2006). In another study based on detecting protease activity in human atheromas, the authors were able to demonstrate the overexpression of

MMP-9 in atherosclerotic plaques using a protease sensing NIRF probe, which was auto-quenched at injection due to closely spaced fluorophores (FRET effect). Upon enzyme-specific protease driven cleavage of the substrate, the fluorophores were successfully dequenched and released fluorescence confirming the presence of MMP-9 at the site of activation (Kim *et al*. 2010). Furthermore, P947-a gadolinium based MMP activatable CA was found to produce signal enhancement *via* MRI of MMP-rich versus MMP-poor regions of a plaque. It was also demonstrated that this probe was not only able to detect MMPs in atherosclerotic plaques *in vivo*, but also other proteases that were overexpressed in unstable plaques (Calcagno *et al*. 2013).

Jastrzebska and colleagues (2009) were successful in synthesizing an enzyme activated, solubility switchable CA for detecting MMP-2 activity *in vivo via* the application of MRI. The contrast agent (Gd-DOTA) was conjugated to the N-terminus of the peptide substrate (SPAY*YTAA) *via* a hydrophobic alkyl chain. The peptide substrate had a favorable cleavage site between Y*Y specifically for MMP-2. The water solubility of the intact probe was maintained by the addition of a PEG chain at the C-terminus of the peptide sequence as well as to lower the rate of elimination from the blood. The authors were able to successfully demonstrate, that upon cleavage by the enzyme the hydrophilic moiety (PEG chain) was separated from the sequence leading to the release of the contrast agent attached to the hydrophobic alkyl unit. This cleavage lead to retention of the contrast agent at the enzyme rich region for a longer time enhancing the contrast of the image produced. Based on this finding an attempt to synthesize a contrast agent covalently coupled to a peptide sequence specific to MMP (2/9) was undertaken. These findings suggest the application of substrate-based probes to be specific and efficient in detecting MMP-2/9 activity within the lesions leading to the imaging of atherosclerotic plaques with more accuracy and clarity.

2.2.1. Synthesis of peptide substrates designed to be specific to MMP-2/9 (JJ series) *via* solid phase peptide synthesis (SPPS)

It is important to reduce/eliminate background noise in order to produce accurate MRI images of the region of interest. This can be achieved by avoiding unspecific binding of probes to areas around the target region. Hence by designing probes that are highly specific to the biomarkers (proteases/enzymes) which are overexpressed in the microenvironment of the disease, non-specific binding could potentially be reduced. Moreover, these probes should be stable and dormant until they reach the cells of interest and then become activated to release the carry load

(therapeutic/prognostic agent). In this section of the research study, emphasis towards the synthesis of peptide substrates that are highly specific to MMP-2/9 was undertaken. In this way, the most desirable substrates for MMP-2/9 identification can be used for imaging atherosclerotic plaques by conjugating the MRI contrast agents to the peptide substrates. The choice of amino acids for the oligopeptide sequences was based on their sequence specificity to MMP-2 and MMP-9 (Xu *et al*. 2005; Kridel *et al*. 2001). The P3 position of the peptide chain corresponding to the S3 sub-pocket of the enzyme having a hydrophobic cavity which created a deviation from linearity into the active site cleft. Hence this feature of the S3 sub-pocket was suitable for the proline residue, which had a ring structure that produced the bend necessary for the ideal substrate binding at this position (Kridel *et al*. 2001).

A library of peptide substrates synthesized by Kridel and colleagues (2001) having arginine at P2 positon of the sequence was found to be readily recognized by MMP-9. This was due to the hydrophilic nature of the Asp-410 side chain located in the S2 subsite, showing a good interaction with the guanidinium group of arginine. It was also found that arginine at this position was less effective with MMP-2 recognition and also served as a significant region of differentiation between the two proteases which demonstrated a lot of sequence similarity (Murphy *et al*. 2002). Additionally small amino acids residues such as Ser, Ala or Leu were found to be better tolerated at the P2 position for MMP-2. Another distinctive feature of MMP-2 subsite was that the presence of Ala at P3 position was significant when compared to proline as in MMP-9 (Chen *et al*. 2002). The S1 subsite was characterized with an extended hydrophobic surface and was found to accommodate a wide range of amino acids (Olson *et al*. 2000). However glycine at this position was highly favored by MMP-9/2 (McGeehan *et al*. 1994). The S1′ sub-pocket was a deep hydrophobic region and also served as another important substrate recognition point. This information supports the reason for the presence of amino acids with side chain residues such as ile, leu, nva and ala (Kridel *et al*. 2002; McGeehan *et al*. 1994). Finally the S2′ sub-pocket was the last point of significant substrate recognition and interaction however interactions outside the significant (P3-P2′) is also common among MMP substrates. The S2′ subsite was found to favor amino acids residues such as val, leu and ile for MMP-2 and MMP-9 tolerated amino acids such as ser and thr. Furthermore, information on MMP-2/9 substrate design was discussed previously in section 1.2.5. (Table:2) which clearly lists the choice of amino acids for synthesizing a peptide substrate which is highly

specific to MMP-2/9. The JJ series of peptide substrates is shown in table 10.

Compound name	*N*	P4	P3	P2	P1	P1'	P2'	P3'	P4'
75	Fmoc	Ala	Pro	Leu	Gly	Ala	Ser	Gly	
76	Fmoc	Ala	Ala	Leu	Ala	Nva	Leu	Gly	
77	Fmoc	Lys	Pro	Ala	Gly	Nva	Ala	Gly	
78	Fmoc	Gly	Pro	Arg	Glu	Ile	Thr	Ala	Gly
79	Fmoc	Ile	Pro	Arg	Thr	Leu	Thr	Ala	Gly
80	Fmoc	Gly	Pro	Arg	Arg	Leu	Thr	Ala	Gly

Table 10: List of amino acids for JJ series of peptide substrates.

I. Synthesis of Fmoc-Ala-Pro-Leu-Gly-Ala-Ser-Gly-OH (75)

(75)
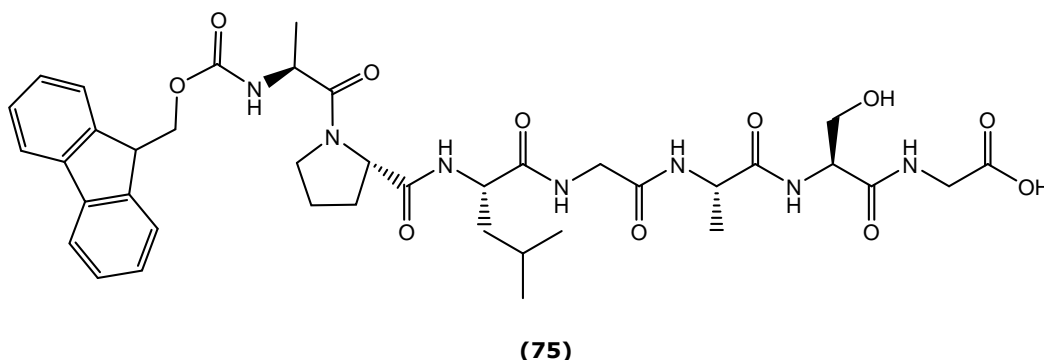
A 2-Chlorotrityl resin preloaded with glycine was reacted in cycles of coupling and deprotection with Fmoc protected amino acids. The amino acids were coupled in the following order: *N*-Fmoc-L-Ser(tBu), *N*-Fmoc-L-Alanine, *N*-Fmoc-Glycine, *N*-Fmoc-L-Leucine, *N*-Fmoc-L-Proline and *N*-Fmoc-D-Alanine. After the final coupling, the peptide sequence was removed from the resin using trifluoroacetic acid (1%) in DCM. The yield of the crude sample was found to be 59%. The crude product **75** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 850.4349 (M+H)+ (figure 68) corresponding to a molecular mass of 849; low resolution ESMS (-) mass spectrometry gave a signal at m/z 848 (M-H)- corresponding to the proton abstraction product.
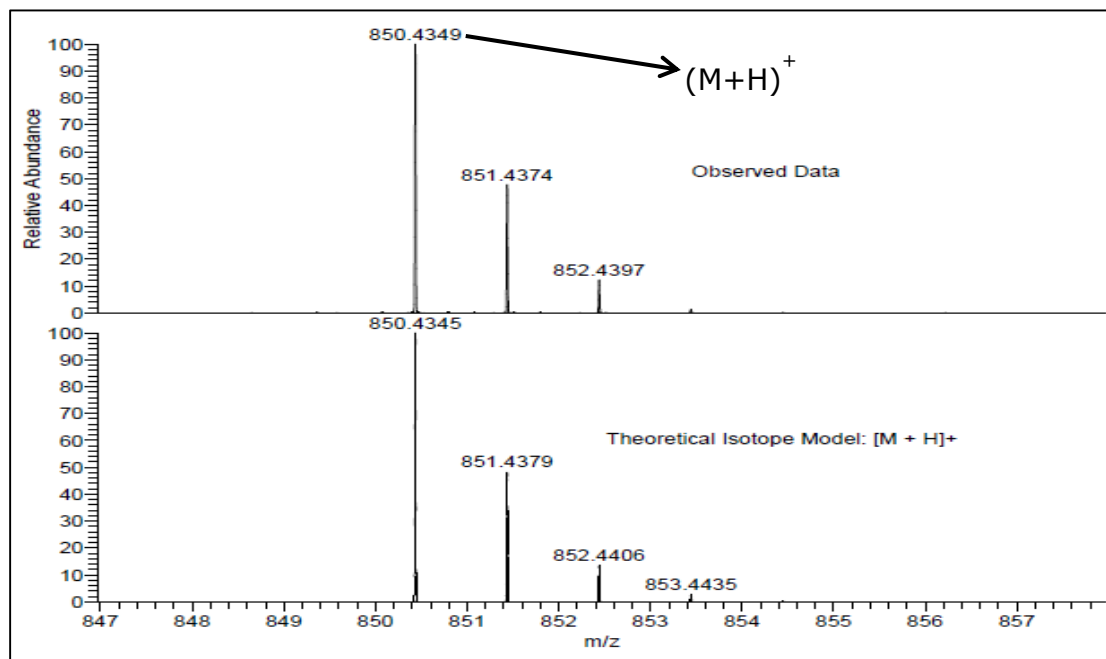
Figure 68: Expansion of the ESI (+) mass spectrum of compound **75** which shows the isotopic pattern of the molecular ion.

All these signals corresponded to the peptide chain with the *tert*-Butyl protecting group attached to the side chain of the serine residue. This was confirmed by low resolution ESMS (+) mass spectrometry on the product retreated with TFA, which gave a signal at m/z 816.4 corresponding to the sodium adduct $(M+Na)^+$. ESMS (-) mass spectrometry also gave a signal at m/z 793 $(M)^-$ corresponding to the molecular mass of the peptide sequence without the *tert*-Butyl protecting group attached to the side chain of the serine residue.

II. Synthesis of Fmoc-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH (76)

(76)
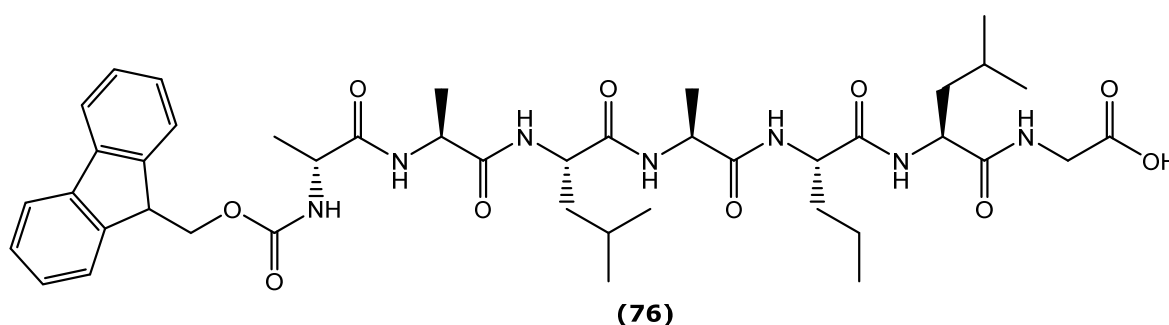
Compound **76** was synthesized following a procedure analogous to compound **75**. The amino acids were coupled in the following order: *N*-Fmoc-L-Leucine, *N*-Fmoc-L-Norvaline, *N*-Fmoc-L-Alanine, *N*-Fmoc-L-Leucine, *N*-Fmoc-L-Alanine and *N*-Fmoc-D-Alanine. The yield of the peptide sequence was found to be 22%. Product **76** was characterized by low resolution ESMS (+) mass spectrometry, which gave

a signal at m/z (+) 836.4 (M+H)$^+$ corresponding to a molecular mass of 835 and also a signal at m/z 859 corresponding to the sodium adduct (M+Na)$^+$.

III. Synthesis of Fmoc-Lys-Pro-Ala-Gly-Nva-Ala-Gly-OH (77)

(77)
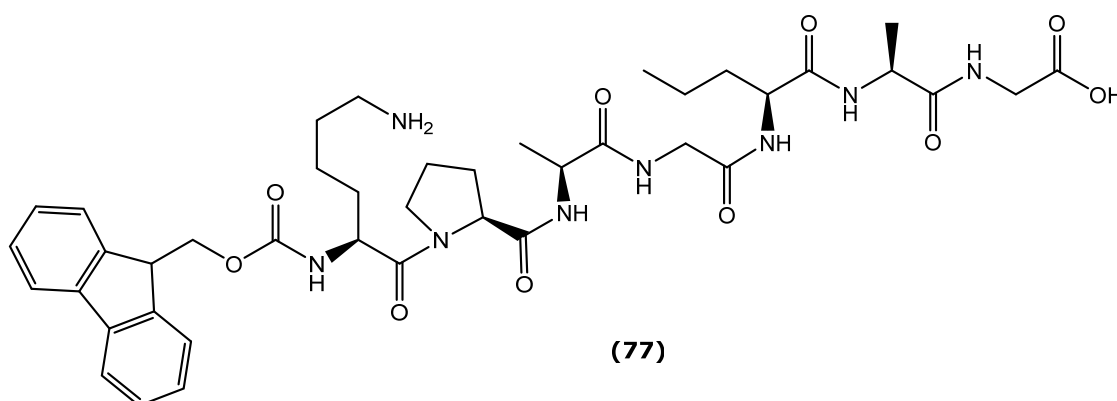
Compound **77** was synthesized following a procedure analogous to compound **75**. The amino acids were coupled in the following order: *N*-Fmoc-L-Alanine, *N*-Fmoc-L-Norvaline, *N*-Fmoc-Glycine, *N*-Fmoc-L-Alanine, *N*-Fmoc-L-Proline and *N*-Fmoc-L-Lysine. The crude yield was found to be 81%. Product **77** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 821 (M+H)$^+$ corresponding to a molecular mass of 820; low resolution ESMS (-) mass spectrometry displayed a signal at m/z 819 (M-H)$^-$ corresponding to the proton abstraction product.

IV. Synthesis of Fmoc-Gly-Pro-Arg(Mtr)-Glu(O^{t}Bu)-Ile-Thr (^{t}Bu)-Ala-Gly-OH (78)

(78)
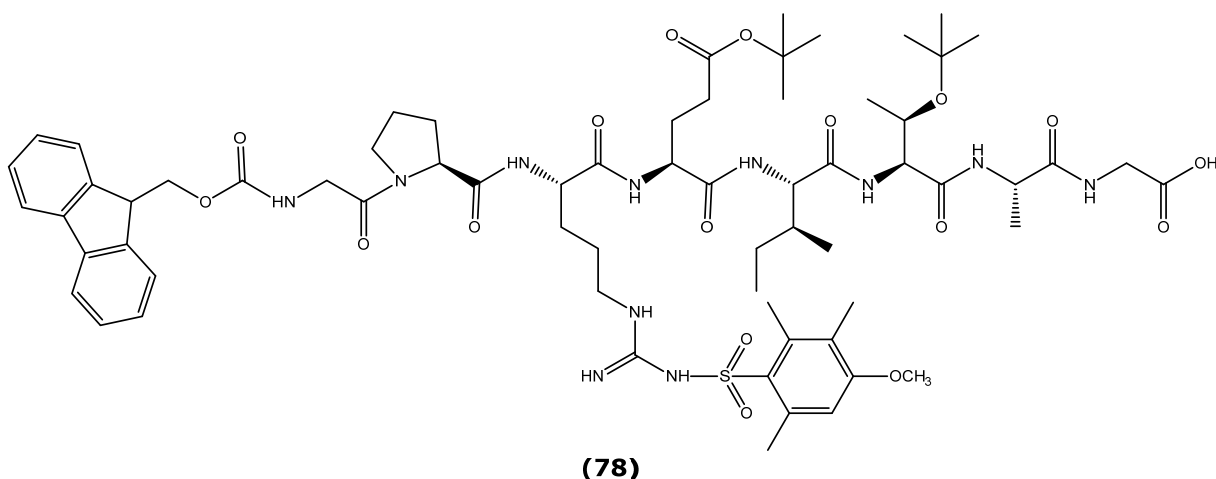
Compound **78** was synthesized following a procedure analogous to compound **75**. The amino acids were coupled in the following order: *N*-Fmoc-L-Alanine, *N*-Fmoc-Threonine (tBu), *N*-Fmoc-L-Isoleucine, *N*-Fmoc-L-Glutamate (OtBu), *N*-Fmoc-L-

arganine (Mtr), *N*-Fmoc-L-Proline and *N*-Fmoc-Glycine. HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column. The purity of the crude sample was found to be 76% and the yield was 58%. Product **78** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1346.6681 (M+H)⁺ corresponding to a molecular mass of 1345 and also a signal at m/z 673 [(M+2H)/2]²⁺; low resolution ESMS (-) mass spectrometry gave a signal at 1344 (M-H)⁻ corresponding to the proton abstraction product.

V. Synthesis of Fmoc-Ile-Pro-Arg(Mtr)-Thr(ᵗBu)-Leu-Thr(ᵗBu)-Ala-Gly-OH (79)

(79)
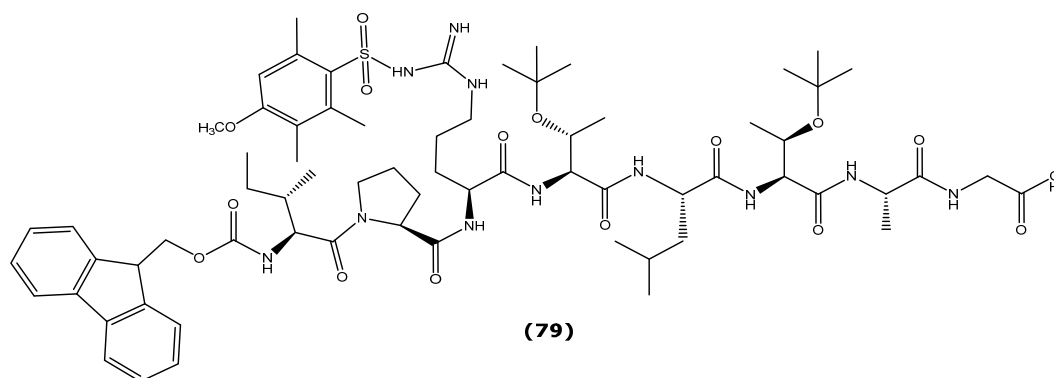
Compound **79** was synthesized following a procedure analogous to compound **75**. The amino acids were coupled in the following order: *N*-Fmoc-L-Alanine, *N*-Fmoc-L-Threonine(ᵗBu), *N*-Fmoc-L-Leucine, *N*-Fmoc-L-Threonine(ᵗBu), *N*-Fmoc-L-Arganine(Mtr), *N*-Fmoc-L-Proline and *N*-Fmoc-L-Isoleucine.
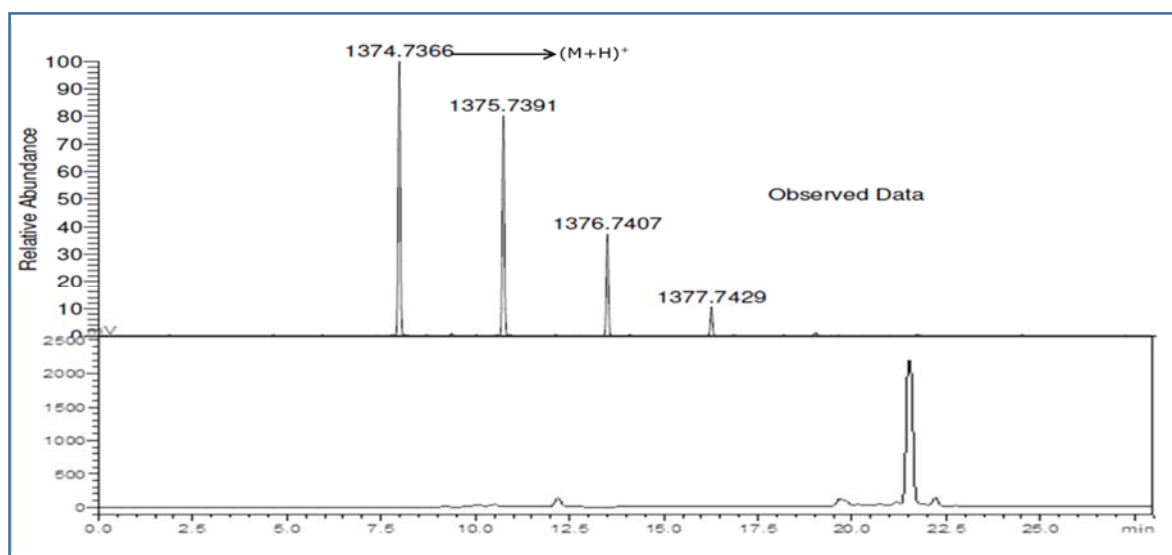
Figure 69: Mass spectrometry result for compound **79**. RP-HPLC elution profile of compound **79** (1mg/mL in acetonitrile); Conditions: Flow rate, 0.5mL/min; Detection, 254nm; stationary phase: Eclipse XDB-C18 RP column; mobile phase A= 0.1%TFA/water; mobile phase B=0.1%TFA/acetonitrile; gradient: low pressure gradient; 28min.

HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector using an eclipse XDB-C18 RP column. The purity of the crude sample was found to be 70% and the yield was 91%. Product **79** was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1374.7366 $(M+H)^+$ (figure 69) corresponding to a molecular mass of 1373 and also a signal at m/z 1396 corresponding to the sodium adduct $(M+Na)^+$; low resolution ESMS (-) mass spectrometry gave a signal at 1372 $(M-H)^-$ corresponding to the proton abstraction product.

VI. Synthesis of Fmoc-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr(^{t}Bu)-Ala-Gly-OH (80)

(80)
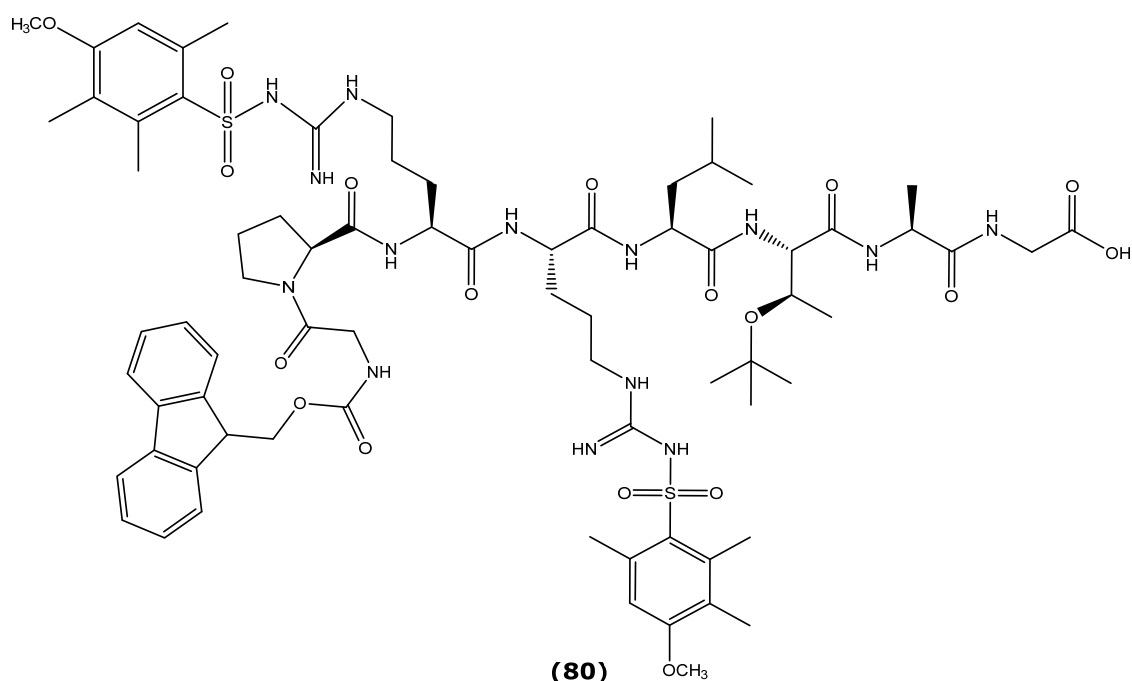
Compound **80** was synthesized following a procedure analogous to compound **75**. The amino acids were coupled in the following order: *N*-Fmoc-L-Alanine, *N*-Fmoc-L-Threonine(tBu), *N*-Fmoc-L-Leucine, *N*-Fmoc-L-Arganine(Mtr), *N*-Fmoc-L-Arganine(Mtr), *N*-Fmoc-L-Proline and *N*-Fmoc-Glycine. HPLC was performed on a Shimadzu Prominence system equipped with UV-Vis detector and fluorescence detector using an eclipse XDB-C18 RP column and the purity of the crude sample was found to be 83% with a yield of 92%. Product (**80**) was characterized by high resolution ESI (+) mass spectrometry, which gave a signal at m/z 1529.7147 $(M+H)^+$ corresponding to a molecular mass of 1528. A signal at m/z 1551

corresponding to the sodium adduct (M+Na)$^+$ and a signal at 765 (M/2+2H)$^{2+}$ was observed on the low resolution ESMS (+) mode; low resolution ESMS (-) mass spectrometry gave a signal at 1527 (M-H)$^-$ corresponding to the proton abstraction product.

Although substrates cleaved by gelatinases are still not fully identified, repetitive units with Pro-XX-Hy-(Ser/Thr), where X represents any amino acid residue and Hy represents a hydrophobic amino acid residue such as Ala, Ile, Leu, Val, Tyr, Trp, Met and Phe (Zitka *et al*. 2010) were found to be often cleaved by the protease. Peptide substrates with arginine residues were recognized by MMP-9 and also sequences with Gly-Leu-(Lys/Arg) were also found to be preferred by gelatinases (Kridel *et al*. 2001). Thus in accordance to the aim of this research study the peptide substrates were then covalently coupled to the paramagnetic contrast agent (Gd-DOTA). These specifically engineered contrast agents are anticipated to deliver the imaging agent site specifically upon cleavage by MMPs (2/9) potentially enhancing the underlying atherosclerotic lesion.

2.2.2. Synthesis of Gadolinium-DOTA-Peptide substrate designed to be specific to MMP-2/9

2.2.2.1. Attempted synthesis of JD-(Gd(III)-DOTA-Gly-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH)

Upon successful synthesis and characterization of the six-peptide substrates (section 3.2.1) designed to be specific to either MMP-9/2, it was decided to couple the substrates to the MRI contrast agent Gd-DOTA *via* the application of both solid and solution phase peptide synthesis. The synthesis of the peptide conjugated MRI contrast agent was undertaken in order to achieve the aim of this research study for specifically and accurately imaging atherosclerotic plaques. The synthesis of compound **JD** was attempted, following an adapted procedure by Jastrzebska and colleagues (2009). The experimental procedure was broken down to four stepwise processes (figure 70):

(i) The first step was the synthesis and characterization of the Fmoc-protected 12-aminododecanonic acid (**81**), a hydrophobic alkyl chain which may potentially increase the retention time of the contrast agent upon cleavage by MMP-9/2.

(ii) The second step was the coupling of compound **81** to the peptide sequence **76** after the removal of the Fmoc protecting group from the peptide sequence using 20% Piperidine in DMF producing compound **82**

(iii) The third step was the coupling of the glycine residue to compound **82** after

the removal of the Fmoc protecting group from 12-Ado acid producing compound **83**.

(iv) The intermediate compound **83** was then covalently attached to the chelating agent Tri-*tert*-butyl-1, 4, 7, 10-tetraazacyclododecane-1, 4, 7, 10-tetra acetate(DOTA-tris[*t*-Bu]) (**84**).
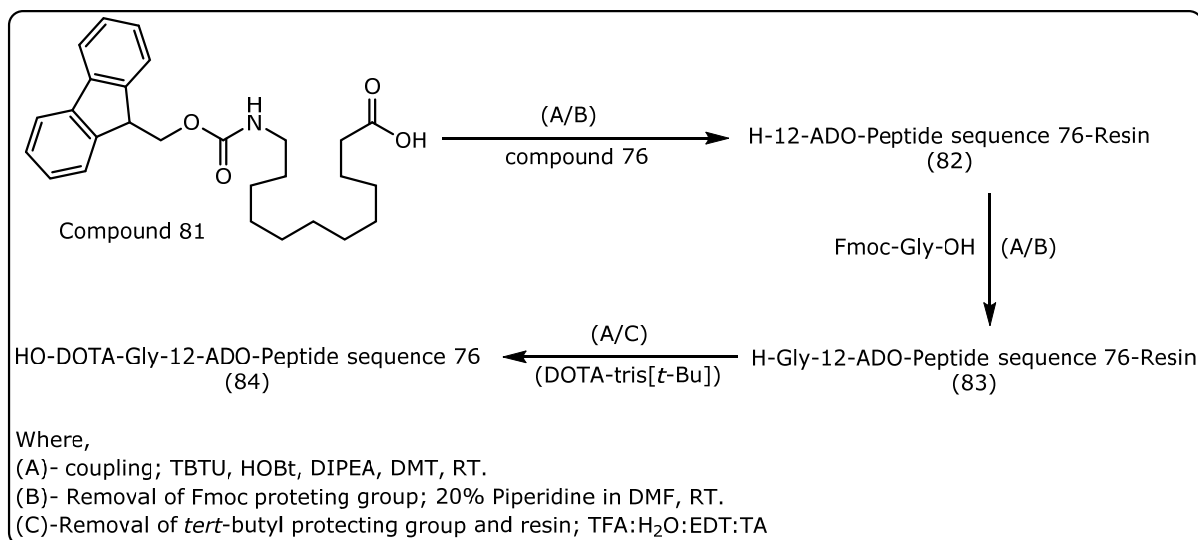
Figure 70: Step involved in the chemical synthesis of compound 84.

I. Synthesis of 12-(((((9H-fluoren-9-yl)methoxy)carbonyl)amino)dodecanoic acid (Compound 81)
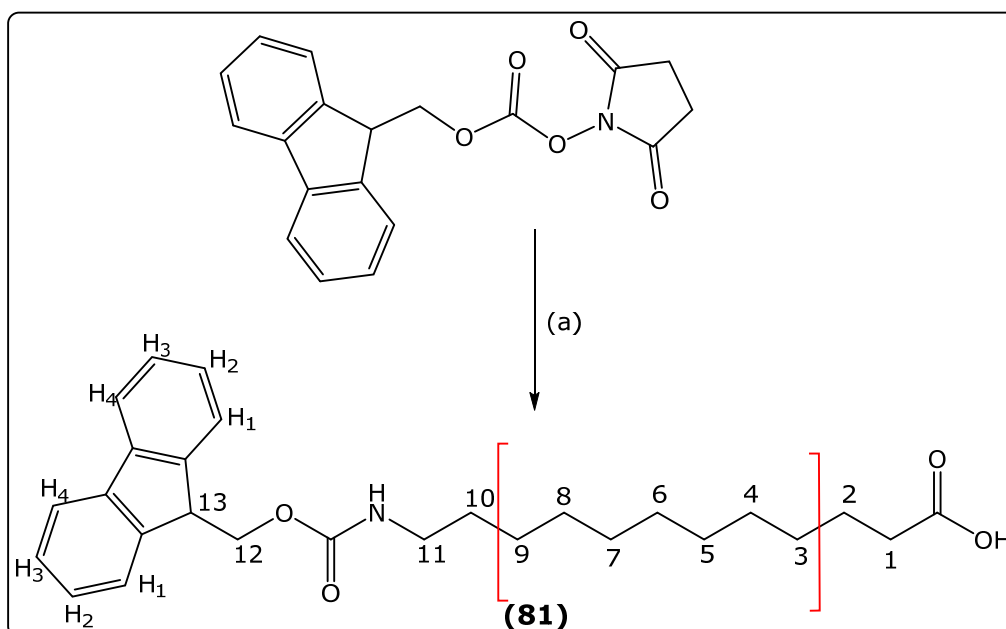
Scheme 8 (a): 12-aminododecanic acid, dioxane: water, 12hrs, rt.

The synthesis of Fmoc-12-aminododecanoic acid **81** was undertaken following an adapted method by Espelt and colleagues (2003). This involved the reaction of Fmoc *N*-hydroxysuccinimide ester (Fmoc-OSu) and 12-aminododecanoic acid

(Ado) in dioxane and water (4:1) (scheme 8). After 12 hours the reaction mixture was concentrated and then extracted into DCM and washed with water (x3), saturated sodium hydrogen carbonate (x3) and dried with anhydrous magnesium sulphate. The organic layer thus obtained was concentrated and purified by column chromatography. The yield of the sample after column chromatography was found to be 33%. The structure of the product was confirmed by low resolution ESMS (+) mass spectrometry which gave a signal at m/z 438 (M+H)$^+$ corresponding to a molecular mass of 437. The ^{1}H NMR spectrum had; for example, a 14-proton broad singlet at 1.30ppm which was assigned to the methylene protons of the spacer arm numbered 3-10 (scheme 8). A 2-proton multiplet between 1.45 and 1.52 and a 2-proton quintet signal at 1.65ppm which were assigned to C$\underline{H}_2$CH$_2$NH and C$\underline{H}_2$CH$_2$COOH respectively. A 2-proton triplet signal at 2.37ppm was assigned to C$\underline{H}_2$COOH and another 2-proton multiplet between 3.15 and 3.24pm was assigned to CH$_2$NHFmoc. A single proton multiplet between 4.22ppm and 4.28ppm was assigned to the proton attached to the Fmoc protecting group and another 2-proton multiplet between 4.42ppm and 4.47ppm was assigned to the methylene group protons of the Fmoc protecting group. The signals for the protons on the aromatic group were assigned as follows: a triplet of doublets for 2-protons at 7.34ppm was assigned to H-2 protons. The H-3 protons gave a triplet signal at 7.43ppm for two protons. A doublet for two protons at 7.62ppm was assigned to the H4 protons followed by another doublet signal at 7.79ppm was assigned to the H1 protons.

II. Synthesis of H-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-2-ClTrt resin (Compound 82)

(82)
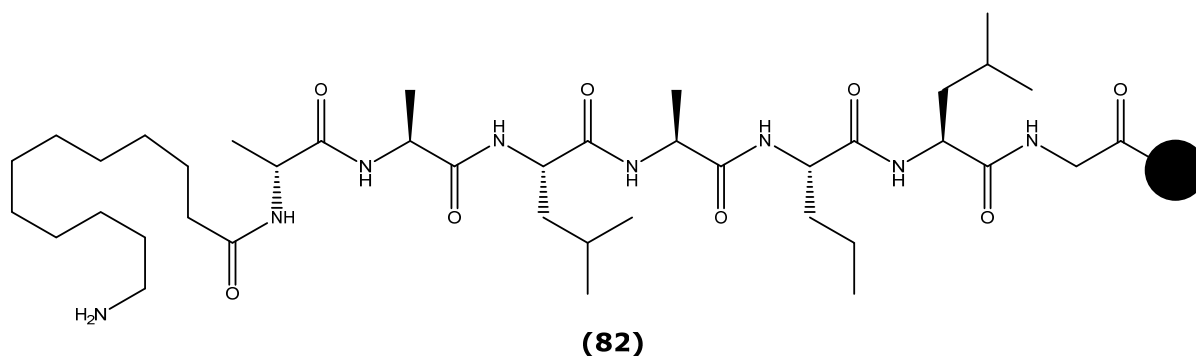
Compound **81** was coupled to compound **76** using the reagents and reaction conditions similar to that of coupling an amino acid *via* SPPS (TBTU, HOBt, DIPEA in DMF reacting at rt). The Kaiser test was performed to confirm the completion of the reaction. Subsequent Fmoc deprotection from 12-aminododecanoic acid was achieved using 20% Piperidine in DMF. Reaction completion was confirmed by the

blue color developed during the Kaiser test determining the presence of the primary amine.

III. Synthesis of H-Gly-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-2-ClTrt resin (Compound 83)

(83)
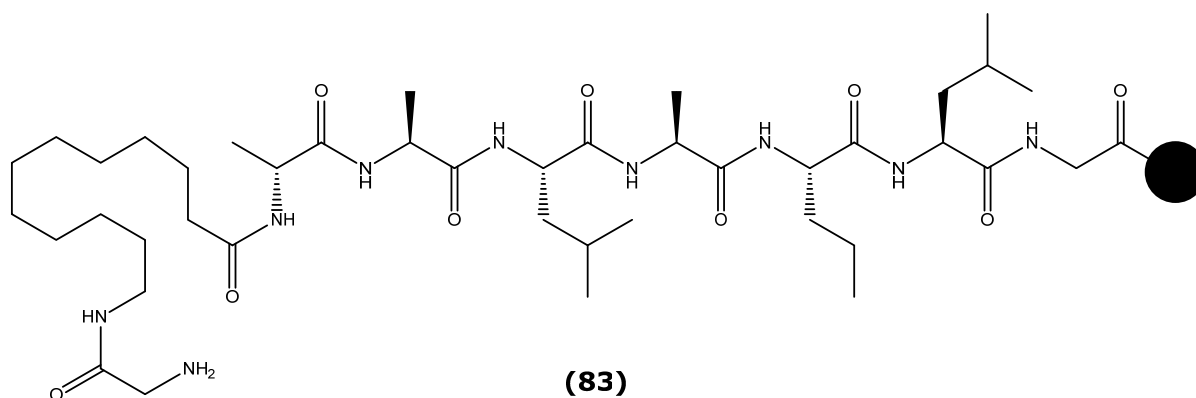
Before coupling the peptide substrate to the chelating agent [DOTA-tris(^{t}Bu)], a glycine residue was coupled to compound **82** in order to enhance the yield and assist in the process of coupling the chelating agent to the alkyl unit (Jastrzebska *et al*. 2009). Therefore, compound **82** was coupled to Fmoc-protected glycine with the aid of coupling agents HOBt/TBTU and DIPEA in DMF and the mixture was left to react for two hours at rt (figure 70). The Kaiser test was undertaken to confirm the completion of the reaction. Upon removal of the Fmoc protecting group from the glycine residue, the chelating agent [DOTA-tris(*t*Bu)] was then coupled to the peptide sequence.

IV. Synthesis of HO-DOTA-Gly-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH (Compound 84)

The next step was the coupling of [DOTA-tris(^{t}Bu)] to compound **83** with the aid of coupling agents HOBt/TBTU and DIPEA in DMF for 2hrs at rt. The Kaiser test was undertaken to confirm the completion of the reaction. DOTA was used as the chelating agent in this research project because, Polyaza polycarboxylic macrocycles such as DOTA or 1, 4, 8, 11-tetraazacyclotetrado cane-N,N',N'',N'''-tetraacetic acid (TETA) are said to be more efficient when compared to similar but noncyclic chelates such as DTPA or EDTA mainly because of their kinetic inertness and thermo stability *in vivo* (Jastrzebska *et al*. 2009). The final cleavage of the resin from the sequence along with the removal of the tert-butyl protecting group from DOTA was achieved using the cleavage mixture containing TFA: H$_2$O: EDT: TA. 1, 2-Ethanedithiol (EDT) and thioanisole (TA) were included in the reaction mixture in order to prevent any side chain reactions by acting as effective

scavengers of carbonium ions and other reactive species. The filtered solution was then added drop wise to diethyl ether by placing the reaction flask on ice to facilitate product precipitation.

(84)
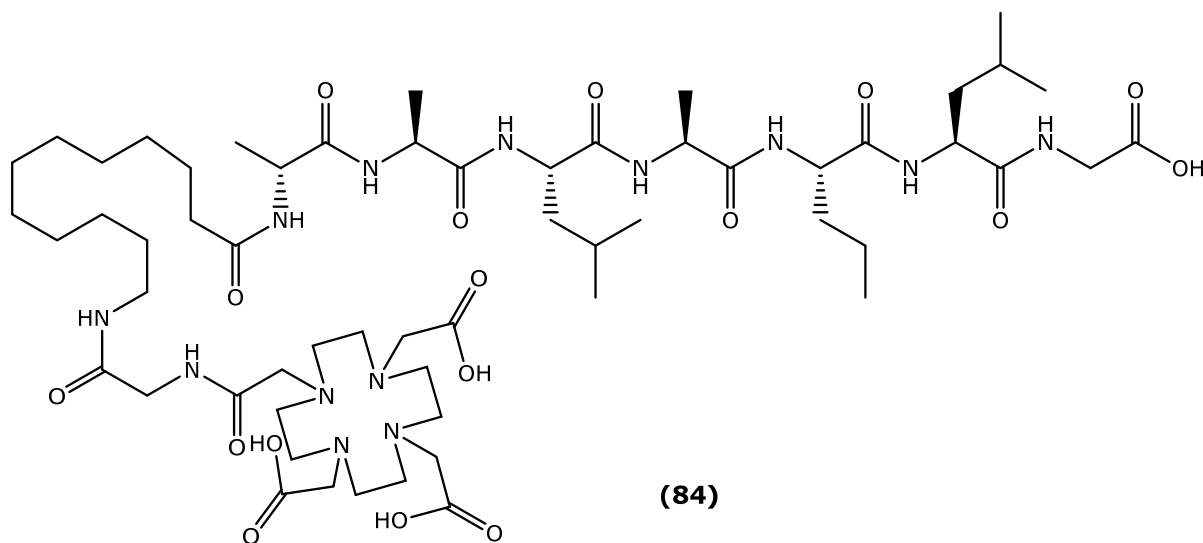
The crude product (**84**) was characterized by low resolution ESMS (+) mass spectrometry which gave a signal at m/z 628 $[(M+2H)/2]^+$ corresponding to a molecular mass of 1253; ESMS (-) mass spectrometry displayed a signal at m/z 1252.4 $(M-H)^-$ corresponding to the proton abstraction product and also a signal at m/z 1253.4 $(M)^-$ corresponding to the molecular mass of compound **84**. However, the final step of complexing the paramagnetic metal ion to the DOTA construct was not possible due to the low yield of compound **84**. Hence, it was decided to attempt another approach using an aminoanthraquinone unit for growing the peptide sequence-DOTA conjugate followed by complexing the metal ion to DOTA, *via* solution phase chemistry to attain a better yield.

2.2.2.2. Synthesis of JL (Compound 87)

Synthesis of the peptide conjugated MRI contrast agent was achieved by coupling peptide substrate **80** to compound **56**, which was then complexed to the paramagnetic metal ion-gadolinium. The design for the JL-87 conjugate consisted of a hydrophobic amino anthraquinone moiety attached to the C-terminal of the peptide sequence *via* a short diamino chain. The rationale for this design was based on the concept, that upon enzymatic hydrolysis of the peptide substrate, the contrast agent remains attached to the hydrophobic aminoanthraquinone unit, which could ultimately enhance the retention time of the macrocyclic gadolinium complex at the site of activation. The experimental procedure was broken down to

three stepwise processes (scheme 9):

Where,
(a)- Coupling; TBTU, HOBt, DIPEA in DMT at RT.
(b)- Removal of *tert*-Butyl group; TFA at RT.
(c)- Gadolinium (III) chloride hexahydrate; water, neutral pH, stirring at RT.

Scheme 9: Steps involved in the chemical synthesis of compound JL-87.

(i) The first step involved coupling of the peptide substrate **80** to compound **56**

to synthesize the AQ-DOTA-peptide conjugate **85**.

(ii) The next step involved the removal of the *tert*-butyl group from the chelating agent (DOTA) **86.**

(iii) The final step involved the synthesis of the target compound AQ-DOTA(Gd)-peptide conjugate **JL-87**.

I. Synthesis of DOTA-tris(tBu)-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr(OtBu)-Ala-Gly-Lys-Spacer-AQ (Compound 85)

(85)
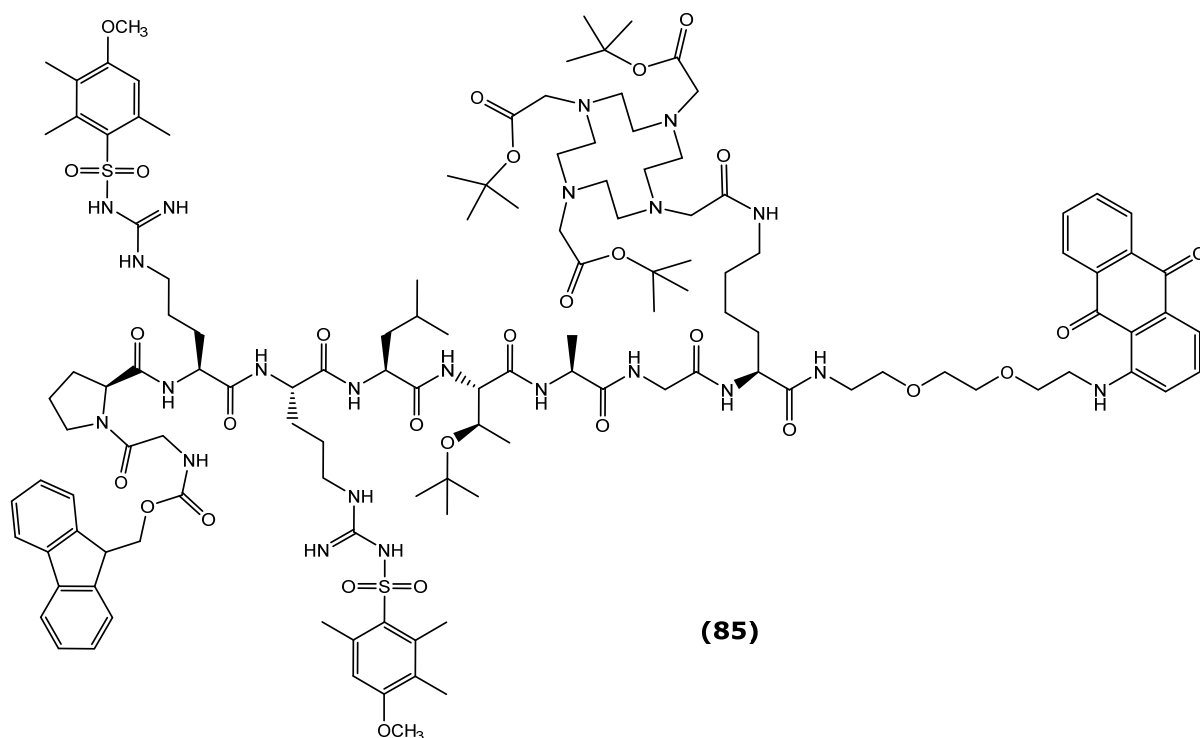
Compound **85** was synthesized by reacting peptide sequence **80** in DMF along with the in the situ-coupling agents TBTU/HOBT and DIPEA with compound **56** (scheme 9a). Reaction progress was monitored by TLC analysis; the crude product was then extracted into DCM and washed with water (x3), saturated sodium hydrogen carbonate (x1) and evaporated to low volume. The product (**85**) was characterized by low resolution ESMS (+) mass spectrometry, which displayed a signal at m/z 1274 $[(M+2H)/2]^{2+}$ corresponding to a molecular mass of 2547 and also a signal at m/z 2547 $(M)^+$ corresponding to the mass of the compound.

II. Synthesis of HO-DOTA-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr-Ala-Gly-Lys-Spacer-AQ (Compound 86)

(86)
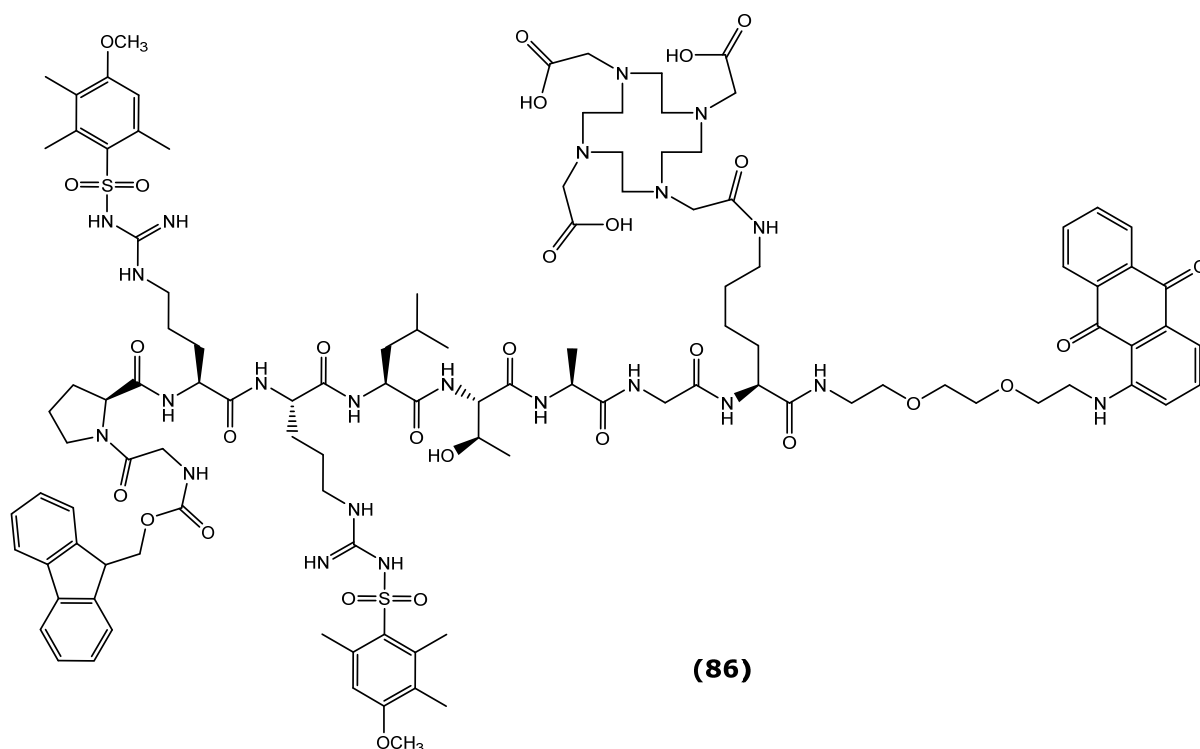
The *tert*-butyl protected compound **85** was deprotected using trifluoroacetic acid (scheme 9b). Deprotection progress was monitored by TLC analysis, which showed the presence of a new component, with a low R_f value due to the polar nature of the deprotected sample. The residual TFA was evaporated under reduced press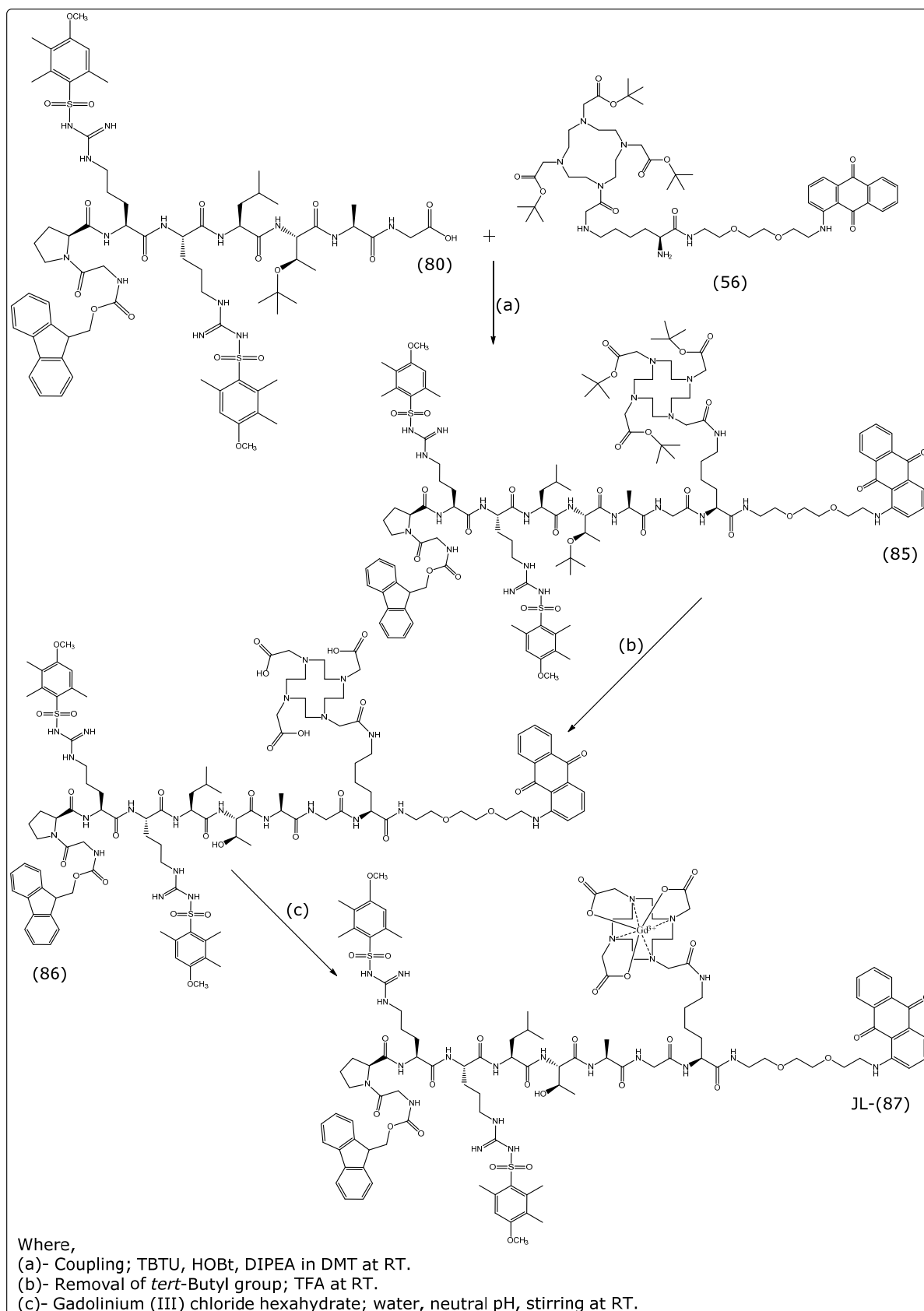ure and any remaining solvent was re-evaporated with ethanol. The product (**86**) was characterized by low resolution ESMS (+) mass spectrometry, which gave a signal at m/z 1162 $[(M+2H)/2]^{2+}$ corresponding to a molecular mass of 2323 and also a signal at m/z 2347 corresponding to the sodium adduct $(M+Na)^+$.

III. Synthesis of Fmoc-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr-Ala-Gly-Lys (DOTA(Gd³⁺))-Spacer-AQ (Compound JL-87)

Synthesis of the target gadolinium complex of DOTA-peptide-aminoanthraquinone spacer conjugate **JL-87** was performed in accordance to the reaction conditions proposed by Digilio and co-workers (2010). The Gd complex was prepared in a neutral aqueous solution by reacting $GdCl_3(6H_2O)$ with compound **86**. The reaction pH was adjusted to 7 by adding a solution of ammonia 0.35% drop wise (scheme 9c).
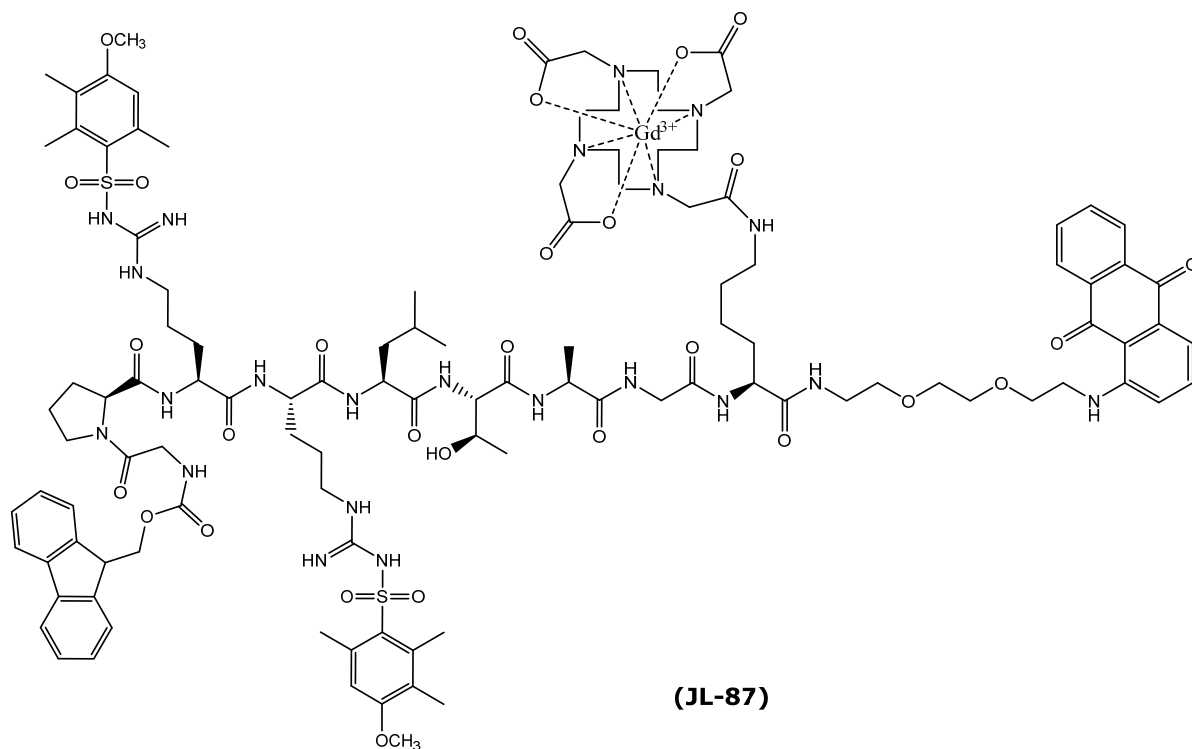
(JL-87)

Reaction progress was monitored by TLC analysis. The reaction volume was then reduced by evaporation at rt. The yield was found to be 56%. The product **JL-87** was characterized by low resolution by ESMS (+) mass spectrometry, which gave a signal at m/z 1239 $[(M+2H)/2]^{2+}$ corresponding to a molecular mass of 2477 and also a signal at m/z 2477 $(M)^{+}$.

2.2.3. Conclusion

The aim of this research study was the synthesis of a library of targeted MRI contrast agents that could have the potential to be used as non-invasive imaging agents of atherosclerotic plaques. It was hoped that this could be achieved by exploiting the role of MMP-2/9 (generally over expression in these plaques). The functioning of these proteases is required for both normal physiological activities and is also a cause for several pathological conditions as discussed in the introduction. During this research study, work was focused on synthesizing a small library of peptide substrates designed to be specific to either MMP-2 or MMP-9. This was achieved by synthesizing six peptide sequences (75, 76, 77, 78, 79 and 80) adopting solid phase peptide synthesis (McGeehan *et al*. 1994; Van Valckenborgh *et al*. 2005; Jastrzebska *et al*. 2009). The choice of amino acids for the peptide substrates was based on the information available from the literature on MMP-2/9 enzyme structure. After removal of the resin from the peptide sequence, the substrates were characterized by high resolution ESI mass spectrometry.

In order to complex the contrast agent, which in this study was Gadolinium with the oligopeptide sequences, work was undertaken to synthesize the paramagnetic gadolinium chelates (Gd-DOTA/ Gd-DTPA). Due to limitations such as poor reactivity, intramolecular cross-linking producing less stable chelation sites, along with clinical conditions such as nephrogenic systemic fibrosis (NSF-caused in renal impaired patients due to the use of Gd-DTPA for MRI scans) it was decided to use DOTA as the chelating agent in this research study. DOTA is one of the most stable lanthanide complexes. It is known for its solid rigidity, symmetric nature and kinetic inertness (Locin *et al*. 1986). Hence the chance of dissociation of the metal ion from the complex was found to be much less even at low pH, post injection (Wang *et al*. 1992; Bousquet *et al*. 1988; Lauffer 1987).

Considering the advantages of DOTA as the ligand of choice an attempted synthesis of Gd-DOTA-peptide substrate was undertaken by conjugating the peptide sequence (76) to a hydrophobic alkyl chain (12-aminododecanoic acid) Ado (8). This was done in order to decrease the solubility of the complex after being cleaved by the target MMP. This sequence was then coupled to DOTA *via* a glycine residue followed by the coupling of DOTA. Due to the low yield and difficulties in purification of the sequence the final step of complexing the gadolinium ion to the peptide substrate was not possible. However, another attempt to synthesize a peptide substrate coupled to the Gd-DOTA (JL-87) complex was achieved following the design: peptide substrate 80-Lys(DOTA-Gd)-Sp-AQ. This sequence was successfully characterized *via* low-resolution mass spectroscopy. Even though the probe JL-87 was successfully synthesized, further optimization of the peptide substrate in terms of efficient MMP-mediated activation to release the contrast agent is essential. From the mass spectrometric analysis it was found that the protecting group in the arg residue was still present and hence a longer deprotection procedure for their removal is essential. This problem can also be addressed by using Arg(Boc$_2$) or Arg(Pbf) which are also attractive alternatives for Arg(Mtr) for the Fmoc/^{t}Bu strategy of SPPS (Stierandova *et al* 1994). Also the protocol for the synthesis of the probe JD has to be altered in order to increase the yield of the intermediates for complexing the paramagnetic metal ion to the peptide conjugate. Finally another major issue would be the purity of the probes (JD and JL87), which would require purification by preparative HPLC. Upon successful purification of the probes-JD and JL87 which are engineered to carry a MMP recognizable substrate coupled to a gadolinium chelate, they could potentially be cleaved by the corresponding MMP-2/9 to release the contrast agent in atherosclerotic tissues or at the site of activation. On substrate cleavage, it is

hoped that the released contrast agent may increase the relaxation time of the water protons at the site of deposition resulting in signal enhancement. Moreover the presence of an anthraquinone moiety in probe JL-87 could possibly increase the hydrophobic nature of the released contrast agent causing an increase in the retention time of the CA at the site of cleavage. This could also influence the contrast of the image produced *via* a MRI scanner.

2.2.4. Limitations

Some of the limitations in this section were that, when the peptide sequences were analyzed for purity *via* analytical RP-HPLC they ranged from 76-97%. There is therefore a requirement to purify all the peptide substrates, intermediates and target compounds by preparative HPLC. The presence of the protecting groups in the peptide sequences 75, 76, 77, 78, 79 and 80, i.e., the *N*-terminal and side chain protecting groups were not removed, so that when coupling the peptide substrate to the contrast agent, as in compound 87, any unspecific binding could be avoided. However compound 87 still showed the presence of the side chain protecting group in Arg even after treatment with TFA. Hence the suggestion to use a different side chain protecting group that can be quickly removed has been stated in the previous section. The final step involving the removal of the Fmoc protecting group from the *N*-terminal of the peptide substrate is desirable.

2.2.5. Future work

Future work on these MMP (2/9) specific peptide substrates should focus on the determination of purity *via* HPLC analysis (75, 76 and 77) and the subsequent purification of target compounds *via* preparative RP-HPLC is recommended. More work on the optimization of experimental conditions for coupling these peptide substrates to the contrast agent (Gd-DOTA) has to be developed. This can be done by following the procedure as demonstrated in scheme 9, although keeping in mind the need for better yield and purity of the target compound. The next step would be to evaluate the substrate specificity of the compounds to MMP-2/9. This could be done by using a FRET system as achieved for the legumain peptide substrates. The fluorogenic probes upon incubation with activated MMP-2/9 enzyme could release fluorescence and in this manner could potentially be used as a screening method for confirming those MMP-2/9 specific compounds. From here the successful substrates could be used for the synthesis of the target peptide-contrast agent conjugate. Finally biological analysis to evaluate the intercellular contrast agent release of the probes along with cytotoxic studies (MTT, LDH) on human

tissue extracts of atherosclerotic lesions is essential. Additionally, it is worth mentioning that both MMPs and legumain are found to be involved in the pathology of other diseases such as cancer and inflammatory diseases and hence the non-invasive contrast agents engineered for atherosclerotic plaques could also be used to evaluate these conditions.

2.3. Structure library
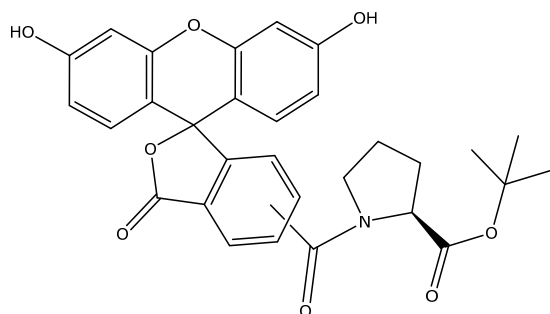
5(6)-CBF-Pro-O^{t}Bu **(1)**
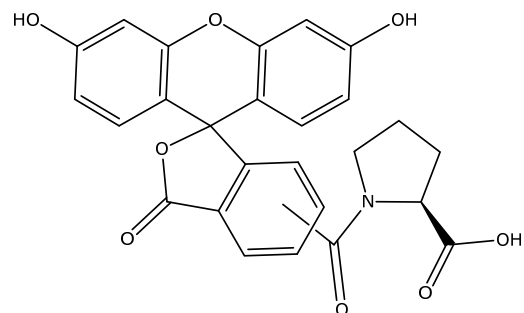
5(6)-CBF-Pro-OH **(2)**
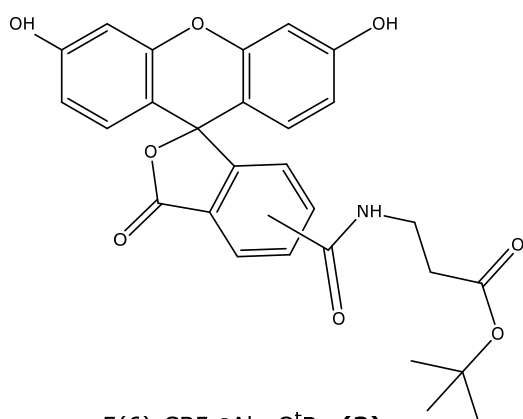
5(6)-CBF-βAla-O^{t}Bu **(3)**
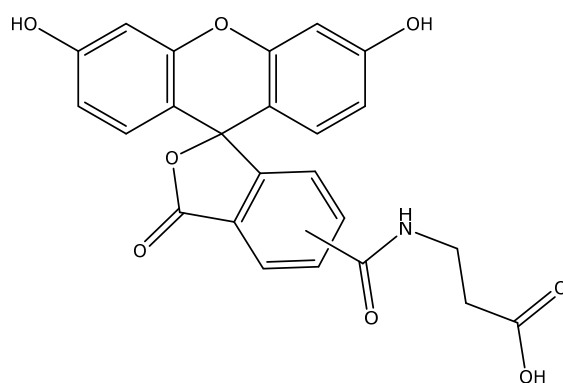
5(6)-CBF-βAla-OH **(4)**

AQ-Spacer-NH$_2$ **(5)**

5(6)-CBF-Pro-Ala-Asn(Trt)-Leu-OH **(6)**

5(6)-CBF-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ **(7)**

5(6)-CBF-Pro-Ala-Asn-Leu-Spacer-AQ **(8)**

5(6)-CBF-Pro-Gly-Asn(Trt)-Leu-OH **(9)**

158
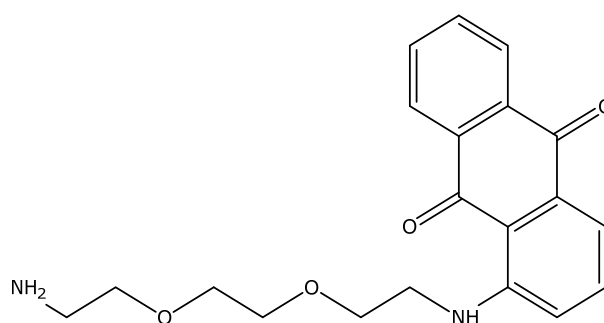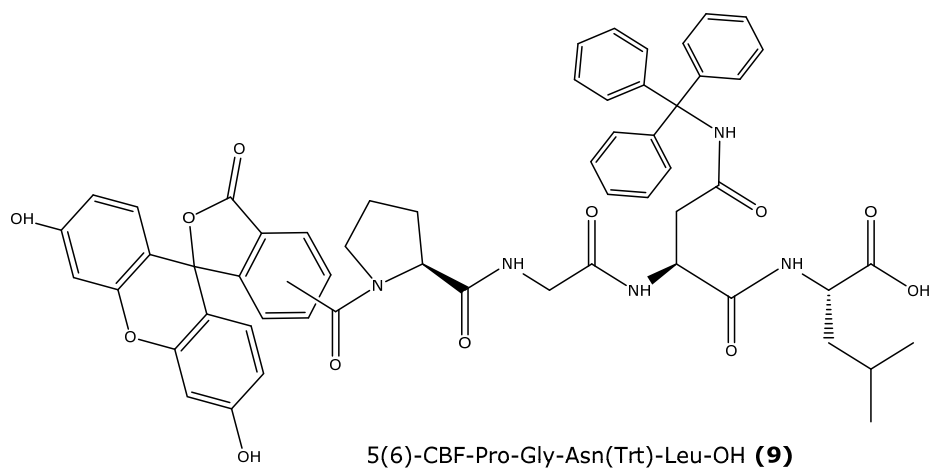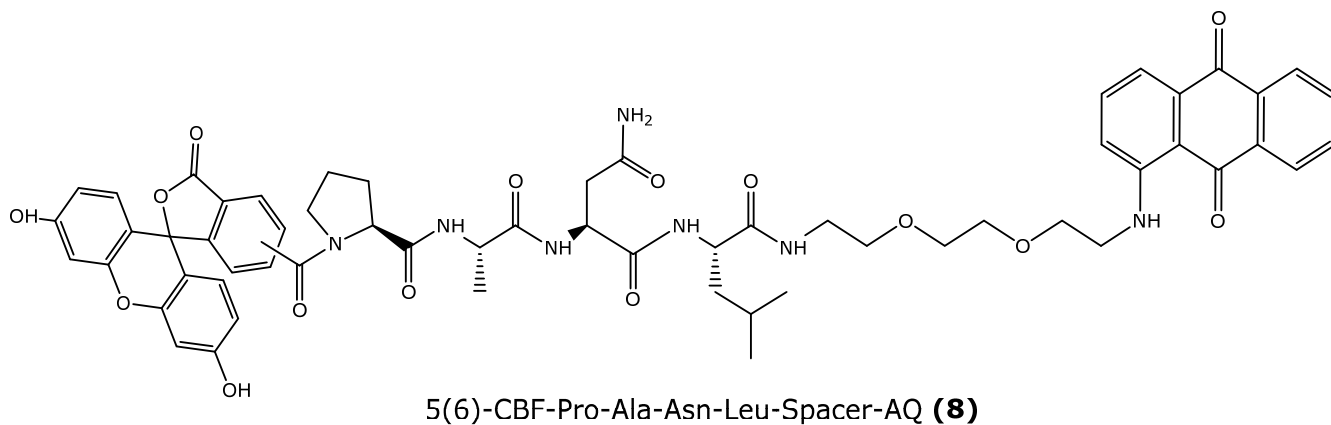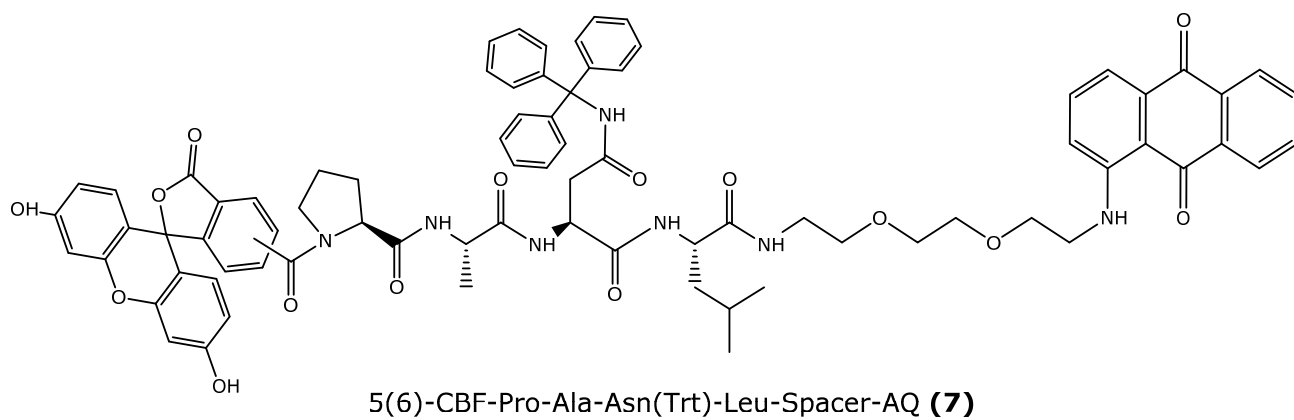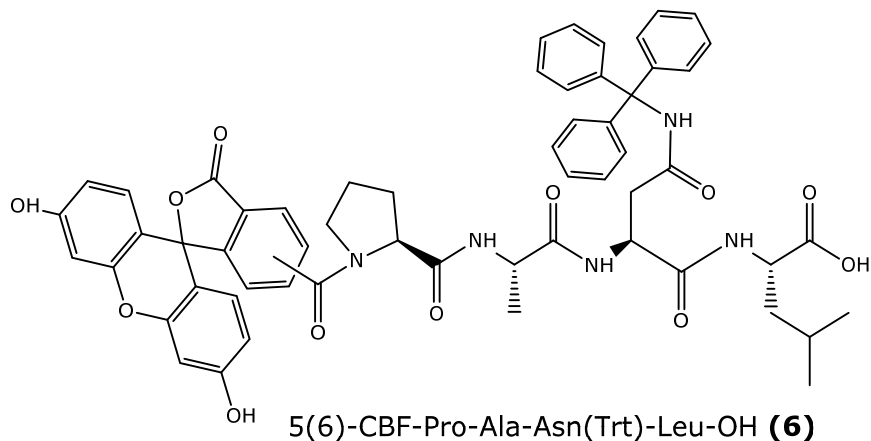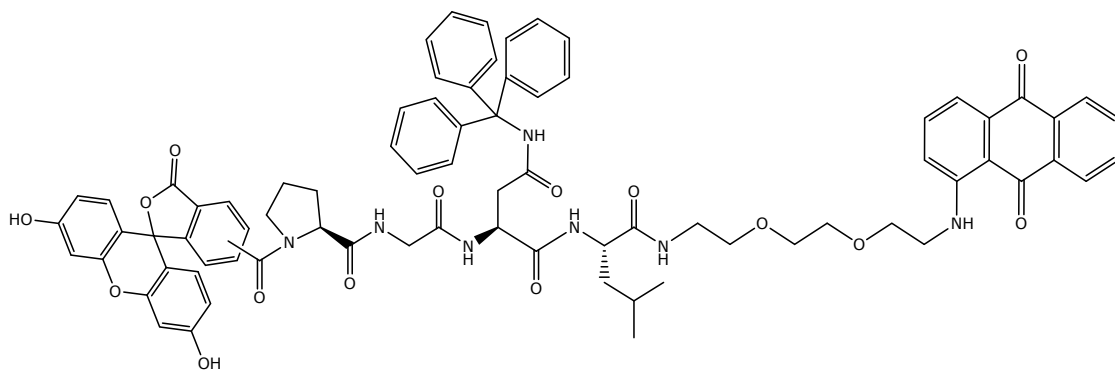
5(6)-CBF-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ **(10)**
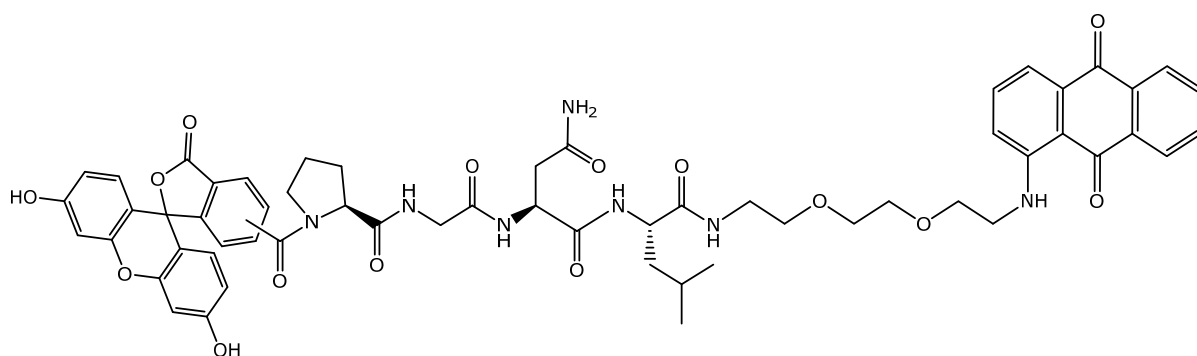
5(6)-CBF-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ **(11)**
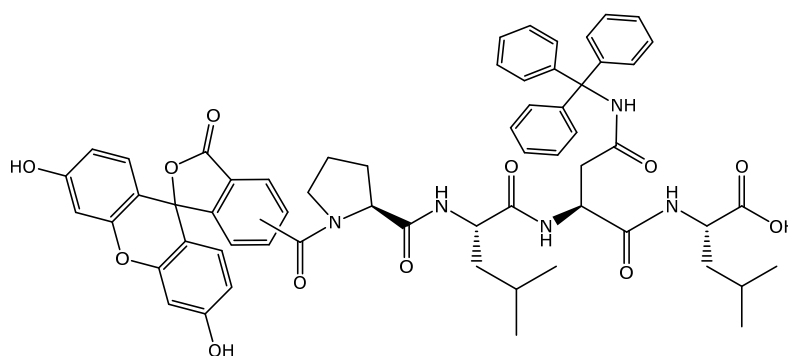
5(6)-CBF-Pro-Leu-Asn(Trt)-Leu-OH **(12)**

5(6)-CBF-Pro-Leu-Asn(Trt)-Leu-Spacer-AQ **(13)**

159
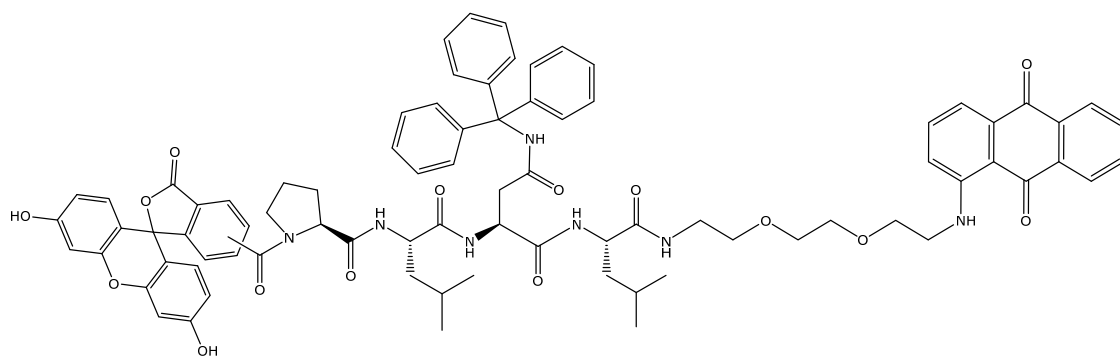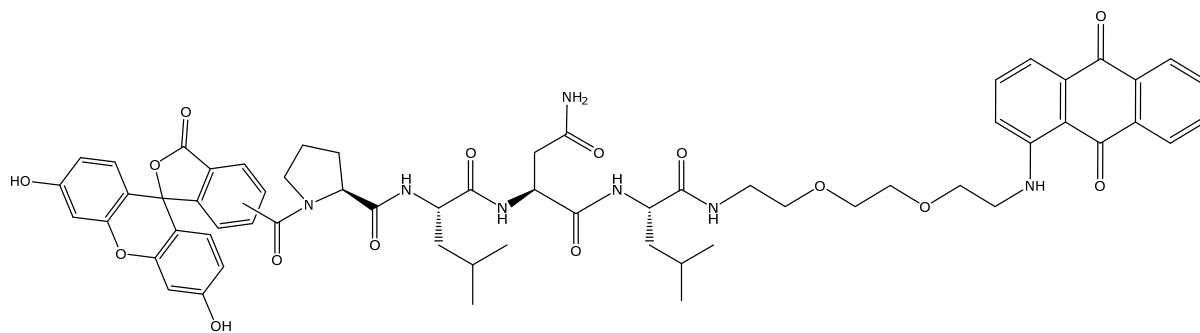
5(6)-CBF-Pro-Leu-Asn-Leu-Spacer-AQ **(14)**
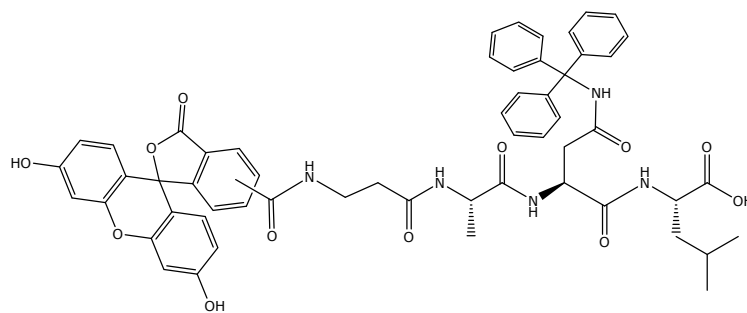
5(6)-CBF-βAla-Ala-Asn(Trt)-Leu-OH **(15)**
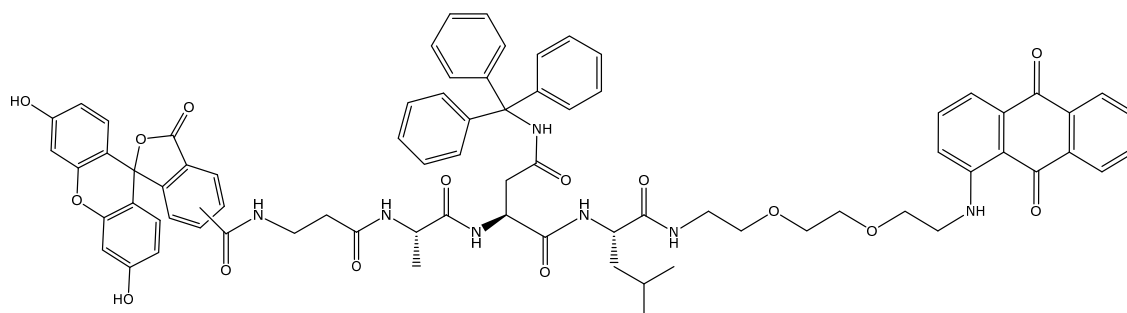
5(6)-CBF-βAla-Ala-Asn(Trt)-Leu-Spacer-AQ **(16)**

5(6)-CBF-βAla-Ala-Asn-Leu-Spacer-AQ **(17)**

160
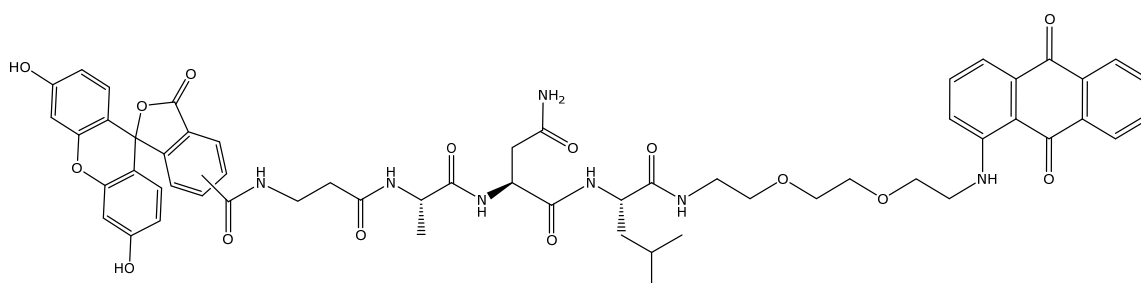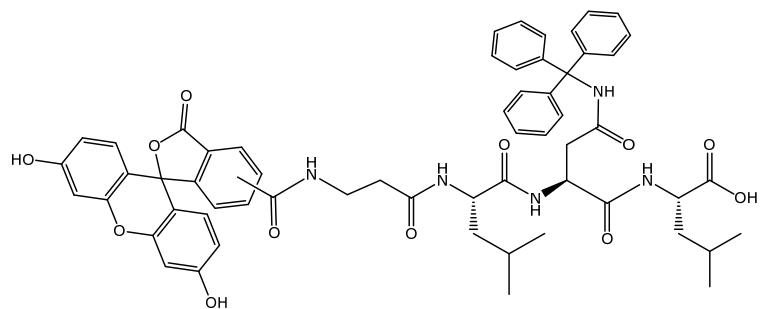
5(6)-Ala-Leu-Asn(Trt)-Leu-OH **(18)**
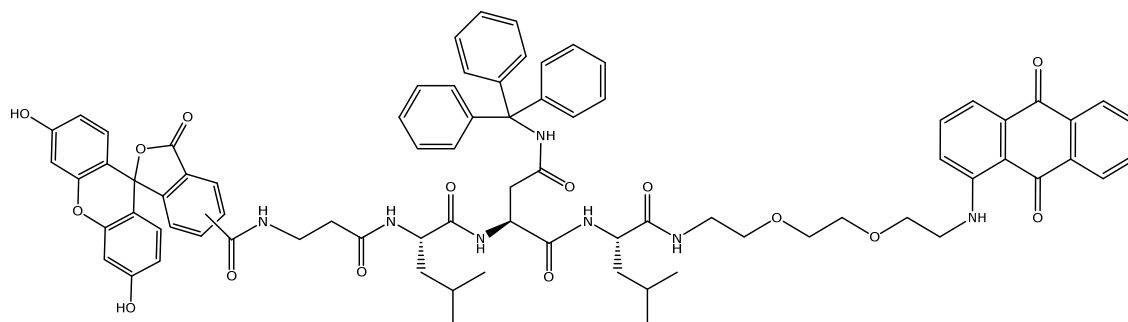
5(6)-Ala-Leu-Asn(Trt)-Leu-Spacer-AQ **(19)**
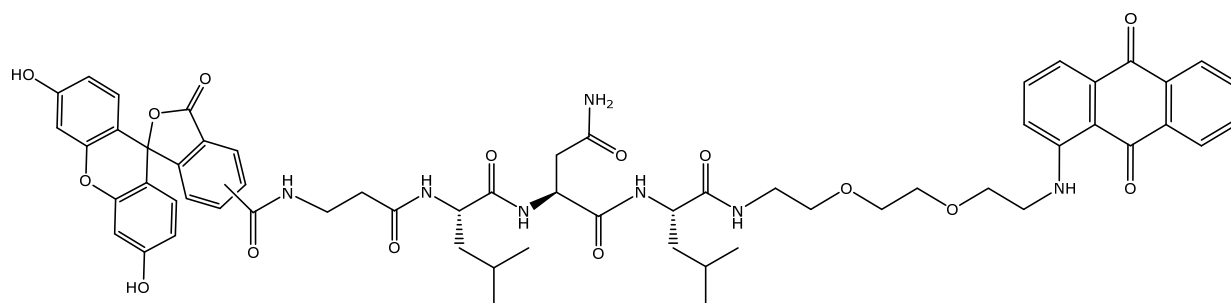
5(6)-Ala-Leu-Asn(Trt)-Leu-Spacer-AQ **(20)**
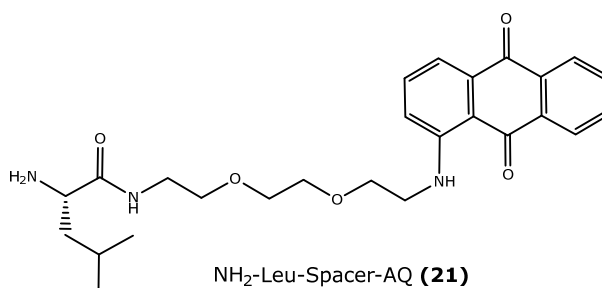
NH$_2$-Leu-Spacer-AQ **(21)**

161
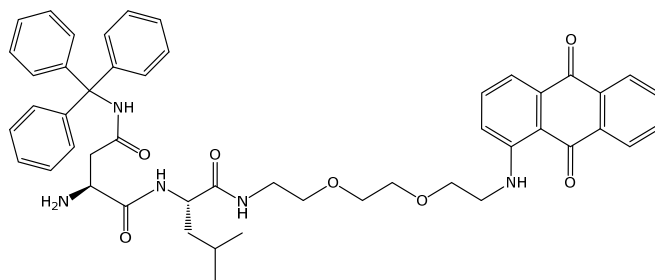
NH$_2$-Asn(Trt)-Leu-Spacer-AQ **(22)**
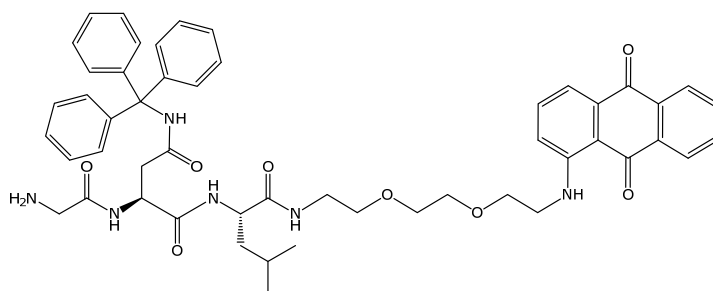
NH$_2$-Gly-Asn(Trt)-Leu-Spacer-AQ **(23)**
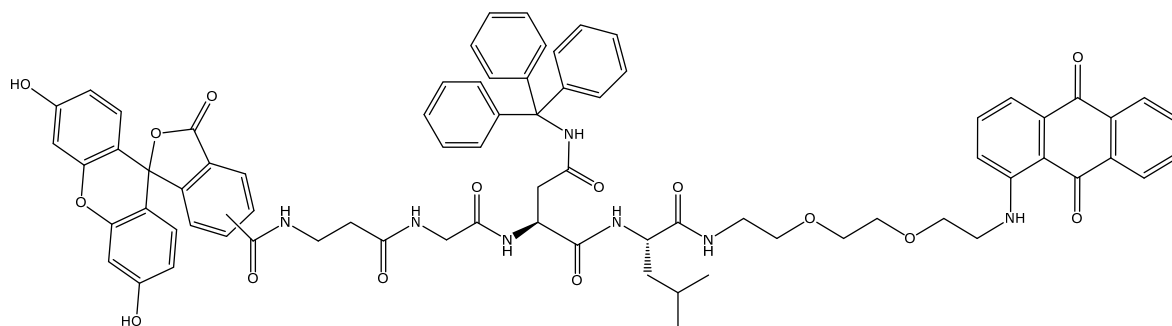
5(6)-CBF-βAla-Gly-Asn(Trt)-Leu-Spacer-AQ **(24)**

5(6)-CBF-βAla-Gly-Asn-Leu-Spacer-AQ **(25)**

162
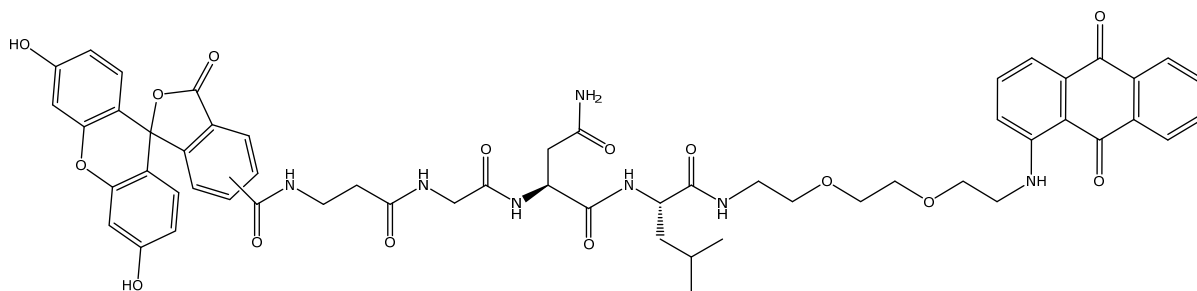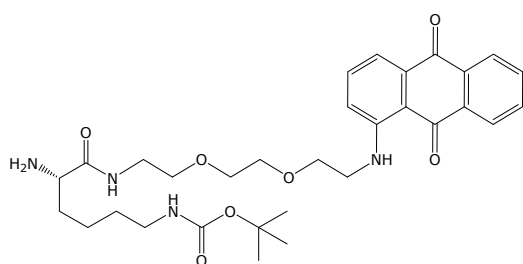
NH₂-Lys(Boc)-Spacer-AQ **(26)**
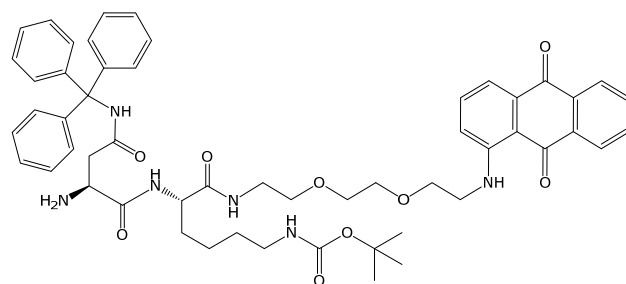
NH₂-Asn(Trt)-Lys(Boc)-Spacer-AQ **(27)**
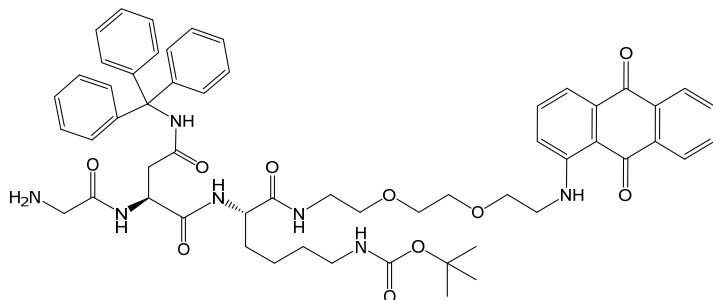
NH₂-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ **(28)**
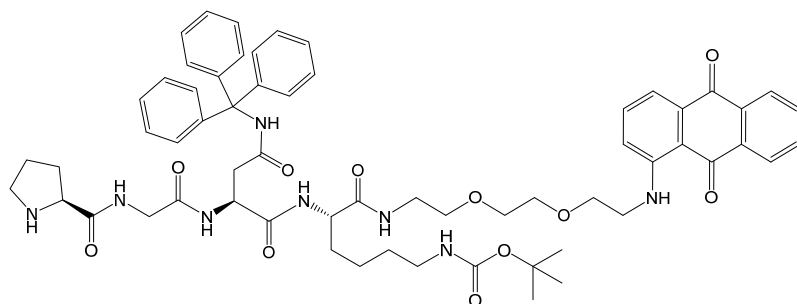
NH-Pro-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ **(29)**
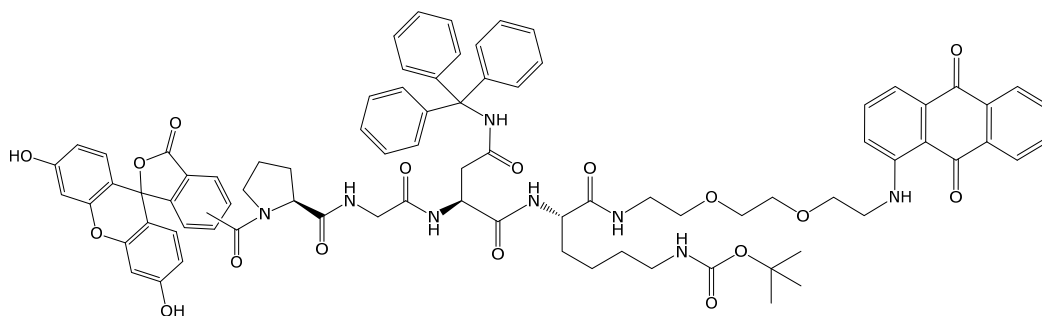
5(6)-CBF-Pro-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ **(30)**

5(6)-CBF-Pro-Gly-Asn-Lys-Spacer-AQ **(31)**

163
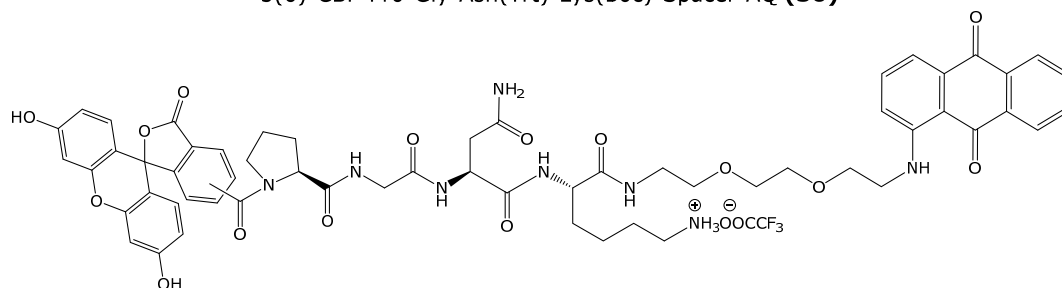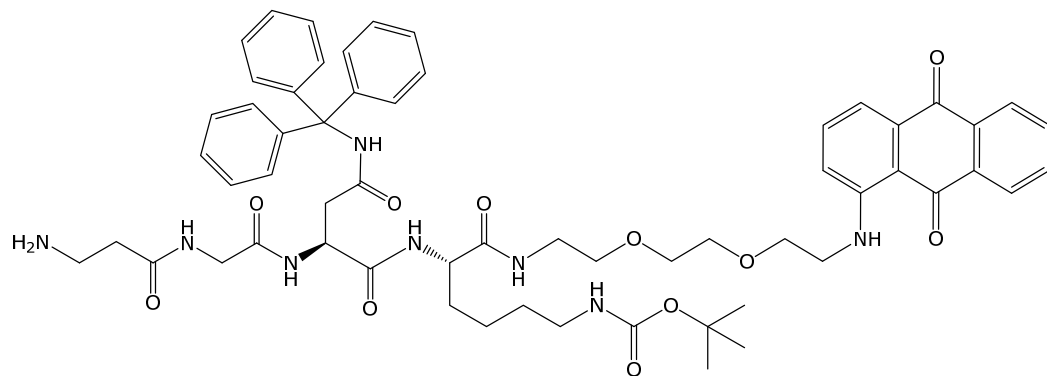
NH$_2$-βAla-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ **(32)**
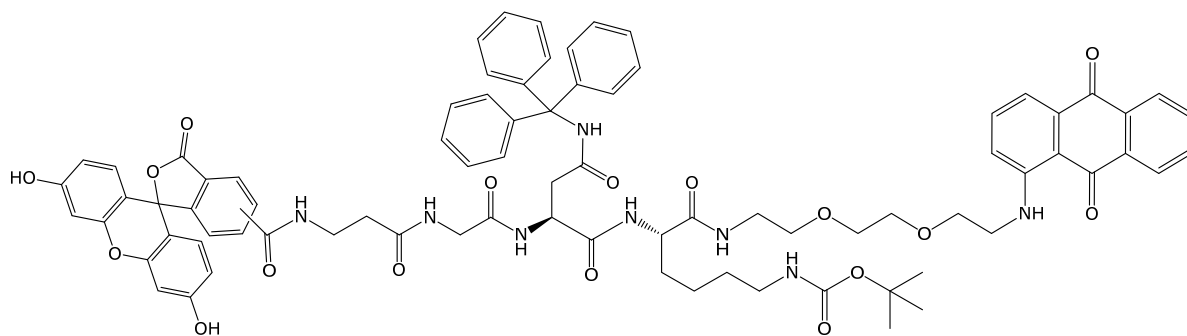
5(6)-CBF-βAla-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ **(33)**
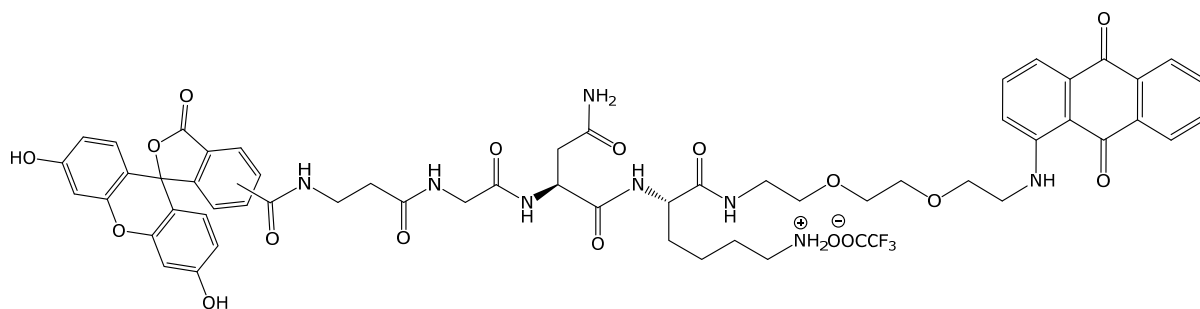
5(6)-CBF-βAla-Gly-Asn-Lys-Spacer-AQ **(34)**
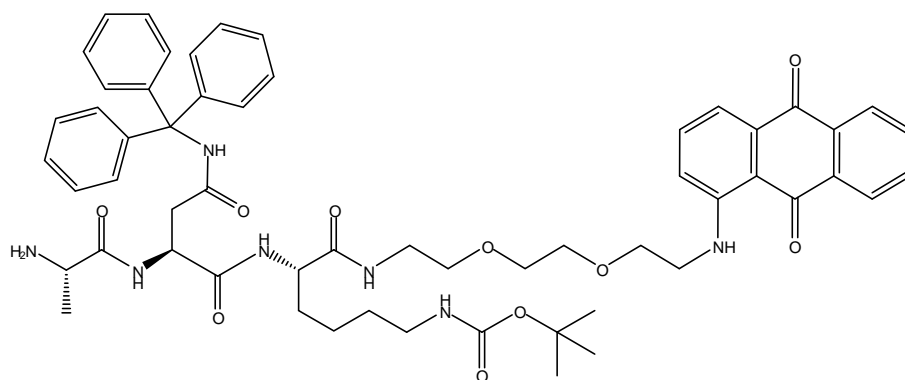
NH$_2$-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ **(35)**

164
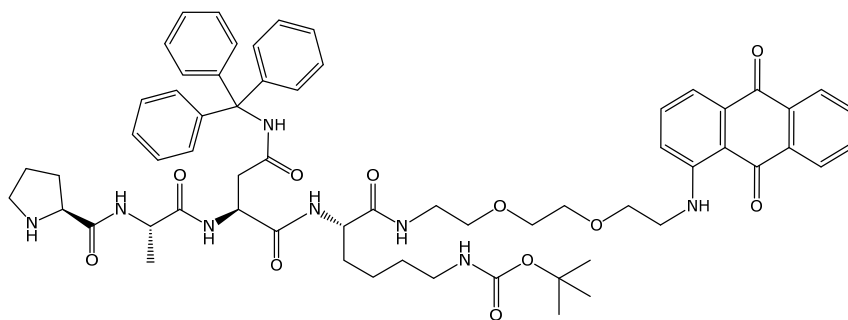
NH-Pro-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ **(36)**
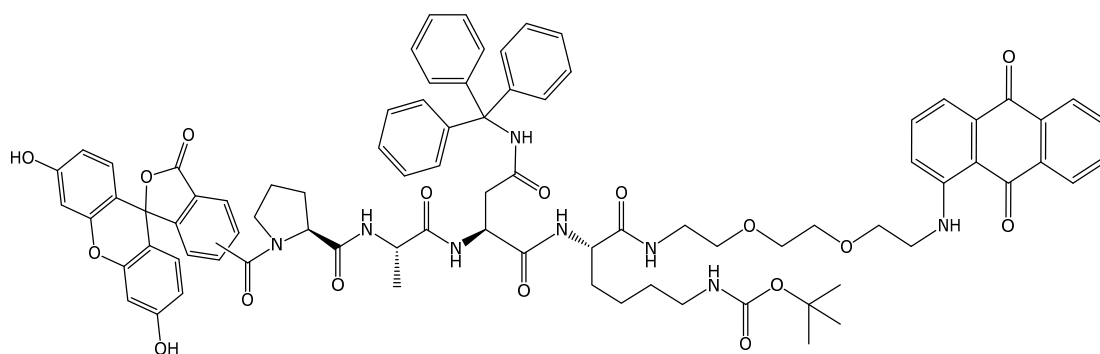
5(6)CBF-Pro-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ **(37)**
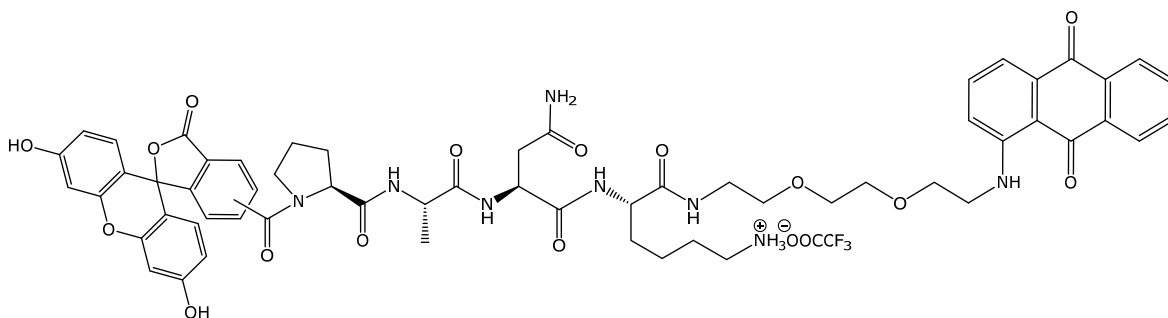
5(6)CBF-Pro-Ala-Asn-Lys-Spacer-AQ **(38)**
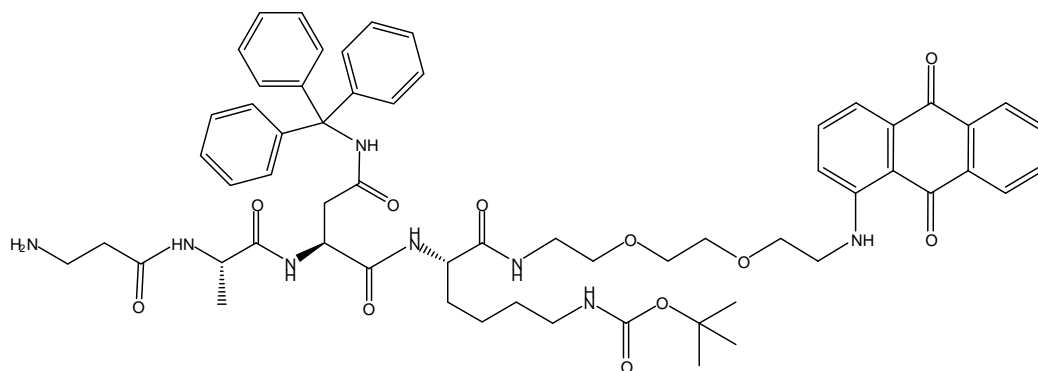
NH$_2$-βAla-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ **(39)**
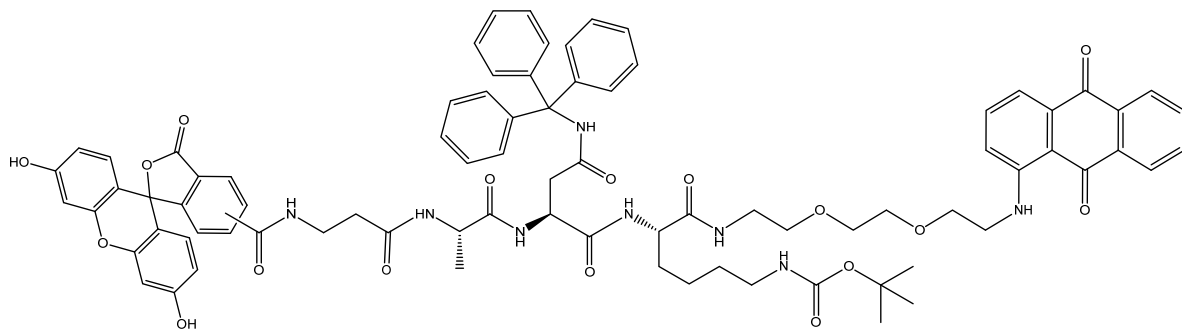
5(6)-CBF-βAla-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ **(40)**
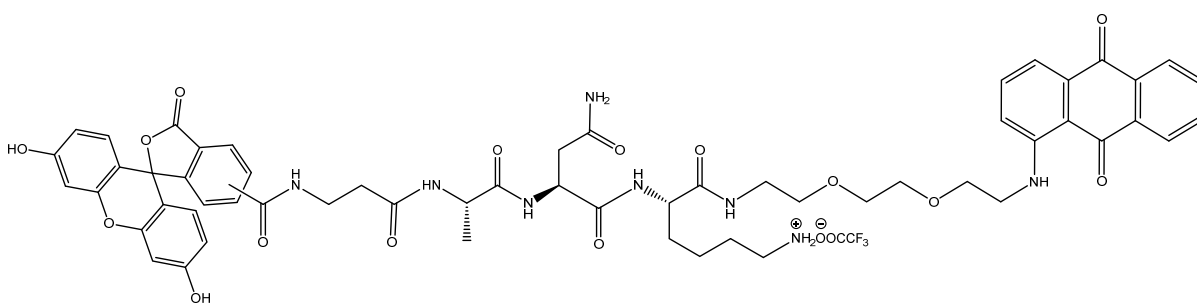
5(6)-CBF-bAla-Ala-Asn-Lys-Spacer-AQ **(41)**
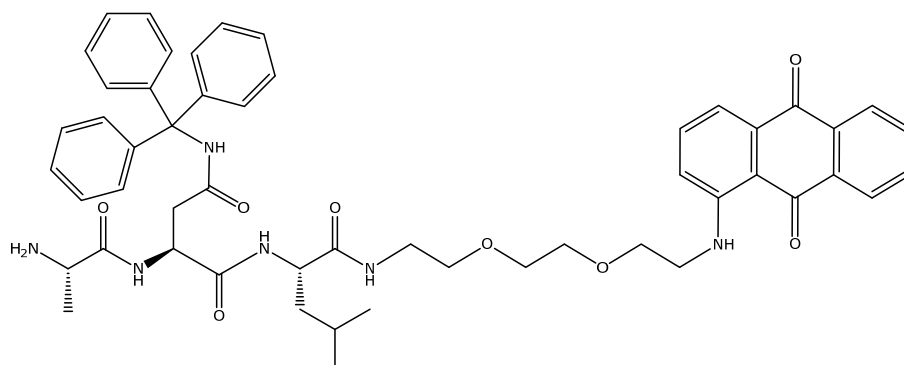
NH$_2$-Ala-Asn(Trt)-Leu-Spacer-AQ **(42)**

NH-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ **(43)**

166

NH$_2$-βAla-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ **(44)**

5(6)-CBF-βAla-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ **(45)**

5(6)-CBF-βAla-Pro-Ala-Asn-Leu-Spacer-AQ **(46)**
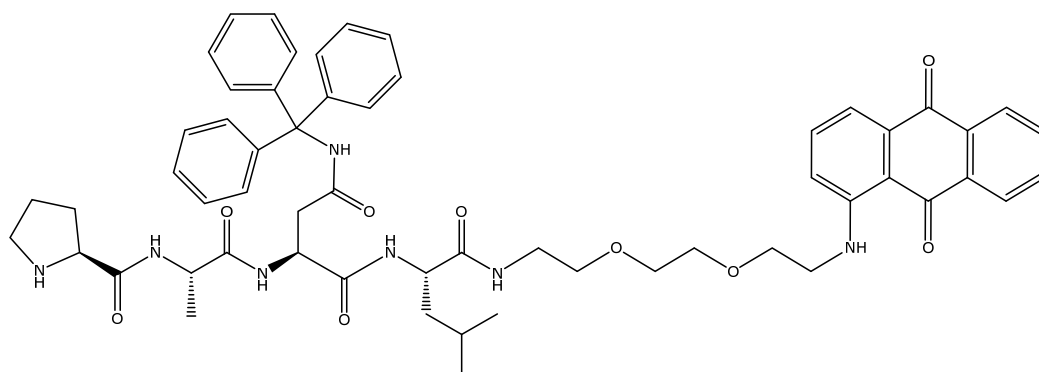
DOTA-tris-^{t}Bu-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ **(47)**
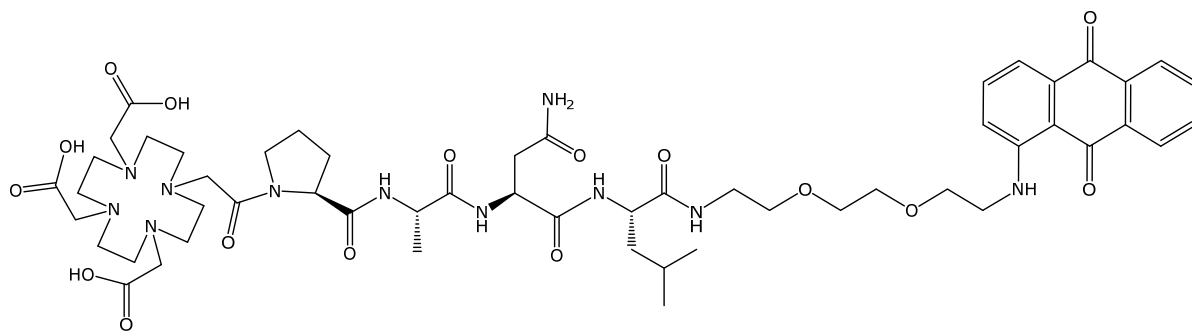
DOTA-Pro-Ala-Asn-Leu-Spacer-AQ **(48)**
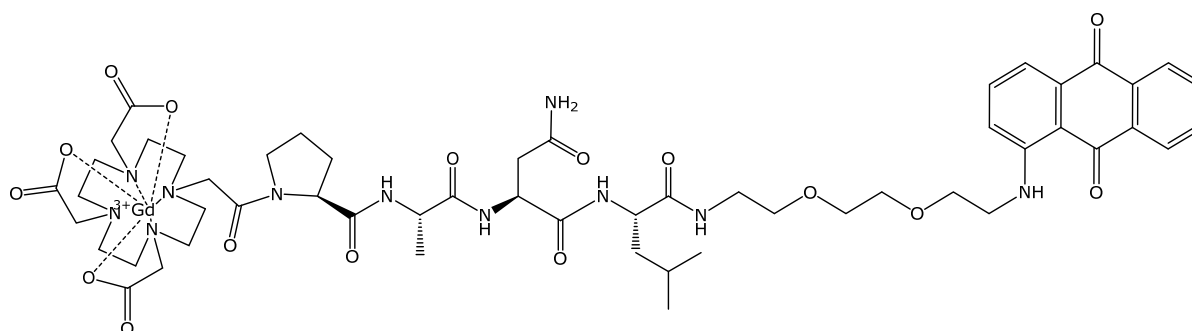
Gd(III)DOTA-Pro-Ala-Asn-Leu-Spacer-AQ **(49)**
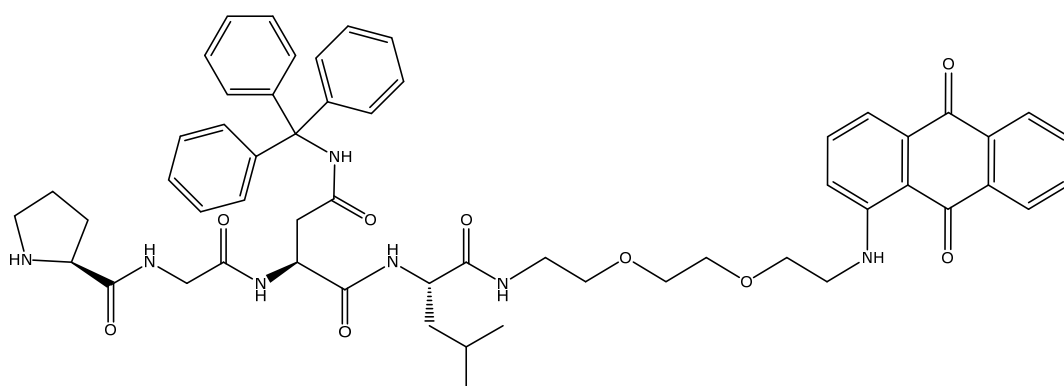
NH-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ **(50)**
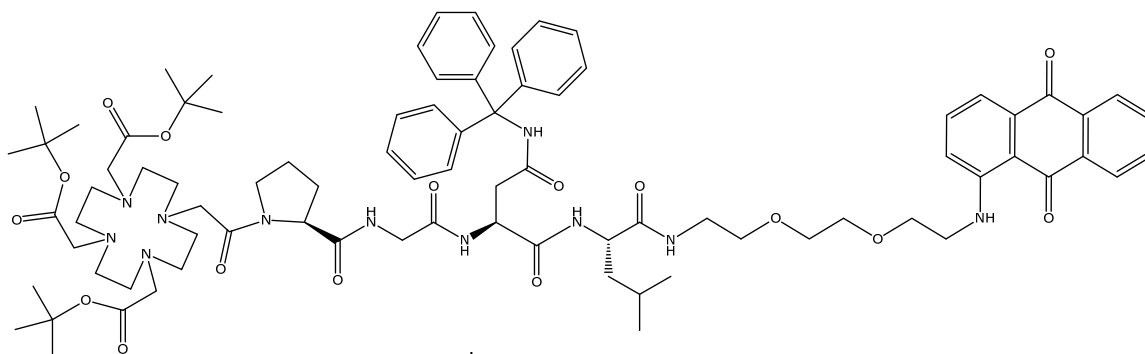
DOTA-tri-^{t}Bu-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ **(51)**
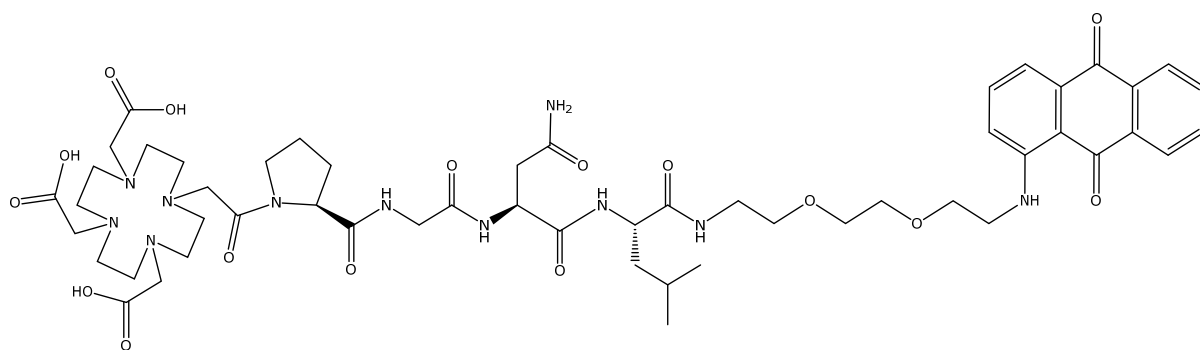
DOTA-Pro-Gly-Asn-Leu-Spacer-AQ **(52)**
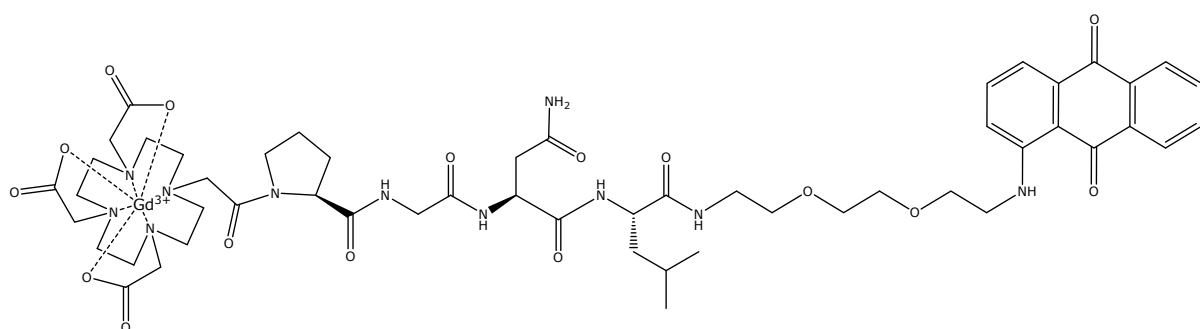
Gd(III)-DOTA-Pro-Gly-Asn-Leu-Spacer-AQ **(53)**
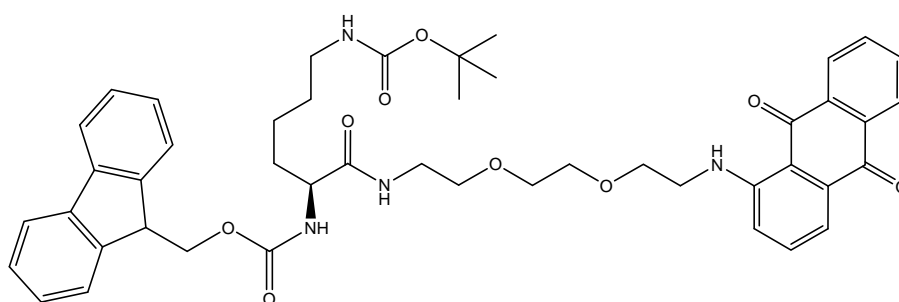
Fmoc-Lys(Boc)-Spacer-AQ **(54)**
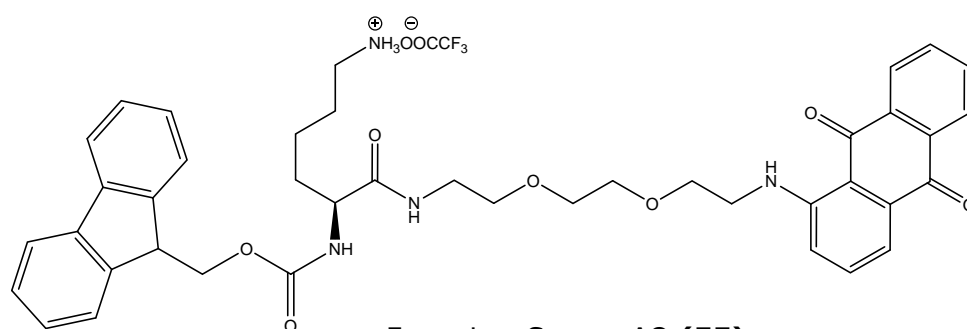
Fmoc-Lys-Spacer-AQ **(55)**
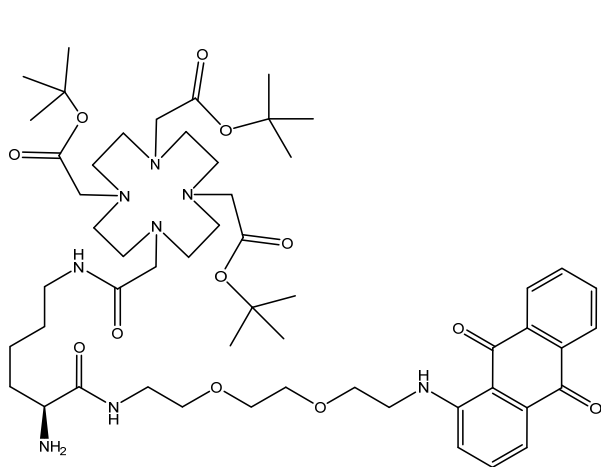
NH$_2$-Lys(DOTA-tri-^{t}Bu)-Spacer-AQ **(56)**
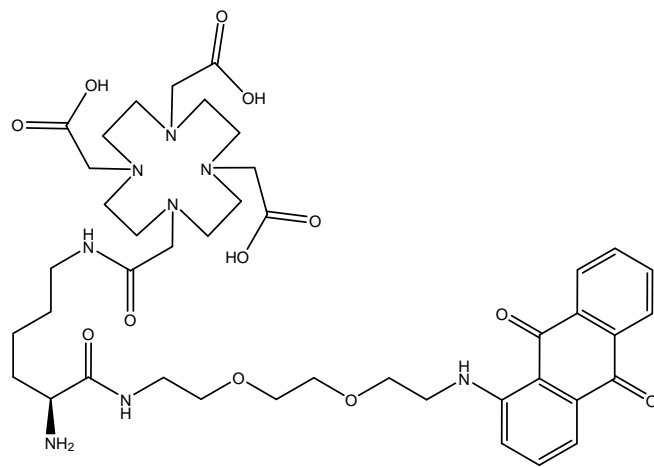
NH$_2$-Lys(DOTA)-Spacer-AQ **(57)**
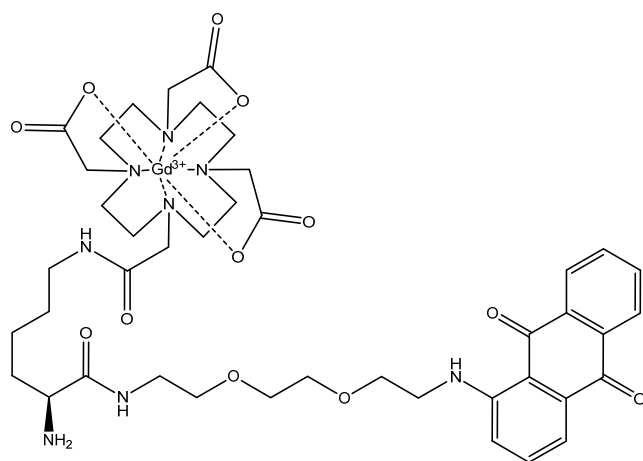
NH$_2$-Lys(DOTA-Gd)-Spacer-AQ **(58)**
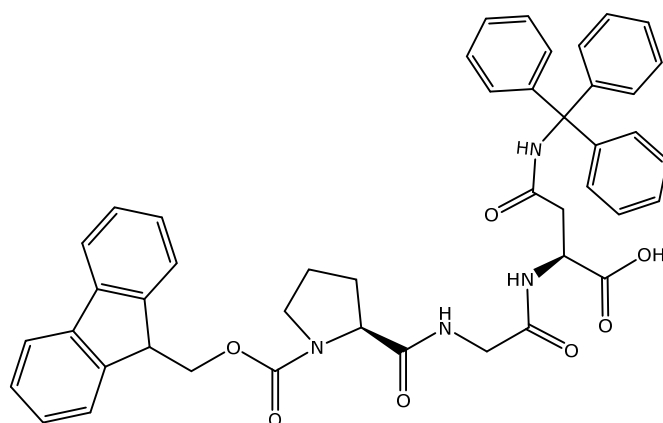
Fmoc-Pro-Gly-Asn(Trt)-OH **(59)**
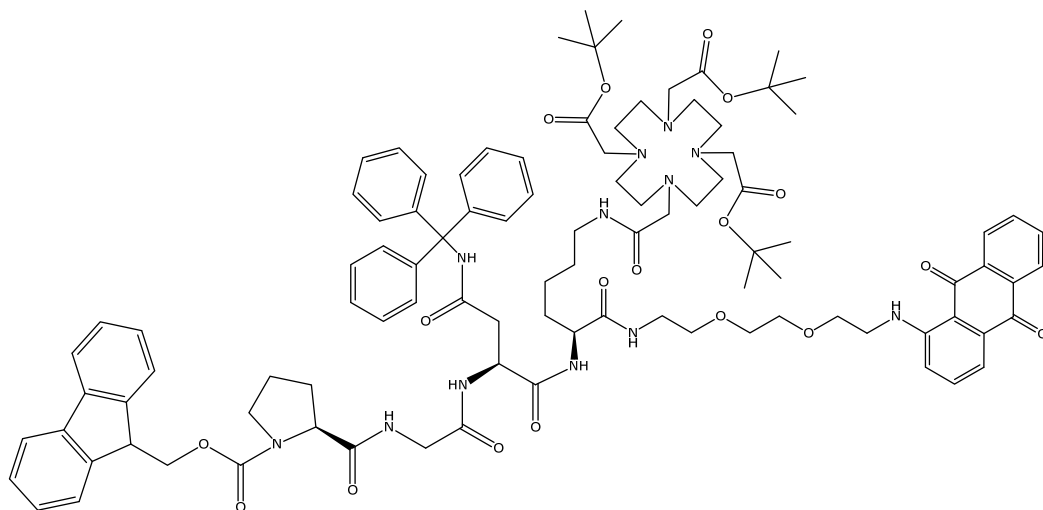
Fmoc-Pro-Gly-Asn(Trt)-Lys(DOTA-tri-^{t}Bu)-Spacer-AQ **(60)**
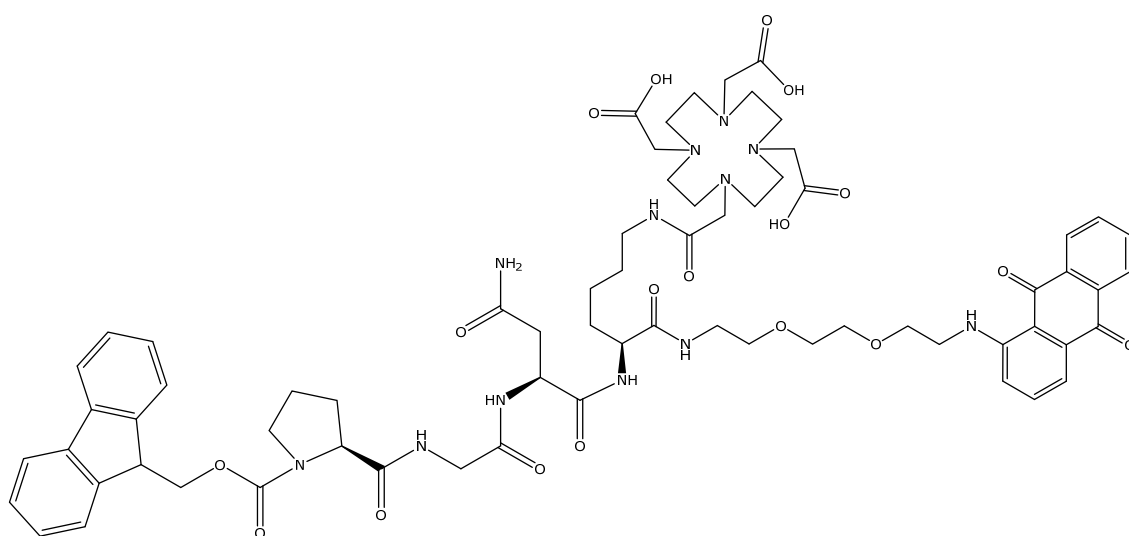
Fmoc-Pro-Gly-Asn-Lys(DOTA)-Spacer-AQ **(61)**
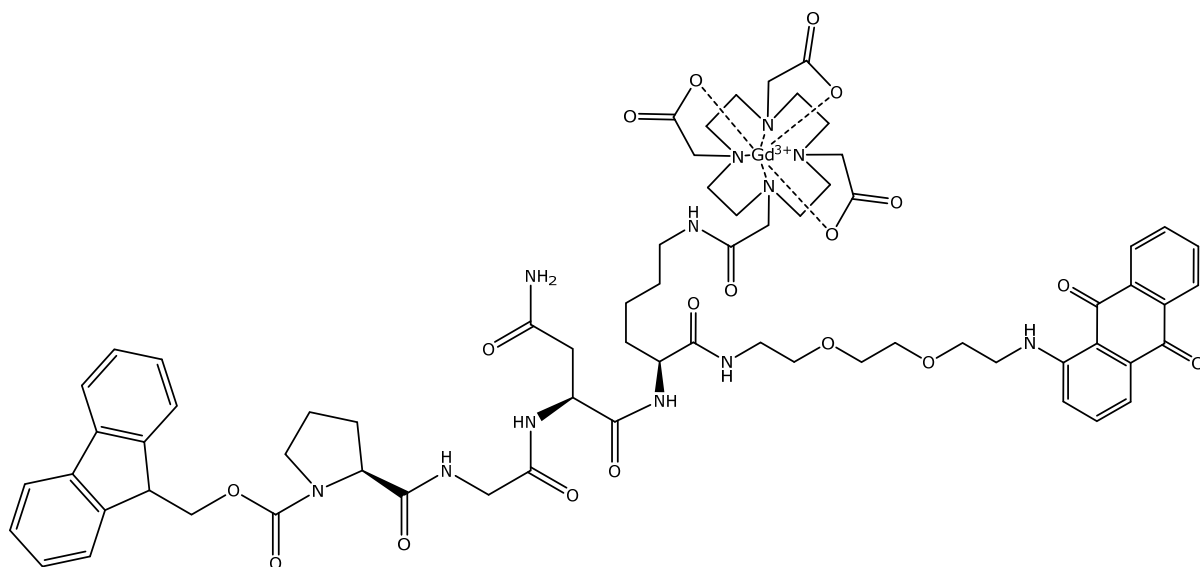
Fmoc-Pro-Gly-Asn-Lys(DOTA-Gd(III))-Spacer-AQ **(62)**
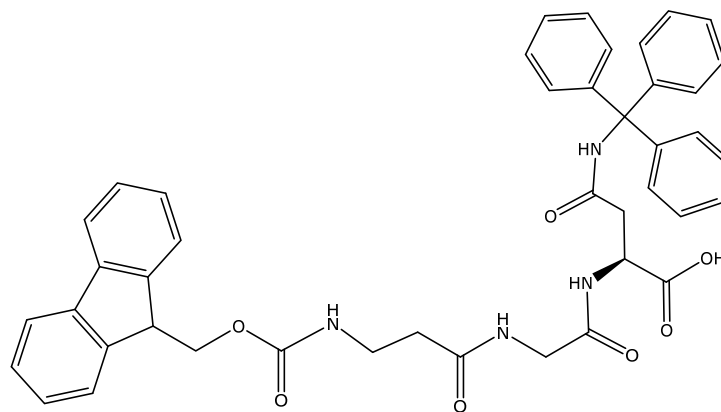
Fmoc-Ala-Gly-Asn(Trt)-OH **(63)**
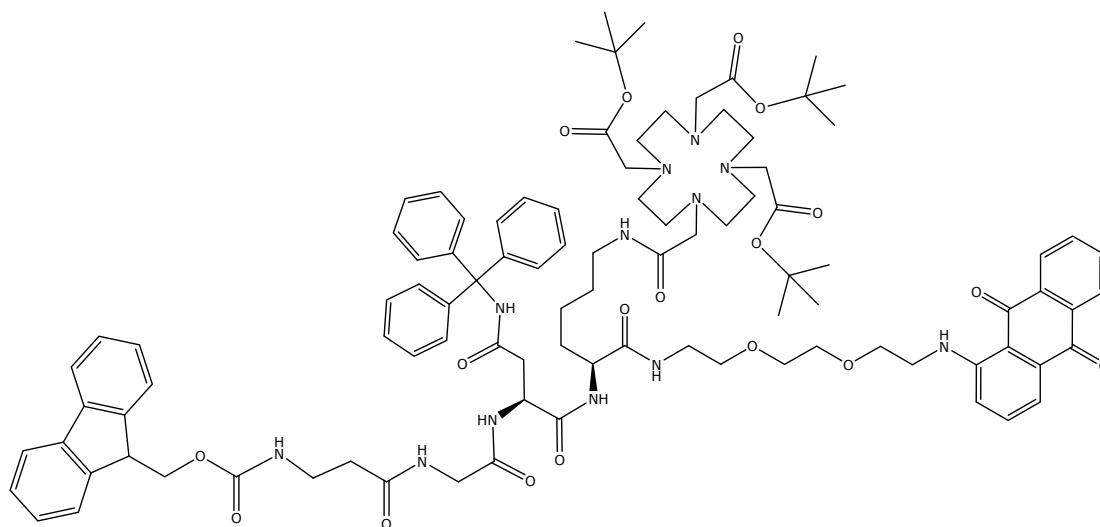
Fmoc-βAla-Gly-Asn(Trt)-Lys(DOTA-tri-^{t}Bu)-Spacer-AQ **(64)**
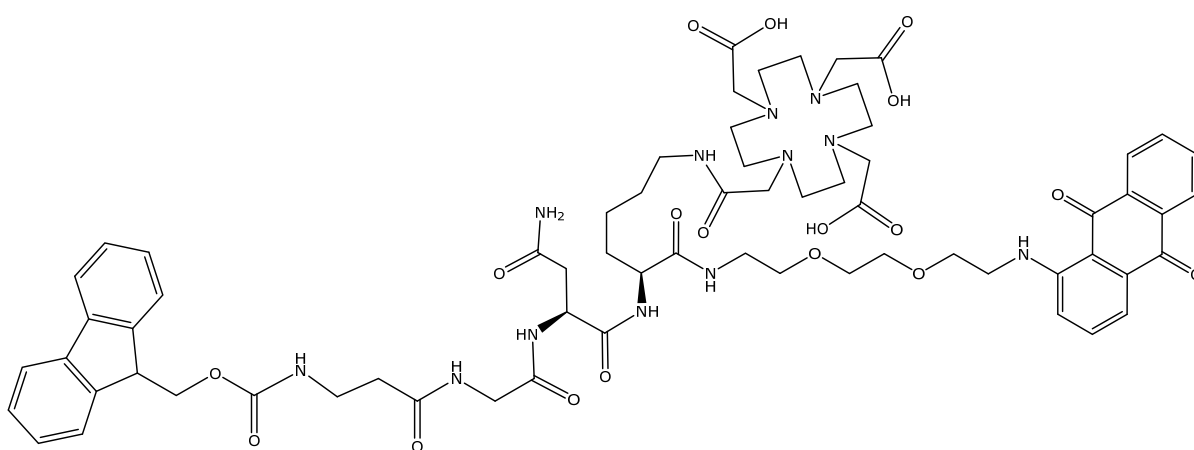
Fmoc-βAla-Gly-Asn-Lys(DOTA)-Spacer-AQ **(65)**
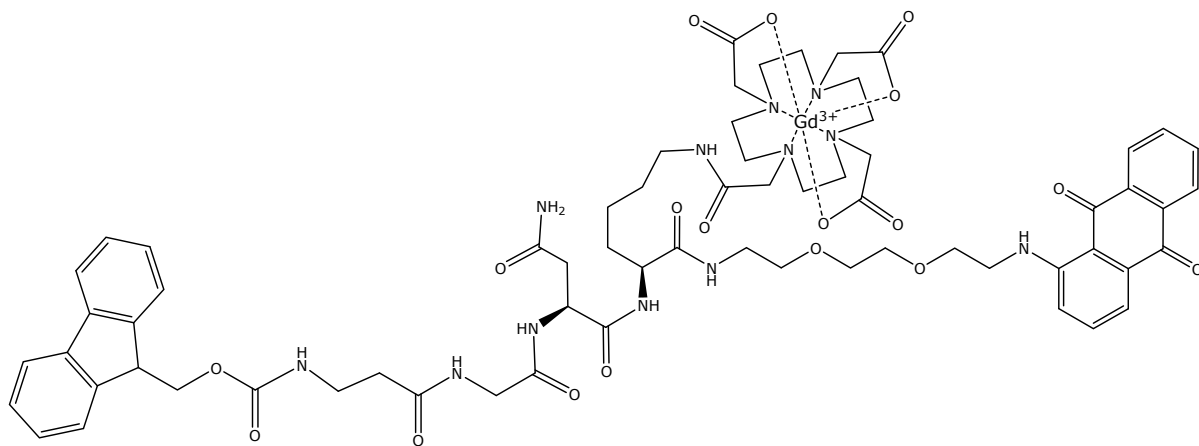
Fmoc-βAla-Gly-Asn-Lys(DOTA-Gd(III))-Spacer-AQ **(66)**
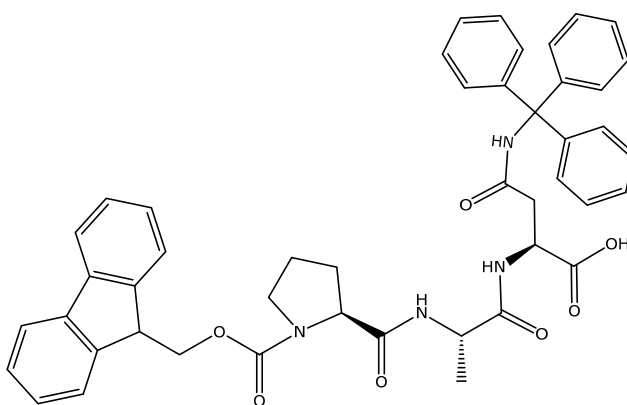
Fmoc-Pro-Ala-Asn(Trt)-OH **(67)**
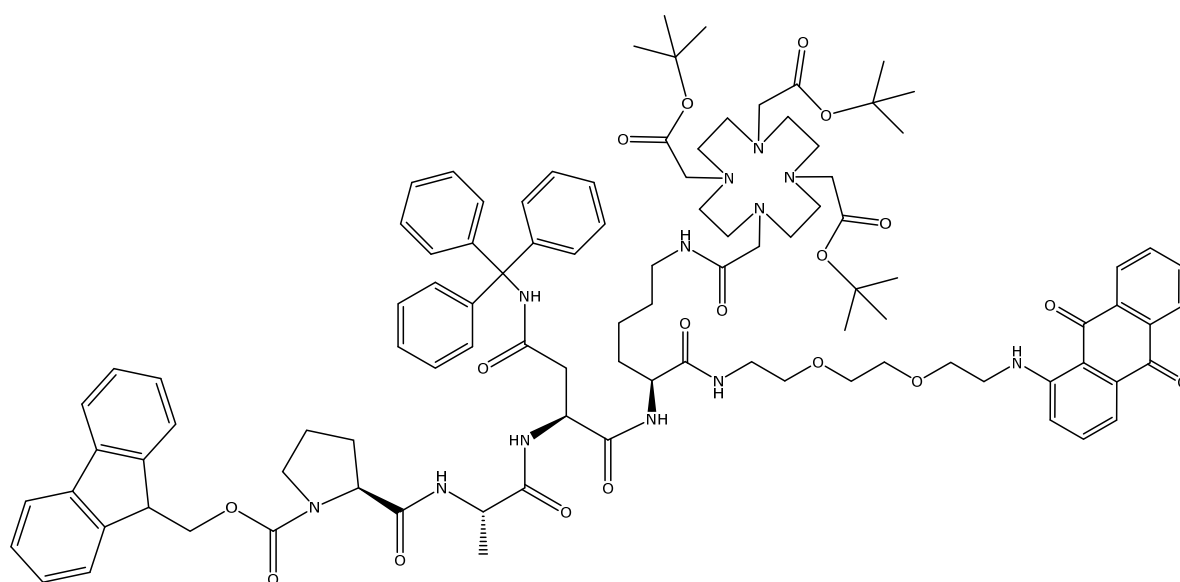
Fmoc-Pro-Ala-Asn(Trt)-Lys(DOTA-tri-^{t}Bu)-Spacer-AQ **(68)**
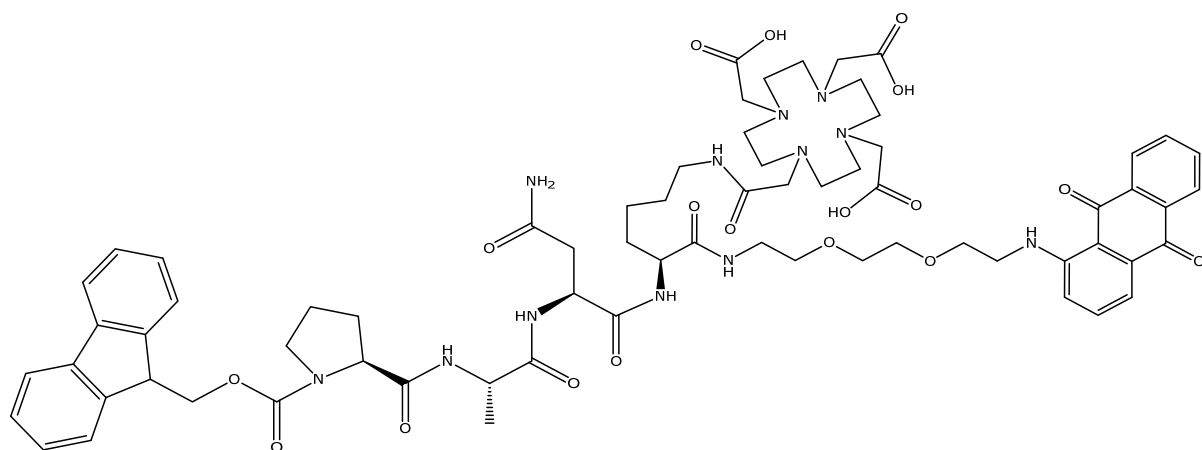
Fmoc-Pro-Ala-Asn-Lys(DOTA)-Spacer-AQ **(69)**
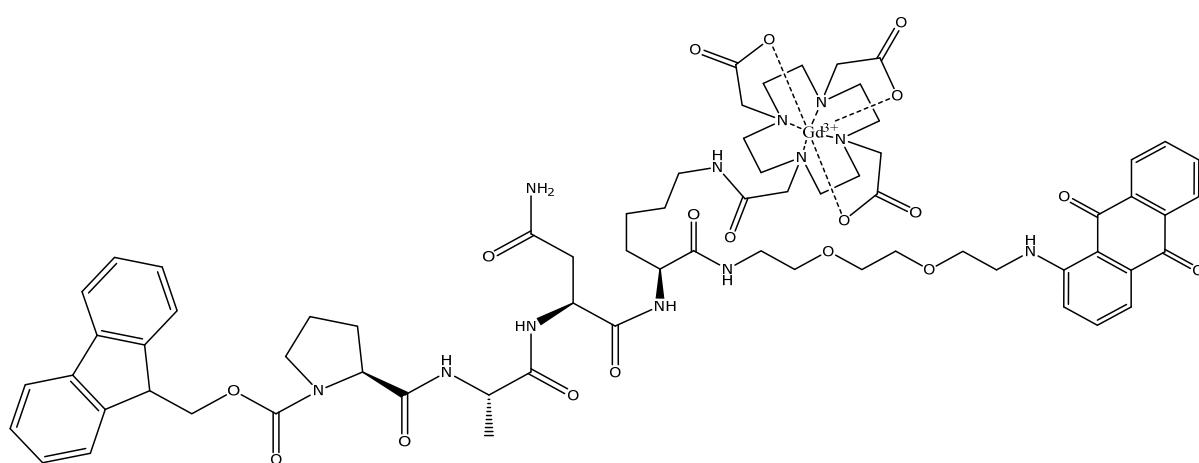
Fmoc-Pro-Ala-Asn-Lys(DOTA-Gd(III))-Spacer-AQ **(70)**

Fmoc-βAla-Ala-Asn(Trt)-OH **(71)**

174
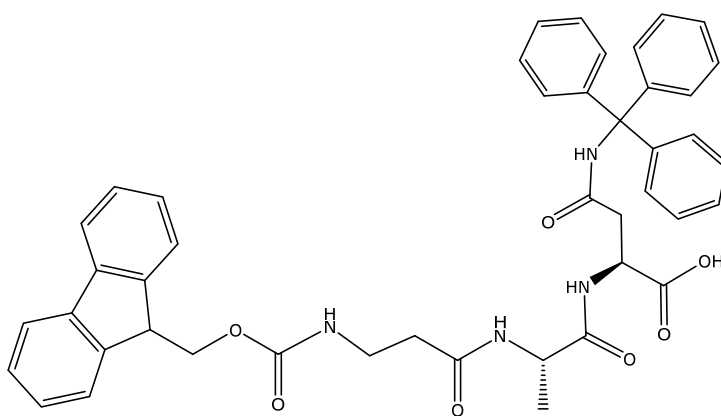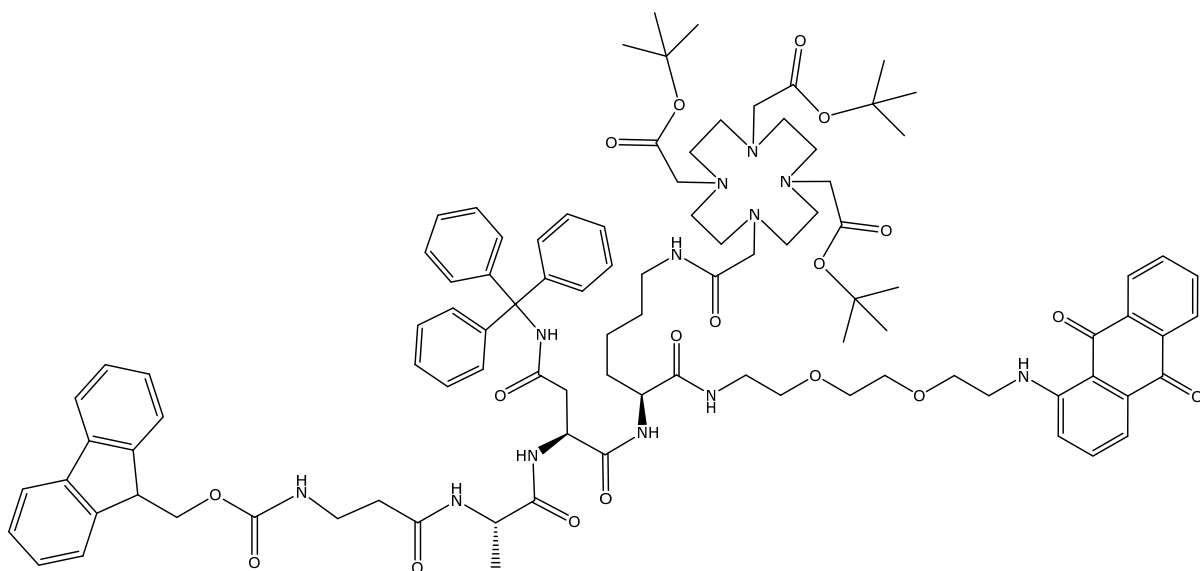
Fmoc-βAla-Ala-Asn(Trt)-Lys(DOTA-tri-^tBu)-Spacer-AQ **(72)**
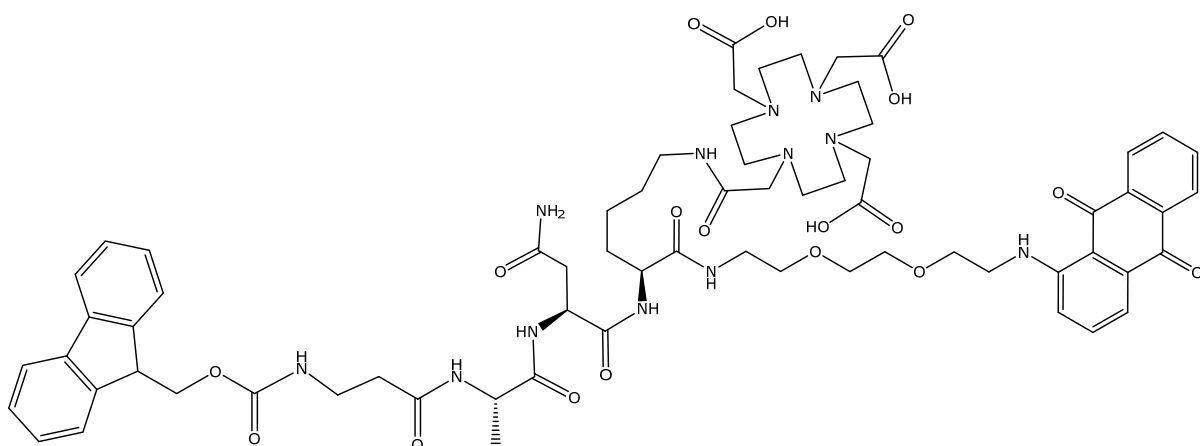
Fmoc-βAla-Ala-Asn-Lys(DOTA)-Spacer-AQ **(73)**
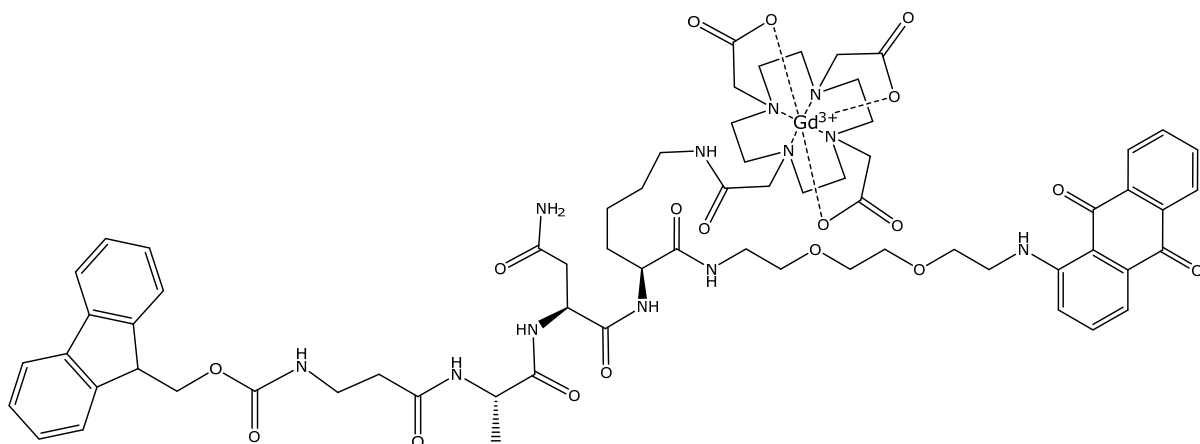
Fmoc-βAla-Ala-Asn-Lys(DOTA-Gd(III))-Spacer-AQ **(74)**

175
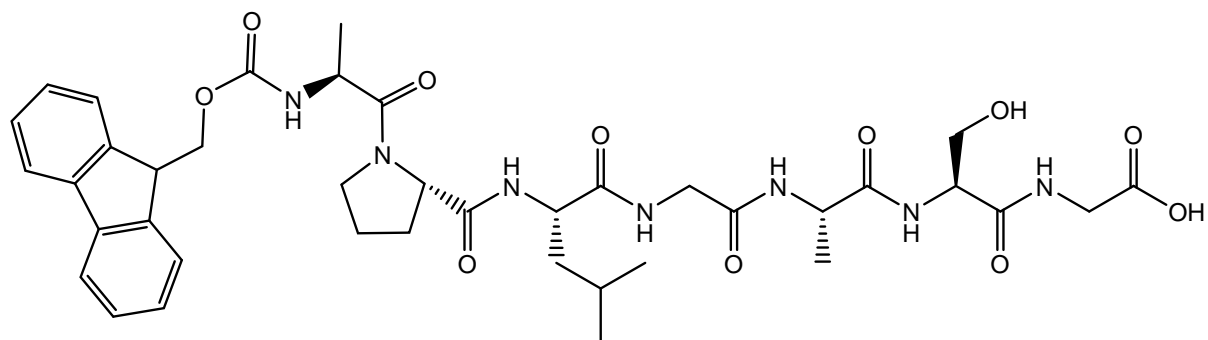
Fmoc-Ala-Pro-Leu-Gly-Ala-Ser-Gly-OH **(75)**
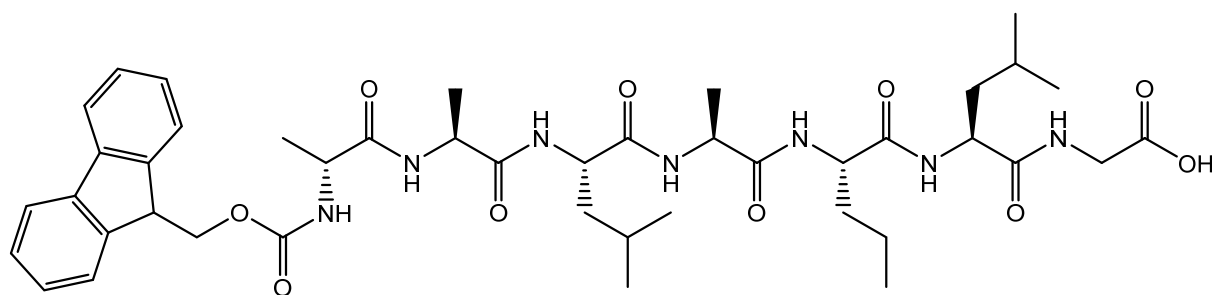
Fmoc-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH **(76)**

Fmoc-Lys-Pro-Ala-Gly-Nva-Ala-Gly-OH **(77)**

176
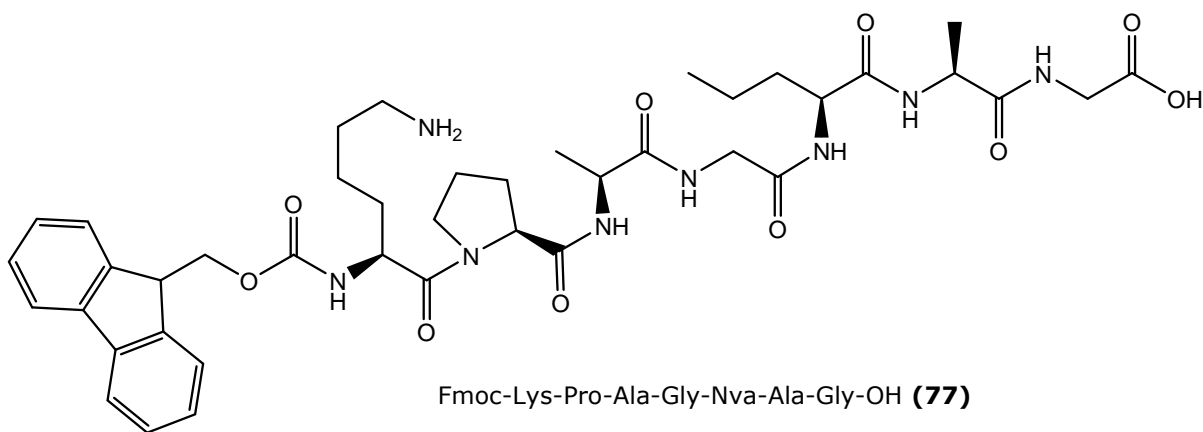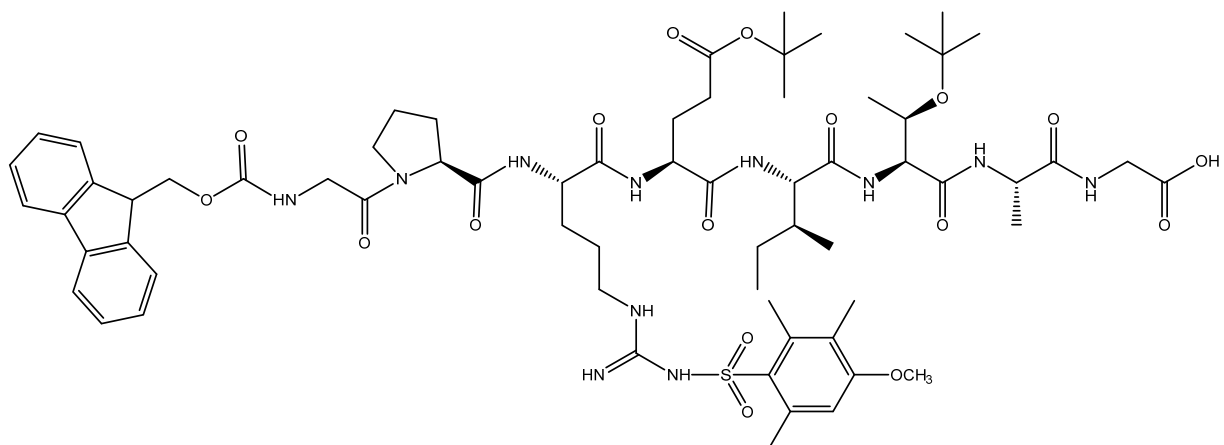
Fmoc-Gly-Pro-Arg(Mtr)-Glu(O^tBu)-Ile-Thr(^tBu)-Ala-Gly-OH**(78)**
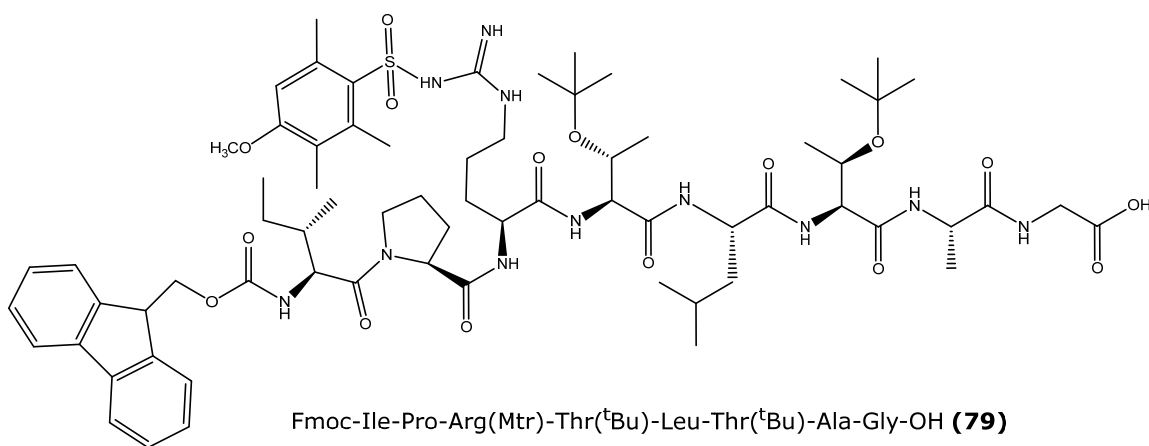
Fmoc-Ile-Pro-Arg(Mtr)-Thr(^tBu)-Leu-Thr(^tBu)-Ala-Gly-OH **(79)**
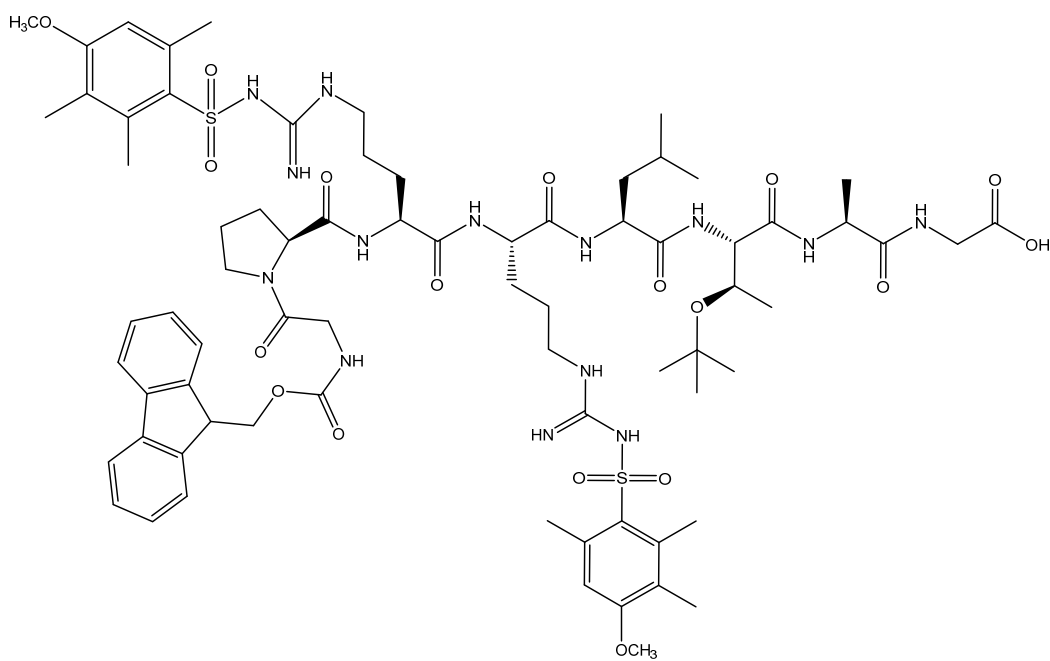
Fmoc-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr(^tBu)-Ala-Gly-OH **(80)**
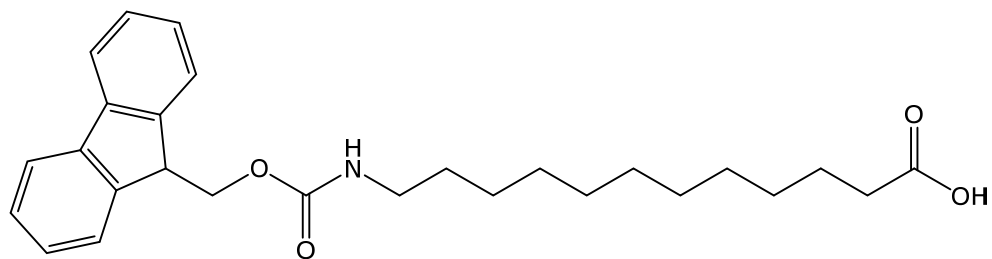
12-((((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)dodecanoic acid **(81)**
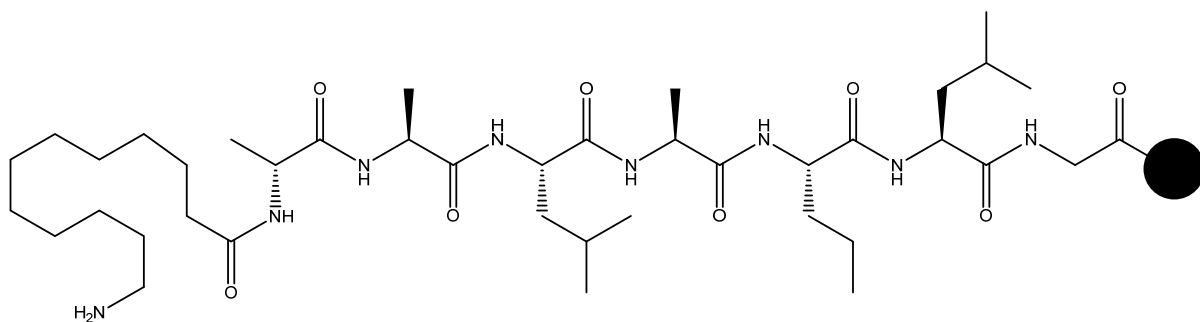
H-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-2-ClTrt resin **(82)**
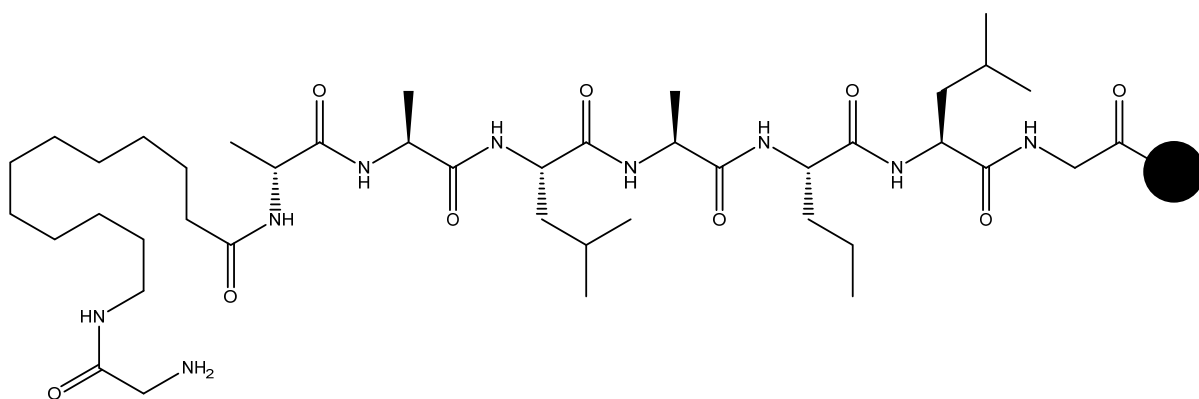
H-Gly-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-2-ClTrt resin **(83)**
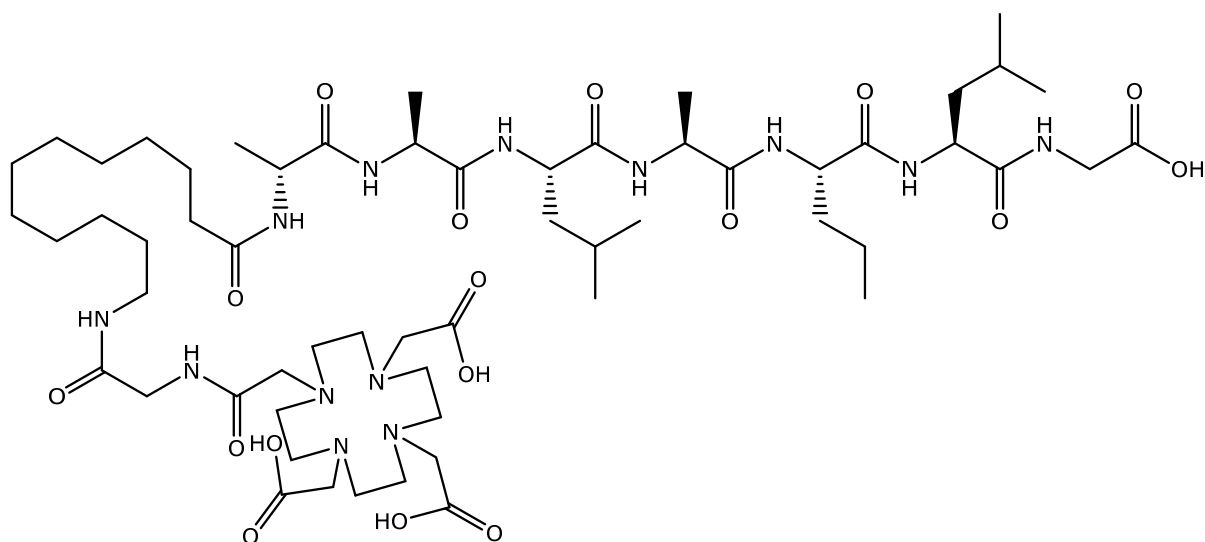
DOTA-Gly-12-aminododecanoic acid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH **(84)**

178
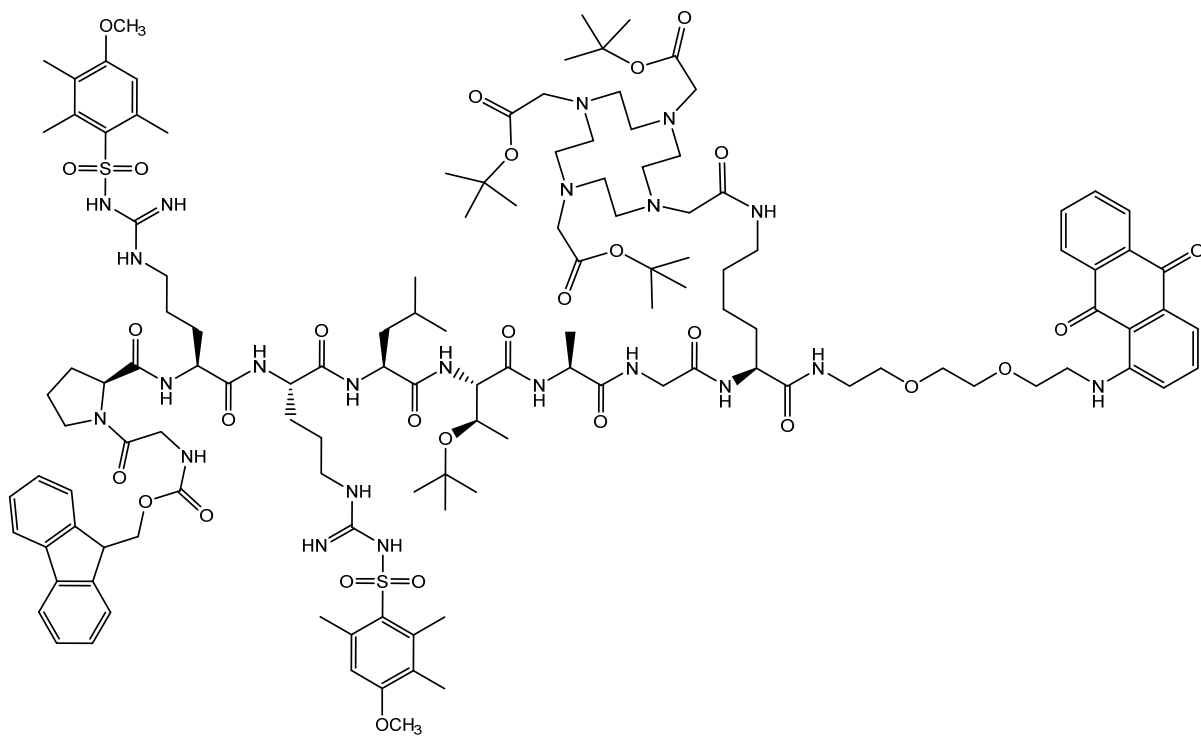
DOTA-tris(^{t}Bu)-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr(O^{t}Bu)-Ala-Gly-Lys-Spacer-AQ **(85)**
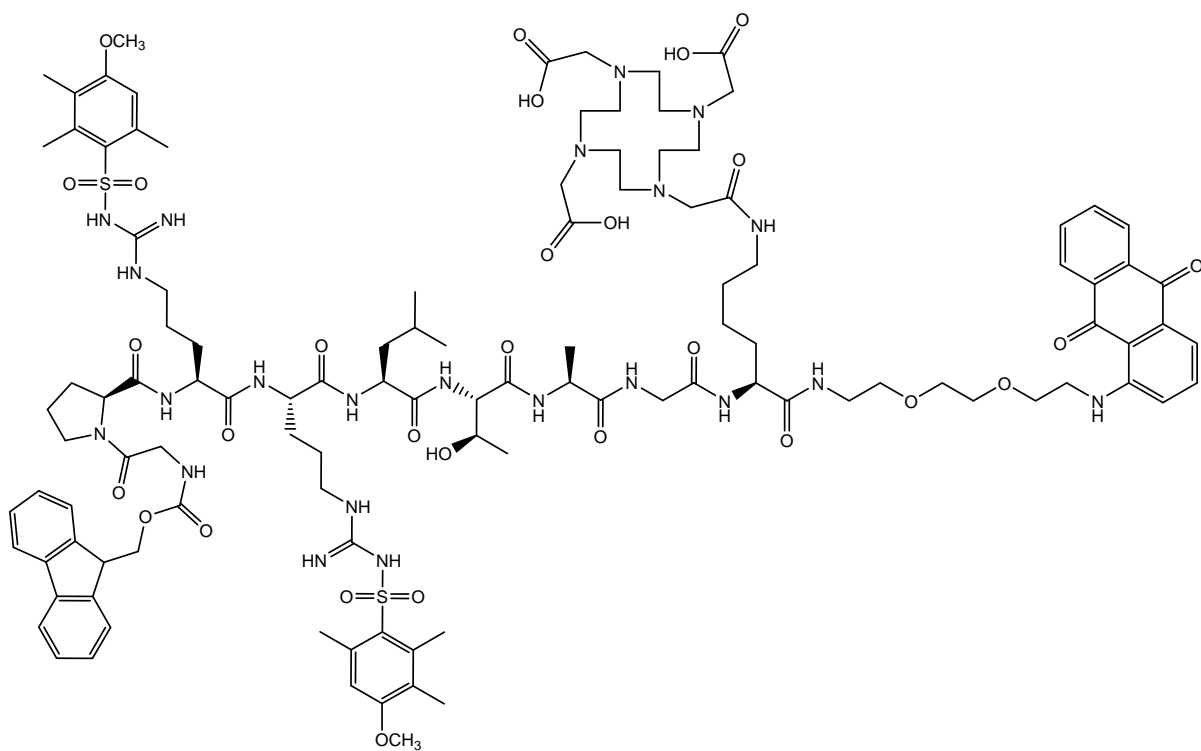
DOTA-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr-Ala-Gly-Lys-Spacer-AQ **(86)**

DOTA(Gd^{3+})-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr-Ala-Gly-Lys-Spacer-AQ **(JL-87)**

180
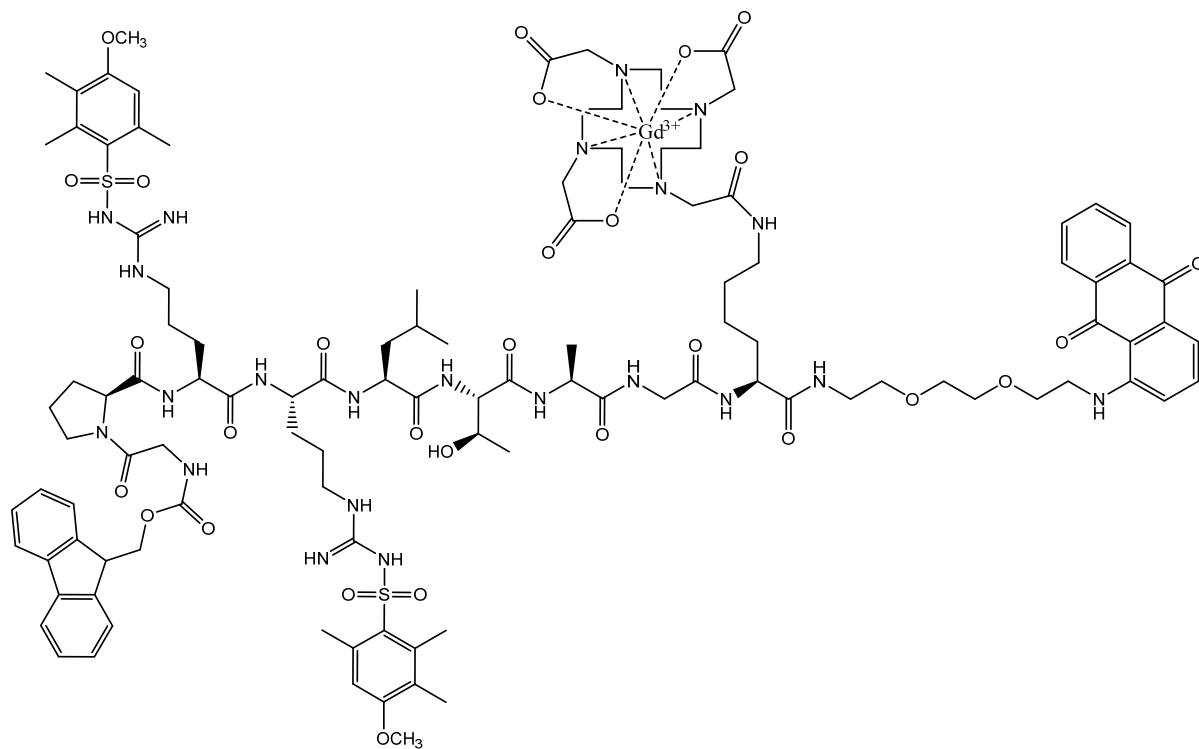
3. Experimental

3.1. Chemicals and reagents

All chemicals and reagents where purchased from Sigma, Molekula, Novabiochem, TCI and Fischer unless otherwise stated; and were used without purification.

3.2. Analytical methods

3.2.1. Chromatography

Thin layer chromatography (TLC) was performed on aluminum sheets (2 cmx 5 cm) pre-coated with silica gel-Kieselgel 60 F_{254}. Most of the compounds synthesized absorbed in the visible range however short wavelength UV light was used whenever additional visualization was required.

Column chromatography was carried out using glass columns of dimensions 30 cmx4.5 cm; 30 cmx2.5 cm; 30 cmx2 cm. The glass columns were packed with silica 60A; particle size ranging between 35-70 micron, suspended in the corresponding solvent system.

The solvent systems used for TLC/ column chromatography were:

Chloroform: Methanol (9:1; 19:1; 5:1; 3:1; 1:1).

Dichloromethane: Methanol (9.5:0.5; 9:1; 19:1; 5:1; 1:1).

Butanol: Acetic acid: Water (4:5:1).

Dichloromethane: Ethyl acetate: Methanol (6:2:1; 4:3:1; 5:2:1; 3:3:2).

Chloroform: Methanol: Pyridine: Ammonia (20:55:15:10).

3.2.2. UV- visible spectrophotometer

Helios-α UV-visible spectrophotometer (Thermo scientific) using VisionPro software, was used for all absorbance measurements with 1cm polystyrene cells for aqueous solutions and quartz cells for organic solvents.

Fluorescence measurements were performed in 1 cm quartz cells using LS-55 Luminescence spectrometer (Perkin Elmer), using FL Win lab software.

3.2.3. Mass spectrometry (MS)

All the samples synthesized (intermediates and final compounds) were characterized using low-resolution (ESMS) liquid column mass spectrometer- using

an Agilent column-ZORBAX 3.5 µm; SB-C18: 2.1 x 30mm guard column samples were prepared in HPLC grade methanol in a 1.5ml vial and analyzed in either positive or negative ion mode.

All final target compounds were also characterized using a high-resolution (HR) positive ion/negative ion nano-electrospray ionization (ESI) as performed on a thermofisher LTQ orbitrap XL mass spectrometer. The high resolution mass spectrum values are mentioned as ESI-HR and the low resolution mass spectrum values are denoted as ESMS-LR.

3.2.4. High Performance Liquid Chromatography (HPLC)

HPLC was performed on a Shimadzu Prominence LC-20AD series fitted with a UV-Vis detector (SPD-20 A), fluorescence detector (RF-10 AXL) and an auto sampler. Although 214nm is the most commomly set UV absorption wavelength for peptides (because the double bonds, such as peptide bonds absord at this wavelength) detection at other wavelengths (254nm, 245nm, 247nm and 300nm) was applied for detecting quenched fluorescent substrates and aromatic residues (Aguilar 2004). Column: Eclipse XDB-C18 RP column (4.6X150 mm; 5 µm particle size). Conditions: solvent A was 0.05% TFA in water and solvent B was 0.05% TFA in acetonitrile or methanol. The compounds were analysed at a concentration of 1 mg/ml with injection volume 20 µL and eluent flow rate=0.5 mL/min. The choice of the gradient conditions varied between samples and this was undertaken according to the separation pattern of each compound under investigation. After attempting a number of gradient conditions the following 8 were used for the compounds and the gradient conditions are as mentioned here:

Gradient A (23 min; 254 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
10	10	90
15	30	70
17	0	100
19	0	100
20	100	0
23 (Stop)	100	0

Gradient B (23 min; 245 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
6	10	90
8	30	70
15	0	100
18	0	100
20	100	0
23 (Stop)	100	0

Gradient C (28 min; 254 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0

Gradient D (28 min; 300 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0

5	10	90
8	30	70
17	0	100
25	0	100
26	100	0
28 (Stop)	100	0

5	10	90
8	30	70
17	0	100
25	0	100
26	100	0
28 (Stop)	100	0

Gradient E (23 min; 254 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
5	10	90
7	23	77
10	22	78
13	24	76
15	27	73
17	26	74
20	0	100
21	100	0
23 (stop)	100	0

Gradient F (20 min; 247 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
3	10	90
5	19	81
7	20	80
10	21	79
12	27	73
15	10	90
17	0	100
19	100	0
20 (stop)	100	0

Gradient G (23 min; 254 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
5	10	90
13	20	80
17	0	100
19	0	100
20	100	0
23 Stop	100	0

Gradient H (23 min; 247 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
7	10	90
15	30	70
17	0	100
19	0	100
20	100	0
23 Stop	100	0

Gradient I (23 min; 247 nm)

Time (min)	Solvent A(%)	Solvent B(%)
0	100	0
7	30	70
15	10	90
17	0	100
19	0	100
20	100	0
23 Stop	100	0

Before analyzing any compound the column was cleaned with 100% solvent B to remove any remaining samples absorbed on the column. This was followed by equilibrating the column with 100% solvent A prior to sample injection. Each eluted peak was manually collection in vials as individual fractions. ESMS on the collected fractions indicated the desired peak corresponding to the mass of interest.

3.2.5. Proton nuclear magnetic resonance (^{1}H NMR)

^{1}H NMR experiments were performed on a Bruker, nuclear magnetic resonance spectrometer (400 MHz/54 mm, software-Topspin). Samples were dissolved in either deuterated CDCl$_3$, deuterated DMSO or deuterated methanol.

Note: Yields quoted in g refer to isolated product(s). Yields quoted in % refer to isolated product(s), based on the limiting reagent, except, when crude starting materials were used where % yield (approximation) is quoted as a guide to the efficiency of the reaction.

3.3. General synthetic methods

3.3.1. Solid phase peptide synthesis
<u>**Method A: Resin swelling**</u>
The preloaded resin (0.5/0.3 g) was added to a SPPS reaction vessel. To this anhydrous DCM (20 mL) was added and the mixture was stirred/shaken (750 rpm) for two hours at room temperature (rt).

<u>**Method B: Amino acid coupling**</u>
Fmoc protected amino acid (7.5 eq), HOBt (7.35 eq), TBTU (7.38 eq) or PyBOP (7.38 eq) were dissolved in DMF (~10 mL) and to this solution DIPEA (11.25 eq) was added. The solution was then poured into the SPPS reaction vessel containing the swollen resin and stirred/shaken at room temperature for 20-30 minutes. Reaction completion was confirmed by performing Kaiser test. Before coupling each amino acid, the resin sample was washed with DMF (3x20 mL).

<u>**Method C: Peptide deprotection (Fmoc group)**</u>
After each coupling cycle the solvent was removed and the resin was washed with DMF (3x20 mL). Piperidine:DMF (20:80 v/v, 20 mL) was added to the resin in the reaction vessel and was stirred/shaken at room temperature for 5-15 minutes. Reaction completion was confirmed by performing Kaiser test.

<u>**Method D: Kaiser test**</u>
After each coupling and deprotection cycle the solvent was removed and the resin was washed with DMF (3x20 mL). To a few resin beads, in a small vial, ninhydrin, 6% in ethanol (2-3 drops), phenol ≈80% in ethanol (2-3 drops) and potassium cyanide in H$_2$O/pyridine (2-3 drops) were added. This mixture was then heated using a heat gun for few minutes. The resin beads along with the solution turned

dark blue when a free primary amine was present indicating a positive result and a successful deprotection.

Method E: Peptide cleavage

The solvent was removed and the resin was washed with DCM (3x20 mL). A solution of TFA in DCM (1% or 5% v/v, 15 mL) was added to the resin in the reaction vessel and stirred/shaken at room temperature for 30 minutes. The resin was then filtered and washed (x3) with TFA: DCM (1% or 5% v/v) and the filtered solution was evaporated under reduced pressure. The resulting concentrate was re-evaporated with ethanol (x3) and the residue was then triturated with diethyl ether and then stored at 4-6 °C for 12 hours. The resulting precipitate was filtered and dried in a vacuum desiccator.

3.3.2. Solution phase peptide synthesis

Method F: Amino acid coupling

The peptide sequence with the substituted aminoanthraquinone conjugate (AQ-spacer-A_n-NH_2, where A represents amino acids) at the carboxyl end was dissolved in DMF (5-15 mL). The solution was stirred at room temperature for 10-15 minutes. To this solution N-protected amino acids (1.1-3.3 eq), TBTU (1.1-3.3 eq), HOBT (1.1-3.3 eq)/ PyBOP (1.1-5.2 eq) dissolved in DMF along with DIPEA (3.2 eq) stirred at room temperature for 10-15 minutes was added. The reaction mixture was stirred at room temperature overnight. Reaction progress was monitored by TLC analysis. The reaction mixture was transferred to a separating flask and DCM/Chloroform was added. The product was then extracted into the organic layer and washed with water (3 x 100 mL), saturated sodium hydrogen carbonate (3 x 100 mL), water (3 x 100 mL), then dried over anhydrous magnesium sulphate ($MgSO_4$), filtered and evaporated at room temperature or using a rotary evaporator (Magnesium sulphate ($MgSO_4$) in the text referred to was anhydrous).

Method G: Removal of Fmoc protecting group

The Fmoc-protecting group was removed by adding 20% v/v piperidine in DMF to the peptide sequence. The reaction mixture was stirred at room temperature for 1 hour. Reaction completion was monitored and confirmed by TLC. The reaction mixture was transferred to a separating flask and DCM was added. The product was then extracted into the organic layer and washed with water (3 x 100 mL), saturated sodium hydrogen carbonate (3 x 100 mL), water (3 x 100 mL), then

dried over anhydrous magnesium sulphate (MgSO₄), filtered and evaporated at room temperature or using a rotary evaporator.

Method H: Removal of ᵗBoc, trityl and _tert_-butyl protecting groups

The side chain-protecting group was removed by treating with trifluoroacetic acid (100%) at room temperature for 30 minutes. Reaction progress was monitored by TLC. TFA was evaporated _in vaccuo_ and the resulting concentrate was re-evaporated with ethanol (5 mL, x3). The residue was then triturated with diethyl ether and stored at 4-6 °C for 12 hours. The resulting precipitate was filtered and dried in a vacuum desiccator.

3.4. Synthesis of 1ˢᵗ batch-FRET tetra peptide substrates (JC series - leucine)

3.4.1. Synthesis of Fluorescent labeled amino acids

I. Synthesis of 5(6)-Carboxyfluorescein-Pro-OᵗBu (1)

5(6)-Carboxyfluorescein (1.00 g, 2.66 mmol) and PyBOP (1.1 eq, 1.52 g, 2.92 mmol) were dissolved in DMF (4 mL). To the solution DIPEA (5.2 eq, 2.38 mL, 13.82 mmol) was added and stirred for 10 minutes. To this mixture, H-Pro-OᵗBu. HCl (1.1 eq, 0.61 g, 2.92 mmol) dissolved in DMF (2 mL) was added and stirred at room temperature for 6 hours. Reaction progress was monitored _via_ TLC. The solution was then washed with citric acid (5%) in water and the resulting solution was partitioned between DCM and water (3 x 100 mL). The organic extract was then evaporated to low volume. The crude product was purified by silica gel chromatography column (7 cm x 4.5 cm) prepared with chloroform: MeOH (4:1). Yield: 1.27 g (90%). TLC (2): R_f 0.9 (DCM: MeOH-4:1). The purity of the compound was 92% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient G, 23 min, 247 nm, Rt.10.23'). The molecular weight of the product (**1**) was 529.17, ESI-HR (+) m/z: 530.1804 (100%) (M+H)⁺. ESMS-LR (-) m/z: 528.0 (100%) (M-H)⁻ , 1057.2 (2M-H)⁻.

II. Synthesis of 5(6)-Carboxyfluorescein-Pro-OH (2)

The OᵗBu protected compound (**1**) (1.12 g, 2.12 mmol) was dissolved in TFA (5 mL) at rt. After 2 hours TFA was evaporated _in vaccuo_ and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (75 mL) and stored at 4-6 °C for 12 hours. The solution was filtered and dried in a vacuum desiccator. Crude yield: 0.76 g (76%). TLC: R_f 0.4 (DCM: MeOH- 9:1). The purity of the compound was 81% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient G,

23 min, 247 nm, Rt. 9.24'). The molecular weight of the product (**2**) was 473.11, ESI-HR (+) m/z: 474.1179 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 237.2 (60%) [(M+2H)/2]$^{2+}$, 474.1179 (100%) (M+H)$^+$. ESMS-LR (-) m/z: 472.0 (100%) (M-H)$^-$, 508.0 (40%) (M+Cl)$^-$.

III. Synthesis of 5(6)-Carboxyfluorescein-β-Ala-O^{t}Bu (3)

5(6)-Carboxyfluorescein (1.00 g, 2.66 mmol) and PyBOP (1.1 eq, 1.52 g, 2.92 mmol) were dissolved in DMF (4 mL). To this solution DIPEA (5.2 eq, 2.38 mL, 13.82 mmol) was added and stirred at room temperature for 10 minutes. To this mixture, H-βAla-O^{t}Bu. HCl (1.1 eq, 0.53 g, 2.92 mmol) dissolved in DMF (2 mL) was added. The reaction mixture was stirred at rt following the same procedure as the synthesis of compound **1**. The crude product was purified by silica gel chromatography column (7 cm x 4.5 cm) prepared with chloroform: MeOH (4:1). Yield: 1.03 g (77%). TLC (2): R$_f$ 0.8 (DCM: MeOH- 4:1). The purity of the compound was 94% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient F, 20 min, 247 nm, Rt. 8.62'). The molecular weight of the product (**3**) was 503.16, ESI-HR (+) m/z:504.1642 (100%) (M+H)$^+$. ESMS-LR (-) m/z: 502.0 (100%) (M-H)$^-$, 1005.0 (100%) (2M-H)$^-$

IV. Synthesis of 5,6-CBF-βAla-OH (4)

The O^{t}Bu protected compound (**3**) (0.79 g, 1.57 mmol) was dissolved in TFA (3 mL) at room temperature following the same procedure as the synthesis of compound **2**. Crude yield: 0.56 g (80%). TLC: R$_f$0.1 (DCM: MeOH- 9:1). The purity of the compound was 51% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient F, 20 min, 247 nm, Rt. 10.70'). The molecular weight of the product (**4**) was 447.10, ESI-HR (+) m/z: 448.1025 (100%) (M+H)$^+$. ESMS-LR (-) m/z: 446.0 (100%) (M-H)$^-$, 482.0 (20%) (M+Cl)$^-$.

3.4.2. Synthesis of anthraquinone-spacer conjugate (5)

1-Chloroanthraquinone (3.00 g, 12.34 mmol) was suspended in DMSO (20 mL). To this mixture 2-2′-(ethylenedioxy)bis(ethylamine) (5.2 eq, 9.39 mL, 64.29 mmol) was added and heated over a water bath for three hours at 60°C. The solution was cooled to rt followed by the addition of a large excess of water. The solution was filtered and the precipitates were air-dried at room temperature. The crude product was dissolved in DCM and then purified by a silica gel chromatography column (16 cmx4.5 cm) prepared with DCM. The eluent was then changed to DCM:MeOH (9:1). The product was finally eluted with the addition of

TEA (5%v/v) to the eluent. Fractions containing the product were combined, filtered and evaporated to dryness using a rotary evaporator. Yield: 1.38 g (31%). TLC (fraction 7): R_f0.3 (DCM:MeOH- 9:1). The purity of the compound was 83% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 13.82′). The molecular mass of the product (**5**) was 354.16, ESMS-LR (+) m/z: 355 (100%) $(M+H)^+$. ESMS-LR (-) m/z: 354 (40%) $(M)^-$.

^{1}H NMR (400MHz, d_6-DMSO), δ ppm: 2.65 (2H, t, J=5.6Hz, C$\underline{H}_2$-NH$_2$), 3.39 (2H, t, J=5.6Hz, C$\underline{H}_2$-CH$_2$-NH$_2$), 3.51-3.58 (4H, m, C$\underline{H}_2$-NH-AQ; C$\underline{H}_2$-O-CH$_2$-CH$_2$-NH$_2$), 3.62-3.64 (2H, m, C$\underline{H}_2$-O-CH$_2$-CH$_2$-NH-AQ), 3.72 (2H, t, J=5.6Hz, C$\underline{H}_2$-CH$_2$-NH-AQ), 7.28 (1H, d, J=8.4Hz, H-2), 7.44 (1H, dd, J_1=7.2Hz, J_2=0.8Hz, H-4), 7.62-7.66 (1H, m, H-3), 7.81-7.92 (2H, m, H-6 and H-7), 8.12 (1H, dd, J_1=7.6Hz, J_2=1.2Hz, H-5), 8.20 (1H, dd, J_1=7.6Hz, J_2=1.2Hz, H-8), 9.78 (1H, t, J=5.2Hz, N$\underline{H}$-AQ).

3.4.3. Synthesis of FRET tetra peptide substrates (JC series-1st batch)

3.4.3.1. Synthesis of 5(6)-CBF-Pro-Ala-Asn-Leu-Spacer-AQ (8)
I. Synthesis of 5(6)-CBF-Pro-Ala-Asn(Trt)-Leu-OH (6)

A leucine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as follows:

- 1st cycle: Removal of Fmoc protecting group was performed in accordance to method C. Presence of free amine due to removal of protecting group was confirmed by Kaiser test following method D. Fmoc-Asn(Trt)-OH (0.11 g, 0.31 mmol) was coupled to the resin according to method B. Primary amine detection confirming the completion of reaction was checked using the Kaiser test kit (method D).

- 2nd cycle: Removal of Fmoc protecting group was performed in accordance to method C. Presence of free amine due to removal of protecting group was confirmed by Kaiser test following method D. Fmoc-Ala-OH (0.10 g, 0.31 mmol) was coupled to the resin according to method B. Primary amine detection confirming the completion of reaction was checked using the Kaiser test kit (method D).

- 3rd cycle: Removal of Fmoc protecting group was performed in accordance to method C. Presence of free amine due to removal of protecting group was confirmed by Kaiser test following method D. 5(6)-CBF-Pro-OH (0.15

g, 0.31 mmol) was coupled to the resin according to method B. Primary amine detection confirming the completion of reaction was checked using the Kaiser test kit (method D).

Upon completion of the third coupling, the tetra peptide was removed from the resin following method E. Crude yield: 0.04 g (36%). The purity of the compound was 44% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 12.75'). The molecular weight of the product (**6**) was 1013.38, ESI-HR(+) m/z: 1014.3917 (100%) $(M+H)^+$. ESMS-LR (-) m/z: 1012.2 (70%) $(M-H)^-$.

II. Synthesis of 5(6)-CBF-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ (7)

Compound **6** (0.04 g, 0.04 mmol) was dissolved in DMF (1-2 mL). To this solution compound **5** (1.1 eq, 0.02 g, 0.04 mmol), HOBt (2 eq, 0.01 g, 0.08 mmol) TBTU (2 eq, 0.03 g, 0.08 mmol), DIPEA (5.8 eq, 0.04 mL, 0.23 mmol) dissolved in DMF (3-5 mL) was added and stirred at rt for 5hrs. The solution was partitioned between DCM and water. The organic extracts were washed with saturated sodium hydrogen carbonate (3x100 mL), then water (3x100 mL), dried with anhydrous $MgSO_4$, filtered and evaporated to a low volume. The crude product was re-dissolved in DCM:MeOH (9.5:0.5) and then applied to a silica gel chromatography column (15 cmx2 cm) prepared with DCM: MeOH (9.5:0.5). The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.04 g (80%). TLC (fraction 8): R_f 0.5 (DCM: MeOH-9:1). The molecular weight of the product (**7**) was 1349.53, ESMS-LR (+) m/z: 1350.4 (15%) $(M+H)^+$, 675.8 (100%) $[(M+2H)/2]^{2+}$. ESMS-LR (-) m/z: 1348.2 (60%) $(M-H)^-$, 1349.4 (80%) $(M)^-$, 1384.4 (15%) $(M+Cl)^-$.

III. Synthesis of 5(6)-CBF-Pro-Ala-Asn-Leu-Spacer-AQ (8)

Compound **7** was dissolved in TFA at rt for 2 hours. TFA was then evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (50 mL) and the precipitate was used without any further characterization. TLC: R_f 0.2 (DCM: MeOH-9.5:0.5).

3.3.3.2. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Leu-Spacer-AQ (11)
I. Synthesis of 5(6)-CBF-Pro-Gly-Asn(Trt)-Leu-OH (9)

A leucine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide removal from the resin according to method E:

- 1st cycle: *N*-Fmoc-L-Asparagine (Trt) (0.47 g, 0.79 mmol).
- 2nd cycle: *N*-Fmoc-L-Glycine (0.23 g, 0.79 mmol).
- 3rd cycle: 5 (6), CBF-Proline (0.37 g, 0.79 mmol).

Upon successful completion of third coupling, the tetra peptide was removed from the resin following method E. Crude yield: 0.08 g (80%). The molecular weight of the product (**9**) was 999.37, ESMS-LR (+) m/z: 1000.2 (100%) $(M+H)^+$, 1022.0 (25%) $(M+Na)^+$. ESI-HR (+) m/z 998.3599 (100%) $(M-H)^-$. ESMS-LR (-) m/z: 498.8 (70%) $([M-2H)/2]^{2-}$, 1113.0 (100%) $(M+TFA)^-$.

II. Synthesis of 5(6)-CBF-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ (10)

Compound **9** (0.08 g, 0.08 mmol) was dissolved in DMF (2 mL). To this solution compound **5** (1.1 eq, 0.03 g, 0.09 mmol), HOBt (2 eq, 0.02 g, 0.16 mmol) TBTU (2 eq, 0.05 g, 0.16 mmol), DIPEA (5.8 eq, 0.08 mL, 0.46 mmol) dissolved in DMF (3-5 mL) following the same procedure as described for the synthesis of compound **7**. The crude product was re-dissolved in DCM:MeOH (9.5:0.5) and then applied to a silica gel chromatography column (15 cmx2 cm) prepared with DCM:MeOH (9.5:0.5). The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: **0.09 g (82%)**. TLC (fraction 3): R_f0.5 (DCM: MeOH-9.5:0.5). The molecular weight of the product (**10**) was 1335.52, ESI-HR (+) m/z: 668.7648 (100%) $[(M+2H/2)]^{2+}$. ESMS-LR (+) m/z: 453.2 (100%) $[(M+2H+Na)/3]^{3+}$. ESMS-LR (-) m/z: 666.7 (100%) $[(M-2H)/2]^{2-}$, 1334.4 (100%) $(M-H)^-$.

III. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Leu-Spacer-AQ (11)

Compound **10** was dissolved in TFA at rt for 2 hours. TFA was then evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was triturated with diethyl ether (50 mL), and the precipitate was used without any further characterization. TLC: R_f0.2 (DCM: MeOH-9.5:0.5).

3.3.3.3. Synthesis of 5(6)-CBF-Pro-Leu-Asn-Leu-Spacer-AQ (14)
I. Synthesis of 5(6)-CBF-Pro-Leu-Asn(Trt)-Leu-OH (12)

A leucine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:
- 1st cycle: *N*-Fmoc-L-Asparagine (Trt) (0.47 g, 0.79 mmol).
- 2nd cycle: *N*-Fmoc-L-Leucine (0.28 g, 0.79 mmol).
- 3rd cycle: 5 (6), CBF-Proline (0.37 g, 0.79 mmol).

Upon successful completion of the third coupling, the tetra peptide was cleaved from the resin following method E. Crude yield: 0.06 g (55%). The purity of the compound was 46% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 13.44'). The molecular weight of the product (**12**) was 1055.43, ESI-HR (+) m/z: 1056.4388 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 539.8 (45%) [(M+2H+Na)/2]$^{2+}$, 1078.2 (25%) (M+Na)$^+$. ESMS-LR (-) m/z: 1054 (100%) (M-H)$^-$.

II. Synthesis of 5(6)-CBF-Pro-Leu-Asn(Trt)-Leu-Spacer-AQ (13)

Compound **12** (0.06 g, 0.06 mmol) was dissolved in DMF (2 mL). To this solution compound **5** (1.1 eq, 0.02 g, 0.06 mmol) along with HOBt (2 eq, 0.02 g, 0.11 mmol), TBTU (1.1 eq, 0.04 g, 0.11 mmol), DIPEA (5.8 eq, 0.06 mL, 0.33 mmol) dissolved in DMF (5-7 mL) following the same procedure as described for the synthesis of compound **7**. The crude product was re-dissolved in DCM:MeOH (9.5:0.5) and then applied to a silica gel chromatography column (14 cmx2 cm) prepared with DCM:MeOH (9.5:0.5). The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 0.05 g (63%). TLC (fraction 2): R$_f$ 0.8 (DCM:MeOH-9:1). The molecular weight of the product (**13**) was 1391.58, ESI-HR (+) m/z: 1392.5854 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 1436.6 (35%) (M+2Na-H)$^+$. ESMS-LR (-) m/z: 1390.4 (35%) (M-H)$^-$.

III. Synthesis of 5(6)-CBF-Pro-Leu-Asn-Leu-Spacer-AQ (14)

Compound **13** was dissolved in TFA at room temperature for 2 hours. TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (75 mL), and the precipitate was used without any further characterization. TLC: R$_f$ 0.1 (DCM:MeOH- 9.5:0.5).

3.3.3.4. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Leu-Spacer-AQ (17)
I. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Leu-OH (15)

A leucine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:

- 1st cycle: *N*-Fmoc-L-Asparagine (Trt) (0.47 g, 0.79 mmol).
- 2nd cycle: *N*-Fmoc-L-Alanine (0.26 g, 0.79 mmol).
- 3rd cycle: 5 (6), CBF-β-Alanine (0.35 g, 0.79 mmol).

Upon successful completion of the third coupling, the tetra peptide was cleaved from the resin following method E. Crude yield: 0.07 g (70%). The purity of the compound was 70% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 275 nm, Rt. 13.23'). The molecular weight of the product (**15**) was 987.37, ESMS-LR (+) m/z: 505.8 (25%) $[(M+2H+Na)/2]^{2+}$, 988.2 (100%) $(M+H)^+$, 1010.0 (40%) $(M+Na)^+$. ESI-HR (-) m/z: 986.3593 (85%) $(M-H)^-$. ESMS-LR (-) m/z: 987.0 (60%) $(M)^-$.

II. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn(Trt)-Leu-Spacer-AQ (16)

Compound **15** (0.06 g, 0.06 mmol) was dissolved in DMF (2 mL). To this solution compound **5** (1.1 eq, 0.02 g, 0.07 mmol) along with HOBt (2 eq, 0.02 g, 0.12 mmol), TBTU (2 eq, 0.04 g, 0.12 mmol), DIPEA (5.8 eq, 0.06 mL, 0.35 mmol) dissolved in DMF (5-7 mL) following the same procedure as described for the synthesis of compound **7**. The crude product was re-dissolved in DCM:MeOH (9.5:0.5) and then applied to a silica gel chromatography column (14 cmx2 cm) prepared with DCM:MeOH (9.5:0.5). The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 0.02 g (25%). TLC (frcation 3): R_f 0.4 (DCM:MeOH-9:1). The molecular weight of the product (**16**) was 1323.52, ESI-MS (+) m/z: 1324.5231 (100%) $(M+H)^+$. ESMS-LR (+) m/z: 662.7 (100%) $[(M+H/2)]^{2+}$, 673.7 (90%) $[(M+2H+Na)/2]^{2+}$.

III. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Leu-Spacer-AQ (17)

Compound **17** was dissolved in TFA at rt for 2 hours. TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (50mL), and the precipitate was used without any further characterization. TLC: R_f 0.3 (DCM:MeOH-9.5:0.5).

3.3.3.5. Synthesis of 5(6)-CBF-β-Ala-Leu-Asn-Leu-Spacer-AQ (20)
I. Synthesis of 5(6)-CBF-β-Ala-Leu-Asn-leu-OH (18)

A leucine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:

- 1st cycle: *N*-Fmoc-L-Asparagine(Trt) (0.47 g, 0.79 mmol).
- 2nd cycle: *N*-Fmoc-L-Leucine (0.29 g, 0.79 mmol).
- 3rd cycle: 5 (6), CBF-β-Alanine (0.35 g, 0.79 mmol).

Upon successful completion of the third coupling, the tetra peptide was cleaved from the resin following method E. Crude yield: 0.07 g (63%). The purity of the compound was 41% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 16.01'). The molecular weight of the product (**18**) was 1029.42, ESMS-LR (+) m/z: 1030.2 (100%) $(M+H)^+$. ESI-HR (-) m/z: 1028.4071 (100%) $(M-H)^-$. ESMS-LR (-) m/z: 1029.2 (80%) $(M)^-$.

II. Synthesis of 5(6)-CBF-β-Ala-Leu-Asn(Trt)-Leu-Spacer-AQ (19)

Compound **18** (0.07 g, 0.07 mmol) was dissolved in DMF (2 mL). To this solution compound **5** (1.1 eq, 0.03 g, 0.07 mmol) along with HOBt (2 eq, 0.02 g, 0.14 mmol), TBTU (2 eq, 0.04 g, 0.14 mmol), DIPEA (5.8 eq, 0.07 mL, 0.39 mmol) dissolved in DMF (5-7 mL) following the same procedure as described for the synthesis of compound **7**. The crude product was re-dissolved in DCM:MeOH (9.5:0.5) and then applied to a silica gel chromatography column (13 cmx2 cm) prepared with the same solvent system. The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 0.06 g (67%). TLC (fraction 5): R_f 0.4 (DCM:MeOH-9:1). The molecular weight of the product (**19**) was 1365.56, ESI-HR (+) m/z: 683.7888 (100%) $[(M+H/2)]^{2+}$. ESMS-LR (+) m/z: 1366.5 (95%) $(M+H)^+$.

III. Synthesis of 5(6)-CBF-β-Ala-Leu-Asn-Leu-Spacer-AQ (20)

Compound **20** was dissolved in TFA at rt for two hours. TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (75 mL), and the precipitate was used without any further characterization. TLC: R_f 0.2 (DCM: MeOH-9.5:0.5).

3.3.3.6. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Leu-Spacer-AQ (25)
I. Synthesis of intermediate compound H₂N-Leu-Spacer-AQ (21) (Method: F,G)

Compound **5** (1.00 g, 2.82 mmol) was dissolved in DMF (10-15 mL) by stirring the mixture for 20 minutes at rt. To this solution Fmoc-Leu-OH (1.1 eq, 1.10 g, 3.11 mmol), TBTU (1.1 eq, 1.00 g, 3.11 mmol) and HOBt (1.1 eq, 0.42 g, 3.11 mol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 1.56 g/ml, 9.04 mmol) was added and at stirred rt. The progress of the reaction was monitored *via* TLC. The solution was then partitioned between DCM and water. The organic extract was washed with saturated sodium hydrogen carbonate (3 x 100 mL), then

water (3 x 100 mL), dried with anhydrous MgSO$_4$, filtered and evaporated to dryness and was used for the following reaction without further purification.

The extracted sample was dissolved in 20% v/v of piperidine in DMF (20 mL) and the amino terminus was deprotected as in method G. The solution was partitioned between DCM and water. The organic extracts were washed with saturated sodium hydrogen carbonate (3 x 100 mL), then water (3 x 100 mL), dried with anhydrous MgSO$_4$, filtered and evaporated to a low volume. The foregoing concentrated solution was then purified by a silica gel chromatography column (13 cmx4.5 cm) prepared with DCM. The eluent was changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated at room temperature overnight. Yield: 1.07 g (81%). TLC (fraction 4): R$_f$ 0.3 (DCM: MeOH-9:1). The purity of the compound was 98% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient H, 23 min; 247 nm; Rt. 11.38'). The molecular mass of the product (**21**) was 467.24, ESI-HR (+) m/z: 468.2488 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 490.2 (20%) (M+Na)$^+$, 935.4 (25%) (2M+H)$^+$.

^{1}H NMR (400MHz, d$_6$-DMSO), δ ppm:0.80-0.85 (6H, m, 2x[δ-C$\underline{H}_3$]), 1.14-1.21 and 1.40-1.33 (2H, m, β-C$\underline{H}_2$-Leu side chain), 1.63-1.70 (1H, m, γ-C$\underline{H}$-Leu side chain), 3.11 (1H, q, J=5.6Hz, α-C$\underline{H}$-Leu side chain), 3.18-3.30 (2H, m, C$\underline{H}_2$-NH-COO), 3.46 (2H, t, J=5.8Hz, C$\underline{H}_2$-CH$_2$-NH-COO), 3.53 (2H, q, J=5.2Hz, C$\underline{H}_2$-NH-AQ), 3.57-3.64 (4H, m, O-C$\underline{H}_2$-C$\underline{H}_2$-O), 3.73 (2H, t, J=5.4Hz, C$\underline{H}_2$-CH$_2$-NH-AQ), 7.29 (1H, dd, J$_1$=8.8Hz, J$_2$=0.8Hz, H-2), 7.45 (1H, dd, J$_1$=7.2Hz, J$_2$=0.8Hz, H-4), 7.63-7.67 (1H, m, H-3), 7.82-7.94 (3H, m, H-6, H-7, N$\underline{H}$-COO), 8.12 (1H, dd, J$_1$=7.6Hz, J$_2$=1.2Hz, H-5), 8.20 (1H, dd, J$_1$=7.6Hz, J$_2$=1.2Hz), 9.79 (1H, t, J=5.2Hz, N$\underline{H}$-AQ).

II. Synthesis of intermediate compound H$_2$N-Asn(Trt)-Leu-Spacer-AQ (22) (Method: F,G)

Compound **21** (1.00 g, 2.14 mmol) was dissolved in DMF (15 mL). To this solution Fmoc-Asn(trt)-OH (1.1 eq, 1.40 g, 2.35 mmol), TBTU (1.1 eq, 0.80 g, 2.35 mmol) and HOBt (1.1 eq, 0.32 g, 2.35 mol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 1.18 g/mL, 6.85 mmol) was added following the same procedure as described for the synthesis of compound **21**. The crude product was then purified by a silica gel chromatography column (16 cmx4.5 cm) prepared with DCM. The eluent was changed to DCM: MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 1.65 g (94%). TLC (fraction 4): R$_f$0.5 (DCM:MeOH-9:1). The purity of the compound was 91% on an analytical RP-HPLC system following the general procedure as mentioned in

4.2.5. (Gradient H, 23 min; 247 nm; Rt. 20.01'). The molecular mass of the product (**22**) was 823.39, ESI-HR (+) m/z: 824.4013 (100%) (M+H)$^+$. ESMS-LR (-) m/z: 822.2 (80%) (M-H)$^-$.

^{1}H NMR (400MHz, d$_6$-DMSO), δ ppm: 0.80-0.85 (6H, m, 2x[δ-C$\underline{H}_3$]), 1.37-1.48 (2H, m, β-C$\underline{H}_2$-Leu side chain), 1.51-1.61 (1H, m, γ-C$\underline{H}$-Leu side chain), 2.38-2.46 (2H, m, C$\underline{H}_2$-COO-NH-Trityl protecting group), 3.07-3.18 (2H, m, C$\underline{H}_2$-NH-Leu), 3.35-3.37 (2H, C$\underline{H}_2$-CH$_2$-NH-Leu), 3.45-3.53 (5H, m, C$\underline{H}_2$-O-CH$_2$-CH$_2$-NH-AQ, C$\underline{H}$-CH$_2$-Trityl group, CH$_2$-NH-AQ), 3.58-3.61 (2H, m, C$\underline{H}_2$-O-CH$_2$-CH$_2$-NH-Leu), 3.71 (2H, t, J=5.2Hz, C$\underline{H}_2$-CH$_2$-NH-AQ), 4.26 (1H, q, J=8.4Hz, C$\underline{H}$-CH$_2$-CH-2[CH$_3$]), 7.17-7.30 (16H, m, 2-H and Trityl protecting group), 7.45 (1H, dd, J_1=7.2Hz, J_2=0.8Hz, H-4), 7.62-7.66 (1H, m, H-3), 7.83 (1H, td, J_1=7.2Hz, J_2=1.2Hz, H-6), 7.89 (1H, td, J_1=7.6Hz, J_2=1.6Hz, H-7), 7.95-8.02 (2H, m, CH$_2$-N$\underline{H}$-COO, CH-N$\underline{H}$-COO), 8.13 (1H, dd, J_1=7.6Hz, J_2=1.2Hz, H-5), 8.20 (1H, dd, J_1=7.6Hz, J_2=1.2Hz), 9.23 (1H, s, N$\underline{H}$-Trityl protecting group), 9.80 (1H, t, J=5.2Hz, N$\underline{H}$-AQ).

III. Synthesis of H$_2$N-Gly-Asn(Trt)-Leu-Spacer-AQ (23) (Method: F,G)

Compound **22** (1.65 g, 2.00 mmol) was dissolved in DMF (15 mL). To this solution Fmoc-Gly-OH (1.1 eq, 0.66 g, 2.20 mmol), TBTU (1.1 eq, 0.71 g, 2.20 mmol) and HOBt (1.1 eq, 0.30 g, 2.20 mmol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 1.12 g/ml, 6.41 mmol) was added following the same procedure as described for the synthesis of compound **21**. The crude product was then purified by silica gel chromatography column (18 cmx4.5 cm) prepared with DCM. The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 1.44g (82%). TLC (fraction 4): R$_f$ 0.3 (DCM: MeOH- 9:1). The purity of the compound was 88% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient I, 23 min; 247 nm; Rt. 19.82'). The molecular mass of the product (**23**) was 880.42, ESI-HR (+) m/z: 881.4232 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 903.2 (100%) (M+Na)$^+$. ESMS-LR (-) m/z: 879.2 (100%) (M-H)$^-$, 915.2 (M+Cl)$^-$, 925.0 (25%) (M+TFA-H)$^-$.

^{1}H NMR (400MHz, d$_6$-DMSO), δ ppm: 0.82 (6H, dd, J=18.8Hz, 6.0Hz, 2x[δ-C$\underline{H}_3$]), 1.38-1.60 (3H, m, C$\underline{H}_2$-C$\underline{H}$[CH$_3$]$_2$), 2.65-2.76 (2H, m, C$\underline{H}_2$-CONH-Trityl protecting group), 2.97-3.10 (2H, m, C$\underline{H}_2$-NH-Leu), 3.13 (2H, s, C$\underline{H}_2$-NH$_2$), 3.26 (2H, t, J=6.2Hz, C$\underline{H}_2$-CH$_2$-NH-Leu), 3.47-3.59 (6H, m, O-C$\underline{H}_2$-C$\underline{H}_2$-O, C$\underline{H}_2$-NH-AQ), 3.71 (2H, t, J=5.2Hz, C$\underline{H}_2$-CH$_2$-NH-AQ), 4.17-4.23 (1H, m, C$\underline{H}$-CH$_2$-CH-2[CH$_3$]), 4.56 (1H, s, C$\underline{H}$-NH-COO), 7.18-7.29 (16H, m, 2-H and Trityl protecting group), 7.45 (1H, dd, J_1=7.2Hz, J_2=0.8Hz, H-4), 7.62-7.66 (1H, m, H-3), 7.81-8.08 (4H, m, H-

6, H-7, N<u>H</u>-Spacer arm, N<u>H</u>-COO-Asn), 8.13 (1H, dd, J_1=7.6Hz, J_2=1.2Hz, H-5), 8.21 (1H, dd, J_1=7.6Hz, J_2=1.2Hz), 9.22 (1H, s, N<u>H</u>-Trityl protecting group), 9.79 (1H, t, J=5.2Hz, N<u>H</u>-AQ).

IV. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn(Trt)-Leu-Spacer-AQ (24) (Method: F)

Compound **23** (1.44 g, 1.64 mmol) was dissolved in DMF (15 mL). To this solution compound **4** (1.1 eq, 0.80 g, 1.80 mmol), TBTU (1.1 eq, 0.58 g, 1.80 mmol) and HOBt (1.1 eq, 0.24 g, 1.80 mol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 0.90 g/ml, 5.22 mmol) was added following the same procedure as described for the synthesis of compound **21**. The crude product was then purified by a silica gel chromatography column (16 cmx4.5 cm) prepared with DCM:MeOH (9.5:0.5). The column was initially eluted with the same solvent system. The eluent was later changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 1.06 g (50%). TLC (fraction 5): R_f 0.4 (DCM: MeOH- 9:1). The molecular mass of the product (**24**) was 1309.50, ESI-HR(+) m/z: 1310.5081 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 1333.5 (85%) (M+H+Na)$^+$. ESMS-LR (-) m/z: 653.3 [(M-2H)/2]$^{2-}$, 1308.2 (65%) (M-H)$^-$.

V. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Leu-Spacer-AQ (25)

Compound **24** (0.05 g, 0.04 mmol) was dissolved in TFA at rt for 2 hours. TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (100 mL), and the precipitate was used without any further characterization. Crude yield: 0.03g (75%) TLC: R_f 0.1 (DCM:MeOH-9.5:0.5). The molecular mass of the product (**25**) was 1067.39, ESMS-LR (+) m/z: 534.8 (60%) [(M+2H)/2]$^{2+}$, 545.7 (40%) [(M+Na)/2]$^{2+}$, 1068.2 (45%) (M+H)$^+$.

3.4.4. Synthesis of 2nd batch-FRET tetra peptide substrates applying solution phase peptide synthesis (JC series-lysine)

3.4.4.1. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Lys-Spacer-AQ (31)
I. Synthesis of intermediate compound H$_2$N-Lys(Boc)-Spacer-AQ (26) (Method: F,G)

Compound **5** (1.00 g, 2.82 mmol) was dissolved in DMF (10 mL). To this solution Fmoc-Lys(Boc)-OH (1.1 eq, 1.46 g, 3.11 mmol), TBTU (1.1 eq, 1.00 g, 3.11 mmol) and HOBt (1.1 eq, 0.42 g, 3.11 mol) dissolved in DMF (10 mL) with the addition

of DIPEA (3.2 eq, 1.56 g/mL, 9.04 mmol) was added and stirred overnight at rt. The solution was then partitioned between DCM and water. The organic extract were washed with saturated sodium hydrogen carbonate (3 x 100 mL), then water (3 x 100 mL), dried with anhydrous $MgSO_4$, filtered and evaporated to dryness and was used for the following reaction without further purification.

The extracted sample was dissolved in 20% v/v of piperidine in DMF (10-15 mL) and the amino terminus was deprotected as in method G. The solution was partitioned between DCM and water. The organic extracts were washed with saturated sodium hydrogen carbonate (3 x 100 mL), then water (3 x 100 mL), dried with anhydrous $MgSO_4$, filtered and evaporated to low volume. The crude product was then purified by silica gel chromatography column (12 cmx4.5 cm) prepared with DCM:MeOH (9.5:0.5). The eluent was then changed to DCM: MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 1.47 g (90%). TLC (fraction 3): R_f 0.3 (DCM:MeOH-9:1). The molecular mass of the product (**26**) was 582.31, ESMS-LR (+) m/z: 583 (80%) $(M+H)^+$, 604.8 (95%) $(M+Na)^+$. ESMS-LR (-) m/z: 581 (90%) $(M-H)^-$, 616 (20%) $(M+Cl)^-$.

^1H NMR (400MHz, d_6-DMSO), δ ppm: 0.82 (6H, dd, J_1=18.8Hz, J_2=6.0Hz, 2x[δ-C$\underline{H}_3$]),1.20-1.52 (15H, m, Boc protecting group; β, γ and δ-C$\underline{H}_2$ of the Lys side chain), 2.86 (2H, q, J=6.4Hz, ε-C$\underline{H}_2$ of the Lys side chain), 3.06 (1H, q, J=5.6Hz, α-C$\underline{H}$ of Lysine), 3.21-3.26 (2H, m, C$\underline{H}_2$-NH-COO), 3.46 (2H, t, J=6Hz, C$\underline{H}_2$-CH$_2$-NH-COO), 3.52-3.64 (6H, m, C$\underline{H}_2$-NH-AQ, O-C$\underline{H}_2$-C$\underline{H}_2$-), 3.73 (2H, t, J=5.6Hz, C$\underline{H}_2$-CH$_2$-NH-AQ), 6.75 (1H, s, N$\underline{H}$-COO-Boc group), 7.30 (1H, d, J=8.8Hz, H-2), 7.46 (1H, dd, J_1=7.6Hz, J_2=0.8Hz, H-4), 7.66 (1H, t, J=7.6Hz, H-3), 7.82-7.92 (3H, m, H-6, H-7, N$\underline{H}$-COO-CH), 8.13 (1H, dd, J_1=7.2Hz, J_2=0.8Hz, H-5), 8.21 (1H, dd, J_1=7.6Hz, J_2=0.8Hz), 9.80 (1H, t, J=5.2Hz, N$\underline{H}$-AQ).

II. Synthesis of intermediate compound H_2N-Asn(Trt)-Lys(Boc)–Spacer–AQ (27)

Compound **26** (1.00 g, 1.72 mmol) was dissolved in DMF (10-15 mL). To this solution of Fmoc-Asn(Trt)-OH (1.1 eq, 1.13 g, 1.89 mmol), TBTU (1.1 eq, 0.61 g, 1.89 mmol) and HOBt (1.1 eq, 0.26 g, 1.89 mol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 0.95 g/mL, 5.50 mmol) was added following the same procedure as described for the synthesis of compound **26**.The crude product was then purified by silica gel chromatography column (13.5 cmx4.5 cm) prepared with DCM:MeOH (9.5:0.5). The product was then eluted with further addition of MeOH (7%) to the eluent. Fractions containing the product were combined, filtered

and evaporated to dryness. Yield: 1.22 g (76%). TLC (fraction 3): R_f 0.2 (DCM:MeOH-9:1). The molecular mass of the product (**27**) was 938.46 ESMS-LR (+) m/z: 938.8 (100%) (M)$^+$, 939.8 (60%) (M+H)$^+$. ESMS-LR (-) m/z: 937.4 (25%) (M-H)$^-$, 1051.6 (20%) (M+TFA-H)$^-$.

III. Synthesis of intermediate compound H₂N-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (28)

Compound **27** (1.00 g, 1.07 mmol) was dissolved in DMF (10-15 mL). To this solution Fmoc-Gly-OH (1.1 eq, 0.35 g, 1.17 mmol), TBTU (1.1 eq, 0.38 g, 1.17 mmol) and HOBt (1.1 eq, 0.16 g, 1.17 mmol) dissolved in DMF (5-10 mL) with the addition of DIPEA (3.2 eq, 0.59 g/mL, 3.41 mmol) was added following the same procedure as described for the synthesis of compound **26**.The crude product was then purified by silica gel chromatography column (14.5 cm x 4.5 cm) prepared with DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 0.71 g (67%). TLC (fraction 3): R_f 0.6 (DCM: MeOH-9:1). The molecular mass of the product (**28**) was 995.48, ESMS-LR (+) m/z: 995.8 (100%) (M)$^+$, 996.6 (M+H)$^+$, 1017.6 (40%) (M+Na)$^+$. ESMS-LR (-) m/z: 993.4 (100%) (M-H)$^-$, 1029.6 (M+Cl)$^-$.

IV. Synthesis of intermediate compound HN-Pro-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (29)

Compound **28** (0.39 g, 0.39 mmol) was dissolved in DMF (5 mL). To this solution Fmoc-Pro-OH (1.1 eq, 0.15 g, 0.43 mmol), TBTU (1.1 eq, 0.14 g, 0.43 mmol) and HOBt (1.1 eq, 0.06 g, 0.43 mmol) dissolved in DMF (5-7 mL) with the addition of DIPEA (3.2 eq, 0.22 g/ml, 1.25 mmol) was added following the same procedure as described for the synthesis of compound **26**.The crude product was then purified by silica gel chromatography column (20 cm x 2.5 cm) prepared with DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 0.37 g (86%). TLC (fraction 3): R_f 0.4 (DCM: MeOH- 9:1). The molecular mass of the product (**29**) was 1092.53, ESMS-LR (+) m/z: 1092.6 (100%) (M)$^+$, 1093.6 (65%) (M+H)$^+$, 1115.6 (M+Na)$^+$. ESMS-LR (-) m/z: 1091.6 (40%) (M-H)$^-$, 1127.6 (65%) (M+Cl)$^-$.

V. Synthesis of intermediate compound 5(6)-CBF-Pro-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (30)

Compound **29** (0.37 g, 0.34 mmol) was dissolved in DMF (3 mL). To this solution 5(6)-Carboxyfluorescein (1 eq, 0.13 g, 0.34 mmol), TBTU (1.1 eq, 0.12 g, 0.37 mmol) and HOBt (1.1 eq, 0.10 g, 0.37 mmol) dissolved in DMF (7 mL) with the

addition of DIPEA (3.2 eq, 0.19 g/mL, 1.08 mmol) was added and stirred at rt. The solution was partitioned between DCM and water. The organic extracts were washed with saturated sodium hydrogen carbonate (2 x 75 mL), then water (3 x 75 mL), dried with anhydrous MgSO$_4$, filtered and evaporated to low volume. The crude product was then purified by a silica gel chromatography column (23 cmx2.5 cm) prepared with chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 0.40 g (82%). TLC (fraction 2): R$_f$ 0.6 (Chloroform: MeOH: Pyridine: Ammonia-20:55:15:10). The molecular mass of the product (**30**) was 1450.58, ESMS-LR (+) m/z: 1451.4 (95%) (M+H)$^+$, 1473.4 (M+Na)$^+$. ESMS-LR (-) m/z: 1449.4 (M-H)$^-$, 1485.4 (M+Cl)$^-$.

V. Synthesis of 5(6)-CBF-Pro-Gly-Asn-Lys-Spacer-AQ (31)

Compound **30** (0.40 g, 0.28 mmol) was dissolved in TFA (2 mL) at rt. After 3 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (100 mL) and then stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. The purity of the compound was 93% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 275 nm, Rt. 10.89'). Crude yield: 0.29 g (97%). TLC: R$_f$ 0.1 (DCM: MeOH- 9:1). The molecular mass of the product (**31**) was 1108.42, ESI-HR (+) m/z: 1109.4252 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 555.2 (100%) [(M+2H)/2]$^{2+}$.

3.4.4.2. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Lys-Spacer-AQ (34)

I. Synthesis of intermediate compound H$_2$N-β-Ala-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (32)

Compound **28** (0.32 g, 0.32 mmoles) was dissolved in DMF (5 mL). To this solution Fmoc-β-Ala-OH (1.1 eq, 0.11 g, 0.35 mmol), TBTU (1.1 eq, 0.11 g, 0.35 mmol) and HOBt (1.1 eq, 0.05 g, 0.35 mol) dissolved in DMF (7 mL) with the addition of DIPEA (3.2 eq, 0.18 g/mL, 1.03 mmol) was added following the same procedure as described for the synthesis of compound **26**.The crude product was then purified by silica gel chromatography column (21 cmx2.5 cm) prepared with DCM:MeOH (9:1). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.30 g (88%). TLC (fraction 3): R$_f$ 0.5 (DCM:MeOH- 9:1). The molecular mass of the product (**32**) was 1066.52, ESMS-LR (+) m/z: 1066.6 (100%) (M)$^+$, 1067.6

(M+H)$^+$, 1089.6 (M+Na)$^+$. ESMS-LR (-) m/z: 1065.6 (60%) (M-H)$^-$, 1101.6 (M+Cl)$^-$.

II. Synthesis of intermediate compound 5(6)-CBF-β-Ala-Gly-Asn(Trt)-Lys(Boc)-Spacer-AQ (33)

Compound **32** (0.30 g, 0.28 mmol) was dissolved in DMF (5 mL). To this solution 5(6)-Carboxyfluorescein (1 eq, 0.11 g, 0.28 mmol), TBTU (1.1 eq, 0.10 g, 0.31 mmol) and HOBt (1.1 eq, 0.04 g, 0.31 mol) dissolved in DMF (5-10 mL) with the addition of DIPEA (3.2 eq, 0.16 g/mL, 0.90 mmol) was added following the same procedure as described for the synthesis of compound **30**. The crude product obtained was then purified a silica gel chromatography column (12 cm×4.5 cm) prepared with chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to dryness *in vacuo*. Yield: 0.31 g (78%). TLC (fraction 2): R$_f$ 0.8(Chloroform: MeOH: Pyridine: Ammonia- 20:55:15:10). The molecular mass of the product (**33**) was 1424.56, ESMS-LR (+) m/z: 713.0 (20%) [(M+2H)/2]$^{2+}$, 1424.4 (100%) (M)$^+$, 1425.4 (95%) (M+H)$^+$. ESMS-LR (-) m/z: 1423.4 (100%) (M-H)$^-$, 1424.2 (60%) (M)$^-$, 1459.4 (60%) (M+Cl)$^-$.

III. Synthesis of 5(6)-CBF-β-Ala-Gly-Asn-Lys-Spacer-AQ (34)

Compound **33** (0.31 g, 0.22 mmol) was dissolved in TFA (3 mL) at rt. After 3 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (50 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. The purity of the compound was 77% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 275 nm, Rt. 10.82'). Crude yield: 0.20 g (83%). TLC: R$_f$ 0.5(DCM: MeOH- 9:1). The molecular mass of the product (**34**) was 1082.40, ESI-HR (+) m/z: 1083.4097 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 542.2 (100%) [(M+2H)/2]$^{2+}$.

3.4.4.3. Synthesis of 5(6)-CBF-Pro-Ala-Asn-Lys-Spacer-AQ (38)

I. Synthesis of intermediate compound H$_2$N-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ (35)

Compound **27** (1.00 g, 1.07 mmol) was dissolved in DMF (10 mL). To this solution Fmoc-Ala-OH (1.1 eq, 0.36 g, 1.17 mmol), TBTU (1.1 eq, 0.38 g, 1.17 mmol) and HOBt (1.1 eq, 0.16 g, 1.17 mol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 0.59 g/mL, 3.41 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then

purified by silica gel chromatography column (15 cmx4.5 cm) prepared with DCM:MeOH (9:1). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to dryness *in vacuo*. Yield: 0.84 g (78%). TLC (fraction 3): R_f 0.3 (DCM: Methanol- 9:1). The molecular mass of the product (**35**) was 1009.49, ESMS-LR (+) m/z: 1009.8 (100%) (M)$^+$, 1010.8 (60%) (M+H)$^+$. ESMS-LR (-) m/z: 1008.6 (65%) (M-H)$^-$, 1044.6 (40%) (M-Cl)$^-$.

II. Synthesis of intermediate compound HN-Pro-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ (36)

Compound **35** (0.42 g, 0.42 mmol) was dissolved in DMF (7 mL). To this solution Fmoc-Pro-OH (1.1 eq, 0.15 g, 0.46 mmol), TBTU (1.1 eq, 0.15 g, 0.46 mmol) and HOBt (1.1 eq, 0.06 g, 0.46 mmol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 0.23 g/mL, 1.33 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then purified by silica gel chromatography column (20 cmx2.5 cm) prepared with DCM: MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of MeOH (5%) to the eluent. Fractions containing the product were combined, filtered and evaporated to dryness *in vacuo*. Yield: 0.37 g (80%). TLC (fraction 3): R_f 0.8 (DCM: MeOH- 9:1). The molecular mass of the product (**36**) was 1106.55, ESMS-LR (+) m/z: 554.0 (15%) [(M+2H)/2]$^{2+}$, 1106.6 (100%) (M)$^+$, 1107.6 (60%) (M+H)$^+$, 1128.6 (15%) (M+Na)$^+$. ESMS-LR (-) m/z: 1105.6 (35%) (M-H)$^-$, 1141.6 (65%) (M+Cl)$^-$

III. Synthesis of intermediate compound 5(6)-CBF-Pro-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ (37)

Compound **36** (0.37 g, 0.33 mmol) was dissolved in DMF (5 mL). To this solution 5(6),carboxyfluorescein (1 eq, 0.13 g, 0.33 mmol), TBTU (1.1 eq, 0.12 g, 0.37 mmol) and HOBt (1.1 eq, 0.05 g, 0.37 mmol) dissolved in DMF (7 mL) with the addition of DIPEA (3.2 eq, 0.18 g/mL, 1.07 mmol) was added following the same procedure as described for the synthesis of compound **30**. The crude product obtained was then purified by a silica gel chromatography column (17.5 cmx2.5 cm) prepared with chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated *in vacuo*. Yield: 0.34 g (69%). TLC (fraction 2): R_f 0.8(Chloroform:MeOH:Pyridine:Ammonia-20:55:15:10). The molecular mass of the product (**37**) was 1464.60, ESMS-LR (+) m/z: 733.0 (15%)

[(M+2H)/2]$^{2+}$, 1464.4 (50%) (M)$^+$, 1486.2 (M+Na)$^+$. ESMS-LR (-) m/z: 1463.2 (95%) (M-H)$^-$, 1498.2 (60%) (M+Cl)$^-$.

IV. Synthesis of 5(6)-CBF-Pro-Ala-Asn-Lys-Spacer-AQ (38)

Compound **37** (0.34 g, 0.23 mmol) was dissolved in TFA (2 mL) at room temperature. After 3 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (50 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.24 g (92%). TLC: R$_f$ 0.5 (DCM: Methanol-9:1). The purity of the compound was 82% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 275 nm, Rt. 10.92'). The molecular mass of the product (**38**) was found to be 1122.43, ESI-HR (+) m/z: 1123.4411 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 562.2 (100%) [(M+2H)/2]$^{2+}$.

3.4.4.4. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Lys-Spacer-AQ (41)

I. Synthesis of intermediate compound H$_2$N-β-Ala-Ala-Asn-Lys-Spacer-AQ (39)

Compound **35** (0.42 g, 0.42 mmol) was dissolved in DMF (5 mL). To this solution Fmoc-β-Ala-OH (1.1 eq, 0.14 g, 0.46 mmol), TBTU (1.1 eq, 0.15 g, 0.46 mmol) and HOBt (1.1 eq, 0.06 g, 0.46 mmol) dissolved in DMF (5-7 mL) with the addition of DIPEA (3.2 eq, 0.23 g/ml, 1.33 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then purified by silica gel chromatography column (20 cmx2.5 cm) prepared with DCM: MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of MeOH (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to dryness *in vacuo*. Yield: 0.42 g (93%). TLC (fraction 3): R$_f$ 0.5 (DCM:MeOH-9:1). The molecular mass of the product (**39**) was 1080.53, ESMS-LR (+) m/z: 1080.6 (100%) (M)$^+$, 1081.6 (75%) (M+H)$^+$, 1103.6 (15%) (M+Na)$^+$. ESMS-LR (-) m/z: 1079.6 (60%) (M-H)$^-$, 1115.6 (70%) (M+Cl)$^-$.

II. Synthesis of intermediate compound 5(6)-CBF-β-Ala-Ala-Asn(Trt)-Lys(Boc)-Spacer-AQ (40)

Compound **39** (0.42 g, 0.39 mmol) was dissolved in DMF (5 mL). To this solution 5(6)-Carboxyfluorescein (1 eq, 0.15 g, 0.39 mmol), TBTU (1.1 eq, 0.14 g, 0.43 mmol) and HOBt (1.1 eq, 0.06 g, 0.43 mol) dissolved in DMF (5-7 mL) with the addition of DIPEA (3.2 eq, 0.21 g/ml, 1.24 mmol) was added following the same

procedure as described for the synthesis of compound **30**. The crude product was then purified by a silica gel chromatography column (17.5 cmx4.5 cm) prepared with chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume *in vacuo*. Yield: 0.30 g (54%). TLC (fraction 2): R_f 0.8 (Chloroform: MeOH: Pyridine: Ammonia- 20:55:15:10). The molecular mass of the product (**40**) was 1438.58, ESMS-LR (+) m/z: 719.8 (40%) $[(M+2H)/2]^{2+}$, 1438.4 (100%) $(M)^+$, 1439.4 (95%) $(M+H)^+$. ESMS-LR (-) m/z: 1437.2 (100%) $(M-H)^-$, 1473.2 $(M+Cl)^-$.

III. Synthesis of 5(6)-CBF-β-Ala-Ala-Asn-Lys-Spacer-AQ (41)

Compound **40** (0.30 g, 0.21 mmol) was dissolved in TFA (2 mL) at room temperature. After 3 hours TFA was evaporated and the resulting concentrate was re-evaporated with ethanol (x3). The residue was then triturated with diethyl ether (50 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.21 g (91%). TLC: R_f 0.5 (DCM:MeOH-4:1). The purity of the compound was 84% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 275 nm, Rt. 10.82'). The molecular mass of the product (**41**) was 1096.42, ESI-HR (+) m/z: 1097.4250 (100%) $(M+H)^+$. ESMS-LR (+) m/z: 549.2 (100%) $[(M+2H)/2]^{2+}$.

3.4.4.5. Synthesis of 5(6)-CBF-Ala-Pro-Ala-Asn-Leu-Spacer-AQ (46)

I. Synthesis of intermediate compound H₂N-Ala-Asn(Trt)-Leu-Spacer-AQ (42)

Compound **22** (2.30 g, 2.79 mmol) was dissolved in DMF (10-15 mL). To this solution Fmoc-Ala-OH (1.1 eq, 0.96 g, 3.21 mmol), TBTU (1.1 eq, 0.99 g, 3.07 mmol), HOBt (1.1 eq, 0.41 g, 3.07 mmol) dissolved in DMF (10 mL) followed by the addition of DIPEA (3.2 eq, 1.54 g/mL, 8.94 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then purified by silica gel chromatography column (15.5 cmx4.5 cm) prepared with DCM:MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of methanol (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated *in vacuo*. Yield: 1.40 g (56%). TLC (fraction 3): R_f 0.4 (DCM: MeOH-9:1). The molecular mass of the product (**42**) was 894.43, ESMS-LR (+) m/z: 894.8 (100%) $(M)^+$, 895.8 (60%) $(M+H)^+$, 916.6 $(M+Na)^+$. ESMS-LR (-) 893.0 (50%) $(M-H)^-$, 928.2 (100%) $(M+Cl)^-$, 1006.6 (40%) $(M+TFA-H)^-$.

II. Synthesis of intermediate compound HN-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ (43)

Compound **42** (1.40 g, 1.57 mmol) was dissolved in DMF (15 mL). To this solution Fmoc-Pro-OH (1.1 eq, 0.58 g, 1.72 mmol), TBTU (1.1 eq, 0.55 g, 1.72 mmoles), HOBt (1.1 eq, 0.23 g, 1.72 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (3.2 eq, 0.86 g/mL, 5.01 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then purified by silica gel chromatography column (14 cmx4.5 cm) prepared with DCM:MeOH (9.5:0.5). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated *in vacuo*. Yield: 0.95 g (61%). TLC (fraction 3): R_f 0.5 (DCM:MeOH-9:1). The molecular mass of the product (**43**) was 991.48, ESMS-LR (+) m/z: 496.4 (15%) $[(M+2H)/2]^{2+}$, 991.8 (100%) $(M)^+$, 992.8 $(M+H)^+$, 1014.6 (20%) $(M+Na)^+$.

III. Synthesis of intermediate compound H₂N-Ala-Pro-Asn(Trt)-Leu-Spacer-AQ (44)

Compound **43** (0.40 g, 0.40 mmol) was dissolved in DMF (7 mL). To this solution Fmoc-Ala-OH (1.1 eq, 0.14 g, 0.44 mmol), TBTU (1.1 eq, 0.14 g, 0.44 mmol), HOBt (1.1 eq, 0.06 g, 0.44 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (3.2 eq, 0.22 g/mL, 1.29 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then purified by silica gel chromatography column (20 cmx2.5 cm) prepared with DCM: MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of MeOH (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to dryness *in vacuo*. Yield: 0.14 g (33%). TLC (fraction 4): R_f 0.6 (DCM: MeOH-9:1). The molecular mass of the product (**44**) was 1062.52, ESMS-LR (+) m/z: 532.0 (40%) $[(M+2H)/2]^{2+}$, 543.0 (20%) $[(M+2H+Na)/2]^{2+}$, 551.0 (15%) $[(M+2H+K)/2]^{2+}$, 1062.8 (100%) $(M)^+$, 1063.8 (70%) $(M+H)^+$, 1084.8 (60%) $(M+Na)^+$.

IV. Synthesis of intermediate compound 5(6)-CBF-Ala-Pro-Asn(Trt)-Leu-Spacer-AQ (45)

Compound **44** (0.14 g, 0.13 mmol) was dissolved in DMF (7 mL). To this solution 5(6)-Carboxyfluorescein (1 eq, 0.05 g, 0.13 mmol), TBTU (1.1 eq, 0.05 g, 0.14 mmol), HOBt (1.1 eq, 0.02 g, 0.14 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (3.2 eq, 0.07 g/ml, 0.42 mmol) was added following the

same procedure as described for the synthesis of compound **30**. The crude product was then purified by silica gel chromatography column (17 cmx2.5 cm) prepared with DCM:MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of methanol (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to a low volume *in vacuo*. Yield: 0.10 g (53%). TLC (fraction 2): R_f 0.7 (DCM: MeOH-9:1). The molecular mass of the product (**45**) was 1420.57, ESMS-LR (+) m/z: 711.0 (15%) $[(M+2H)/2]^{2+}$, 1420.4 (45%) $(M)^+$, 1421.4 (40%) $(M+H)^+$, 1443.2 (95%) $(M+Na)^+$. ESMS-LR (-) m/z: 1419.8 (80%) $(M-H)^-$.

V. Synthesis of 5(6)-CBF-Ala-Pro-Asn-Leu-Spacer-AQ (46)

Compound **45** (0.10 g, 0.07 mmol) was dissolved in TFA (2 mL) at room temperature. After 3 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (50 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.06 g (75%). TLC: R_f 0.5 (DCM:MeOH-9:1). The purity of the compound was 64% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 275 nm, Rt. 17.16'). The molecular mass of the product (**46**) was 1178.46, ESI-HR (+) m/z: 1179.4674 (100%) $(M+H)^+$. ESMS-LR (+) m/z: 590.2 (100%) $[(M+2H)/2]^{2+}$. ESMS-LR (-) m/z: 1177.6 (60%) $(M-H)^-$, 1213.6 (20%) $(M+Cl)^-$.

3.5. Synthesis of Gadolinium-DOTA-Peptide substrates [Leucine-JS Linear sequences]

A. Synthesis of Gd(III)DOTA-Pro-Ala-Asn-Leu-Spacer-AQ (49)
I. Synthesis of intermediate compound DOTA(*t*-Bu)-Pro-Ala-Asn(Trt)-Leu-Spacer-AQ (47)

Compound **43** (0.30 g, 0.30 mmol) was dissolved in DMF (5 mL). To this solution Tri-*tert*-butyl1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetate (DOTA-tri-*t*-Bu-ester) (1.1 eq, 0.19 g, 0.33 mmol), TBTU (1.1 eq, 0.11 g, 0.33 mmol), HOBt (1.1 eq, 0.04 g, 0.33 mmol) dissolved in DMF (7 mL) followed by the addition of DIPEA (3.2 eq, 0.17 g/ml, 0.97 mmol) was added and stirred at rt overnight. The solution was then partitioned between DCM and water. The organic extract was washed with saturated sodium hydrogen carbonate (1x75 mL), then water (3x 75 mL), dried with anhydrous MgSO₄, filtered and evaporated to low volume. The crude product was then purified by a silica gel chromatography column (20 cmx2.5 cm) prepared with DCM:MeOH (9.5:0.5) and the major product was eluted with

the further addition of methanol (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to a low volume *in vacuo*. Yield: 0.20 g (43%). TLC (fraction 2): R_f 0.8 (DCM:MeOH-9:1). The molecular mass of the product (**47**) was 1545.85, ESI-HR (+) m/z: 1568.8402 (100%) (M+Na)⁺. ESMS-LR (+) m/z: 784.9 (100%) $[(M+Na)/2]^{2+}$.

II. Synthesis of intermediate compound DOTA-Pro-Ala-Asn-Leu-Spacer-AQ (48)

Compound **47** (0.20 g, 0.13 mmol) was dissolved in TFA (2 mL) at rt. After 4 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (75 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.13 g (87%). TLC: R_f 0.2 (DCM:MeOH-9:1). The molecular mass of the product (**48**) was 1135.55, ESI-HR (+) m/z: 568.7844 (100%) $[(M+2H)/2]^{2+}$. ESMS-LR (+) m/z: 587.7(100%) $[(M+2H+K)/2]^{2+}$.

III. Synthesis Gd(III)DOTA-Pro-Ala-Asn-Leu-Spacer-AQ (49)

Compound **48** (0.10 g, 0.09 mmol) was dissolved in water (3 mL). To this solution Gadolinium(III)chloridehexahydrate (2.8 eq, 0.09 g, 0.25 mmol) dissolved in water (2 mL) was added. The pH of the reaction mixture was adjusted to 7 by adding a solution of ammonia 0.35% drop wise. Reaction progress was monitored by TLC analysis. The solution was evaporated to low volume. The crude product was then purified by a silica gel chromatography column (10 cmx2.5 cm) prepared with chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume *in vacuo*. Yield: 0.10 g (91%). TLC (fraction 3): R_f 0.2 (chloroform:MeOH:pyridine: ammonia-20:55:15:10). The purity of the compound was 72% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient B, 23 min, 245 nm, Rt. 14.77'). The molecular mass of the product (**49**) was 1290.45, ESI-HR (+) m/z: 646.2351 (100%) $[(M+2H)/2]^{2+}$. ESMS-LR (+) m/z: 657.2 (20%) $[(M+2H+Na)/2]^{2+}$, 1291.4 (90%) (M+H)⁺, 1313.4 (100%) (M+Na)⁺.

B. Synthesis of Gd(III)DOTA-Pro-Gly-Asn-Leu-Spacer-AQ (53)
I. Synthesis of intermediate compound HN-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ (50)

Compound **23** (0.54 g, 0.61 mmol) was dissolved in DMF (10 mL). To this solution Fmoc-Pro-OH (1.1 eq, 23.00 g, 0.67 mmoles), TBTU (1.1 eq, 0.22 g, 0.67 mmol),

HOBt (1.1 eq, 0.09 g, 0.67 mmol) dissolved in DMF (10 mL) followed by the addition of DIPEA (3.2 eq, 0.34 g/mL, 1.96 mmol) was added following the same procedure as described for the synthesis of compound **26**. The crude product was then purified by silica gel chromatography column (12 cmx4.5 cm) prepared with DCM: MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of MeOH (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to a low volume *in vacuo*. Yield: 0.51 g (85%). TLC (fraction 3): R_f 0.2 (DCM:MeOH-9:1). The molecular mass of the product (**50**) was 977.47, ESMS-LR (+) m/z: 489.8 (15%) $[(M+2H)/2]^{2+}$, 978.4 (100%) $(M+H)^+$, 1000.2 (25%) $(M+Na)^+$. ESMS-LR (-) m/z: 976.2 (100%) $(M-H)^-$, 1012.2 (15%) $(M+Cl)^-$.

II. Synthesis of intermediate compound DOTA(*t*-Bu)-Pro-Gly-Asn(Trt)-Leu-Spacer-AQ (51)

Compound **50** (0.20 g, 0.20 mmol) was dissolved in DMF (5 mL). To this solution DOTA-tri-*t*-Bu-ester (1.1 eq, 0.13 g, 0.23 mmol), TBTU (1.1 eq, 0.07 g, 0.23 mmol), HOBt (1.1 eq, 0.03 g, 0.23 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (3.2 eq, 0.11 g/mL, 0.65 mmol) was added following the same procedure as described for the synthesis of compound **47**. The crude product obtained was then purified by a silica gel chromatography column (16.5 cmx2.5 cm) prepared with DCM:MeOH (9.5:0.5) and the major product was eluted with the further addition of MeOH (5%-15% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.16 g (52%). TLC (fraction 2): R_f 0.6 (DCM:MeOH-9:1). The molecular mass of the product (**51**) was 1531.84, ESI-HR (+) m/z: 788.9055 (100%) $[(M+2H+Na)/2]^{2+}$.

III. Synthesis of intermediate compound DOTA-Pro-Gly-Asn-Leu-Spacer-AQ (52)

Compound **51** (0.16 g, 0.10 mmol) was dissolved in TFA (2 mL) at rt. After 7 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (50 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.10 g (83%). TLC: R_f 0.3 (DCM:MeOH-9:1). The molecular mass of the product (**52**) was 1121.54, ESI-HR (+) m/z: 583.7 (100%) $[(M+2Na)/2]^{2+}$.

IV. Synthesis of Gd(III)DOTA-Pro-Gly-Asn-Leu-Spacer-AQ (53)

Compound **52** (0.10 g, 0.09 mmol) was dissolved in water (3 mL). To this solution Gadolinium(III)chloridehexahydrate (2.8 eq, 0.09 g, 0.25 mmol) dissolved in water (2 mL) was added following the same procedure as described for the synthesis of compound **49**. The crude product was then purified by a silica gel chromatography column (9cmx2.5cm) prepared with chloroform:MeOH: pyridine: 35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.02 g (18%). TLC (4): R_f 0.3 (chloroform: MeOH:pyridine:ammonia-20:55:15:10). The purity of the compound was 77% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient B, 23 min, 245 nm, Rt. 15.43'). The molecular mass of the product (**53**) was 1276.44, ESI-HR (+) m/z: 661.2134 (100%) $[(M+2Na)/2]^{2+}$. ESMS-LR (+) m/z: 659.7 (100%) $(M/2+ACN+H)^{2+}$, 1299.4 (100%) $(M+Na)^+$.

3.6. Synthesis of Gadolinium-DOTA-Peptide substrates- JS Lysine sequence

3.6.1. Synthesis of Gadolinium complex using 1,4,7,10-Tetraazacyclodo decane-1,4,7,10-tetraacetic acid (58)

I. Synthesis of Fmoc-Lys(Boc)-Spacer-AQ (54)

Compound **5** (1.00 g, 2.82 mmol) was dissolved in DMF (10 mL). To this solution Fmoc-Lys(Boc)-OH (1.1 eq, 1.46 g, 3.11 mmol), TBTU (1.1 eq, 1.00 g, 3.11 mmol) and HOBt (1.1 eq, 0.42 g, 3.11 mol) dissolved in DMF (10 mL) with the addition of DIPEA (3.2 eq, 1.56 g/mL, 9.04 mmol) was added and stirred at rt. Reaction progress was monitored by TLC analysis. The solution was partitioned between DCM and water. The organic extract was washed with saturated sodium hydrogen carbonate (3 x 100 mL), then water (3 x 100 mL), dried with anhydrous $MgSO_4$, filtered and evaporated to a low volume using a rotary evaporator. The crude product was purified by a silica gel chromatography column (17 cmx4.5 cm) with 5% methanol in dichloromethane. The fraction containing the required product was filtered and evaporated to dryness. Yield: 1.85 g (81%). TLC (fraction 2): R_f 0.8 (DCM:MeOH-9:1). The molecular mass of the product (**54**) was 804.37, ESI-HR (+) m/z: 805.3811 (100%) $(M+H)^+$. ESMS-LR (+) m/z: 827 $(M+Na)^+$.

^{1}H NMR (400MHz, d_6-DMSO), δ ppm: 1.17-1.40 (13H, m, ^{t}Boc group, γ-CH$_2$, δ-CH$_2$ of lys side chain), 1.46-1.62 (2H, m, β-CH$_2$-Lys side chain), 2.82-2.89 (2H, m, ε-CH$_2$ of lys side chain), 3.17-3.29 (2H, m, C<u>H</u>$_2$-NH-COO), 3.45 (2H, t, J=5.8Hz,

CH$_2$-CH$_2$-NH-COO), 3.49-3.61 (6H, m, CH$_2$-NH-AQ, O-CH$_2$-CH$_2$-O), 3.70 (2H, t, J=5.2Hz, CH$_2$-CH$_2$-NH-AQ), 3.90-3.95 (1H, m, α-CH-Lys side chain), 4.17-4.29 (3H, m, CH$_2$-CH-Fmoc group), 6.76 (1H, t, J=5.6Hz, NH-COO-^{t}Boc group), 7.26-7.33 (3H, m, H-2, H-10, H-15), 7.36-7.45 (3H, m, H-4, H-11, H-14), 7.62-7.66 (1H, m, H-3), 7.70-7.73 (2H, m, H-9, H-16), 7.81-7.95 (5H, m, H-6, H-7, H12, H13, NH-Spacer arm), 8.12 (1H, dd J_1=7.6Hz, J_2=1.2Hz, H-5), 8.20 (1H, dd, J_1=8Hz, J_2=1.2Hz, H-8), 9.78 (1H, t, J=5Hz, NH-AQ).

II. Synthesis of Fmoc-Lys-Spacer-AQ trifluoroacetate salt (55)

Compound **54** (0.42 g, 0.52 mmoles) was dissolved in TFA (3 mL) at rt. Reaction progress was monitored by TLC analysis. TFA was then evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was then triturated with diethyl ether (100 mL) and stored at 4-6°C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.20g (54%). TLC: R$_f$ 0.1 (DCM: MeOH- 9:1). The purity of the compound was 95% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 16.59'). The molecular mass of the product (**55**) was 704.32, ESI-HR (+) m/z: 705.3280 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 1409.64 (100%) (M+2H)$^+$.

^{1}H NMR (400MHz, d$_6$-DMSO), δ ppm: 1.27-1.32 (2H, m, γ-CH$_2$-Lys side chain), 1.51-1.61 (4H, m, β-CH$_2$, δ-CH$_2$ of lys side chain), 2.74-2.76 (2H, m, ε-CH$_2$), 3.24 (2H, q, J=6.0Hz, CH$_2$-NH-COO), 3.46 (2H, t, J=5.8Hz, CH$_2$-CH$_2$-NH-COO), 3.51-3.61 (6H, m, CH$_2$-NH-AQ, O-CH$_2$-CH$_2$-O), 3.70 (2H, t, J=5.2Hz, CH$_2$-CH$_2$-NH-AQ), 3.92-3.98 (1H, m, α-CH-Lys side chain), 4.19 (1H, t, J=7.0Hz, CH-Fmoc group), 4.26 (2H, d, J=6.4Hz, CH$_2$-CH-Fmoc group), 7.26-7.33 (3H, m, H-2, H-10, H-15), 7.38-7.48 (3H, m, H-4, H-11, H-14), 7.62-7.73 (5H, m, H-3, H-9, H-16, NH$_2$-Lys side chain), 7.81-7.91 (4H, m, H-12, H-13, H-6, H-7), 7.94 (1H, t, J=5.4Hz, NH-Spacer arm), 8.13 (1H, dd, J_1=7.6Hz, J_2=1.2Hz, H-5), 8.19 (1H, d, J=7.6Hz, H-8), 9.77 (1H, t, J=4.4Hz, NH-AQ).

III. Synthesis of H$_2$N-Lys(DOTA^{t}Bu)-Spacer-AQ (56)

Compound **55** (0.20 g, 0.28 mmol) was dissolved in DMF (4 mL). To this solution DOTA-tri-t-Bu-ester (1.1 eq, 0.18 g, 0.31 mmol), TBTU (1.1 eq, 0.10 g, 0.31 mmol) and HOBt (1.1 eq, 0.04 g, 0.31 mol) dissolved in DMF (2 mL) with the addition of DIPEA (3.2 eq, 0.17 g/mL, 0.91 mmol) was added and stirred at rt. Reaction progress was monitored by TLC analysis. The solution was then partitioned between DCM and water. The organic extracts were washed with

saturated sodium hydrogen carbonate (3 x 75 mL), then water (3 x 75 mL) and dried with anhydrous MgSO₄, filtered and evaporated to low volume. The crude product was then purified using a silica gel chromatography column (12 cmx2.5 cm) prepared with DCM:MeOH (9.5:0.5). The eluent was then changed to DCM:MeOH (9:1). Fractions containing the product were combined, filtered and evaporated to low volume. Then followed by treating the pure sample with 20% v/v of piperidine in DMF (3 mL) and the amino terminus was deprotected as in method G. The solution was partitioned between DCM and water. The organic extracts were washed with saturated sodium hydrogen carbonate (2 x 75 mL), then water (2 x 75 mL) and evaporated to low volume. The crude product was then purified by silica gel chromatography column (15 cmx2.5 cm) prepared with DCM:MeOH (9.5:0.5). The eluant was changed to DCM:MeOH (9:1) and the major product was eluted with the further addition of MeOH (5%-10% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to low volume. Yield: 0.08 g (28%). TLC: R_f 0.2 (DCM:MeOH-9:1).

The purity of the compound was 90% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient F, 20 min, 247 nm, Rt. 8.34′). The molecular mass of the product (**56**) was 1036.62, ESI-HR (+) m/z: 1059.6071 (100%) $(M+Na)^+$. ESMS-LR (+) m/z: 519.4 (100%) $(M/2+H)^+$, 1037.4 (85%) $(M+H)^+$, ESMS (-) m/z: 1035.2 (50%) $(M-H)^-$.

IV. Synthesis of H₂N-Lys(DOTA-OH)-Spacer-AQ (57)

Compound **56** (0.08 g, 0.08 mmoles) was dissolved in trifluoroacetic acid (2 ml) at rt. Reaction completion was monitored by TLC analysis. The solvent was then evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was triturated with diethyl ether (70 mL) and stored at 4-6 °C for 12 hours, filtered and dried in a vacuum desiccator. Crude yield: 0.03 g (43%). TLC: R_f 0.2 (DCM:MeOH-9:1). The purity of the compound was 95% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient F, 20 min, 247 nm, Rt. 8.82′). The molecular mass of the product (**57**) was 868.43, ESMS-LR (+) m/z: 435.2 (100%) $[(M+2H)/2]^{2+}$, 869.2 (95%) $(M+H)^+$. ESI-HR (-) m/z: 867.4259 (100%) $(M-H)^-$.

V. Synthesis of H₂N-Lys(DOTA-Gd³⁺)-Spacer-AQ (58)

Compound **57** (0.03 g, 0.03 mmol) was dissolved in water (3 mL). To this solution Gadolinium(III)chloridehexahydrate (2.8 eq, 0.04 g, 0.10 mmol) dissolved in water (2 mL) was added. The pH of the reaction mixture was adjusted to 7 by

adding a solution of ammonia 0.35% drop wise. Reaction progress was monitored by TLC, which showed a single new component and no presence of starting material. Crude yield: 0.02 g (50%). TLC: R_f 0.4 (Chloroform:MeOH:Pyridine:Ammonia-20:55:15:10). The molecular mass of the product (**58**) was 1023.33, ESMS-LR (+) m/z: 1024.2 (30%) (M+H)$^+$. ESMS-LR (-) m/z: 1022 (60%) (M-H)$^-$ 1057.2 (20%) (M-Cl)$^-$.

3.6.2. Synthesis of JS-Lysine series of tetra peptide substrates

A. Synthesis of Fmoc-Pro-Gly-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (62)

I. Synthesis of Fmoc-Pro-Gly-Asn(Trt)-OH (59)

An asparagine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:

- 1st cycle: *N*-Fmoc-Glycine (0.23 g, 0.79 mmol)
- 2nd cycle: *N*-Fmoc-L-Proline (0.27 g, 0.79 mmol)

Upon successful coupling of the second amino acid, the tri peptide was cleaved from the resin following method E. Crude yield: 0.06 g (75%). The purity of the compound was 70% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 15.36'). The molecular weight of the product (**59**) was 750.30, ESI-HR (+) m/z: 751.3120 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 750.8 (100%) (M)$^+$, 772.8 (100%) (M+Na)$^+$. ESMS-LR (-) m/z: 250.2 (80%) [(M-3H)/3]$^{3-}$, 749.4 (15%) (M-H)$^-$.

II. Synthesis of Fmoc-Pro-Gly-Asn(Trt)-Lys(DOTA-*t*Bu)-Spacer-AQ (60)

Peptide sequence **59** (0.06 g, 0.08 mmol) was dissolved in DMF (3 mL). To this solution compound **56** (1.1 eq, 0.09 g, 0.09 mmol), TBTU (3.3 eq, 0.08 g, 0.26 mmol), HOBt (3.3 eq, 0.04 g, 0.26 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (9.6 eq, 0.13 g/mL, 0.77 mmol) was added and stirred at rt for 5 hrs. Reaction progress was monitored *via* TLC analysis. The solution was then partitioned between DCM and water. The organic extracts were washed with saturated sodium hydrogen carbonate (1 x 75 mL), then water (3 x 75 mL), dried with anhydrous MgSO$_4$, filtered and evaporated to low volume. The crude product was then purified by a silica gel chromatography column (15 cmx2.5 cm) prepared with DCM:MeOH (9.3:0.8) and the major product was eluted with the further addition of MeOH (10%-15% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated *in vacuo*. Yield: 0.06 g (43%).

TLC (fraction 3): R_f 0.7 (DCM:MeOH-9:1). The molecular weight of the product (**60**) was 1770.92, ESMS-LR (+) m/z: 885.0 (100%) $[(M+2H)/2]^{2+}$, 1792.4 (15%) $(M+Na)^+$. ESMS-LR (-) m/z: 1803.6 (15%) $(M+Cl)^-$.

III. Synthesis of Fmoc-Pro-Gly-Asn-Lys(DOTA-OH)-Spacer-AQ (61)

Compound **60** (0.06 g, 0.03 mmol) was dissolved in TFA (2 mL) at rt. After 3 hours TFA was evaporated and the resulting concentrate was re-evaporated with EtOH(x3). The residue was then triturated with diethyl ether (50 mL) and stored at 4-6 °C for 12 hours. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.04 g (80%). TLC: R_f 0.5 (DCM:MeOH-9:1). The molecular mass of the product (**61**) was 1358.62, ESMS-LR (+) m/z: 680.0(95%) $[(M+2H)/2]^{2+}$, 1358.4 (100%) $(M)^+$, 1359.4 (85%) $(M+H)^+$, 1381.4 (40%) $(M+Na)^+$. ESMS-LR (-) m/z: 1357.2 (20%) $(M-H)^-$.

IV. Synthesis of Fmoc-Pro-Gly-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (62)

Compound **61** (0.04 g, 0.03 mmol) was dissolved in water (2 mL). To this solution Gadolinium(III)chloridehexahydrate (2.8 eq, 0.03 g, 0,08 mmol) dissolved in water (2 mL) was added. The pH of the reaction mixture was adjusted to 7 by adding a solution of ammonia 0.35% drop wise. Reaction progress was monitored by TLC analysis. The solution was evaporated to a low volume. The crude product obtained was then purified by a silica gel chromatography column (15 cmx2.5 cm) prepared with chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.03 g (75%). TLC (fraction 2): R_f 0.6 (Chloroform:MeOH:Pyridine:Ammonia-20:55:15:10). The purity of the compound was 45% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient E, 23 min, 254 nm, Rt. 11.81'). The molecular mass of the product (**62**) was 1513.52, ESI-HR (+) m/z: 1536.5082 (100%) $(M+Na)^+$. ESMS-LR (+) m/z: 768.7 (60%) $[(M+2H+Na)/2]^{2+}$, 779.7 (100%) $[(M+2Na)/2]^{2+}$, 1514.5 (30%) $(M+H)^+$.

B. Synthesis of of Fmoc-β-Ala-Gly-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (66)
I. Synthesis of Fmoc-β-Ala-Gly-Asn(Trt)-OH (63)

An asparagine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:
- 1st cycle: *N*-Fmoc-Glycine (0.23 g, 0.79 mmol).

- 2nd cycle: *N*-Fmoc-β-Alanine (0.25 g, 0.79 mol).

Upon successful coupling of the second amino acid, the tri peptide was cleaved from the resin following method E. Crude yield: 0.06 g (75%). The purity of the compound was 83% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 14.41'). The molecular weight of the product (**63**) was 724.29, ESI-HR (+) m/z: 725.2967 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 724.8(40%) (M)$^+$, 748.8 (100%) (M+Na)$^+$. ESMS-LR (-) m/z: 723.2 (20%) (M-H)$^-$.

II. Synthesis of Fmoc-β-Ala-Gly-Asn(Trt)-Lys(DOTA-^{t}Bu)-Spacer-AQ (64)

Peptide sequence **63** (0.06 g, 0.08 mmol) was dissolved in DMF (3 mL). To this solution compound **56** (1.1 eq, 0.09 g, 0.09 mmol), TBTU (3.3 eq, 0.09 g, 0.27 mmol), HOBt (3.3 eq, 0.04 g, 0.27 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (9.6 eq, 0.14 g/mL, 0.80 mmol) was added following the same procedure as described for the synthesis of compound **60**. The crude product obtained was then purified by a silica gel chromatography column (15 cmx2.5 cm) prepared with DCM:MeOH (9.3:0.8) and the major product was eluted with the further addition of MeOH (10%-15% gradients) to the eluant. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.08 g (57%). TLC (fraction 2): R$_f$ 0.7 (DCM:MeOH-9:1). The molecular weight of the product (**64**) was 1742.90, ESMS-LR (+) m/z: 872.4 (100%) [(M+2H)/2]$^{2+}$, 1742.4 (30%) (M)$^+$, 1764.4 (15%) (M+Na)$^+$. ESMS-LR (-) m/z: 1777.6 (40%) (M+Cl)$^-$.

III. Synthesis of Fmoc-β-Ala-Gly-Asn-Lys(DOTA-OH)–Spacer-AQ (65)

Compound **64** (0.06 g, 0.03 mmol) was dissolved in TFA (2 mL) at rt following the same procedure as described for the synthesis of compound **61**. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.03 g (60%). TLC: R$_f$ 0.4 (DCM:MeOH-9:1). The molecular mass of the product (**65**) was 1332.60, ESMS-LR (+) m/z: 667.0 (100%) [(M+2H)/2]$^{2+}$, 1333.6 (40%) (M+H)$^+$. ESMS-LR (-) m/z: 1331.4 (55%) (M-H)$^-$.

IV. Synthesis of of Fmoc-β-Ala-Gly-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (66)

Compound **65** (0.03 g, 0.02 mmol) was dissolved in water (3 mL). To this solution Gadolinium(III)chloridehexahydrate (2.8 eq, 0.02 g, 0.06 mmol) dissolved in water (2 mL) was added following the same procedure as described for the synthesis of compound **62**. The crude product obtained was then purified by a silica gel chromatography column (13.5 cmx2.5 cm) prepared with

chloroform:MeOH:pyridine:35%ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.01 g (33%). TLC (fraction 2): R_f 0.2 (Chloroform:MeOH:Pyridine Ammonia-20:55:15:10). The molecular mass of the product (**66**) was 1487.50, ESMS-LR (+) m/z: 1520.9 (100%) $(M+CH_3OH+H)^+$, 1533.9 (85%) $(M+2Na-H)^+$.

C. Synthesis of of Fmoc-Pro-Ala-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (70)

I. Synthesis of Fmoc-Pro-Ala-Asn(Trt)-OH (67)

An asparagine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:

- 1st cycle: *N*-Fmoc-L-Alanine (0.25 g, 0.79 mmol).
- 2nd cycle: *N*-Fmoc-L-Proline (0.27 g, 0.79 mmol).

Upon successful coupling of the second amino acid, the tri peptide was cleaved from the resin following method E. Crude yield: 0.05 g (63%). The purity of the compound was 66% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 14.68'). The molecular weight of the product (**67**) was 764.32, ESI-HR (+) m/z: 765.3280 (100%) $(M+H)^+$, ESMS-LR (+) m/z: 764.8(40%) $(M)^+$, 786.8 (100%) $(M+Na)^+$. ESMS-LR (-) m/z: 763.6 (20%) $(M-H)^-$.

II. Synthesis of Fmoc-Pro-Ala-Asn(Trt)-Lys(DOTA-^{t}Bu)-Spacer-AQ (68)

Peptide sequence **67** (0.05 g, 0.07 mmol) was dissolved in DMF (3 mL). To this solution compound **56** (1.1 eq, 0.07 g, 0.07 mmol), TBTU (3.3 eq, 0.07 g, 0.22 mmol), HOBt (3.3 eq, 0.03 g, 0.22 mmol) dissolved in DMF (5-10 mL) followed by the addition of DIPEA (9.6 eq, 0.11 g/mL, 0.63 mmol) was added following the same procedure as described for the synthesis of compound **60**. The product was used for subsequent reaction without further purification. Crude yield: 0.09 g (75%). TLC: R_f 0.7 (DCM:MeOH-9:1). The molecular weight of the product (**68**) was 1782.93, ESMS-LR (+) m/z: 892.0 (100%) $[(M+2H/2)]^{2+}$, 1782.4 (40%) $(M)^+$. ESMS-LR (-) m/z: 1816.0 (40%) $(M+Cl)^-$.

III. Synthesis of Fmoc-Pro-Ala-Asn-Lys(DOTA-OH)-Spacer-AQ (69)

Compound **68** (0.09 g, 0.05 mmol) was dissolved in TFA (2 mL) at rt following the same procedure as described for the synthesis of compound **61**. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.06 g (86%). TLC: R_f

0.5 (DCM:MeOH-9:1). The molecular mass of the product (**69**) was 1372.63, ESMS-LR (+) m/z: 687.2 (100%) $[(M+2H)/2]^{2+}$, 698.0 (40%) $[(M+2H+Na)/2]^{2+}$, 1373.0 (15%) $(M+H)^+$.

IV. Synthesis of of Fmoc-Pro-Ala-Asn-Lys(DOTA-Gd³⁺)-Spacer-AQ (70)

Compound **69** (0.06 g, 0.04 mmol) was dissolved in water (2 mL). To this solution Gadolinium (III) chloride hexahydrate (2.8 eq, 0.05 g, 0.12 mmol) dissolved in water (2 mL) was added following the same procedure as described for the synthesis of compound **62**. The crude product obtained was then purified by a silica gel chromatography column (13 cmx2.5 cm) prepared with chloroform: MeOH: pyridine: 35% ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.02 g (29%). TLC (fraction 2): R_f 0.3 (Chloroform:MeOH:Pyridine:Ammonia-20:55:15:10). The molecular mass of the product (**70**) was 1527.53, ESI-HR (+) m/z: 764.7751 (100%) $[(M+2H)/2]^{2+}$. ESMS-LR (+) m/z: 775.7 (100%) $[(M+2H+Na)/2]^{2+}$, 1527.5 (100%) $(M)^+$, 1528.5 (90%) $(M+H)^+$, 1550.5 (60%) $(M+Na)^+$.

D. Synthesis of of Fmoc-β-Ala-Ala-Asn-Lys(DOTA-Gd³⁺)-Spacer-AQ (74)

I. Synthesis of Fmoc-β-Ala-Ala-Asn(Trt)-OH (71)

An asparagine pre-loaded resin for solid phase peptide synthesis with a 4-carboxytrityl linker (0.5 g, 0.21 mmol/g) was swollen in DCM according to method A and the coupling and deprotection cycles were performed as per methods C, D, B followed by peptide cleavage from the resin according to method E:

- 1st cycle: *N*-Fmoc-L-Alanine (0.25 g, 0.79 mmol).
- 2nd cycle: *N*-Fmoc-β-Alanine (0.25 g, 0.79 mmol).

Upon successful coupling of the second amino acid, the tri peptide was cleaved from the resin following method E. Crude yield: 0.07 g (88%). The purity of the compound was 63% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient A, 23 min, 254 nm, Rt. 14.68'). The molecular weight of the product (**71**) was 738.31, ESI-HR (+) m/z: 739.3124 (100%) $(M+H)^+$. ESMS-LR (+) m/z: 738.8 (60%) $(M)^+$, 760.8 (100%) $(M+Na)^+$. ESMS-LR (-) m/z: 737.4 (20%) $(M-H)^-$.

II. Synthesis of Fmoc-β-Ala-Ala-Asn(Trt)-Lys(DOTA-ᵗBu)-Spacer-AQ (72)

Peptide sequence **71** (0.07 g, 0.09 mmol) was dissolved in DMF (3 mL). To this solution compound **56** (1.1 eq, 0.11 g, 0.10 mmol), TBTU (3.3 eq, 0.10 g, 0.31 mmol), HOBt (3.3 eq, 0.04 g, 0.31 mmol) dissolved in DMF (5-10 mL) followed by

the addition of DIPEA (9.6 eq, 0.16 g/mL, 0.91 mmol) was added following the same procedure as described for the synthesis of compound **60**. The product was used for subsequent reaction without further purification. Crude yield: 0.15 g (88%). TLC: R_f 0.8 (DCM:MeOH-9:1). The molecular weight of the product (**72**) was 1758.92, ESMS-LR (+) m/z: 879.0 (100%) $[(M+2H)/2]^{2+}$, 1758.4 (40%) $(M)^+$. ESMS-LR (-) m/z: 1790.4 (45%) $(M+Cl)^-$.

III. Synthesis of Fmoc-β-Ala-Ala-Asn-Lys(DOTA-OH)-Spacer-AQ (73)

Compound **72** (0.15 g, 0.09 mmol) was dissolved in TFA (3 mL) at rt following the same procedure as described for the synthesis of compound **61**. The precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.08 g (73%). TLC: R_f 0.08 (DCM:MeOH-9:1). The molecular mass of the product (**73**) was 1346.62 ESMS-LR (+) m/z: 674.0 (100%) $[(M+2H)/2]^{2+}$, 1346.6 (80%) $(M)^+$, 1347.6 (60%) $(M+H)^+$, 1368.6 (20%) $(M+Na)^+$. ESMS-LR (-) m/z: 1345.6 (80%) $(M-H)^-$

IV. Synthesis of of Fmoc-β-Ala-Ala-Asn-Lys(DOTA-Gd^{3+})-Spacer-AQ (74)

Compound **73** (0.08 g, 0.06 mmol) was dissolved in water (2 mL). To this solution Gadolinium (III) chloride hexahydrate (2.8 eq, 0.06 g, 0.17 mmol) dissolved in water (2 mL) was added following the same procedure as described for the synthesis of compound **62**. The crude product obtained was then purified by a silica gel chromatography column (20 cmx2.5 cm) prepared with chloroform: MeOH:pyridine:35% ammonia (20:55:15:10). The column was eluted with the same solvent system. Fractions containing the product were combined, filtered and evaporated to a low volume. Yield: 0.07 g (78%). TLC (fraction 2): R_f 0.4 (Chloroform: MeOH: Pyridine: Ammonia- 20:55:15:10). The purity of the compound was 65% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient G, 23 min, 254 nm, Rt. 11.81'). The molecular mass of the product (**74**) was 1501.52, ESI-HR (+) m/z: 1502.5225 (100%) $(M+H)^+$. ESMS-LR (+) m/z: 751.7 (100%) $[(M+2H)/2]^{2+}$, 762.7 (90%) $[(M+2H+Na)/2]^{2+}$, 1524.5 (55%) $(M+Na)^+$.

3.7. MMP (2/9) targeted gadolinium based contrast agents for MR imaging of atherosclerotic plaques

3.7.1. Synthesis of peptide substrates designed to be specific to MMP-2/9 *via* solid phase peptide synthesis (JJ series)

I. Synthesis of *N*-Fmoc-Ala-Pro-Leu-Gly-Ala-Ser-Gly-OH (75)

Glycine pre-loaded resin for solid phase peptide synthesis with a 2-chlorotrityl

linker (0.5 g, 0.64 mmol/g) was swollen in DCM as in method A and coupling/deprotection cycles were performed as follow:

- 1st cycle: *N*-Fmoc-L-Ser (tBu) (0.92 g, 2.40 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit following method D. Removal of the subsequent Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

- 2nd cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit following method D. Removal of the subsequent Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

- 3rd cycle: *N*-Fmoc-Glycine (0.71 g, 2.40 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit, following method D. Removal of the subsequent Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

- 4th cycle: *N*-Fmoc-L-Leucine (0.85 g, 2.40 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit, following method D. Removal of the subsequent Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

- 5th cycle: *N*-Fmoc-L-Proline (0.81 g, 2.40 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit, following method D. Removal of the subsequent Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

- 6th cycle: *N*-Fmoc-D-Alanine (0.74 g, 2.40 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit, following method D.

On successful completion of the sixth coupling the polypeptide was cleaved from

the resin following the procedure in method E. Crude yield: 0.16 g (59%). Molecular mass of the peptide sequence (**75**) with *tert*-Butyl protecting group on serine was 849.97, ESI-HR (+) m/z: 850.4349 (99%) (M+H)$^+$. ESMS-LR (-) m/z: 849.2 (50%) (M)$^-$.

To a few resin beads TFA (100%) was added and the procedure as in method E was followed. Molecular mass of peptide sequence without the *tert*-Butyl protecting group on serine was 793.99, ESMS-LR (+) m/z: 816.4 (99%) (M+Na)$^+$. ESMS-LR (-) m/z: 793.2 (50%) (M)$^-$.

II. Synthesis of *N*-Fmoc-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH (76)

Glycine pre-loaded resin for solid phase peptide synthesis with a 2-chlorotrityl linker (0.5 g, 0.64 mmol/g) was swollen in DCM as in method A and coupling/deprotection cycles were performed as follow:

- 1st cycle: *N*-Fmoc-L-Leucine (0.85 g, 2.40 mmol)
- 2nd cycle: *N*-Fmoc-L-Norvaline (0.81 g, 2.40 mmol)
- 3rd cycle: *N*-Fmoc-Alanine (0.74 g, 2.40 mmol)
- 4th cycle: *N*- Fmoc-L-Leucine (0.85 g, 2.40 mmol)
- 5th cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)
- 6th cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)

On successful completion of the sixth coupling the polypeptide was cleaved from the resin following method E. Crude yield: 0.06 g (22%). Molecular mass of the peptide chain (**76**) was 835.39, ESMS-LR (+) m/z: 836.4 (100%) (M+H)$^+$, 859 (70%) (M+Na)$^+$.

III. Synthesis of *N*-Fmoc-Lys-Pro-Ala-Gly-Nva-Ala-Gly-OH (77)

Glycine pre-loaded resin for solid phase peptide synthesis with a 2-chlorotrityl linker (0.5 g, 0.64 mmol/g) was swollen in DCM as in method A and coupling/ deprotection cycles were performed as follow:

- 1st cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)
- 2nd cycle: *N*-Fmoc-L-Norvaline (0.81 g, 2.40 mmol)
- 3rd cycle: *N*-Fmoc-Glycine (0.71 g, 2.40 mmol)
- 4th cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)
- 5th cycle: *N*-Fmoc-L-Proline (0.81 g, 2.40 mmol)
- 6th cycle: *N*-Fmoc-L-Lysine(Boc) (1.12 g, 2.40 mmol)

On successful completion of the seventh coupling the polypeptide was cleaved from the resin following method E. Crude yield: 0.21 g (81%). Molecular mass of the peptide chain (**77**) was 820.93, ESI-HR (+) m/z: 821.4191 (100%) (M+H)$^+$ ESMS-

LR (-) m/z: 819.2 (99%) (M-H)⁻.

IV. Synthesis of *N*-Fmoc-Gly-Pro-Arg(Mtr)-Glu(OᵗBu)-Ile-Thr(ᵗBu)-Ala-Gly-OH (78)

Glycine pre-loaded resin for solid phase peptide synthesis with a 2-chlorotrityl linker (0.5 g, 0.64 mmol/g) was swollen in DCM as in method A and coupling/deprotection cycles were performed as follow:

- 1ˢᵗ cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)
- 2ⁿᵈ cycle: *N*-Fmoc-L-Theronine (ᵗBu) (0.95 g, 2.40 mmol)
- 3ʳᵈ cycle: *N*-Fmoc-L-Isoleucine (0.84 g, 2.40 mmol)
- 4ᵗʰ cycle: *N*- Fmoc-L-Glutamate (OᵗBu) (1.02 g, 2.40 mmol)
- 5ᵗʰ cycle: *N*-Fmoc-L-Arginine (Mtr) (1.46 g, 2.40 mmol)
- 6ᵗʰ cycle: *N*-Fmoc-L-Proline (0.81 g, 2.40 mmol)
- 7ᵗʰ cycle: *N*-Fmoc-Glycine (0.71 g, 2.40 mmol)

On successful completion of the seventh coupling the polypeptide was cleaved from the resin following method E. Crude yield: 0.25 g (58%). The purity of the compound was 76% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient D, 28 min, 300 nm, Rt. 19.76'). Molecular mass of the peptide chain (**78**) was 1345.40, ESI-HR (+) m/z: 1346.6681 (100%) (M+H)⁺. ESMS-LR (+) m/z: 673.8 (90%) [(M+2H)/2]²⁺. ESMS-LR (-) M/Z: 1344.4 (99%) (M-H)⁻; 1345.2 (75%) (M)⁻.

V. Synthesis of *N*-Fmoc-Ile-Pro-Arg(Mtr)-Thr(ᵗBu)-Leu-Thr(ᵗBu)-Ala-Gly-OH (79)

Glycine pre-loaded resin for solid phase peptide synthesis with a 2-chlorotrityl linker (0.5 g, 0.64 g/mmol) was swollen in DCM as in method A and coupling/deprotection cycles were performed as follow:

- 1ˢᵗ cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)
- 2ⁿᵈ cycle: *N*-Fmoc-L-Theronine(ᵗBu) (0.95 g, 2.40 mmol)
- 3ʳᵈ cycle: *N*-Fmoc-L-Leucine (0.84 g, 2.40 mmol)
- 4ᵗʰ cycle: *N*- Fmoc-L-Theronine(ᵗBu) (0.95 g, 2.40 mmol)
- 5ᵗʰ cycle: *N*-Fmoc-L-Arginine(Mtr) (1.46 g, 2.40 mmol)
- 6ᵗʰ cycle: *N*-Fmoc-L-Proline (0.81 g, 2.40 mmol)
- 7ᵗʰ cycle: *N*-Fmoc-L-Isoleucine (0.84 g, 2.40 mmol)

On successful completion of the seventh coupling the polypeptide was cleaved from the resin following method E. Crude yield: 0.40 g (91%). The purity of the compound was 97% on an analytical RP-HPLC system following the general

procedure as mentioned in 4.2.5. (Gradient C, 28 min, 254 nm, Rt. 21.52'). Molecular mass of the peptide chain (**79**) was 1373.15, ESI-HR (+) m/z: 1374.7366 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 1396.7 (20%) (M+Na)$^+$. ESMS-LR (-) m/z: 1372.2 (99%) (M-H)$^-$; 1373.2 (85%) (M)$^-$.

VI. Synthesis of *N*-Fmoc-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr(^{t}Bu)-Ala-Gly-OH (80)

Glycine pre-loaded resin for solid phase peptide synthesis with a 2-chlorotrityl linker (0.5 g, 0.64 g/mmol) was swollen in DCM as in method A and coupling/ deprotection cycles were performed as follow:

- 1st cycle: *N*-Fmoc-L-Alanine (0.74 g, 2.40 mmol)
- 2nd cycle: *N*-Fmoc-L-Theronine(^{t}Bu) (0.95 g, 2.40 mmol)
- 3rd cycle: *N*-Fmoc-L-Leucine (0.84 g, 2.40 mmol)
- 4th cycle: *N*- Fmoc-L-Arganine(Mtr) (1.46 g, 2.40 mmol)
- 5th cycle: *N*-Fmoc-L-Arginine(Mtr) (1.46 g, 2.40 mmol)
- 6th cycle: *N*-Fmoc-L-Proline (0.81 g, 2.40 mmol)
- 7th cycle: *N*-Fmoc-Glycine (0.71 g, 2.40 mmol)

On successful completion of the seventh coupling the polypeptide was cleaved from the resin following method E. Crude yield: 0.45 g (92%). The purity of the compound was 83% on an analytical RP-HPLC system following the general procedure as mentioned in 4.2.5. (Gradient C, 28 min, 254 nm, Rt. 19.94'). Molecular mass of the peptide chain (**80**) was 1528.35, ESI-HR (+) m/z: 1529.7147 (100%) (M+H)$^+$. ESMS-LR (+) m/z: 765 (70%) [(M+2H)/2]$^{2+}$; 1551 (25%) (M+Na)$^+$. ESMS-LR (-) m/z: 1527 (85%) (M-H)$^-$.

3.7.2. Synthesis of Gadolinium-DOTA-Peptide substrate designed to be specific to MMP-2/9

3.7.1.1. Attempted synthesis of JD- (Gd(III)-DOTA-Gly-12-aminododecanoicacid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-OH)

I. Synthesis of 12-({[(9*H*-fluoren-9-yl)methoxy]carbonyl}amino)dodecanoicacid (81)

Fmoc *N*-hydroxysuccinimide ester (Fmoc-OSu) (3.00 g, 8.89 mmol) was dissolved in dioxane:water (60 mL, 4:1). To this solution 12-aminododecanoic acid (1.1 eq, 2.11 g, 9.78 mmol) dissolved in dioxane:water (50 mL, 4:1) with triethylamine (TEA) (2 eq, 2.48 mL, 17.79 mmol) was added and stirred at rt for 12hrs. The solution was concentrated and the mixture was partitioned between DCM and

water. The DCM extracts were washed with saturated sodium hydrogen carbonate (3 x 100 mL), water (3 x 100 mL), dried, filtered and evaporated to low volume. The crude product obtained was purified by column chromatography. The column (13 cmx2.5 cm) was eluted initially with DCM. The eluent was then changed to DCM:MeOH (9.5:0.5). Fractions containing the product were combined, filtered and evaporated to dryness. Yield: 1.29 g (33%). TLC (fraction 3): R_f 0.2 (DCM:MeOH [9.5:0.5]). Molecular mass of compound (**81**) was 437.26, ESMS-LR (+) m/z: 438.2 (99%) $(M+H)^+$.

H^1 NMR (400MHz, d_1-CDCl$_3$), δ ppm: 1.30 (14H, s, 7xCH$_2$), 1.45-1.52 (2H, m, C$\underline{H}_2$-CH$_2$-NH), 1.65 (2H, quin, J=7.2Hz, C$\underline{H}_2$-CH$_2$-COOH), 2.37 (2H, t, J=7.4Hz, C$\underline{H}_2$-COOH), 3.20 (2H, m, C$\underline{H}_2$-NH), 4.22-4.28 (1H, m, C$\underline{H}$-Fmoc protecting group), 4.42-4.47 (2H, m, C$\underline{H}_2$-CH-Fmoc protecting group), 7.34 (2H, td, J_1=7.6Hz, J_2=1.2Hz, 2xH-2), 7.43 (2H, t, J=7.4Hz, 2xH-3), 7.62 (2H, d, J=7.2Hz, 2x H-4), 7.79 (2H, d, J=7.6Hz, 2xH-1).

II. Synthesis of HO-DOTA-Gly-12-aminododecanoicacid-Ala-Ala-Leu-Ala-Nva-Leu-Gly-2-ClTrt resin (84)

1stcycle: Compound 81 (0.15 g, 0.35 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit, following method D. Removal of the corresponding Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

2ndcycle (82): *N*-Fmoc-Glycine (0.10 g, 0.35 mmol) was coupled to the resin according to method B. Primary amine detection confirming the success of the reaction was checked using the Kaiser test kit, following method D. Removal of the corresponding Fmoc protecting group was performed in accordance to method C. Presence of free amine due to deprotection was confirmed by Kaiser test (method D).

3rdcycle (83): DOTA-tris-(^{t}Bu) (0.20 g, 0.35 mmol) was coupled to the resin according to method B. The reaction mixture was stirred for 2hrs at rt.

On completion of the three coupling the polypeptide was cleaved from the resin using the cleavage solution, TFA: H$_2$O: EDT: TA (87.5:2.5:4:6). After 3.5 hrs the solvent was evaporated and the concentrate was re-evaporated with EtOH. The resulting product (84) was triturated in diethyl ethyl and stored at 6°C overnight. The molecular mass of the compound (**84**) was 1253.57, ESMS-LR (-) m/z: 1252.4 (99%) $(M-H)^-$; 1253 (75%) $(M)^-$.

3.7.1.2. Synthesis of JL (87)

I.   Synthesis of DOTA-tris(^{t}Bu)-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr(O^{t}Bu)-Ala-Gly-Lys-Spacer-AQ (85)

Compound **56** (0.10 g, 0.10 mmol) was dissolved in DMF (5 mL). To this solution peptide sequence **80** (1.1 eq, 0.16 g, 0.11 mmol), TBTU (2.2 eq, 0.07 g, 0.21 mmol), HOBt (2.2 eq, 0.03 g, 0.21 mmol), DIPEA (6.4 eq, 0.11 g/mL, 0.62 mmol) dissolved in DMF (5 mL) was added and stirred at room temperature overnight. The solution was then partitioned between DCM and water. The organic solution was washed with saturated sodium hydrogen carbonate (3x 100 mL), then water (3x100 mL), dried, filtered and evaporated to low volume. Crude yield: 0.17 g (68%). TLC: R_f 0.8 (DCM:MeOH-9:1). The molecular weight of the product (**85**) was 2547.48, ESMS-LR (+) m/z: 857.0 (10%) $[(M+2H+Na)/3]^{3+}$, 1274.2 (100%) $[(M+2H)/2]^{2+}$, 1285.2 (15%) $[(M+2H+Na)/2]^{2+}$, 2547.8 (15%) $(M)^+$.

II. Synthesis of HO-DOTA-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr-Ala-Gly-Lys-Spacer-AQ (86)

Compound **85** (0.17 g, 0.07 mmol) was dissolved in TFA (2 mL) at rt for 7hours. TFA was then evaporated and the resulting concentrate was re-evaporated with EtOH (x3). The residue was triturated with diethyl ether (100 mL) and then stored at 4-6 °C for 12 hours. The resulting precipitate was filtered and dried in a vacuum desiccator. Crude yield: 0.08 g (50%). TLC: R_f 0.05 (DCM:MeOH-9:1). The molecular mass of the product (**86**) was 2323.07, ESMS-LR (+) m/z: 775.0 (75%) $[(M+3H)/3]^{3+}$, 1162.2 (100%) $[(M+2H)/2]^{2+}$, 1184.0 (15%) $[(M+2Na)/2]^{2+}$, 2323.2 (15%) $(M)^+$, 2347.0 (10%) $(M+Na)^+$.

III. Synthesis of DOTA(Gd^{3+})-Gly-Pro-Arg(Mtr)-Arg(Mtr)-Leu-Thr-Ala-Gly-Lys-Spacer-AQ (JL-87)

Compound **86** (0.08 g, 0.03 mmol) was dissolved in water (2 mL). To this solution Gadolinium (III) chloride hexahydrate (2.8 eq, 0.04 g, 0.10 mmol) dissolved in water (2 mL) was added. The pH of the reaction mixture was adjusted to 7 by adding a solution of ammonia (0.35%) drop wise and stirred at rt for 12hrs. Reaction progress was monitored *via* TLC analysis. The solution was then evaporated at rt. Crude yield: 0.05 g (56%). TLC: R_f 0.2 (Chloroform:MeOH:Pyridine:Ammonia-20:55:15:10). The molecular mass of the product (**JL-87**) was 2477.97, ESMS-LR (+) m/z: 826.4 (60%) $[(M+3H)/3]^{3+}$, 1239.0 (100%) $[(M+2H)/2]^{2+}$, 1250.0 (35%) $[(M+2H+Na)/2]^{2+}$, 2477.4 (15%) $(M)^+$.

3.8. UV-Vis Absorption Spectra

3.8.1. Materials

Water was used as the control; compound: **9**, 1 mg/mL stock in water, similarly stock concentration was prepared for compounds 11, 31 and A (5(6)-CBF-Pro-Gly-Asn-OH); UV-Vis absorbance spectra were recorded on a Helios-alpha UV/Vis spectrometer (Thermo Scientific, software-VisionPro,UK); 3 mL quartz cuvette.

3.8.2. Method

UV-Vis absorption spectra was performed to determine the absorption area between wavelength 200-600 nm. The compounds were dissolved in water and the final concentration was 50 µM (for each compound) in the cuvette.

3.9. Fluorescence measurements

3.9.1. Materials

Water was used as the control; compounds: **9**, **11**, **31** and A (1 mg/mL stock in water); Fluorescence spectra were recorded on a LS-55 luminescence spectrophotometer (Perkin Elmer, software-FL Win Lab, UK) using a quartz cell of 1cm path length at rt.

3.9.2. Method

Fluorescent measurements are done in order to evaluate fluorescence intensity of FRET compounds. This is a quick and simple procedure to compare fluorescence intensities between the FRET compounds and its fluorophore. The fluorescence spectra for the samples were measured using an excitation wavelength of 490nm and emission wavelength scan over the range 500nm-600nm. In order to have a final concentration of 0.5 µM of the fluorophore 5(6)-CBF 15 µL of 5(6)-CBF (stock conc. 100 µM) was mixed with 2985 µL of water; and 16.6 µL of compound 11 was mixed with 2983.4 µL of water to make a final concentration of 5 µM and the results were recorded.

3.10. Legumain activity assay

The legumain assay for the FRET compounds (JC series) was done in order to evaluate the substrate specificity of the compounds when incubated with legumain. This can be analyzed by the gradual increase in fluorescence when cleaved by the activated enzyme (legumain).

3.10.1. Materials

Recombinant human legumain/asparaginyl endopeptidase (rh Legumain) was purchased from R&D Systems (UK) as a pro-form 49 kDa enzyme. Reagents and materials required for the legumain assay were: activation buffer consisting of 50 mM sodium acetate, 100 mM sodium chloride, pH 4.0; assay buffer consisting of 50 mM MES (2-(*N*-morpholino)ethanesulfonic acid), 250 mM NaCl, pH 5.0; 96well black maxisorp F16 plate; FluoStar Omega multi-mode microplate reader and test compounds (selective compounds from the JC and JS series).

3.10.2. Legumain assay on JC series of peptide substrates

A stock solution was produced by taking rh-Legumain 10 µg and diluting with 100 µL activation buffer (pH 4.0) and stored at -78°C. The enzyme was activated by incubating the aliquots for 2hrs at 37ºC. The activated enzyme was then diluted to 1 ng/µL using assay buffer at pH 5.0 Each compound (8, 11, 31, 34, 38, 41 and 46) in the substrate library (stock solution: 1mg/ml dissolved in DMSO) was assayed in triplicate at a final concentration of 10 µM per well against the activated enzyme (40 µl/well).

For e.g., in each well: A1-C1: 100µL assay buffer (B).
A2-C2: 90µL assay buffer (B) + 10µL 100µM 5(6)Carboxyfluorescein (F).
A3-C3: 98.9µL assay buffer (B) + 1.1µL 902.1µM probe (P).
A4-C4: 60 µL assay buffer (B) + 40µL 1ng/µL rh Legumain stock solution (L).
A5-C5: 50µL (B) + 10µL 100µM (F) + 40µL 1ng/µL (L).
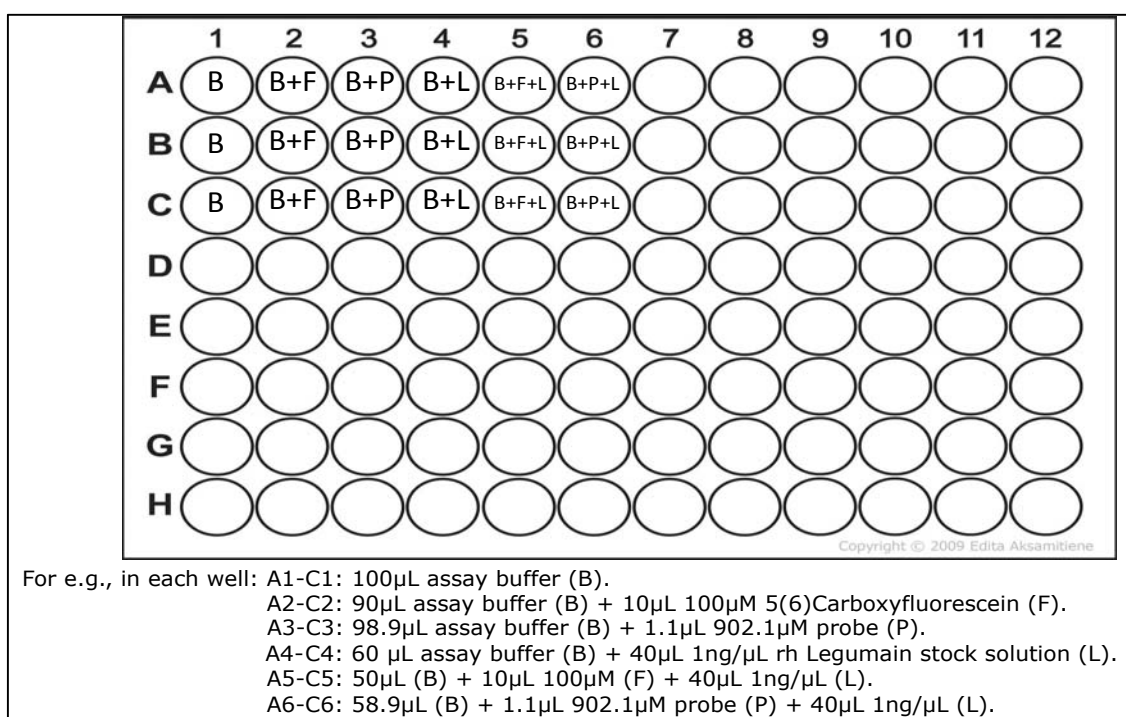A6-C6: 58.9µL (B) + 1.1µL 902.1µM probe (P) + 40µL 1ng/µL (L).

Figure 64: Template for assay conditions.

Enzyme assay was performed in a 96 well microtiter plate and the final volume in each well was made 100µl using assay buffer (figure 64). The gradual increase in fluorescence was measured at 5 minute intervals using FLUOstar omega multi-mode microplate reader (BMG Labtech, software-Omega) with excitation at 485

nm and emission at 520 nm based on the fluorescence properties of free 5(6)-CBF.

3.10.3. Legumain assay for JS linear/JS lysine series of peptide substrates

The same procedure as mentioned in section 3.9.2. was followed for rh-Legumain storage, preparation activation and incubation with compounds. The major difference was, the assay was performed in sterile eppendorf tubes with the final volume in each tube being 600 μL (e.g., Assay buffer: 352.3 μL, Substrate (10 μM): 7.7 μL, rh Legumain: 240 μL) for compounds 49, 53, 62, 66, 70 and 74. The compounds were incubated with the activated legumain for 2 hours at 37°C. The resulting enzyme digested product was neutralized with triethylamine and then extracted into chloroform (500 μL) and washed with water (500 μL) (x10). Both the organic and aqueous solutions were air-dried at room temperature and characterized by high-resolution ESI-MS.

3.11. Cytotoxicity assay

3.11.1. MDA-MB-231 cells

Human breast epithelial cancer cells (MDA-MB-231) were purchased from European Collection of Cell Cultures (ECACC) (ECACC: 92020424). MDA-MB-231 ($1x10^4$ cells/well) cells were cultured in Roswell Park Memorial Institute 1640 medium (RPMI-1640) containing GlutaMAX^{TM}-1 with 25 mM HEPES supplemented with Fetal calf serum (FCS) (10% v/v) and penicillin (1%,100 U/mL)/ streptomycin (100 μg/mL).

3.11.2. Maintenance of MDA-MB-231 cells

The MDA-MB-231 cells were incubated (HERAcell 150, Thermo scientific, UK) at 37°C in a humidified atmosphere of 5% CO_2. Cells were passaged by washing twice with phosphate buffered saline (PBS) and detached for 5 minutes with trypsin-EDTA at 37°C. Trypsinised cells were neutralized with RPMI medium and centrifuged (Heraeus Primo R, Thermo scientific, UK) for 5 minutes (1500 rpm) at room temperature. The supernatant was discarded and the cell pellet resuspended in RPMI medium. The cells were then counted using a Neubauer Haemocytometer under a Leica DMIL- Inverted microscope (Leica Micro systems ltd, UK).

3.1.2.3. Colourimetric 3-(4,5-dimethylthiazol-2-yl)-2,5-(diphenyltetrazo liumbromide) (MTT) assay

MDA-MB-231 ($1x10^4$ cells/well) cells were seeded in a 96 well plates. After 24 hour of cell attachment, four compounds- 62, 66, 70 and 58 were added to each well to produce a final concentration of 0.01, 0.1, 1, 5, 10, 15, 20 and 40Mm in DMSO

(100 µL/well) and incubated for 24 and 72 hours (figure 65). DMSO was used as the control solvent. After the preferred incubation time, the cell culture medium was removed and replaced with sterile filtered MTT solution (1 mg/mL in medium, 100 µL/well). The cells were incubated for 4 hours at 37°C.

R- Blank (RPMI medium)
H1- Control 1 (e.g., dH$_2$O)
H2- Control 2 (e.g., DMSO at highest test solution concentration)
S1-S8 – Compound solution (each compound at 8 different concentrations- 0.01, 0.1, 1, 5, 10, 15, 20 and 40µM)

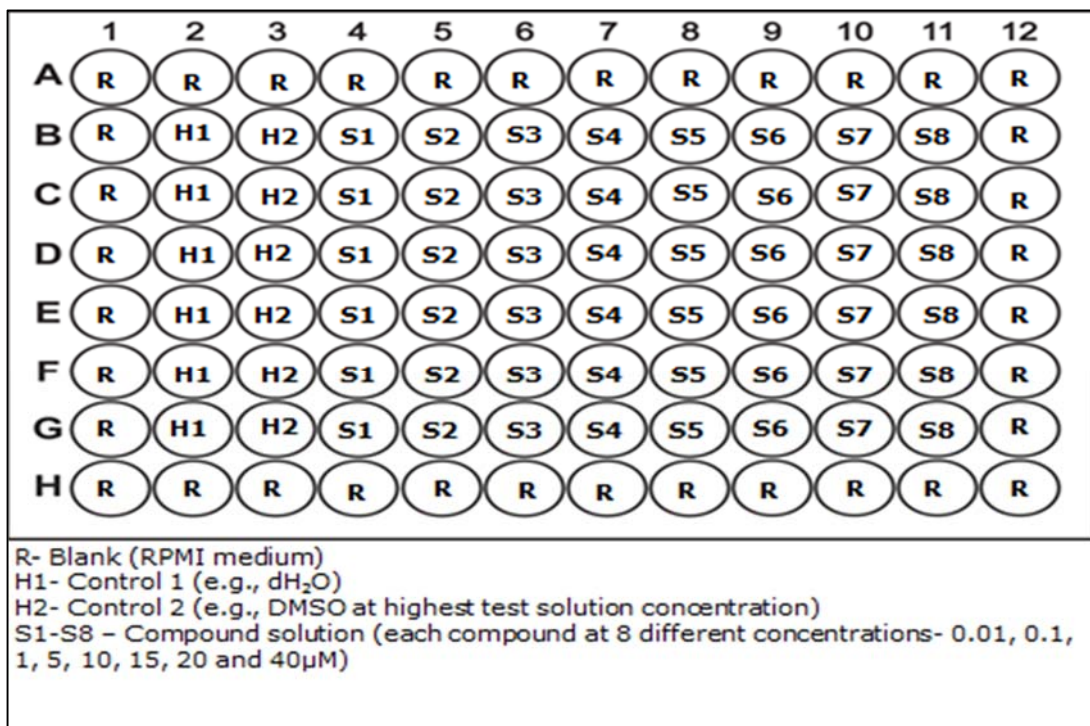Figure 65: MTT assay condition for the drug compounds.

After the incubation time, MTT solution was replaced with DMSO (200 µL/well), which solubilizes the metabolized MTT product (formazan). The plates were shaken (Polymax 1040 wave platform shaker, Heidolph) for 15-20 minutes at room temperature and the absorbance was measured at 560 nm (Synergy HT microplate reader, BioTek, software-Gen 52.04).

5. Reference

Albericio F (2000). Orthogonal protecting groups for N (alpha)-amino and C-terminal carboxyl functions in solid-phase peptide synthesis. *Biopolymers*. **55**, 2: 123-39.

Alberto P, Francesca I, Chiara S, and Ranuccio N (2006). Acute Coronary Syndromes: From the Laboratory Markers to the Coronary Vessels. *Biomarker Insights*. **1**, 123–130.

American Heart Association (2017). Heart Disease and Stroke Statistics - 2017 Update: A report from American heart association. *Circulation*. **135,** 10.

Anichkov NN. Experimental arteriosclerosis in animals. In: Cowdry EV, editor, Arteriosclerosis: A survey of the problem. New York: MacMillan Publishing; 1933. p. 271–322.

Atlas D, Levit S, Schechter I and Berger A (1970). On the active site of elastase: Partial mapping by means of specific peptide substrates. *FEBS Letters*. **11**, 4: 281-283.

Banks PR and Paquette DM (1995). Comparison of three common amine reactive fluorescent probes used for conjugation to biomolecules by capillary zone electrophoresis. *Bioconjugate Chemistry*. **6**, 4: 447-458.

Backes BJ, Harris JL, Leonetti F, Craik CS, Ellman JA (2000). Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. *Nature Biotechnology*. **18**, 2: 187-93.

Bagshawe KD (1993) Antibody-Directed Enzyme Prodrug Therapy (ADEPT). *Advances in Pharmacology*. **24**, 99-121.

Barchiesi F, Jackson EK, Fingerle J, Gillespie DG, Odermatt B, Dubey RK (2006). 2-Methoxyestradiol, an estradiol metabolite, inhibits neointima formation and smooth muscle cell growth via double blockade of the cell cycle. *Circulation Research*. **99**, 266-274.

Barlos K, Gatos D, Kallitsis J, Papaphotiu G, Sotiriu P, Wenging Y, Schafer W (1989). Darstellung geschutzter Peptid-Fragmente unter Einsatz substituierter Triphenylmethyl-Harze. *Tetrahedron Letters*. **30**, 3943.

Berridge MV and Tan AS (1993). Characterization of the cellular reduction of 3-(4, 5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): subcellular

localization, substrate dependence, and involvement of mitochondrial electron transport in MTT reduction. *Archives of Biochemistry and Biophysics*. **303**, 2:474-82.

Barron GA, Bermano G, Gordon A and Lin PKT (2010). Synthesis, cytotoxicity and DNA-binding of novel bisnaphthalimidopropyl derivatives in breast cancer MDA-MB-231 cells. *European Journal of Medicinal Chemistry*. **45**, 4: 1430-1437.

Bidmanova S, Hlavacek A, Damborsky J, Prokop Z (2012). Conjugation of 5(6)-carboxyfluorescein and 5(6)-carboxynaphthofluorescein with bovine serum albumin and their immobilization for optical pH sensing. *Sensors and Actuators B*: *Chemical*. **161**, 93-99.

Blum G, Mullins SR, Keren K, Fonovic M, Jedeszko C, Rice MJ, Sloane BF, Bogyo M. (2005) Dynamic imaging of protease activity with fluorescently quenched activity-based probes. *Natural Chemical Biology*. **1**, 203-209.

Blum G, von DG, Merchant MJ, Blau HM, Bogyo M (2007). Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. *Natural Chemical Biology*. **3**, 668-677.

Borkakoti N (2004). Matrix metalloproteinase inhibitors: design from structure. *Biochemical Society Translation*. **32**, 17-20.

Botnar RM, Stuber M, Kissinger KV, Kim WY, Spuentrup E, Manning WJ (2000). Non-invasive coronary vessel wall and plaque imaging with magnetic resonance imaging. *Circulation*. **102**, 21: 2582-2587.

Boudi FB and Ahsan CH (2014). Atherosclerosis. [Online]. Available from: http://emedicine.medscape.com/article/150916-overview. [Accessed 22 November 2014].

Bousquet JC, Saini S, Stark DD, Hahn PF, Nigam M, Wittenberg J, Ferrucci JT Jr (1988). Gd-DOTA: characterization of a new paramagnetic complex. *Radiology*. **166**, 3: 693-8.

Brosnihan KB, Senanayake PS, Li P and Ferrario CM (1999). Bi-directional actions of estrogen on the renin-angiotensin system. *Brazilian Journal of Medical and Biological Research*. **32**, 4: 373-81.

Buja LM and Willerson JT (1994). Role of inflammation in coronary plaque disruption. *Circulation*. **89**, 503–505.

Buntru A, Zimmermann T and Hauck CR (2009). Fluorescence resonance energy transfer (FRET) - based subcellular visualization of pathogen - induced host receptor signalling. *BioMed Central Biology*. **25**, 7: 81.

Cailleau R, Young R, Olivé M, Reeves WJ Jr (1974). Breast tumor cell lines from pleural effusions. *Journal of the National Cancer Institute*. **53**, 3: 661-74.

Calcagno C, Ramachandran S, Millon A, Robson PM, Mani V, Fayad Z (2014). Gadolinium-Based Contrast Agents for Vessel Wall Magnetic Resonance Imaging (MRI) of Atherosclerosis. *Current Cardiovascular Imaging Reports*. **6**, 1: 11–24.

Camici PG, Rimoldi OE, Gaemperli O, Libby P (2012). Non-invasive anatomic and functional imaging of vascular inflammation and unstable plaque. *European Heart Journal*. **33**, 11:1309-17.

Caravan P, Ellison JJ, McMurry TJ, Lauffer RB (1999). Gadolinium (III) Chelates as MRI Contrast Agents: Structure, Dynamics, and applications. *Chemical Reviews*. **99**, 2293-2352.

Caravan P (2006). Strategies for increasing the sensitivity of gadolinium based MRI contrast agents. *Chemical Society Reviews*. **35**, 512-523.

Caravan P (2009). Protein targeted gadolinium based magnetic resonance imaging (MRI) contrast agents: design and mechanism of action. *Accounts of chemical Research*. **42**, 7: 851-862.

Chan WC and White PD (2000). Fmoc solid phase peptide synthesis: a practical approach. Oxford, Oxford University Press. 341.

Chandrudu S, Simerska P and Toth I (2013). Chemical methods for peptide and protein production. *Molecules*. **18**, 4: 4373-88.

Chen JM, Dando PM, Rawlings ND, Brown MA, Young NE, Stevens RA, Hewitt E, Watts C, Barrett AJ (1997). Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *Journal of Biological Chemistry*. **272**, 8090–8098.

Chen JM, Dando PM, Stevens RA, Fortunato M, Barrett AJ (1998). Cloning and expression of mouse legumain, a lysosomal endopeptidase. *Biochemical Journal*. **335**, 111–117.

Chen JM, Rawlings ND, Stevens RAE, Barrett AJ (1998). Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. *FEBS Letters*. **441**, 361-365.

Chen JM, Fortunato M, Barrett AJ (2000). Activation of human prolegumain by cleavage at a C-terminal asparagine residue. *Biochemical Journal*. **352**, 2: 327–334.

Chen JM, Fortunato M, Barrett AJ, Stevens RAE (2001). Activation of pro-gelatinase A by mammalian legumain, a recently discovered cysteine proteinase. *The Journal of Biological Chemistry*. **382**, 777-83.

Chen EI, Kridel SJ, Howard EW, Li W, Godzik A, Smith JW (2002). A unique substrate recognition profile for matrix metalloproteinase-2. *The Journal of Biological Chemistry*. **277**, 6: 4485-91.

Chen YJ, Wu SC, Chen CY, Tzou SC, Cheng TL, Huang YF, Yuan SS, Wang YM (2014). Peptide-based MRI contrast agent and near-infrared fluorescent probe for intratumoral legumain detection. *Biomaterials*. **35**, 304-315.

Chesler NC, Ku DN, Galis ZS (1999). Transmural pressure induces matrix-degrading activity in porcine arteries ex vivo. *American Journal of Physiology Heart and Circulatory Physiology*. **277**, H2002-H2009.

Choi SJ, Reddy SV, Devlin RD, Menaa C, Chung H, Boyce BF, Roodman GD (1999). Identification of human asparaginyl endopeptidase (legumain) as an inhibitor of osteoclast formation and bone resorption. *Journal of Biological Chemistry*. **274**, 39:27747-53.

Choi SJ, Kurihara N, Oba Y, Roodman GD (2001). Osteoclast Inhibitory Peptide 2 Inhibits Osteoclast Formation *via* Its C-Terminal Fragment. *Journal of Bone and Mineral Research*. **16**, 1804-11.

Choi KY, Swierczewska M, Lee S and Chen X (2012). Protease-activated drug development. *Theranostics*. **2**, 2: 156-78.

Choudhury RP, Fisher EA (2009). Molecular imaging in atherosclerosis, thrombosis, and vascular inflammation. *Arteriosclerosis, Thrombosis and Vascular Biology*. **29**, 983–91.

Chu B, Ferguson MS, Chen H, Hippe DS, Kerwin WS, Canton G, Yuan C, Hatsukami TS (2010). MRI Features of the disruption-prone and the disrupted carotid plaque: A pictorial essay. *A Journal of the American College of Cardiology: Cardiovascular Imaging*. **2**, 7: 883-896.

Chyu K-Y, Nilsson J and Shah PK (2011). Immune mechanisms in atherosclerosis and potential for an atherosclerosis vaccine. *Discovery Medicine*. **11**, 60: 403-412.

Cipollone F, Fazia M and Mezzetti A (2005). Novel determinants of plaque instability. *Journal of Thrombosis and Haemostasis*. **3**, 1962-75.

Clerin V, Shih HH, Deng N, Hebert G, Resmini C, Shields KM, Feldman JL, Winkler A, Albert L, Maganti V, Wong A, Paulsen JE, Keith JC Jr, Vlasuk GP, Pittman DD (2008). Expression of the cysteine protease legumain in vascular lesions and functional implications in atherogenesis. *Atherosclerosis*. **201**, 1: 53-66.

Connors TA and Knox RJ (1995). Prodrugs in cancer chemotherapy, *Stem Cells*. **13**, 501-511.

Corti R (2006). Non-invasive imaging of atherosclerotic vessels by MRI for clinical assessment of the effectiveness of therapy. *Pharmacology and Therapeutics*. **110**, 57-50.

Corti R, Fuster V, Badimon JJ, Hutter R, Fayad ZA (2001). New understanding of atherosclerosis (clinically and experimentally) with evolving MRI technology *in vivo*. *Annals of the New York Academy of Sciences*. **947**, 181-195.

Coulden RA, Moss H, Graves MJ, Lomas DJ, Appleton DS, Weissberg PL (2000). High resolution magnetic resonance imaging of atherosclerosis and the response to balloon angioplasty. *Heart*. **83**, 2: 188-191.

Cook JA and Mitchell JB (1989) Viability measurements in mammalian cell systems. *Analytical Biochemistry*. **179**, 245–253.

Creemers EJM, Cleutjens JPM, Smits JFM, Daemen MJAP (2001). Matrix metalloproteinase inhibition after myocardial infraction: a new approach to prevent heart failure? *Journal of American heart association*. **89**, 201-210.

Crouch SP, Kozlowski R, Slater KJ, Fletcher J (1993). The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity. *Journal of Immunological Methods*. **160**, 81–88.

Cullen P, Rauterberg J and Lorkowski S (2005). The pathogenesis of atherosclerosis. *Handbook of Experimental Pharmacology*. **170**, 3-70.

Cybulsky MI and Gimbrone Jr MA (1991). Endothelial expression of a mononuclear leukocyte adhesion molecule during atherogenesis. *Science*. **251**, 788–91.

Dall E and Brandstetter H (2012). Activation of legumain involves proteolytic and conformational events, resulting in a context- and substrate- dependent activity profile. *Acta Crystallographica*. **68**, 24-31.

Dall E and Brandstetter H (2013). Mechanistic and structural studies on legumain explain its zymogenicity, distinct activation pathways, and regulation. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 27: 10940-10945.

Dall E and Brandstetter H (2015). Structure and function of legumain in health and disease. *Biochimie*. **122**, 126-50.

Dando PM, Fortunato M, Smith L, Knight CG, McKendrick JE, Barrett AJ (1999). Pig kidney legumian: an asparaginyl endopeptidase with restricted specificity. *Biochemical Journal*. **339**, 743-749.

Decker T and Lohmann-Matthes ML (1988). A quick and simple method for the quantitation of lactate dehydrogenase release in measurements of cellular cytotoxicity and tumor necrosis factor (TNF) activity. *Journal of Immunological Methods*. **115**, 1: 61-9.

Deguchi JO, Aikawa M, Tung CH, Aikawa E, Kim DE, Ntziachristos V, WeisslederR, Libby P (2006). Inflammation in atherosclerosis: visualizing matrix metalloproteinase action in macrophages in vivo. *Circulation*. **114**, 1: 55-62.

Denis LJ and Verweij J (1997). Matrix metalloproteinase inhibitors: present achievements and future prospects. *Investigational New Drugs*. **15**, 3: 175-185.

Ding Y (2014). Design, synthesis and evaluation of novel endoprotease-activated prodrugs and fluorogenic probes with theranostic applications in chemotherapy PhD Thesis, Edinburgh Napier University.

Doonan (2002). Peptides and proteins. Cambridge, UK: The Royal Society of Chemistry.

Duff GF and McMillan GC (1951). Pathology of atherosclerosis. *American Journal of Medicine*. **11**, 92-108.

Duncan JR, Franano FN, Edward WB, Welch MJE (1994). Evidence of gadolinium dissociation from protein-DTPA-gadolinium complexes. *Investigative Radiology*. **29**, S58–S61.

Dunn BM (2015). Peptide chemistry and drug design. New Jersey, USA. John Wiley and Sons, Inc.

Edgington LE, Verdoes M and Bogyo M (2011). Functional imaging of proteases: recent advances in the design and application of substrate-based and activity-based probes. *Current Opinion in Chemical Biology*. **15**, 798-805.

Edwards D (2008). The cancer degradome: proteases and cancer biology. *Springer*. 840.

Erickson BW and Merrifield RB (1976). "The Proteins". H. Neurath and R. Hill, Eds. Academic Press, New York. 2, 3rd ed., 255-527.

Espelt L, Parella T, Bujons J, Solans C, Joglar J, Delgado A, Clapes P (2003). Stereoselective aldol additions catalyzed by dihydroxyacetone phosphate-dependent aldolases in emulsion systems: preparation and structural characterization of linear and cyclic iminopolyols from aminoaldehydes. Chemistry. *A European Journal*. **9**, 4887-4889.

Fayad ZA and Fuster (2000). Characterization of atherosclerotic plaques by magnetic resonace imaging. *Annals of the New York Academy of Sciences*. **902**, 173-186.

Fayad ZA and Fuster V (2001). Clinical imaging of the high-risk or vulnerable atherosclerotic plaque. *Circulation Research*. **89**, 305-316.

Fayad ZA, Fuster V, Fallon JT, Jayasundera T, Worthley SG, Helft G (2000). Non-invasive in vivo human coronary artery lumen and wall imaging using black-blood magnetic resonance imaging. *Circulation*. **102**, 5: 506-510.

Fernandez-Carneado and Giralt E (2004). An efficient method for the solid-phase synthesis of fluorescently labelled peptides. *Tetrahedron Letters*. **45**, 31: 6079-81.

Fischer R, Mader O, Jung G, Brock R (2003). Extending the applicability of carboxyfluorescein in solid-phase synthesis. *Bioconjugate Chemistry*. **14**, 653-

660.

Flacke S, Fischer S, Scott MJ, Fuhrhop RJ, Allen JS, McLean M, Winter P, Sicard GA, Gaffny PJ, Wickline SA, Lanza GM (2001). Novel MRI contrast agent for molecular imaging of fibrin: implications for detecting vulnerable plaques. *Circulation*. **104**, 1280.

Fotakis G and Timbrell JA (2006). In vitro cytotoxicity assays: comparison of LDH, neutral red, MTT and protein assay in hepatoma cell lines following exposure to cadmium chloride. *Toxicology Letters*. **160**, 2:171-7.

Frias JC, Williams KJ, Fisher EA, Fayad ZA (2004). Recombinant HDL-like nanoparticles: a specific contrast agent for MRI of atherosclerotic plaques. *Journal of the American Chemical Society*. **126**, 50:16316-7.

Friedman M (2004). Applications of the ninhydrin reaction for analysis of amino acids, peptides, and proteins to agricultural and biomedical sciences. *Journal of Agricultural and Food Chemistry*. **52**, 3:385-406.

Frullani L and Meade TJ (2007). Multimodal MRI contrast agents. *Journal of Biology and Inorganic Chemistry*. **12**, 939-949.

Fu X, Parks WC and Heinecke JW (2008). Activation and silencing of matrix metalloproteinases. *Seminars in Cell and Developmental Biology*. **19**, 2-13.

Fuentes-Prior P and Salvesen GS (2004). The protein structures that shape caspase activity, specificity, activation and inhibition. *Biochemical Journal*. **1**, 384 (Pt 2):201-32.

Fuster V, Fayad ZA, Moreno PR, Poon M, Corti R, Badimon JJ (2005). Atherothrombosis and high-risk plaque: Part II: approaches by non-invasive computed tomographic/magnetic resonance imaging. *Journal of American College of Cardiology*. **46**, 7:1209-18.

Gabriel D, Zuluaga MF, van den Bergh H, Gurny R and Lange N (2011). It is all about proteases: from drug delivery to in vivo imaging and photomedicine. *Current Medicinal Chemistry*. **18**, 1785-1805.

Galis ZS, Sukhova GK, Lark MW, Libby P (1994). Increased expression of matrix metalloproteinases and matrix degrading activity in vulnerable regions of human atherosclerotic plaques. *The Journal of Clinical Investigation*. **94**, 6:2493-503.

Galis ZS, Sukhova GK, Kranzhofer R, Clark S, Libby P (1995). Macrophage foam cells from experimental atheroma constitutively produce matrix- degrading proteinases. *Proceedings of the National Academy of Science U S A*. **92**, 402–406.

Galis ZS and Khatri JJ (2002). Matrix metalloproteinases in vascular remodelling and atherogenesis: the good, the bad, and the ugly. *Circulation Research*. **90**, 251–262.

Galis ZS (2004). Vulnerable plaque: the devil is in the details. *Circulation*. **110**, 244–246.

Galkina E and Ley K (2007). Leukocyte influx in atherosclerosis. *Current Drug Targets*. **8**, 1239–48.

Grignon N, Touraine B and Durand M (1989). 6(5)Carboxyfluorescein as a Tracer of Phloem Sap Translocation. *American Journal of Botany*. **76**, 871-877.

Guo H, Shi Y, Liu L, Sun A, Xu F, Chi J (2009). Rosuvastatin inhibits MMP-2 expression and limits the progression of atherosclerosis in LDLR-deficient mice. *Archives of Medical Research*. **40**, 5: 345-51.

Gutte B 1995. Peptides: Synthesis, structures and applications. Revised ed. United States of America: Academic Press, Inc.

Guzman F, Barberis S and Illanes A (2007). Peptide synthesis: chemical or enzymatic. *Journal of Biotechnology*. **10**, 2: 279-314.

Hackam DG and Anand SS (2003). Emerging risk factors for atherosclerotic vascular disease. A critical review of the evidence. *The Journal of the American Medical Association*. **290**, 932-940.

Haidaria M, Wydec PR, Litovsky S, Vela D, Ali M, Casscells SW, Madjid M (2010). Influenza virus directly infects, inflames, and resides in the arteries of atherosclerotic and normal mice. *Atherosclerosis*. **208**, 90-96.

Halfon S, Patel S, Vega F, Zurawski S, Zurawski G (1998). Autocatalytic activation of human legumain at aspartic acid residues. *FEBS Letters*. **438**, 114-118.

Hamid R, Rotshteyn Y, Rabadi L, Parikh R, Bullock P (2004). Comparison of alamar blue and MTT assays for high through-put screening. *Toxicology In Vitro*. **18**, 5: 703-10.

Han SY and Kim YA (2004). Recent development of peptide coupling reagents in organic synthesis. *Tetrahedron*. **60**, 2447-67.

Hansen BH, Moore WGI, Bodden MK, Windsor LJ, DeCarlo A, Engler JA (1993). Matrix metalloproteinase: a review. *Critical Reviews in Oral Biology and Medicine*. **4**, 197-250.

Hansson GK, Robertson AK and Soderberg-Naucler C (2006). Inflammation and atherosclerosis. *Annual Review of Pathology*. **1**, 297-329.

Hara-Nishimura I, Inoue K and Nishimura M (1991). A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. *FEBS Letters*. **294**, 89–93.

Hasegawa BH and Zaidi H (2006). Dual-Modality Imaging: More than the Sum of its Components. *Quantitative Analysis in Nuclear Medicine Imaging*. **35**, 81.

Hashimoto S, Suzuki T, Dong HY, Yamazaki N, Matsushima K (1999). Serial analysis of gene expression in human monocytes and macrophages. *Blood*. **94**, 3: 837-44.

Hayes C, Padhani AR and Leach MO (2002). Assessing changes in tumour vascular function using dynamic contrast-enhanced magnetic resonance imaging. *NMR in Biomedicine*. **15**, 2: 154-63.

Hayward RD, Goguen JD and Leong JM (2010). No better time to FRET: shedding light on host pathogen interaction. *Journal of Biology*. **18**, 9: 2-12.

Hicks B (2009). Management of atherosclerotic symptoms and treatment. [Online]. Available from: www.clivir.com. [Accessed on 19th October 2011].

Hidalgo M and Eckhardt SG (2001). Development of matrix metalloproteinase inhibitors in cancer therapy. *Journal of National Cancer Institute*. **93**, 3:178-193.

Himelstein BP, Canete-Soler R, Bernhard EJ, Dilks DW, Muschel RJ (1994- 1995). Metalloproteinase in tumour progression: the contribution of MMP-9. *Invasion Metastasis*. **14**, 1-6: 246-258.

Homeister JW and Willis MS (2010). Atherosclerosis: pathogenesis, genetics and experimental models. In: eLS. John Wiley & Sons lnc., www.els.net.

Hong YM (2010). Atherosclerotic cardiovascular disease beginning in childhood. *Korean Circulation Journal*. **40**, 1:1-9.

Huang J, Wang L, Lin R, Wang AY, Yang L, Kuang M, Qian W, Mao H (2013). Casein-coated iron oxide nanoparticles for high MRI contrast enhancement and efficient cell targeting. *ACS Applied Materials and Interfaces*. **5**, 11: 4632-9.

Huttunen KM, Raunio H and Rautio J (2011). Prodrugs--from serendipity to rational design. *Pharmacological Reviews*. **63**, 3: 750-71.

Hwang BH, Kim MH and Chnag K (2011). Molecular imaging of high risk atherosclerotic plaques: is it clinically translatable? *Korean Circulation Journal*. **41**, 9: 497-502.

Ibanez B, Badimon JJ and Garcia MF (2009). Diagnosis of atherosclerosis by imaging. *The American Journal of Medicine*. **122**, S15-S25.

Ishizaki T, Erickson A, Kuric E, Shamloo M, Hara-Nishimura I, Inácio AR, Wieloch T, Ruscher K (2010). The asparaginyl endopeptidase legumain after experimental stroke. *Journal of Cerebral Blood Flow Metabolism*. **30**, 10: 1756-66.

Isobe I, Yanagisawa K and Michikawa M (2001). 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) causes Akt phosphorylation and morphological changes in intracellular organelles in cultured rat astrocytes. *Journal of Neurochemistry*. **77**, 274-280.

Jaffer FA, Sosnovik DE, Nahrendorf M, Weissleder R (2006). Molecular imaging of

myocardial infarction. *Journal of Molecular and Cellular Cardiology*. **41**, 6:921-33.

Jaffer FA, Libby P and Weissleder R (2007). Molecular imaging of cardiovascular disease. *Journal of American Heart Society*. **116**, 1052- 1061.

Jamous M, Haberkorn U and Mier W (2013). Synthesis of peptide radiopharmaceuticals for the therapy and diagnosis of tumour diseases. *Molecules*. **18**, 3:3379-409.

Jastrzebska B, Lebel R, Therriault H, McIntyre JO, Escher E, Guérin B, Paquette B, Neugebauer WA, Lepage M (2009). New enzyme-activated solubility-switchable contrast agent for magnetic resonance imaging: from synthesis to in vivo imaging. *Journal of Medicinal Chemistry*. **52**, 6: 1576-81.

Jernigan S and Eddy AA (2000). Experimental insights into the mechanisms of tubulo-interstitial scarring in: Mechanisms and Clinical Management of Chronic Renal Failure (El Nahas A, Harris K, Anderson S, Eds.), pp. 104–145, Oxford University Press, Oxford.

Jurasz ASKD (2003). Molecular imaging in vivo: an introduction. *The British Journal of Radiology*. **76**, s98-s109.

Kaiser E, Colescott RL, Bossiriger CD, CookPI (1970). Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Analytical Biochemistry*. **34**, 595.

Kalab P and Soderholm J (2010). The design of Foster (fluorescence) resonance energy transfer (FRET) - based molecular sensor for Ran GTPase. *Methods*. **51**, 2: 220-232.

Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J (1961). Factors of risk in the development of coronary heart disease-six year follow-up experience: the Framingham study. *Annals of International Medicine: Journal*. **55**, 33-50.

Kates SA, McGuinness BE, Blackburn C, Griffin GW, Sole NA, Barany G, Albericio E (1998). "High-load" polyethylene glycol-polystyrene (PEG-PS) graft supports for solid-phase synthesis. *Biopolymers*. **47**, 365-380.

Kei PL and Chan LP (2008). Gadolinium chelate-associated nephrogenic systemic fibrosis. *Singapore Medical Journal*. **49**, 181-5.

Kent SBH (1988). Chemical synthesis of peptides and proteins. *Annual Reviews in Biochemistry*. **57**, 957-989.

Klinkert MQ, Felleisen R, Link G, Ruppel A, Beck E (1989). Primary structure of Sm31/32 diagnostic proteins of Schistosoma mansoni and their identification as proteases. *Molecular and Biochemical Parasitology*. **33**, 113 –122.

Kembhavi AA, Buttle DJ, Knight CG, Barrett AJ (1993). The two cysteine endopeptidases of legume seeds: purification and characterization by use of

specific fluorometric assays. *Archives of Biochemistry and Biophysics*. **303**, 2: 208-13.

Kim DE, Kim JY, Schellingerhout D, Kim EJ, Kim HK, Lee S, Kim K, Kwon IC, Shon SM, Jeong SW, Im SH, Lee DK, Lee MM, Kim GE (2010). Protease imaging of human atheromata captures molecular information of atherosclerosis, complementing anatomic imaging. *Arteriosclerosis, Thrombosis and Vascular Biology*. **30**, 3: 449-56.

Kooi ME, Cappendijk VC, Cleutjens KB, Kessels AGH, Kitslaar PJEHM, Daemen MJAP, van Engelshoven JMA (2003). Accumulation of ultrasmall paramagnetic particles of iron oxide in human atherosclerotic plaques can be detected by in vivo magnetic resonance imaging. *Circulation*. **107**, 2453-2458.

Kodama S, Asano S, Moriguchi T, Sawai H and Shinozuka K (2006). Novel fluorescent oligoDNA probe bearing a multi-conjugated nucleoside with a fluorophore and a non-fluorescent intercalator as a quencher. *Bioorganic and Medicinal Chemistry Letters*. **16**, 10: 2685-8.

Korngold EC, Jaffer FA, Weissleder R, Sosnovik DE (2008). Non-invasive imaging of apoptosis in cardiovascular disease. *Heart Failure Reviews*. **13**, 163–73.

Kridel SJ, Chen E, Kotra LP, Howard EW, Mobashery S, Smith JW. Substrate hydrolysis by matrix metalloproteinase-9 (2001). *The Journal of Biological Chemistry*. **276**, 23: 20572-8.

Kruper WJ, Rudolf PR and Langhoff CA (1993). Unexpected Selectivity in the Alkylation of Polyazamacrocycles. *The Journal of Organic Chemistry*. **58**, 15: 3869-76.

Kuennemeyer J, Terborg L, Nowak S, Telgmann L, Tokmak F, Kraemer BK, Guensel A, Wiesmue ller GA, Waldeck J, Bremer C, Karst U (2009). Analysis of the Contrast Agent Magnevist and Its Transmetalation Products in Blood Plasma by Capillary Electrophoresis/ Electrospray Ionization Time-of-Flight Mass Spectrometry. *Analytical Chemistry*. **81**, 3600-3601.

Kullo IJ and Ballantyne CM (2005). Conditional risk factors for atherosclerosis. *Mayo Clinic Proceeding*. **80**; 2: 219-230.

Kundra R and Kornfeld S (1999). Asparagine-linked oligosaccharides protect Lamp- 1 and Lamp-2 from intracellular proteolysis. *The Journal of Biological Chemistry*. **274**, 31039-31046.

Kurizaki T, Toi M and Tominaga T (1998). Relationship between matrix metalloproteinase expression and tumor angiogenesis in human breast carcinoma. *Oncology Reports*. **5**, 3: 673-7.

Laikhter A, Behlke M, Walder J, Roberts K, Rose S and Huang L (2004). Anthraquinone quencher dyes, their methods of preparation and use. 20040110308.

Lancelot E, Amirbekian V, Brigger I, Raynaud JSB, Ballet SB, David C, Rousseaux O, Greneur SL, Port M, Lijnen HR, Bruneval P, Michel JB, Ouimet T, Roques B, Amirbekian S, Hyafil F, Vucic E, Aguinaldo LGS, Corot F, Fayad ZA (2008). Evaluation of matrix metalloproteinase in atherosclerosis using a novel noninvasive imaging approach. *Journal of the American Heart Association*. **28**, 425-432.

Lauffer RB (1987). Paramagnetic metal complexes as water proton relaxation agents for NMR imaging: theory and design. *Chemical Reviews*. **87**, 5: 901-927.

Leary T (1943). Experimental atherosclerosis in the rabbit compared with human (coronary) atherosclerosis. *Archives Pathology*. **17**, 453-492.

Lee J and Bogyo M (2010). Development of near-infrared fluorophore (NIRF)-labeled activity-based probes for *in vivo* imaging of legumain. *ACS Chemical Biology*. **5**, 2: 233–243.

Lee J and Bogyo M (2012). Synthesis and evaluation of aza-peptidyl inhibitors of the lysosomal asparaginyl endopeptidase, legumain. *Bioorganic and Medicinal Chemistry Letters*. **22**, 3: 1340-3.

Lepage M, Dow WC, Melchior M, You Y, Fingleton B, Quarles CC, Pépin C, Gore JC, Matrisian LM, McIntyre JO. (2007). Noninvasive detection of matrix metalloproteinase activity in vivo using a novel magnetic resonance imaging contrast agent with a solubility switch. *Molecular Imaging*. **6**, 6: 393-403.

Leuschner F and Nahrendorf M (2011). Molecular imaging of coronary atherosclerosis and myocardial infarction: Considerations for the bench and perspectives for the clinic. *Circulation Research*. **108**, 5: 593-606.

Lewen S, Zhou H, Hu HD, Cheng T, Markowitz D, Reisfeld RA, Xiang R, Luo Y (2008). A legumain based minigene vaccine targets the tumor stroma and suppresses breast cancer growth and angiogenesis. *Cancer Immunology, Immunotherapy*. **57**, 507-515.

Leznoff CC (1978). The Use of Insoluble Polymer Supports in General Organic Synthesis. *Accounts of Chemical Research*. **11**, 327-333.

Li K, Quisling R, Armitage F, Richardson D, Mladinich C (1992). In vivo MR evaluation of Gd-DTPA conjugated to dextran in normal rabbits. *Magnetic Resonance Imaging*. **10**, 3: 439-444.

Li L, Tsai SW, Anderson AL, Keire DA, Raubitschek AA, Shively JE (2002). Vinyl sulfone bifunctional derivatives of DOTA allow sulfhydryl- or amino- directed coupling to antibodies. Conjugates retain immunoreactivity and have similar biodistributions. *Bioconjugate Chemistry*. **13**, 110–115.

Li W, Li Z, Jing F, Deng Y, Wei L, Liao P, Yang X, Li X, Pei F, Wang X, Lei H (2008). Synthesis and evaluation of Gd-DTPA labelled arabinogalactans as potential MRI contrast agent. *Carbohydrate Research*. **343**, 685-694.

Li ZB, Niu G, Wang H, He L, Yang l, Ploug M, Chen X (2008). Imaging of Urokinase-Type Plasminogen Activator Receptor Expression Using a 64Cu-Labeled Linear Peptide Antagonist by microPET. *Clinical Cancer Research*. **14**, 4758-66.

Libby P (2002). Inflammation in atherosclerosis. *Nature*. **420**, 6917: 868-74.

Libby P, DiCarli M and Weissleder (2010). The vascular biology of atherosclerosis and imaging targets. *The Journal of Nuclear Medicine*. **51**, 33-37.

Lim NH, Kashiwagi M, Visse R, Jones J, Enghild JJ, Brew K, Nagase H (2010). Reactive-site mutants of N-TIMP-3 that selectively inhibit ADAMTS-4 and ADAMTS-5: biological and structural implications. *Biochemical Journal*. **431**, 1: 113-22.

Ling H, Ardjomand P, Samvakas S, Simm A, Busch GL, Lang F, Sebekova K, Heidland A (1998). Mesangial cell hypertrophy induced by NH_4Cl: role of depressed activities of cathepsins due to elevated lysosomal pH. *Kidney International*. **53**, 1706- 1712.

Lipinski MJ, Amirbekian V, Frias JC, Aguinaldo JG, Mani V, Briley-Saebo KC, Fuster V, Fallon JT, Fisher EA, Fayad ZA (2006). MRI to detect atherosclerosis with gadolinium-containing immunomicelles targeting the macrophage scavenger receptor. *Magnetic Resonance in Medicine*. **56**, 3: 601-610.

Liu C, Sun C, Huang H, Janda K, Edgington T (2003). Overexpression of legumain in tumors is significant for invasion/metastasis and a candidate enzymatic target for prodrug therapy. *Cancer Research*. **63**, 2957–2964.

Liu J, Sukhova GK, Yang JT, Sun J, Ma L, Ren A, Xu WH, Fu H, Dolganov GM, Hu C, Libby P, Shi GP (2006). Cathepsin L expression and regulation in human abdominal aortic aneurysm, atherosclerosis, and vascular cells. *Atherosclerosis*. **184**, 2: 302-11.

Liu T, Qian Y, Hu X, Ge Z, Liu S (2012). Mixed polymeric micelles as multifunctional scaffold for combined magnetic resonance imaging contrast enhancement and targeted chemotherapeutic drug delivery. *Journal of Materials Chemistry*. **22**, 5020-30.

Liu VH, Vassiliou CC, Imaad SM, Cima MJ (2014). Solid MRI contrast agents for long-term, quantitative in vivo oxygen sensing. *Proceedings of the National Academy of Sciences of the United States of America*. **111**, 18: 6588-93.

Loftus IM, Naylor AR, Goodall S, Crowther M, Jones L, Bell PRF, Thompson MM (2000). Increased matrix metalloproteinase-9 activity in unstable carotid plaques: a potential role in acute plaque disruption. *Stroke*. **31**, 40–47.

Loncin MF, Desreux JF and Merciny E (1986). Coordination of lanthanides by two

polyamino polycarboxylic macrocycles: formation of highly stable lanthanide complexes. *American Chemical Society*. **25**, 2646-2648.

López-Otín C and Bond JS (2008). Proteases: multifunctional enzymes in life and disease. *The Journal of Biological Chemistry*. **283**, 45: 30433-7.

Lorkowski S and Cullen P (2007). Atherosclerosis: pathogenesis, clinical features and treatment. *Encyclopaedia of Life Science*. **1**, 11.

Lozito TP, Jackson WM, Nesti LJ, Tuan RS (2014). Human mesenchymal stem cells generate a distinct pericellular zone of MMP activities via binding of MMPs and secretion of high levels of TIMPs. *Matrix Biology*. **34**, 132-143.

Luo Y, Zhou H, Krueger J, Kaplan C, Lee SH, Dolman C, Markowitz D, Wu W, Liu C, Reisfeld RA, Xiang R (2006). Targeting tumor-associated macrophages as a novel strategy against breast cancer. *Journal of Clinical Investigation*. **116**, 2132–2141.

Maehr R, Hang HC, Mintern JD, Kim YM, Cuvillier A, Nishimura M, Yamada K, Shirahama-Noda K, Hara-Nishimura I, Ploegh HL (2005). Asparagine endopeptidase is not essential for class II MHC antigen presentation but is required for processing of cathepsin L in mice. *Journal of Immunology*. **174**, 7066–7074.

Manoury B, Hewitt EW, Morrice N, Dando PM, Barrett AJ, Watts C (1998). An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature*. **396**, 6712: 695-9.

Marras SAE, Kramer FR and Tyagi S (2002). Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acid Research*. **30**, 21: 122.

Massova I, Fridman R and Mobashery S (1997) Structural insights into the catalytic domains of human matrix metalloprotease-2 and human matrix metalloprotease-9: implications for substrate specificities. *Journal of Molecular Modeling*. **3**, 17-30.

Mathieu MA, Bogyo M, Caffrey CR, Choe Y, Lee J, Chapman H, Sajid M, Craik CS, McKerrow JH (2002). Substrate specificity of schistosome versus human legumain determined by P1-P3 peptide libraries. *Molecular and Biochemical Parasitology*. **121**, 1: 99-105.

Mattock KL, Gough PJ, Humphries J, Burnand K, Patel L, Suckling KE, Cuello F, Watts C, Gautel M, Avkiran M, Smith A (2010). Legumain and cathepsin-L expression in human unstable carotid plaque. *Atherosclerosis*. **208**, 1: 83-9.

May JP, Brown LJ, van Delft I, Thelwell N, Harley K, Brown T (2005). Synthesis and evaluation of a new non-fluorescent quencher in fluorogenic oligonucleotide probes for real-time PCR. *Organic and Biomolecular Chemistry*. **3**, 14: 2534-42.

McAteer MA, Akhtar AM, Muhlen C, Choudhury RP (2010). An approach to molecular imaging of atherosclerosis, thrombosis and vascular inflammation using

microparticles of iron oxide. *Atherosclerosis*. **209**, 18-27.

McCawley LJ and Matrisian LM (2001). Matrix metalloproteinase: they're not just for matrix anymore. *Current Opinion in Cell Biology*. 13, 5: 534-540.

McGeehan GM, Bickett DM, Green M, Kassel D, Wiseman JS, Berman J (1994). Characterization of the peptide substrate specificities of interstitial collagenase and 92-kDa gelatinase. *The Journal of Biological Chemistry*. **269**, 52: 32814-32820.

McMurry TJ, Brechbiel M, Kumar K, Gansow OA (1992). Convenient synthesis of bifunctional tetraaza macrocycles. *Bioconjugate Chemistry*. **3**, 2:108-17.

Meade TJ and Allen M J (2004). *Functional MRI agents for cancer imaging*. [Online]. PCT Int. App. WO Patent No. 2004037210. Available from: MedicineNet.com.http://www.medicinenet.com/peripheral_vascular_disease/artic le.htm. [Accessed 22 November 2010].

Meding J, Urich M, Licha K, Reinhardt M, Misselwitz B, Fayad ZA, Weinmann HJ (2007).Magnetic resonance imaging of atherosclerosis by targeting extracellular matrix deposition with Gadofluorine M. *Contrast Media and Molecular Imaging*. **2**, 3: 120-9.

Medintz IL, Clapp AR, Mattoussi H, Goldman ER, Fisher B, Mauro JM (2003). Self-assembled nanoscale biosensors based on quantum dot FRET donors. *Natural Materials*. **2**, 9: 630-8.

Meier CR, Derby LE, Jick SS, Vasilakis C, Jick H (1999). Antibiotics and risk of subsequent first-time acute myocardial infarction. *The Journal of the American Medical Association*. **281**, 5: 427-31.

Merrifield RB (1963). Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society*. **85**, 14: 2149-2154.

Merrifield RB (1963). Solid phase peptide synthesis I: the synthesis of a tetrapeptide. *Journal of American Chemical Society*. **85**, 2149-2154.

Miranda LP and Alewood PF (2000). Challenges for protein chemical synthesis in the 21st century: Bridging genomics and proteomics. *Biopolymers*. **55**, 217-226.

Mishra A, Pfeuffer J, Mishra R, Engelmann J, Mishra AK, Ugurbil K, Logothetis NK. (2006). A new class of Gd-based DO3A-ethylamine-derived targeted contrast agents for MR and optical imaging. *Bioconjugate Chemistry*. **17**, 3: 773-80.

Mohs AM, Nguyen T, Jeong EK, Feng Y, Emerson L, Zong Y, Parker DL, Lu ZR (2007). Modification of Gd-DTPA cystine copolymers with PEG-1000 optimizes pharmacokinetics and tissue retention for magnetic resonance angiography. *Magnetic Resonance in Medicine*. **58**, 1: 110-8.

Mordon S, Maunoury V, Devoisselle JM, Abbas Y, Coustaud D (1992). Characterization of tumorous and normal tissue using a pH-sensitive fluorescence indicator (5(6)-carboxyfluorescein) in vivo. *Journal of Photochemistry and*

Photobiology. **13**, 3-4: 307-14.

Mordon S, Devoisselle JM and Maunoury V. In vivo pH measurement and imaging of tumor tissue using a pH-sensitive fluorescent probe (5,6-carboxyfluorescein): instrumental and experimental studies (1994). *Photochemistry and Photobiology*. **60**, 3: 274-9.

Moreno PR, Falk E, Palacios IF, Newell JB, Fuster V, Fallon JT (1994). Macrophage infiltration in acute coronary syndromes. Implications for plaque rupture. *Circulation*. **90**, 2: 775-8.

Morgan DC, Mills CK, Lefkowitz LD, Lefkowitz SS (1991). An improved colorimetric assay for tumor necrosis factor using WEHI 164 cells cultured on novel microtiter plates. *Journal of Immunological Methods*. **145**, 259–262.

Morita Y, Araki H, Sugimoto T, Takeuchi K, Yamane T, Maeda T, Yamamoto Y, Nishi K, Asano M, Shirahama-Noda K, Nishimura M, Uzu T, Hara-Nishimura I, Koya D, Kashiwagi A, Ohkubo I (2007). Legumain/asparaginyl endopeptidase controls extracellular matrix remodeling through the degradation of fibronectin in mouse renal proximal tubular cells. *FEBS Letters*. **581**, 1417–1424.

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.

Mulder WJ, Strijkers GJ, Habets JW, Bleeker EJ, van der Schaft DW, Storm G, Koning GA, Griffioen AW, Nicolay K (2005). MR molecular imaging and fluorescence microscopy for identification of activated tumor endothelium using a bimodal lipidic nanoparticle. *FASEB Journal*. **19**, 14: 2008-10.

Muntz K, Blattner1 FR, Shutov AD (2002). Legumain – a family of asparagine-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. *Journal of Plant Physiology*. **159**, 1281–1293 (2002).

Murphy G and Nagase H (2008). Progress in matrix metalloproteinase research. *Molecular Aspects of Medicine*. **29**, 290-308.

Nagase H (1997). Activation mechanisms of matrix metalloproteinases. *Biological Chemistry*. **378**, 3-4: 151-60.

Nagase H and Woessner JF Jr (1999). Matrix metalloproteinase. *The Journal of Biological Chemistry*. **274**, 31: 21491-21949.

Naghavi M, John R, Naguib S, Siadaty MS, Grasu R, Kurian KC, van Winkle WB, Soller B, Litovsky S, Madjid M, Willerson JT, Casscells W (2002). pH Heterogeneity of human and rabbit atherosclerotic plaques; a new insight into detection of vulnerable plaque. *Atherosclerosis*, **164**, 1:27-35.

Nakayama H, Arakaki A, Maruyama K, Takeyama H, Matsunaga T (2003). Single-nucleotide polymorphism analysis using fluorescence resonance energy transfer between DNA-labeling fluorophore, fluorescein isothiocyanate, and DNA

intercalator, POPO-3, on bacterial magnetic particles. *Biotechnology and Bioengineering*. **84**, 1: 96-102.

Neefjes J and Dantuma NP (2004). Fluorescent probes for proteolysis: tools for drug discovery. *Natural Review Drug Discovery*. **3**, 58-69.

Nelson CH (2012). Proteolytic profiling with imaging agents for rational design of target-activated peptide prodrugs. PhD Thesis. The University of Michigan.

Newby AC (2005). Dual role of matrix metalloproteinases (matrixins) in intimal thickening and atherosclerotic plaque rupture. *Physiological Reviews*. **85**, 1- 31.

Nguyen M, Arkell J and Jackson CJ (2001). Human endothelial gelatinases and angiogenesis. *The International Journal of Biochemistry and Cell Biology*. **33**, 10: 960- 970.

Nikkari ST, O'Brien KD, Ferguson M, Hatsukami T, Welgus HG, Alpers CE, Clowes AW (1995). Interstitial collagenase (MMP-1) expression in human carotid atherosclerosis. *Circulation*. **92**, 1393–1398.

Nicoletti A, Kaveri S, Caligiuri G, Bariety J, Hansson GK (1998). Immunoglobulin treatment reduces atherosclerosis in apo E knockout mice. *Journal of Clinical Investigation*. **102**, 910-918.

Nunn AD, Linder KE and Tweedle MF (1997). Can receptors be imaged with MRI agents? *Quarterly Journal of Nuclear Medicine*. **41**, 2:155-62.

Nwe K, Bernardo M, Regino CAS, Williams M, Brechbiel MW (2010). Comparison of MRI properties between derivatized DTPA and DOTA gadolinium-dendrimer conjugates. *Bioorganic and Medicinal Chemistry*. **18**, 5925-5931.

Nwe K, Milenic D, Bryant LH, Regino CA, Brechbiel MW (2011). Preparation, characterization and in vivo assessment of Gd-albumin and Gd-dendrimer conjugates as intravascular contrast-enhancing agents for MRI. *Journal of Inorganic Biochemistry*. **105**, 5:722-727.

Nyfeler R (1994). Peptide synthesis via fragment condensation. *Methods in Molecular Biology*. **35**: 303-16.

O'Brien ER, Garvin MR, Dev R, Stewart DK, Hinohara T, Simpson JB Schwartz SM (2004) Angiogenesis in human coronary atherosclerotic plaques. *American Journal of Pathology*. **145**: 883–894.

Ostergaard L, Hochberg FH, Rabinov JD, Sorensen AG, Lev M, Kim L, Weisskoff RM,Gonzalez RG,Gyldensted C, Rosen BR (1999). Early changes measured by magnetic resonance imaging in cerebral blood flow, blood volume, and blood-brain barrier permeability following dexamethasone treatment in patients with brain tumors. *Journal of Neurosurgery*. **90**: 2; 300-305.

Overall CM and Kleifeld O (2006). Validating matrix metalloproteinase as drug targets and anti-targets for cancer therapy. *Nature Reviews Cancer*. **6**, 227- 239.

Ozawa T and Umezawa Y (2002). Peptide assemblies in living cells. Methods for detecting protein-protein interactions. *Journal of Supramolecular Chemistry*. **14**, 271.

Papaspyridonos M, Smith A, Burnand KG, Taylor P, Padayachee S, Suckling KE, James CH, Greaves DR, Patel L (2006). Novel candidate genes in unstable areas of human atherosclerotic plaques. *Arteriosclerosis, Thrombosis and Vascular Biology*. **26**, 8:1837-44.

Parizel PM, Degryse HR, Gheuens J, Martin JJ, Van Vyve M, De La Porte C, Selosse P, Van de Heyning P, De Schepper AM (1989). Gadolinium-DOTA enhanced MR imaging of intracranial lesions. *Journal of Computer Assisted Tomography*. **13**, 3: 378-85.

Perez JM, Simeone FJ, Saeki Y, Joseohson L, Weissleder R (2003). Viral-induced self-assembly of magnetic nanoparticles allows the detection of viral particles in biological medial. *Journal of the American Chemical Society*. **125**, 10192-10193.

Pople JA, Schneider WG and Bernstein HJ (1959). High-resolution nuclear magnetic resonance. New York: McGraw-Hill: 209.

Pozzolini M, Scarfi S, Benatti U, Giovine M (2003). Interference in MTT cell viability assay in activated macrophage cell line. *Analytical Biochemistry*. **313**, 2: 338-41.

Puthenedam M (2010). *Fluorescence + MR imaging probe can guide cancer surgery*. [Online]. Available from: http://www.Healthimaging.com/topics/ molecular-imaging/pnas-fluorescence-mr-imaging-probe-can-guide-cancer-surgery. [Accessed on 20 September 2014].

Quillard T, Croce K, Jaffer FA, Weissleder R, Libby P (2011). Molecular imaging of macrophage protease activity in cardiovascular inflammation in vivo. *Thrombosis and Haemostasis*. **105**, 828-836.

Rajoria S, Suriano R, George A, Shanmugam A, Schantz SP Geliebter J, Tiwari RK (2011). Estrogen induced metastatic modulators MMP-2 and MMP-9 are targets of 3,3'-Diindolylmethane in thyroid cancer. *Plos One*. **6**, 1: 1-11.

Rautio J, Kumpulainen H, Heimbach T, Oliyai R, Oh D, Järvinen T, Savolainen J (2008). Prodrugs: design and clinical applications. *Nature Reviews Drug Discovery*. **7**, 3: 255-70.

Rawlings ND and Barrett AJ (1999). MEROPS: the peptidase database. *Nucleic Acids Research*. **27**, 325-331.

Rawlings ND, Barrett AJ and Bateman A (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research.* **40**, D343-D350.

Raymond N and Pierre VC (2005). Next generation, high relaxivity gadolinium MRI agents. *Bioconjugate Chemistry*. **16**, 3-8.

Ross R (1993). The pathogenesis of atherosclerosis: A perspective for the 1990s. *Nature*. **362**, 801.

Ross R (1999). Atherosclerosis—an inflammatory disease. *The New England Journal of Medicine*. **340**, 115-126.

Sadeghi MM, Glover DK, Lanza GM, Fayad ZA, Johnson LL (2010). Imaging atherosclerosis and vulnerable plaque. *The Journal of Nuclear Medicine*. **51**, 51-65.

Sandros MG, Sarraf CB and Tabrizian M (2008). Prodrugs in cardiovascular therapy. *Molecules*. **13**, 5: 1156-78.

Sands MJ and Levitin A (2004). Basics of magnetic resonance imaging. *Seminars in Vascular Surgery*. **17**, 66-82.

Sapsford KE, Berti L and Medintz IL (2006). Materials for fluorescence resonance energy transfer analysis: Beyond traditional donor-acceptor combinations. *Angewandte Chemie International Edition*. **45**, 4562-88.

Saskia CA de Jager and Kuiper J (2011). Vaccination strategies in atherosclerosis. *Thrombosis and Haemostasis*. **106**, 5: 796-803.

Schmid JA and Sitte HH (2003). Fluorescence resonance energy transfer in the study of cancer pathway. *Current Opinion in Oncology*. **15**, 55-64.

Schonbeck U, Mach F, Sukhova GK, Atkinson E, Levesque E, Herman M, Graber P, Basset P, Libby P (1999). Expression of stromelysin-3 in atherosclerotic lesions: regulation via CD40-CD40 ligand signalling in vitro and in vivo. *The Journal of Experimental Medicine*. **189**, 843-853.

Schwarz G, Brandenburg J, Reich M, Burster T, Driessen C and Kalbacher H. Characterization of legumain. *The Journal of Biological Chemistry*. **383**, 11: 1813-6.

Scott MP, Jung R, Müntz K, Nielsen NC (1992). A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major storage protein of soybean. *Proceedings of the National Academy of Sciences USA*. **89**, 658–662.

Selvin PR (2000). The renaissance of fluorescence resonance energy transfer. *Natural structural and Molecular Biology*. **7**, 9: 730-4.

Sexton KB, Wittea MD, Bluma G, Bogyoa M (2007). Design of cell permeable, fluorescent activity based probes for the lysosomal cysteine protease asparaginyl endopeptidase (AEP)/ legumain. *Bioorganic and Medicinal Chemistry Letters*. **17**, 3, 649–653.

Shah PK (2003). Mechanisms of plaque vulnerability and rupture. *Journal of the American College of Cardiology*. **41**, 15s-22.

Sheppeck JE, Kar H and Hong H (2000). A convenient and scalable procedure for removing the Fmoc group in solution. *Tetrahedron Letters*. **41**, 28: 5329-33.

Shirahama-Noda K, Yamamoto A, Sugihara K, Hashimoto N, Asano M, Nishimura M, Hara-Nishimura I (2003). Biosynthetic processing of cathepsins and lysosomal degradation are abolished in asparaginyl endopeptidase-deficient mice. *The Journal of Biological Chemistry*. **278**, 33194–33199.

Shiraishi K, Kawano K, Maitani Y, Yokoyama M (2010). Polyion complex micelle MRI contrast agents from poly(ethylene glycol)-b-poly(l-lysine) block copolymers having Gd-DOTA; preparations and their control of T(1)-relaxivities and blood circulation characteristics. *Journal of Controlled Release*. **148**, 2:160-7.

Singer RM 2012. *A review of gadolinium based contrast agents in magnetic resonance imaging*. [Online]. Available from: CEwebsourse.com. [Accessed 14 October 2012].

Sirol M, Itskovich VV, Mani V, Aguinaldo JG, Fallon JT, Misselwitz B, Weinmann HJ, Fuster V, Toussaint JF, Fayad ZA (2004). Lipid-rich atherosclerotic plaques detected by gadofluorine-enhanced *in vivo* magnetic resonance imaging. *Circulation*. **109**, 23: 2890-96.

Smith R, Johansen HT, Nilsen H, Haugen MH, Pettersen SJ, Maelandsmo GM, Abrahamson M, Solberg R (2012). Intra and extracellular regulation of activity and processing of legumain by cystatin E/M. *Biochimie*. **94**, 12:2590-99.

Smith RL, Astrand OAH, Nguyen L, Elvestrand T, Hagelin G, Solberg R, Johansen HT and Rongved P (2014). Synthesis of a novel legumain-cleavable colchicine prodrug with cell-specific toxicity. *Bioorganic and Medicinal Chemistry*. **22**, 13: 3309-15.

Smith RL (2014). Regulation of legumain in human cells. PhD. University of Oslo.

Sosnovik DE and Caravan P (2009). Molecular MRI of atherosclerotic plaque with targeted contrast agents. *Current Cardiovascular Imaging Reports*. **2**, 2: 87-94.

Spanoghe M, Lanens D, Dommisse R, Van der Linden A, Alderweireldt F (1992). Proton relaxation enhancement by means of serum albumin and poly-L-lysine labeled with DTPA-Gd3+: relaxivities as a function of molecular weight and conjugation efficiency. *Magnetic Resonance Imaging*. **10**, 6: 913-7.

Spuentrup E, Botnar RM, Wiethoff AJ, Ibrahim T, Kelle S, Katoh M, Ozgun M, Nagel E, Vymazal J, Graham PB, Günther RW, Maintz D (2008). MR imaging of thrombi using EP-2104R, a fibrin-specific contrast agent: initial results in patients. *European Radiology*. **18**, 9: 1995-2005.

Stawikowski M and Fields G B (2002). Introduction to Peptide Synthesis. *Current Protocols in Protein Science.* CHAPTER: Unit–18.1.

Steinberg D (2005). An interpretive history of the cholesterol controversy, part III: mechanistically defining the role of hyperlipidemia. *Journal of Lipid Research*. **46**, 2037-2051.

Stern L, Perry R, Ofek P, Many A, Shabat D, Satchi-Fainaro R (2009). A Novel Antitumor Prodrug Platform Designed to Be Cleaved by the Endoprotease Legumain. *Bioconjugate Chemistry*. **20**, 500–510.

Stierandová A, Sepetov NF, Nikiforovich GV, Lebl M (1994). Sequence-dependent modification of Trp by the Pmc protecting group of Arg during TFA deprotection. *International Journal of Peptide Protein Research*. **43**, 1: 31-8.

Strijkers GJ, Mulder WJ, van Tilborg GA, Nicolay K (2007). MRI contrast agents: current status and future perspectives. *Anticancer Agents in Medicinal Chemistry*. **7**, 291-305.

Tate CR, Rhodes LV, Segar HC, Driver JL, Pounder FN, Burow ME, Collins-Burow BM (2012). Targeting triple-negative breast cancer cells with the histone deacetylase inhibitor panobinostat. *Breast Cancer Research*. **14**, 3: R79.

The Eurowinter Group (1997). Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. *Lancet*. **349**, 1341–1346.

Thompson RW, Holmes DR, Mertens RA, Liao S, Botney MD, Mecham RP, Welgus HG, Parks WC (1995). Production and localisation of 92-kD gelatinase in abdominal aortic aneurysms. *The Journal of Clinical Investigation*. **96**, 318 –326.

Tolfrey K (2002). Intraindividual variability of children's blood, lipid and lipoprotein concentrations: a review. *Preventive Cardiology*. **3**, 145-151.

Touryan LA, Baneyx G, Vogel V (2009). Exploiting fluorescence resonance energy transfer to probe structural changes in a macromolecule during adsorption and incorporation into a growing biomineral crystal. *Colloids and Surface B: Biointerfaces*. **74**, 2: 401-9.

Tsuji M, Ueda S, Hirayama T, Okuda K, Sakaquchi Y, Isono A, Nagasawa H (2013). FRET- based imaging of transbilayer movement of pepducin in living cells by novel intracellular bio reductively activatable fluorescent probes. *Organic Bimolecular Chemistry*. **11**, 3030-37.

Tung CH (2004). Fluorescent peptide probes for in vivo diagnostic imaging. *Biopolymers*. **76**, 5:391-403.

Turnbull, A. (2003). Design and development of novel DNA topoisomerase inhibitors. PhD, Edinburgh Napier University.

Van Valckenborgh E, Mincher D, Di Salvo A, Van Riet I, Young L, Van Camp B, Vanderkerken K (2005). Targeting an MMP-9-activated prodrug to multiple myeloma-diseased bone marrow: a proof of principle in the 5T33MM mouse model. *Leukemia*. **19**, 9: 1628-33.

Van Wart HE and Birkedal-Hansen H (1990). The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire

matrix metalloproteinase gene family. *Proceedings of the National Academy of Science USA*. **87**, 14: 5578-82.

Virmani R, Kolodgie FD, Burke AP, Farb A, Schwartz SM (2000). Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. *Arteriosclerosis Thrombosis and Vascular Biology*. **20**, 5: 1262-75.

Wagner S, Breyholz HJ, Faust A, Höltke C, Levkau B, Schober O, Schäfers M, Kopka K (2006). Molecular imaging of matrix metalloproteinases in vivo using small molecule inhibitors for SPECT and PET. *Current Medicinal Chemistry*.**13**, 23: 2819-38.

Wagsater D, Zhu C, Bjorkegren J, Skogsberg J, Eriksson P (2011). MMP-2 and MMP-9 are prominent matrix metalloproteinases during atherosclerosis development in the Ldlr-/-Apob100/100 mouse. *International Journal of Molecular Medicine*. **28**, 2: 247-253.

Walter NG and Burke JM (1997). Real-time monitoring of hairpin ribozyme kinetics through base-specific quenching of fluorescein-labeled substrates. *RNA*. **3**, 4: 392-404.

Wang X, Jin T, Comblin V, Lopez-Mut A, Merciny E, Desreux JF (1992). A kinetic investigation of the lanthanide DOTA chelates. Stability and rates of formation and of dissociation of a macrocyclic gadolinium(III) polyaza polycarboxylic MRI contrast agent. *Inorganic Chemistry*. **31**, 1095-1099.

Webb JA, Thomsen HS and Morcos SK; Members of Contrast Media Safety Committee of European Society of Urogenital Radiology (ESUR) (2005). The use of iodinated and gadolinium contrast media during pregnancy and lactation. *European Radiology*. **15**, 6: 1234-40.

Weber PJ, Bader JE, Folkers G, Beck-Sickinger AG (1998). A fast and inexpensive method for N-terminal fluorescein labelling of peptides. *Bioorganic and Medicinal Chemistry Letters*. **8**, 6:597-600.

Weissleder R and Mahmood U (2001). Molecular imaging. *Radiology*. **219**, 2: 316-333.

Welgus HG, Campbell EJ, Curry JD, Eisen AZ, Senior RM, Wilhelm SM, Golberg GI (1990). Neutral metalloproteinases produced by human mononuclear phagocytes: enzyme profile, regulation and cellular differentiation. *Journal of Clinical Investigation*. **86**, 1496 –1502.

Westermarck J and Kahari VM (1999). Regulation of matrix metalloproteinase expression in tumour invasion. *The FASEB Journal*. **13**, 781- 792.

Weyermann J, Lochmann D and Zimmer A (2005). A practical note on the use of cytotoxicity assays. *International Journal of Pharmaceutics*. **288**, 2: 369-76.

Widengren J (2010). Fluorescence-based transient state monitoring for biomolecular spectroscopy and imaging. *Journal of the Royal Society Interface*. **7**, 49:1135-44.

Wu W, Luo Y, Sun C, Liu Y, Kuo P, Varga J, Xiang R, Reisfeld R, Janda KD, Edgington TS, Liu C (2006). Targeting cell-impermeable prodrug activation to tumor microenvironment eradicates multiple drug-resistant neoplasms. *The Journal of Cancer Research*. **66**, 2:970-80.

Xu X, Chen Z, Wang Y, Yamada Y, Steffensen B (2005). Functional basis for the overlap in ligand interactions and substrate specificities of matrix metalloproteinases-9 and -2. *Biochemical Journal*. **392**, 1: 127-34.

Yanagi H, Sasaguri Y, Sugama K, Morimatsu M, Nagasi H (1992). Production of tissue collagenase (matrix metalloproteinase 1) by human aortic smooth muscle cells in response to platelet- derived growth factor. *Atherosclerosis*. **91**, 207-216.

Yang CT and Chuang KH (2012). Gd(III) chelates for MRI contrast agents: from high relaxivity to "smart", from blood pool to blood–brain barrier permeable. *Medicinal Chemistry Communications*. **3**, 552.

Yang CT, Chandrasekharan P, He T, Poh Z, Raju A, Chuang KH, Robins EG (2014). An intravascular MRI contrast agent based on Gd(DO3A-Lys) for tumor angiography. *Biomaterials*. **35**, 1: 327-36.

Yuan C, Mitsumori LM, Beach KW, Maravilla KR (2001). Carotid atherosclerotic plaque: non-invasive MR characterization and identification of vulnerable lesions. *Radiology*. **221**, 2: 285-299.

Yuan C, Wang J, Balu N (2012). High-field atherosclerotic plaque magnetic resonance imaging. *Neuroimaging Clinics of North America*. **22**, 2: 271-84.

Zaccolo M (2004). Use of chimeric fluorescent proteins and fluorescence resonance energy transfer to monitor cellular response. *Circulation Research*. **16**, 94: 866-73.

Zhang S, Merritt M, Woessner DE, Lenkinski RE, Sherry AD (2003). PARACEST agents: modulating MRI contrast via water proton exchange. *Accounts of Chemical Research*. **36**, 10: 783-90.

Zitka O, Kukacka J, Krizkova S, Huska D, Adam V, Masarik M, Prusa R, Kizek R (2010). Matrix metalloproteinases. *Current Medicinal Chemistry*. **17**, 31:3751-68.

Zucker S and Cao J (2001). Imaging metalloproteinase activity *in vivo*. *Nature Medicine*. **7**, 6: 655-6.

Zuo X, Xia F, Xiao Y, Plaxco KW (2010). Sensitive and selective amplified fluorescence DNA detection based on exonuclease III-aided target recycling. *Journal of the American Chemical Society*. **132**, 6: 1816-8.

5. Appendix

Appendix 1
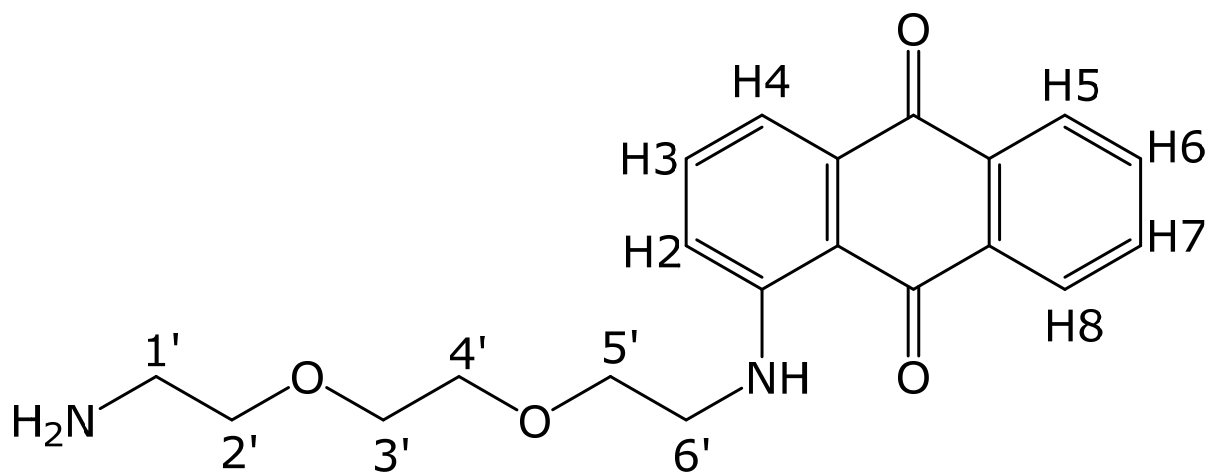
Compound (5)
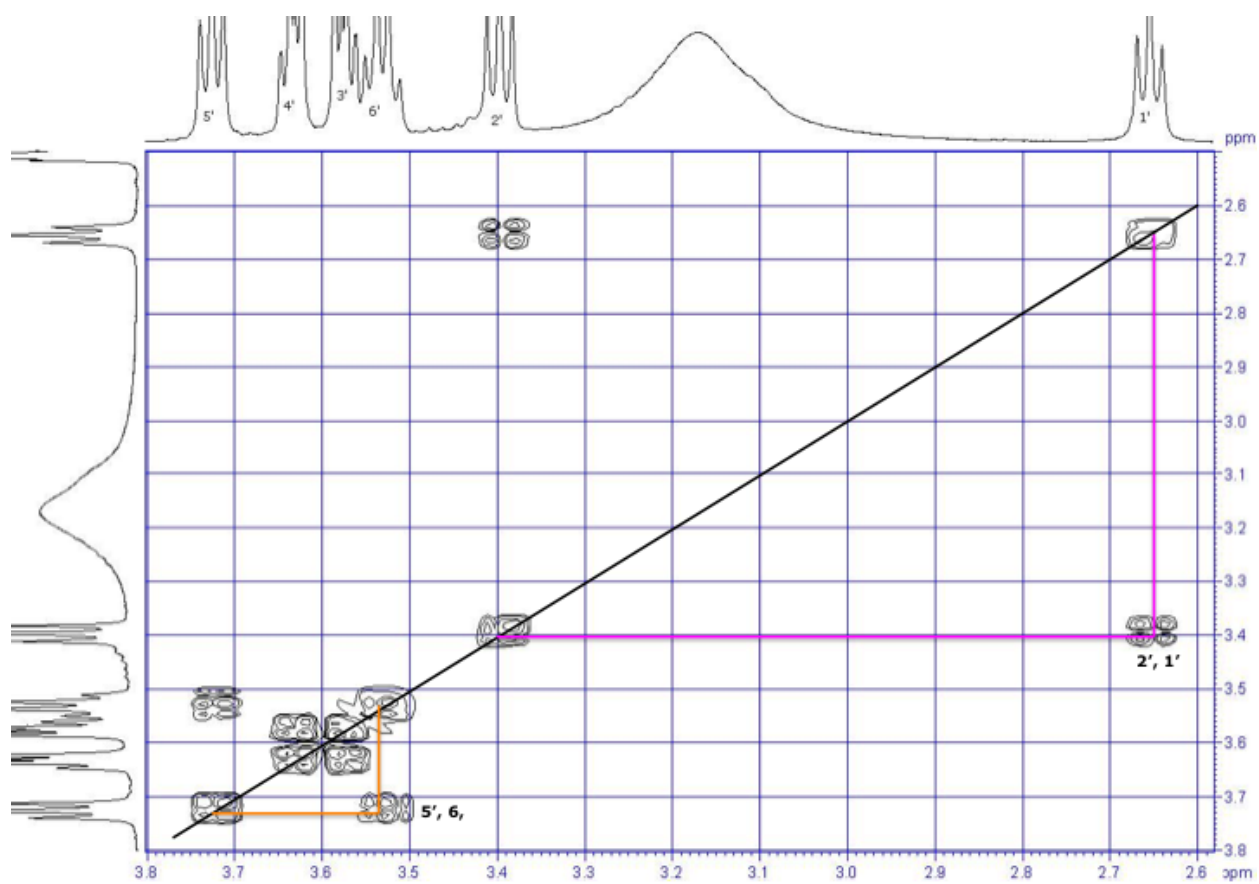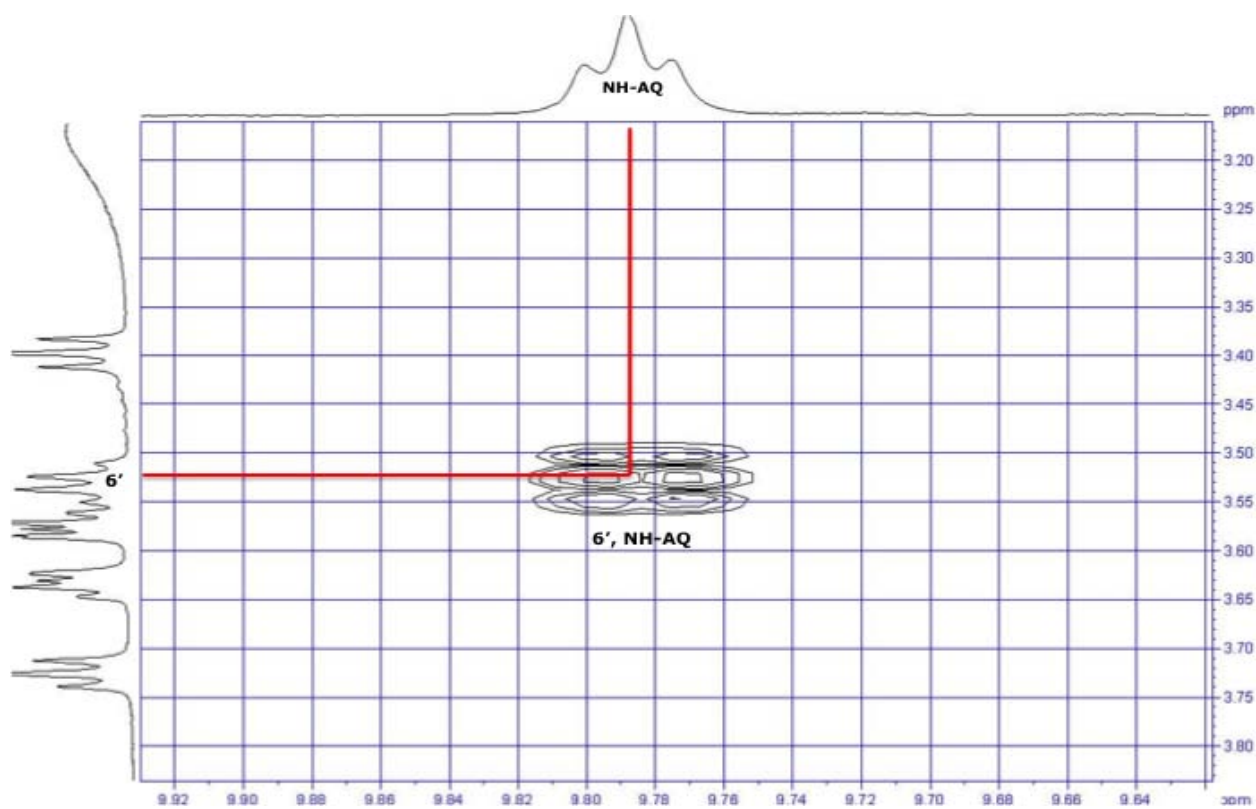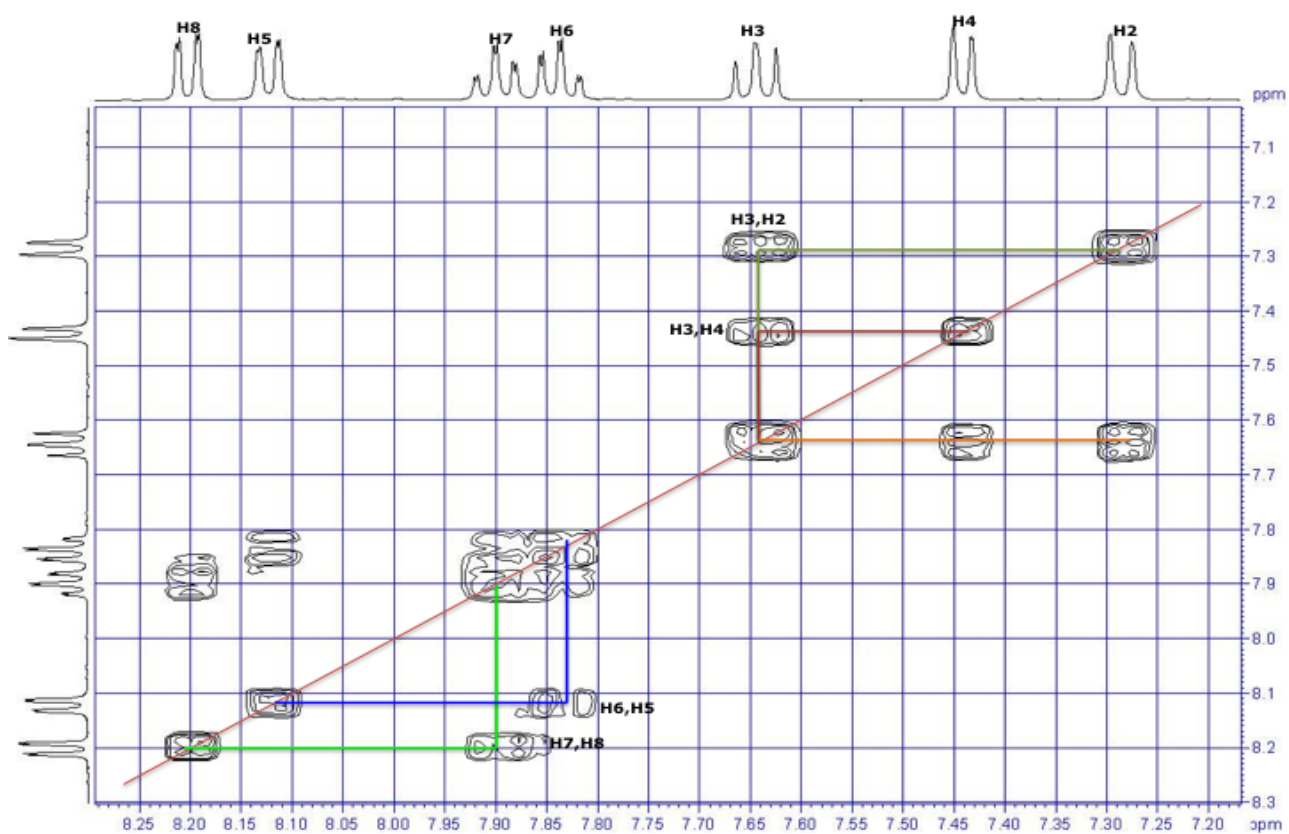
Figure S1: ¹H, ¹H-COSY NMR spectra of compound 5 in d-DMSO.
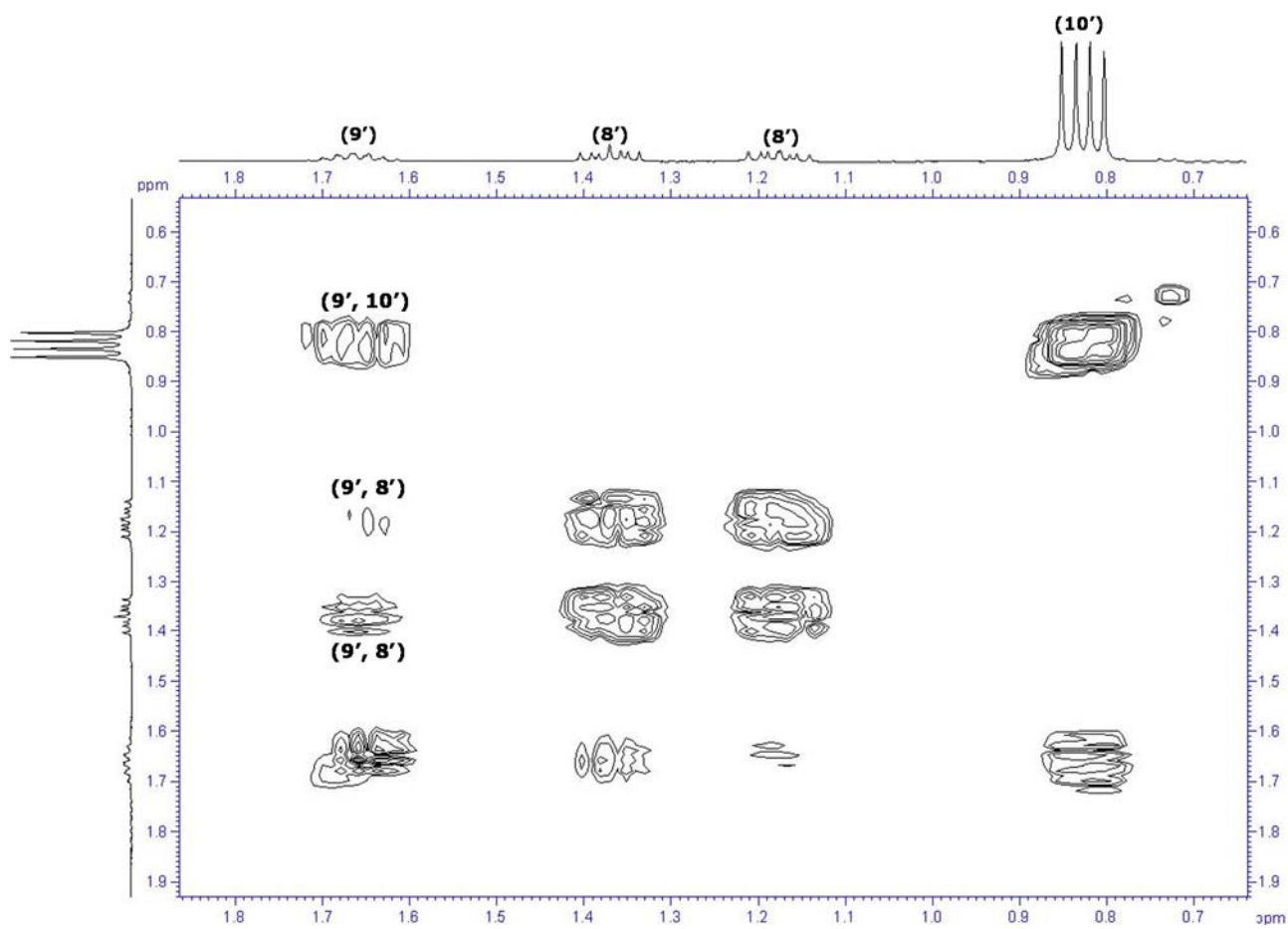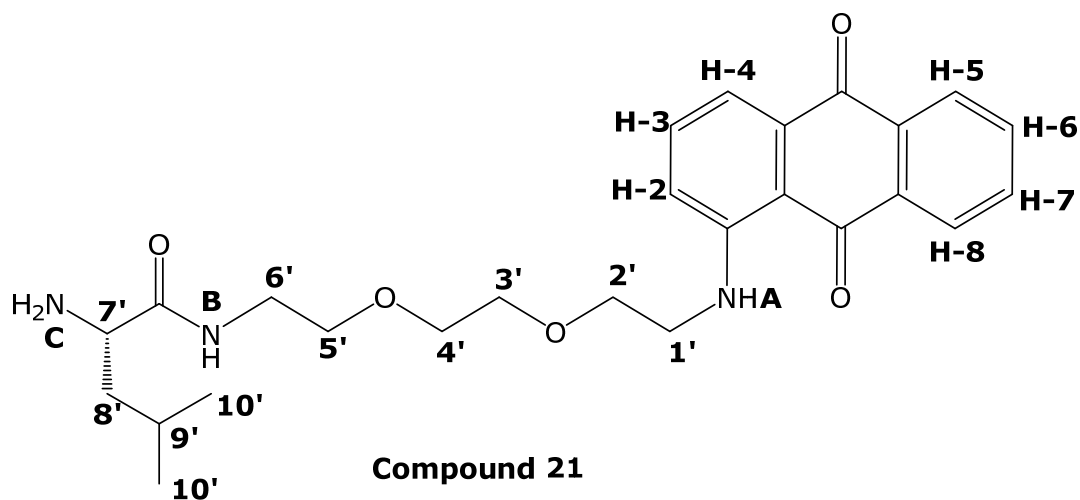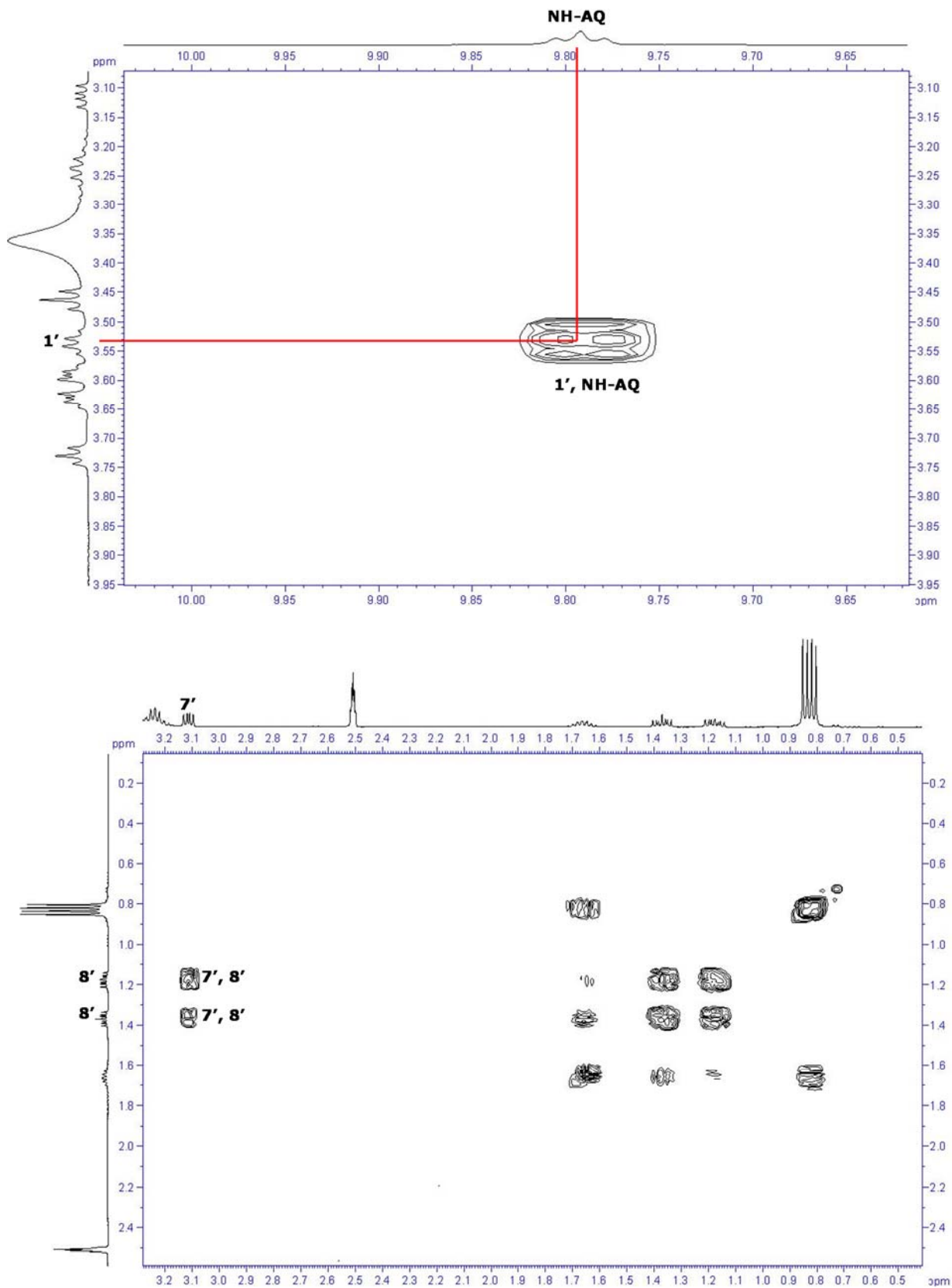
Compound 21

Figure S2: ^{1}H, ^{1}H-COSY NMR spectra of compound 21 in d-DMSO confirming the coupling between δ-methine (9') and β-methylene (8'); amino group proton of AQ with 1' and the α proton of leucine with β-methylene (8').
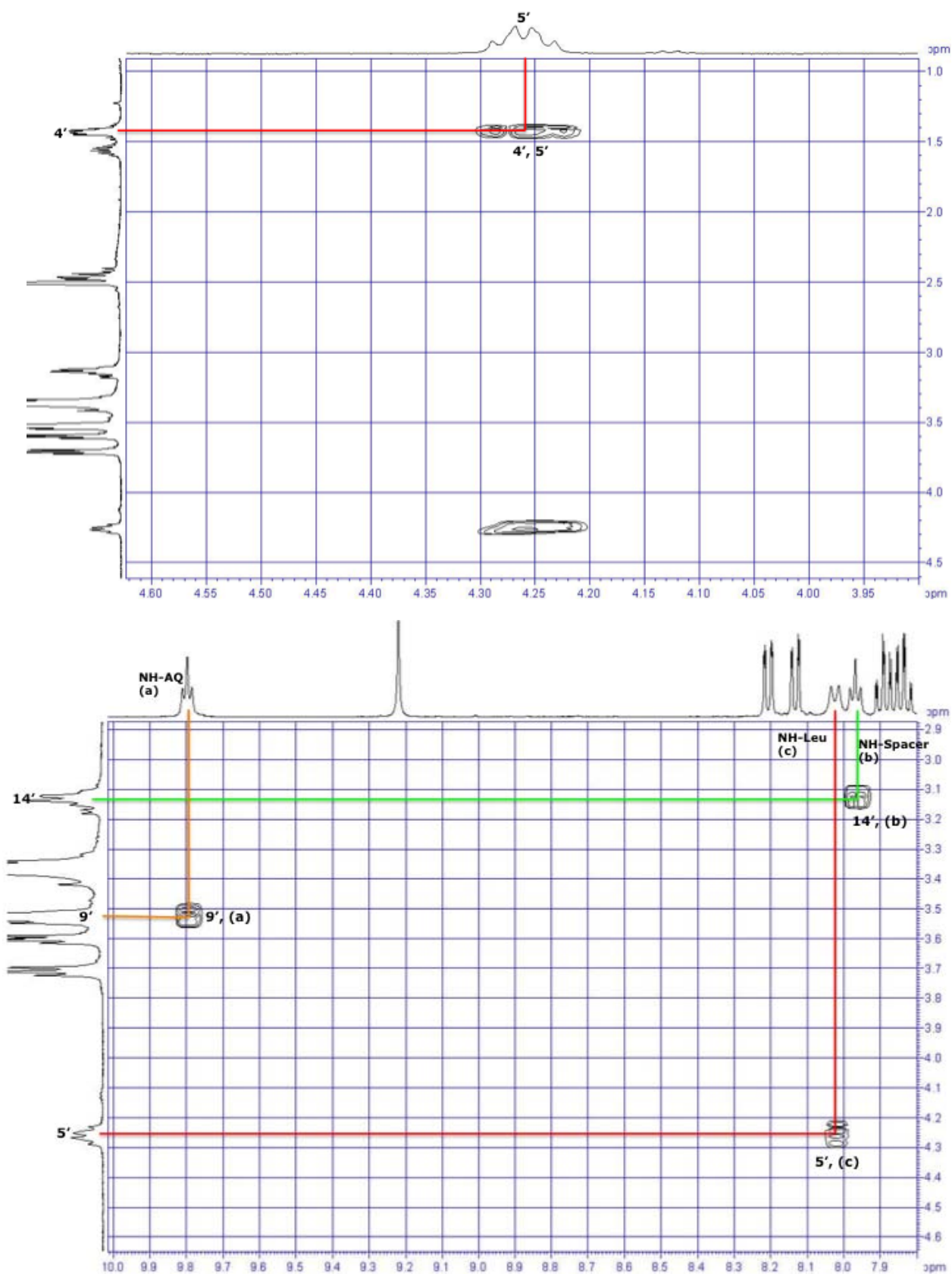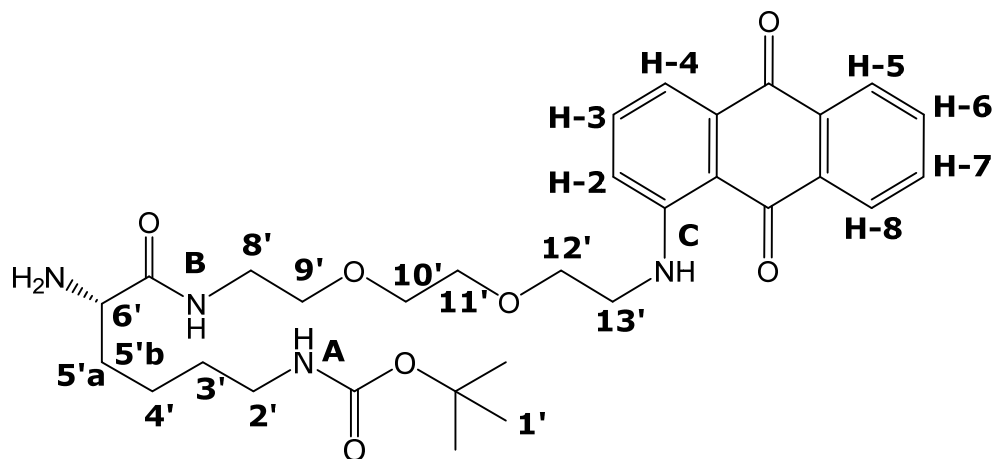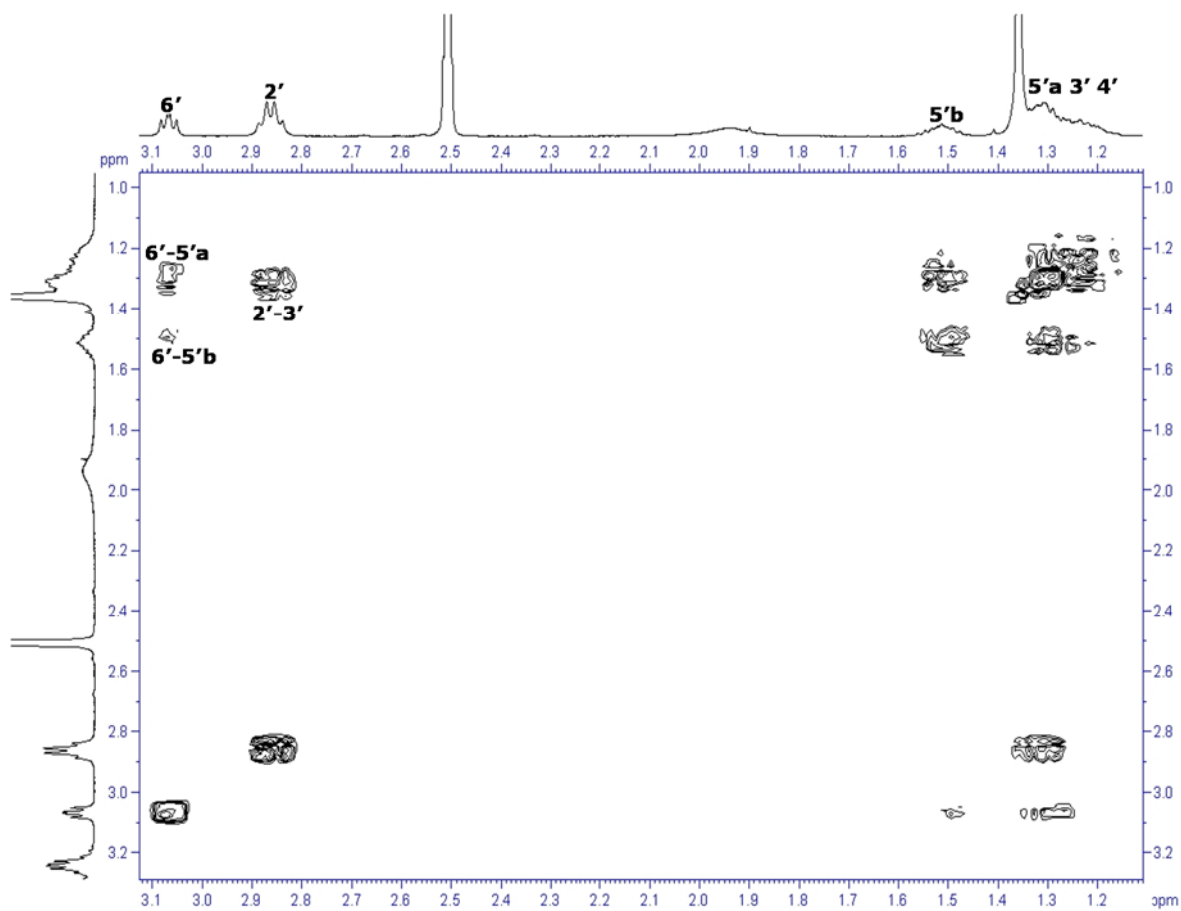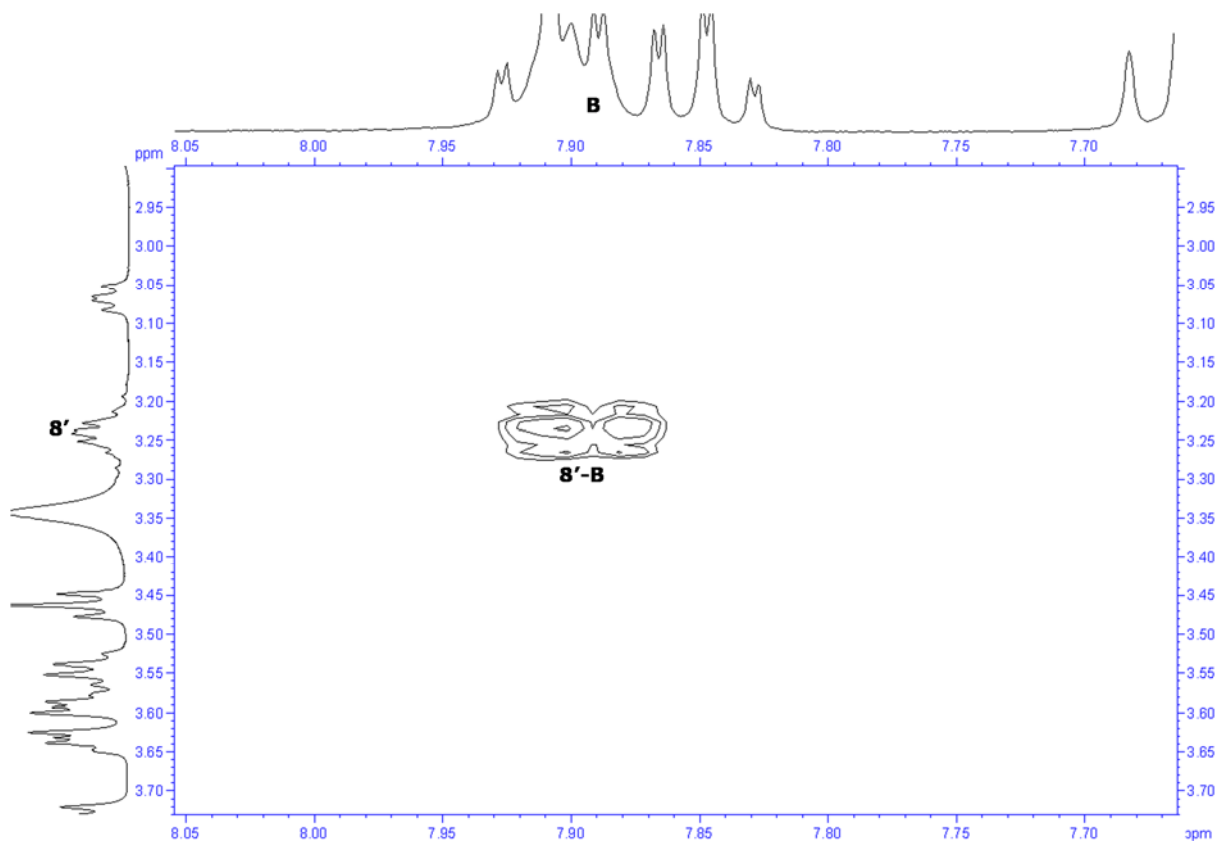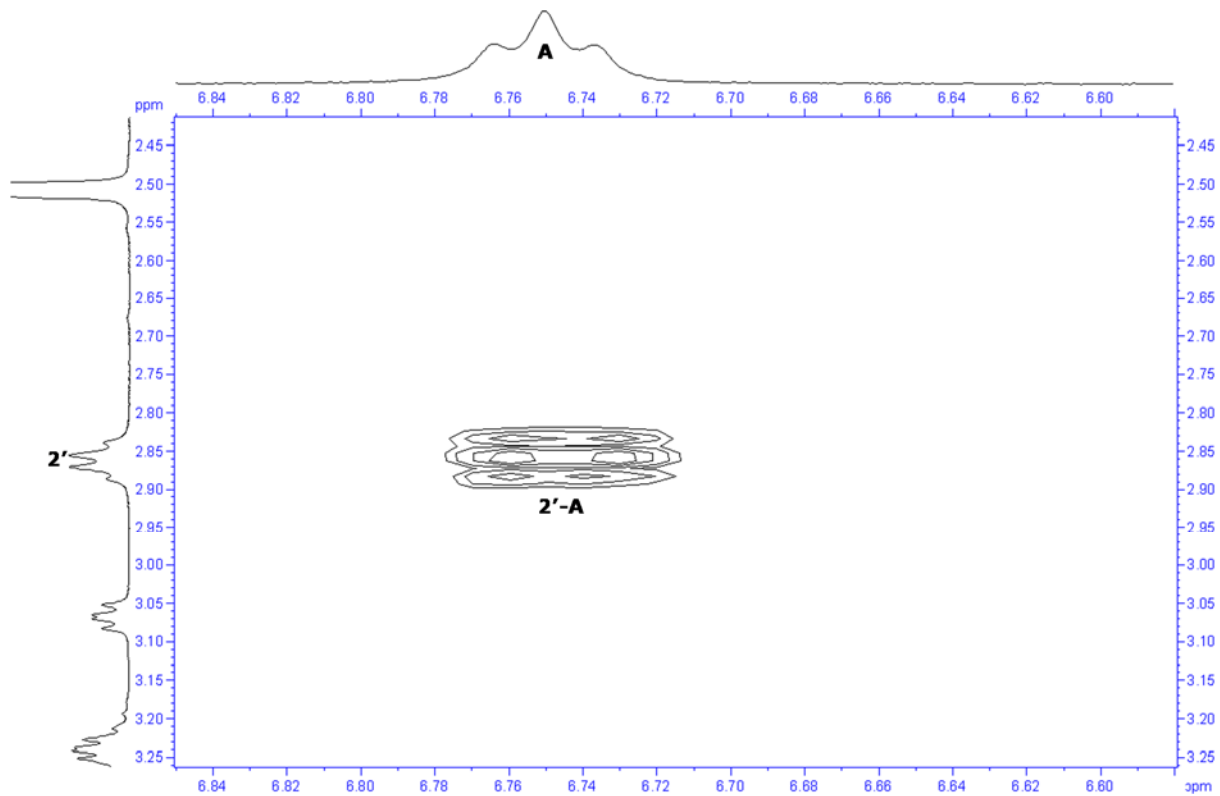
Trityl protecting group protons
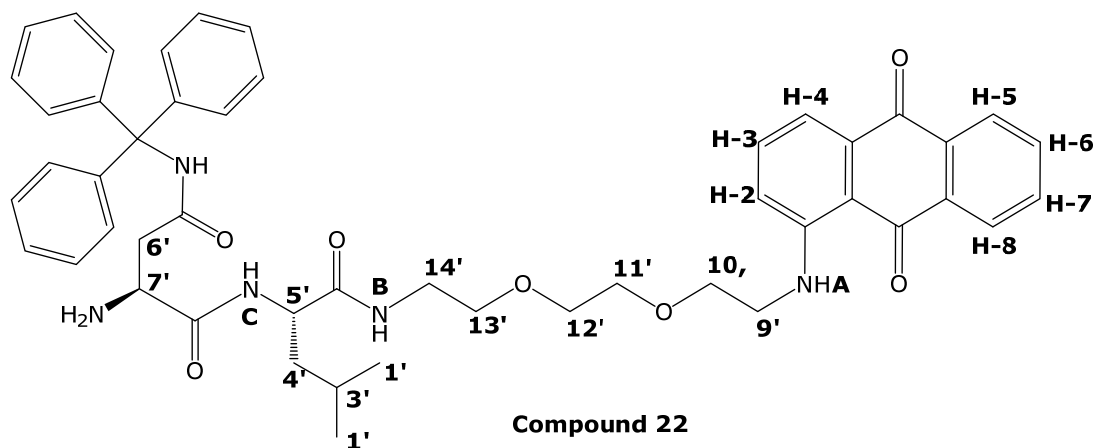
Compound 22
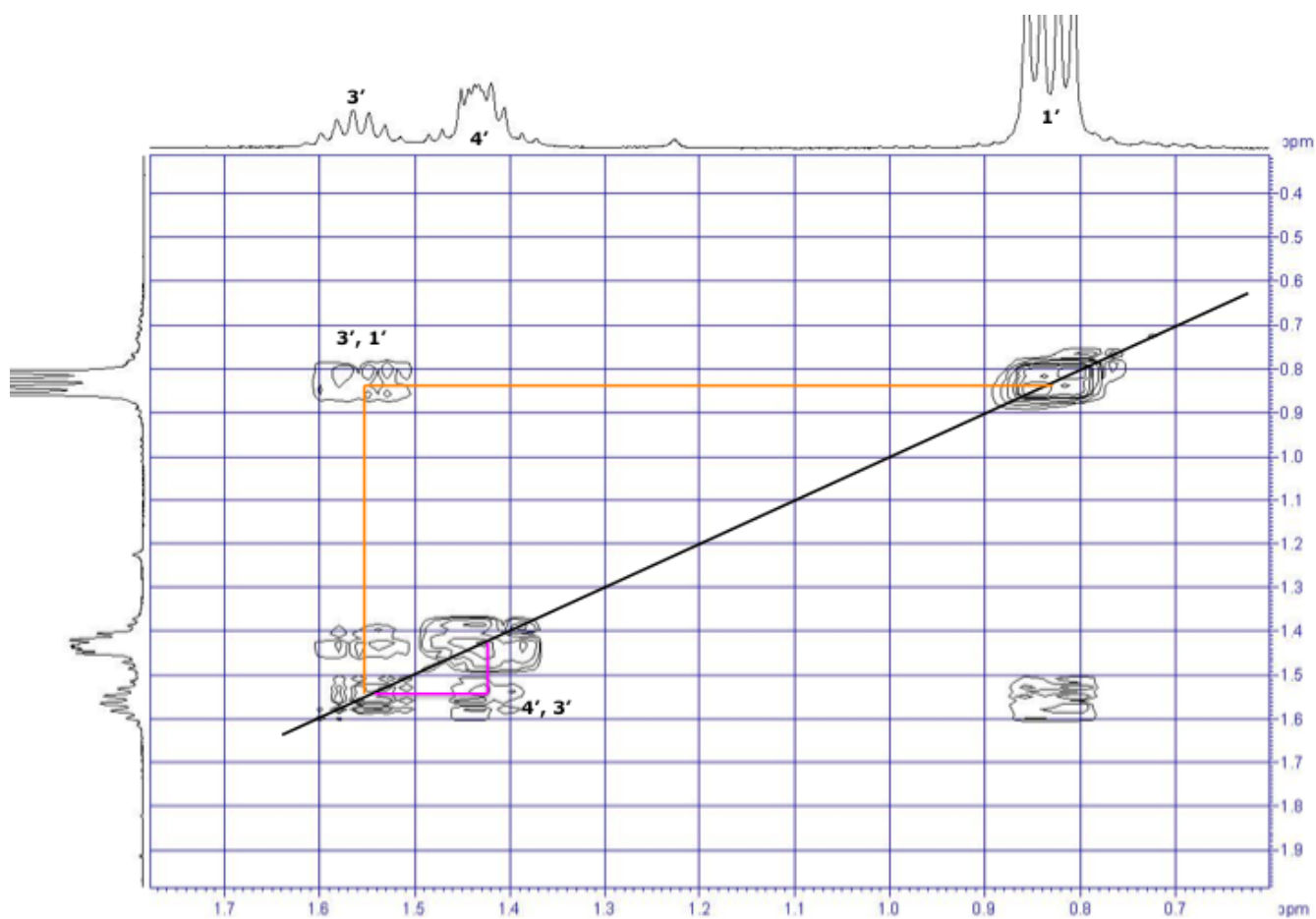
Figure S3: ^{1}H, ^{1}H-COSY NMR spectra of compound 22 in d-DMSO.
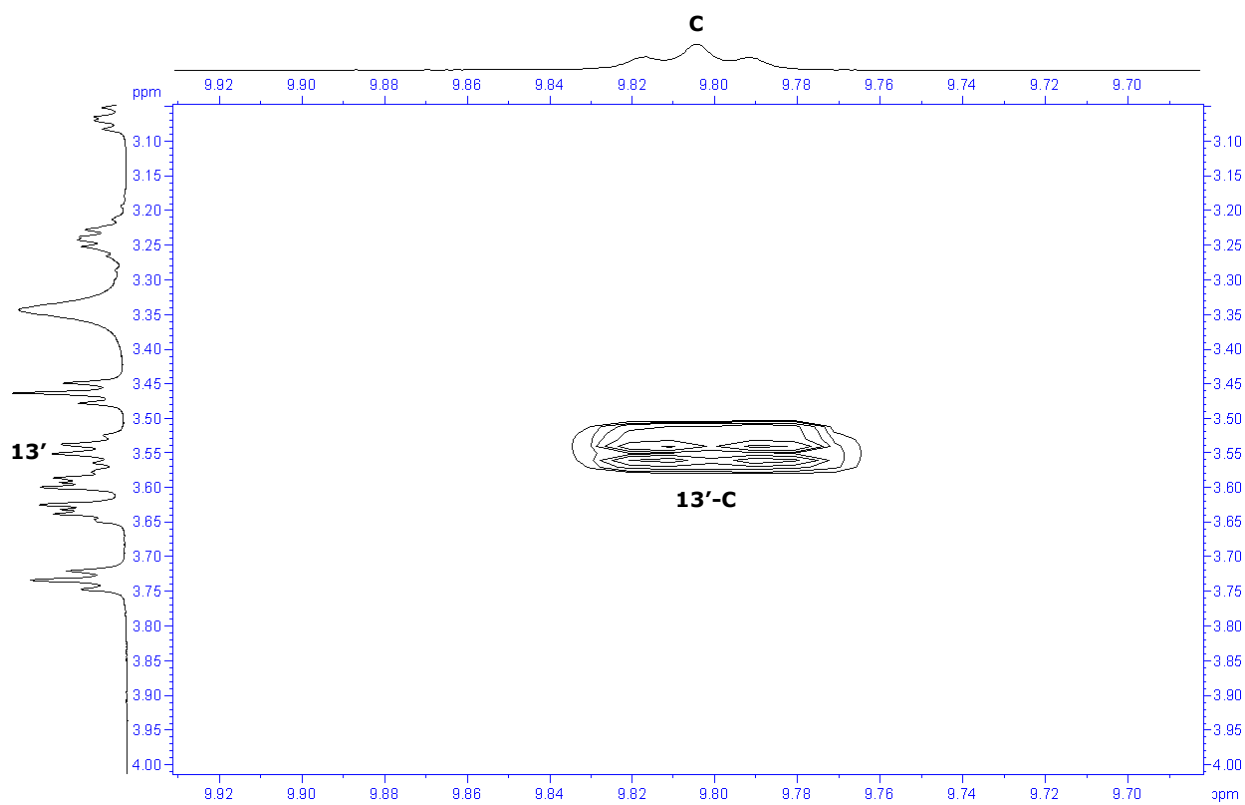
Compound 26

Figure S4: ^{1}H, ^{1}H-COSY NMR spectra of compound 26 in d-DMSO.

Appendix 5

High resolution mass spectrometry results for compounds were obtained *via* positive ion/negative ion nano-electrospray ionization (ESI) as performed on a thermofisher LTQ orbitrap XL mass spectrometer.

Compound 9

Molecular weight: 999.37
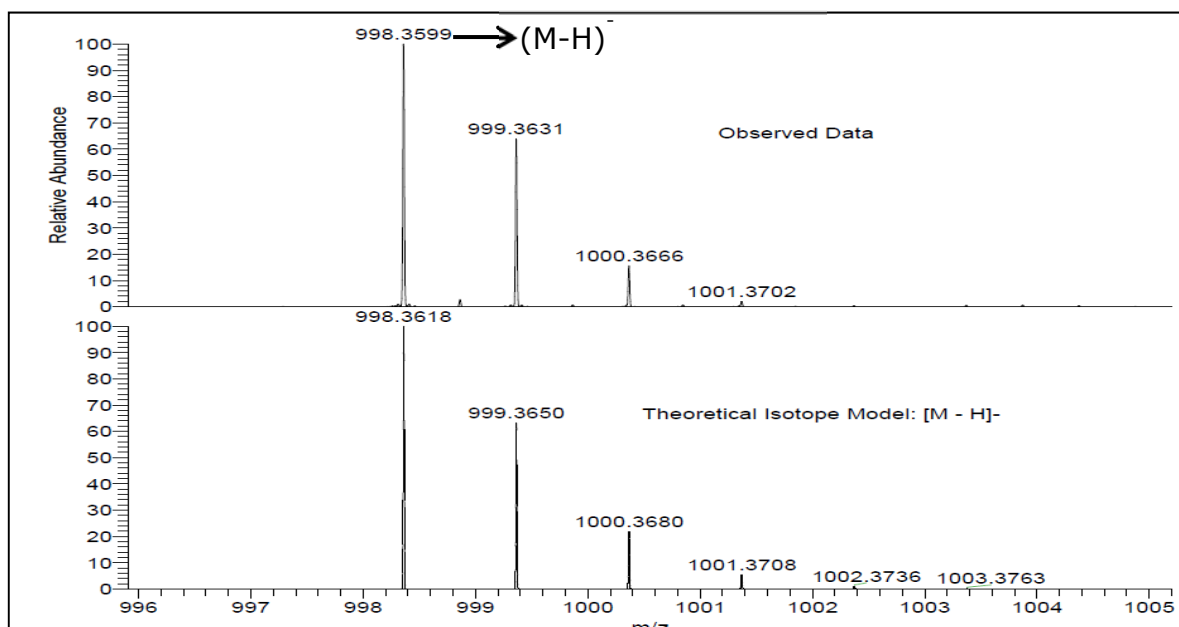
Compound 12

Molecular weight: 1055.43
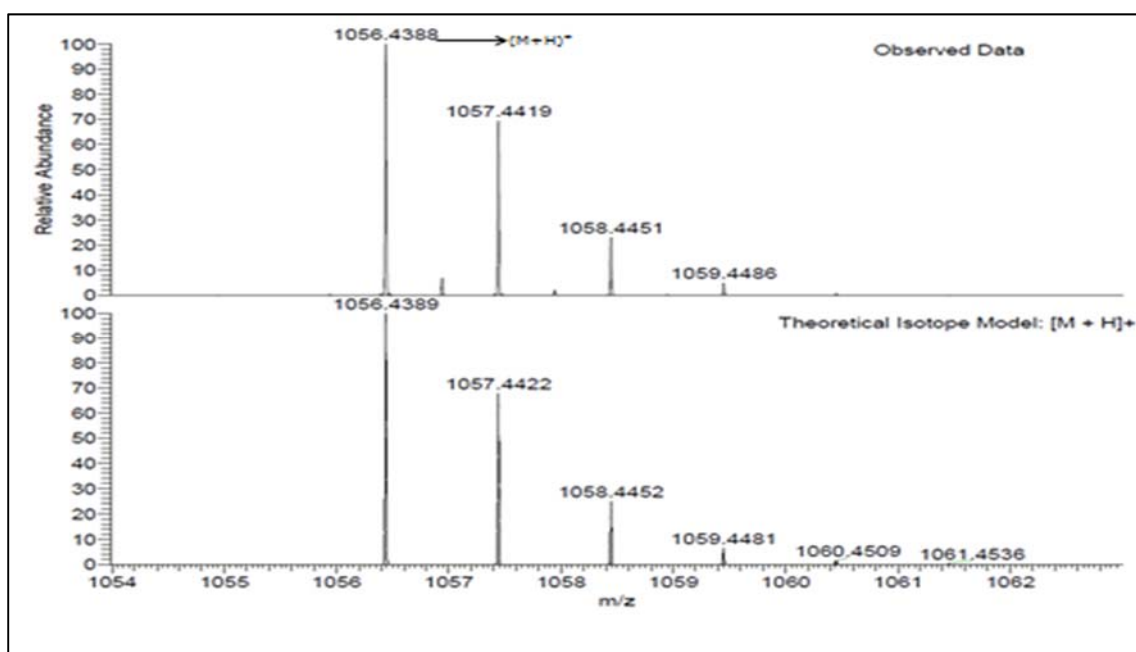
Compound 15

Molecular weight 987.37
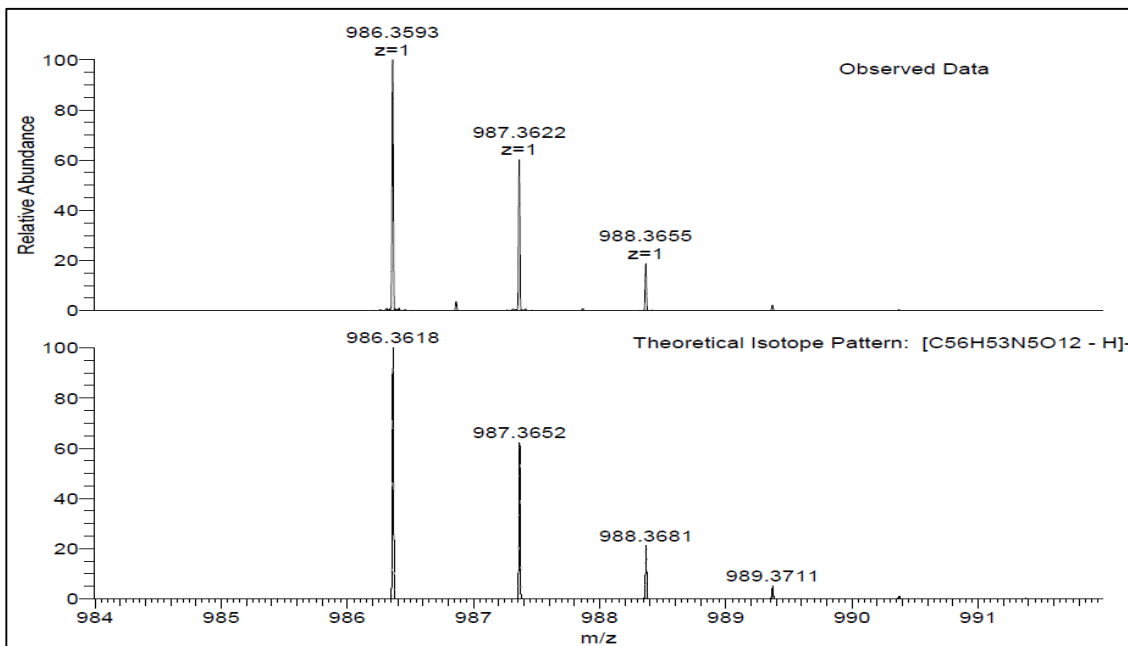
Compound 18

Molecular weight: 1029.42

260
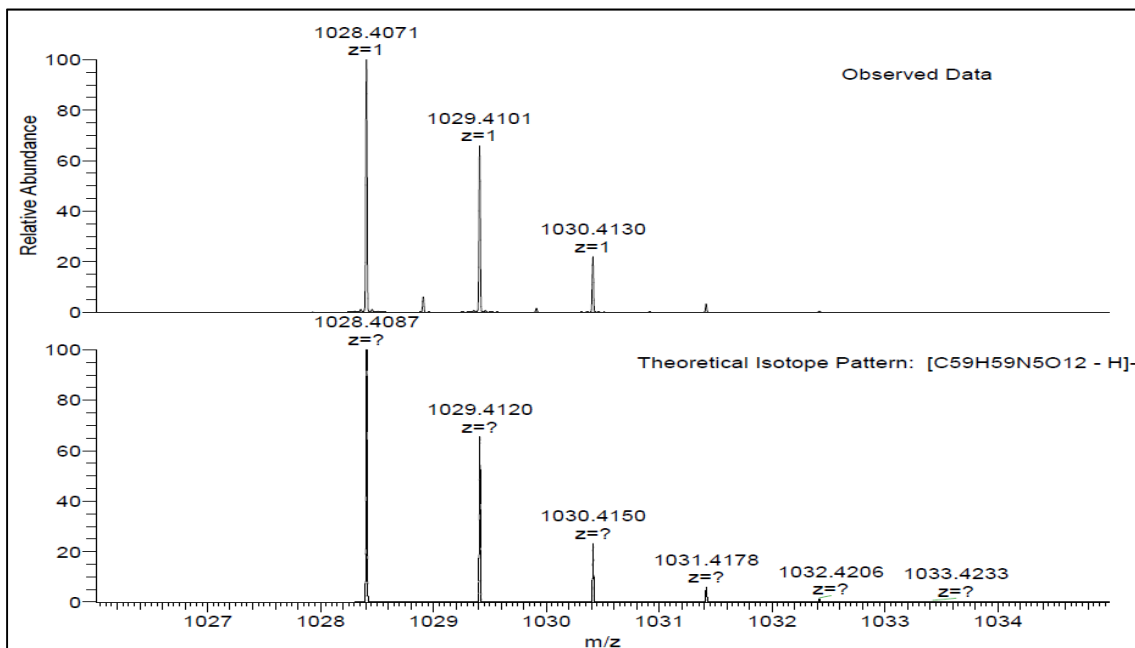
Compound 34

Molecular weight: 1082.40
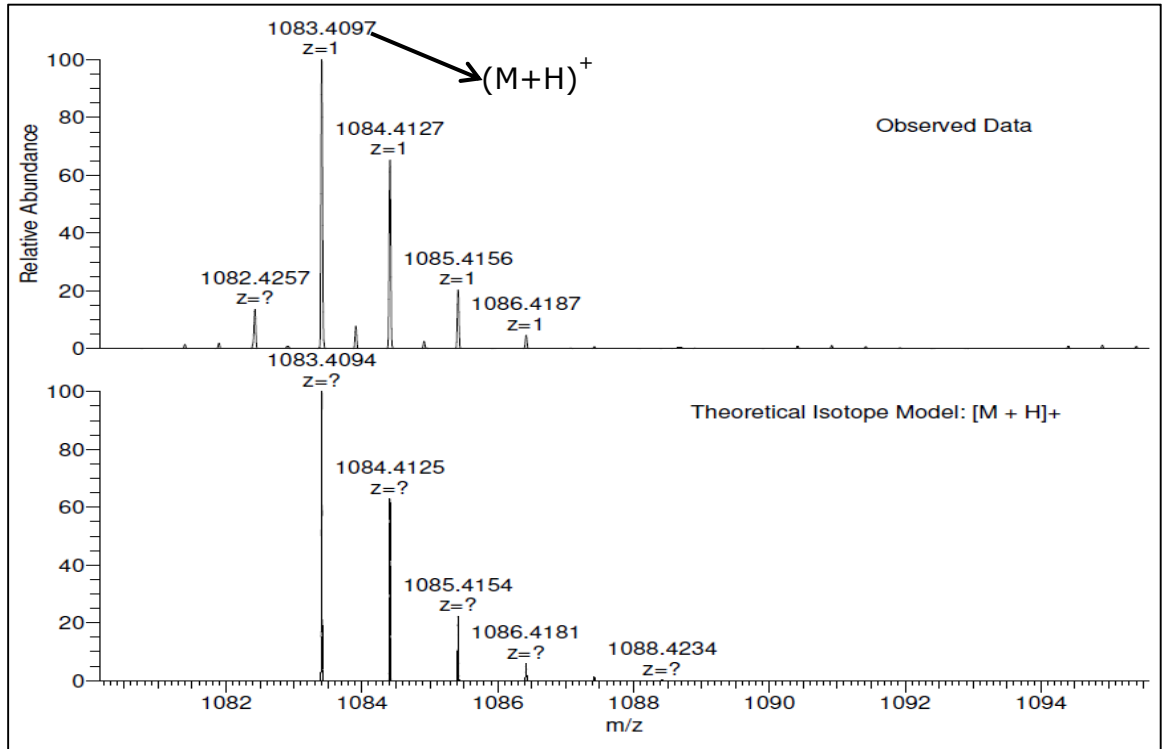
Compound 38

Molecular weight: 1122.43
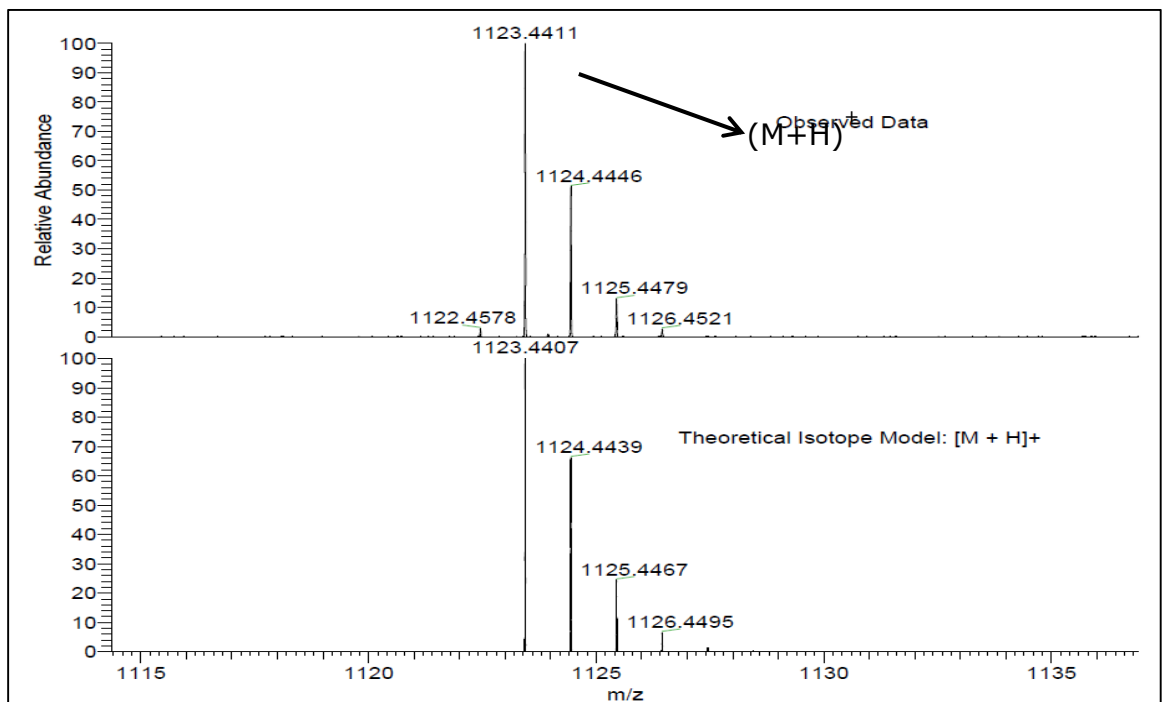
Compound 41

Molecular weight: 1096.42
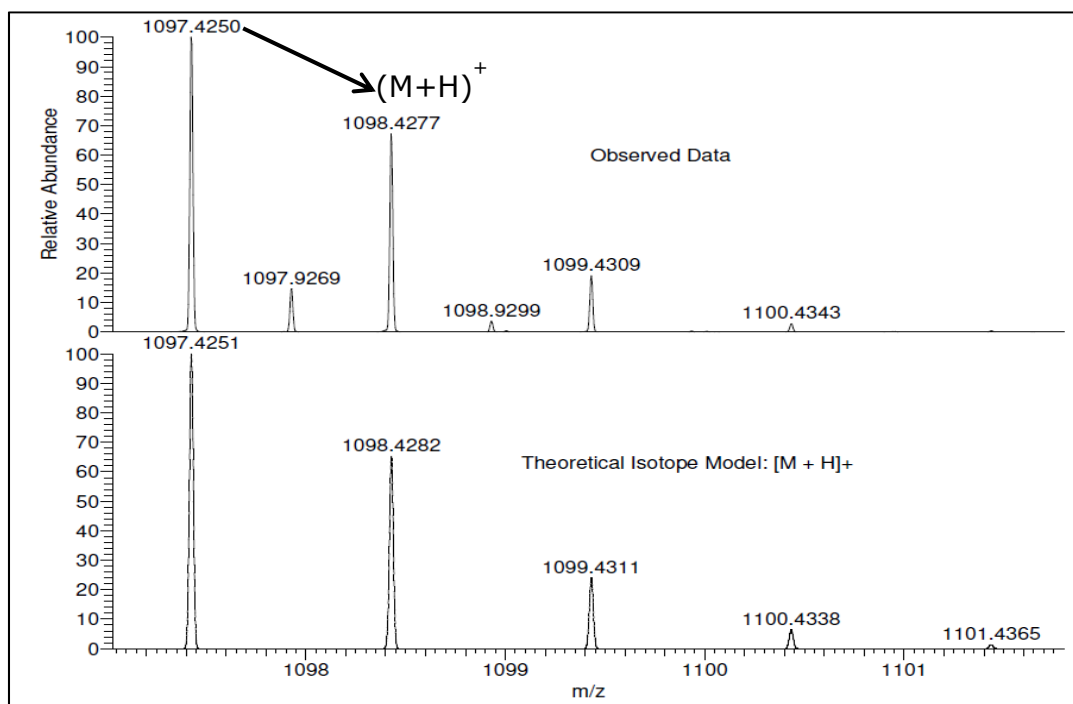
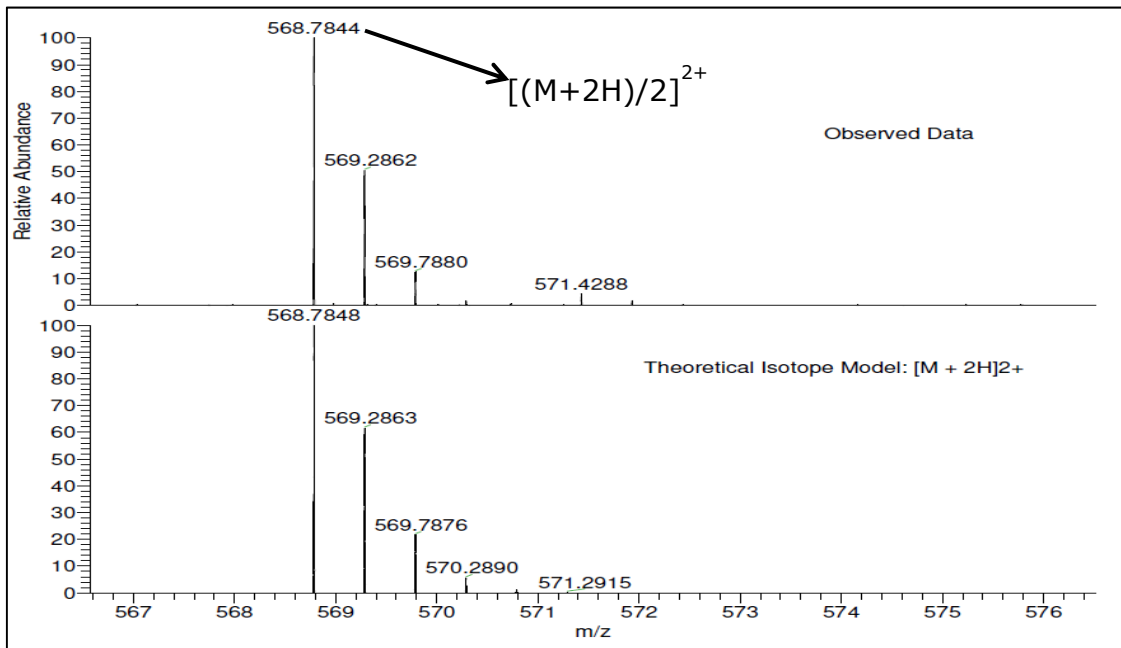## Compound 46

Molecular weight: 1178.46
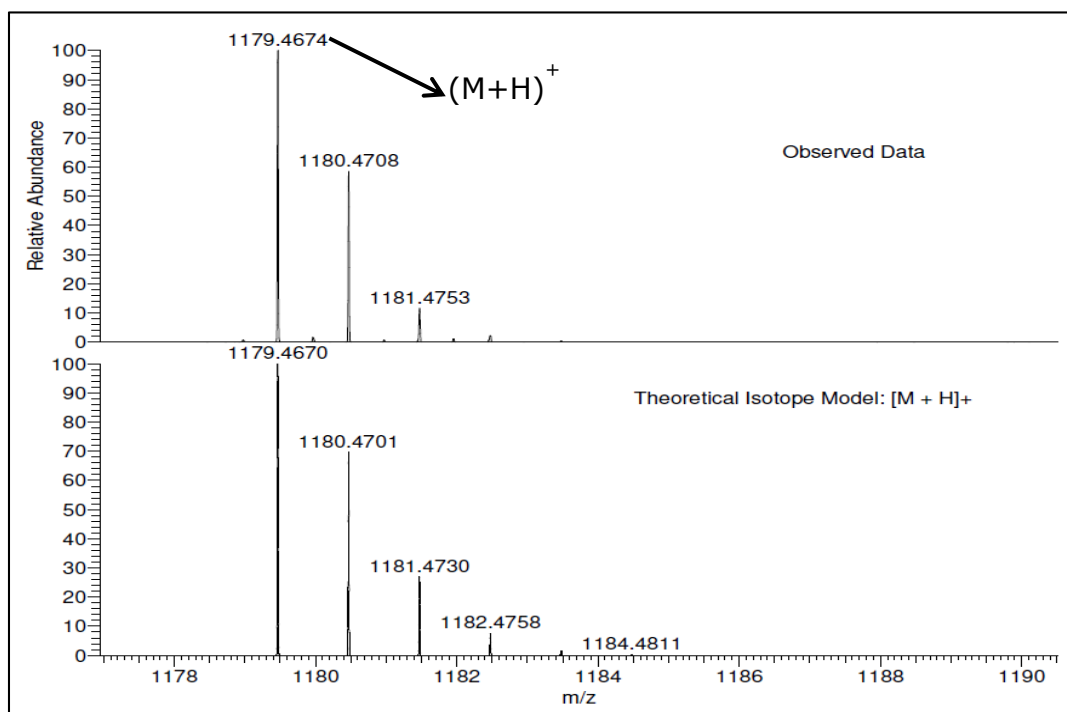
Compound 47

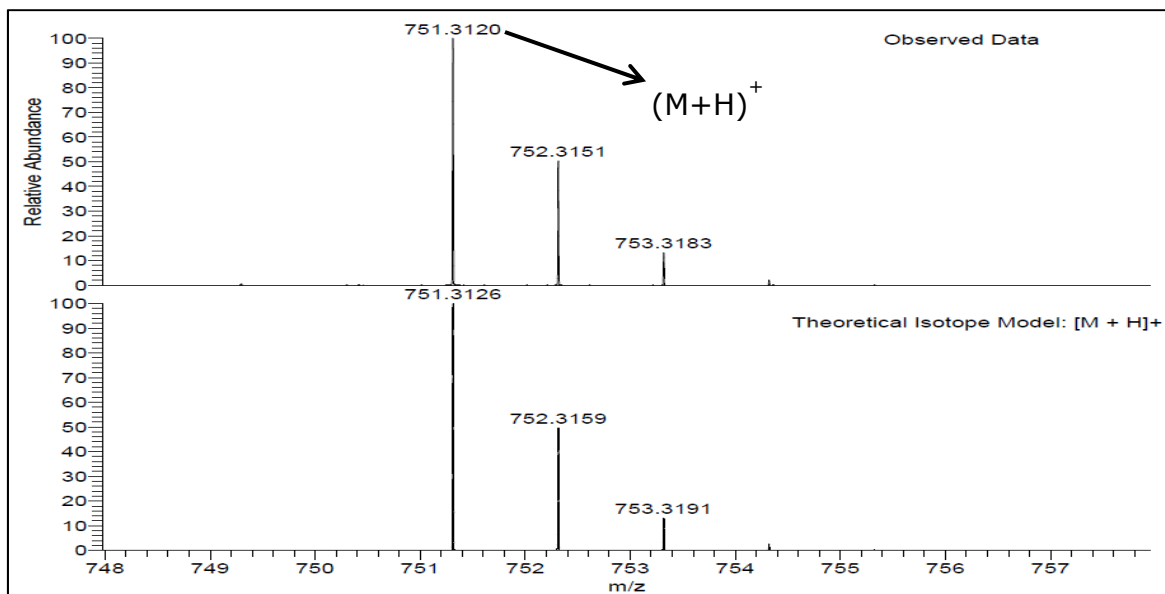Molecular weight: 1545.85
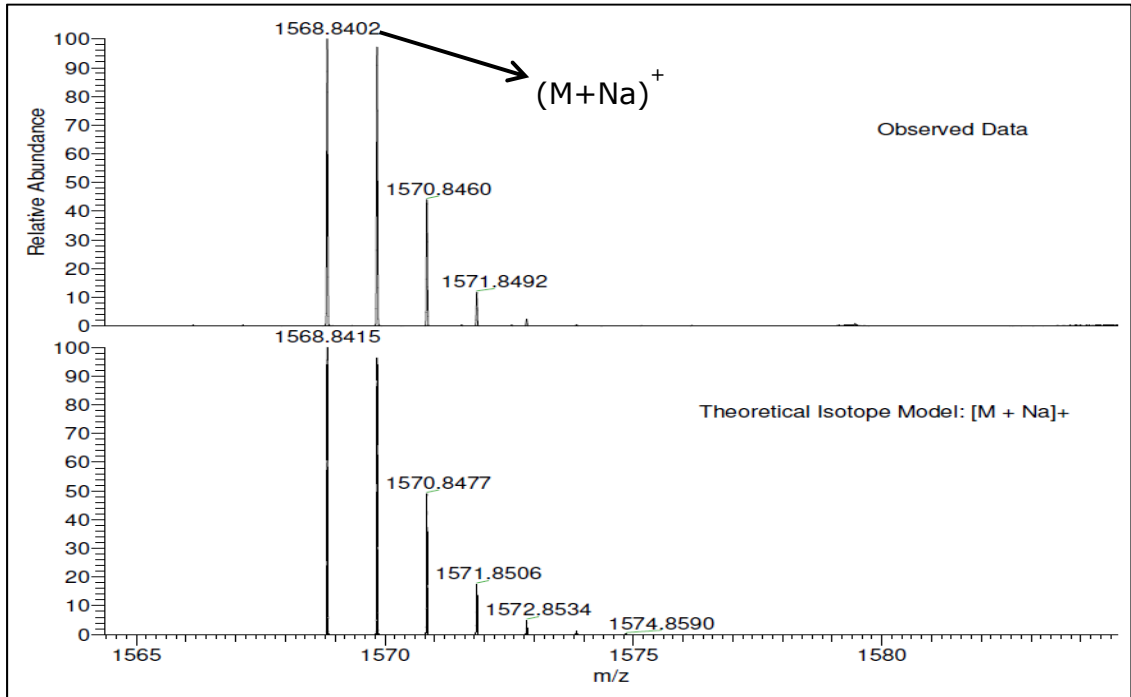
Compound 48
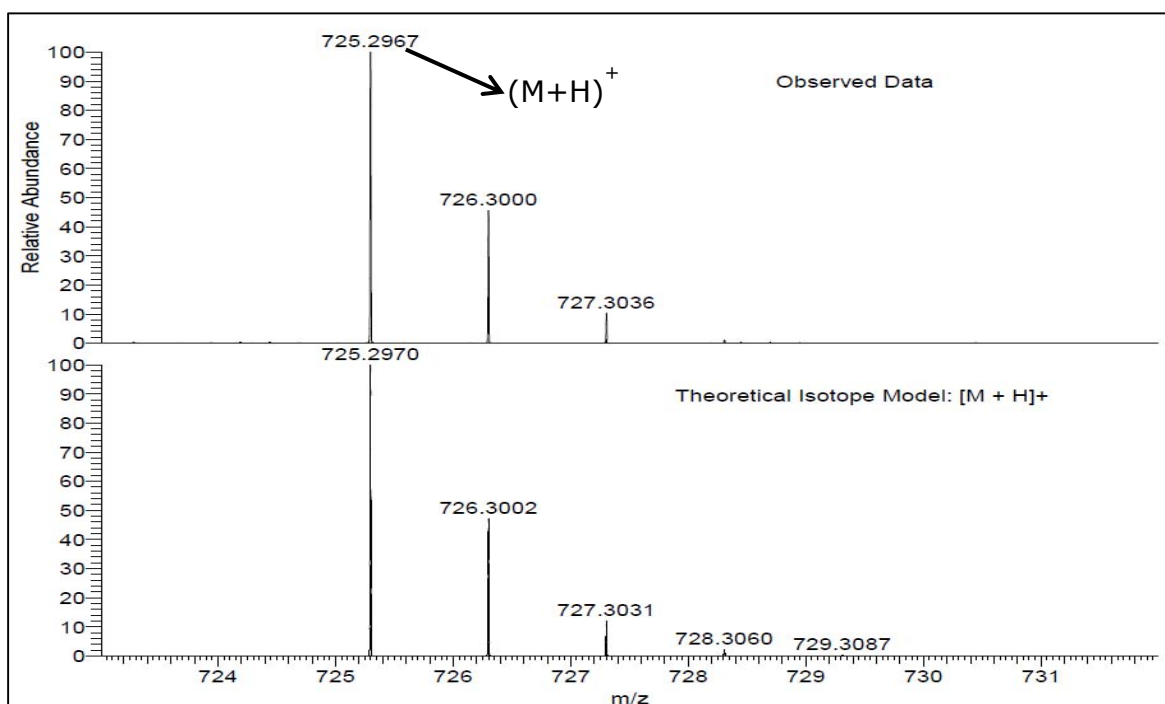
Molecular weight: 1135.55

Compound 59

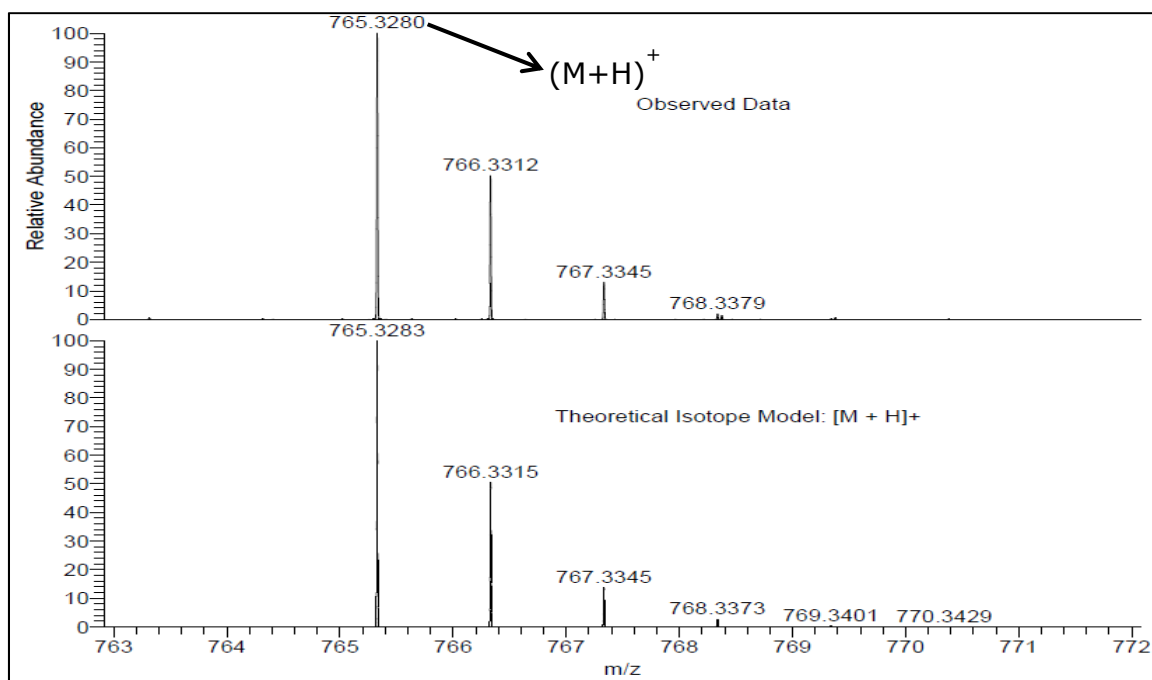Molecular weight: 750.30

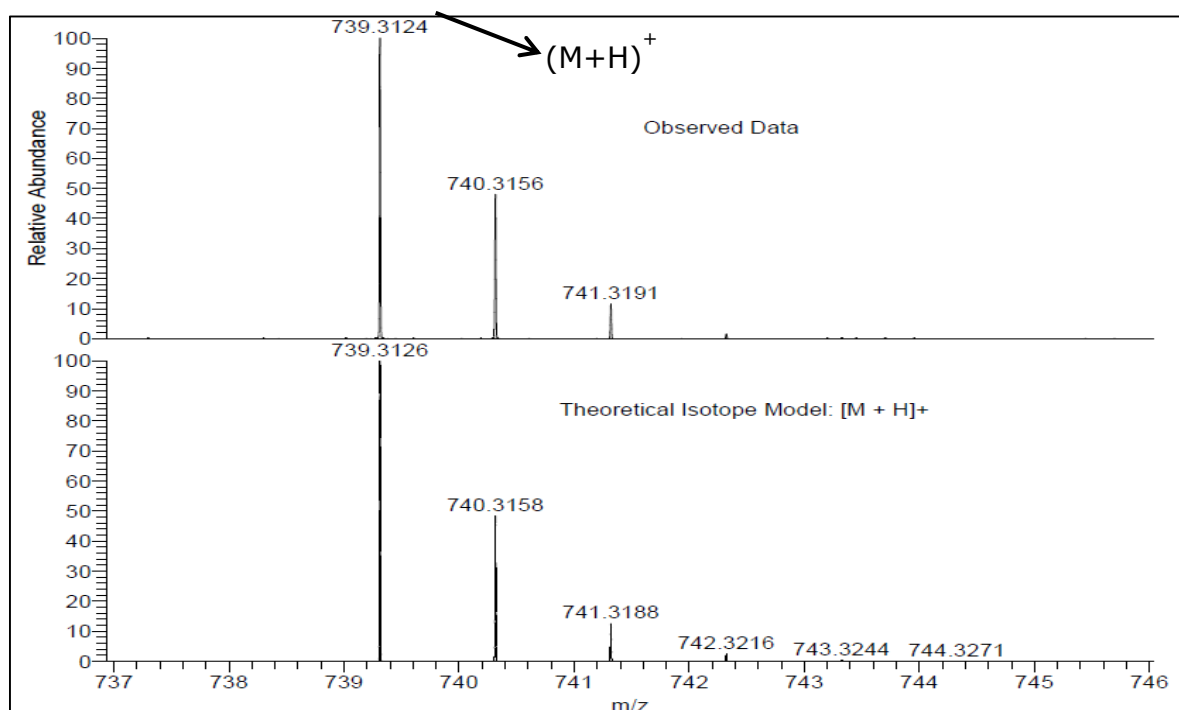## Compound 63

Molecular weight: 724.29

Compound 67

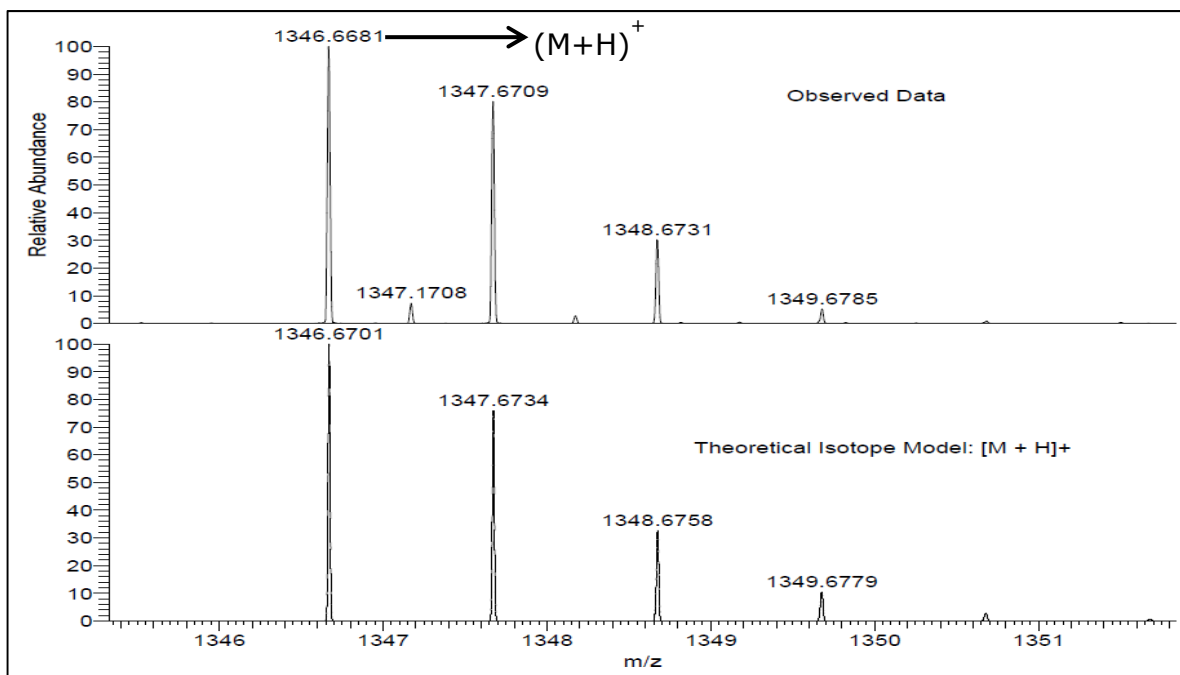Molecular weight: 764.32

Compound 71
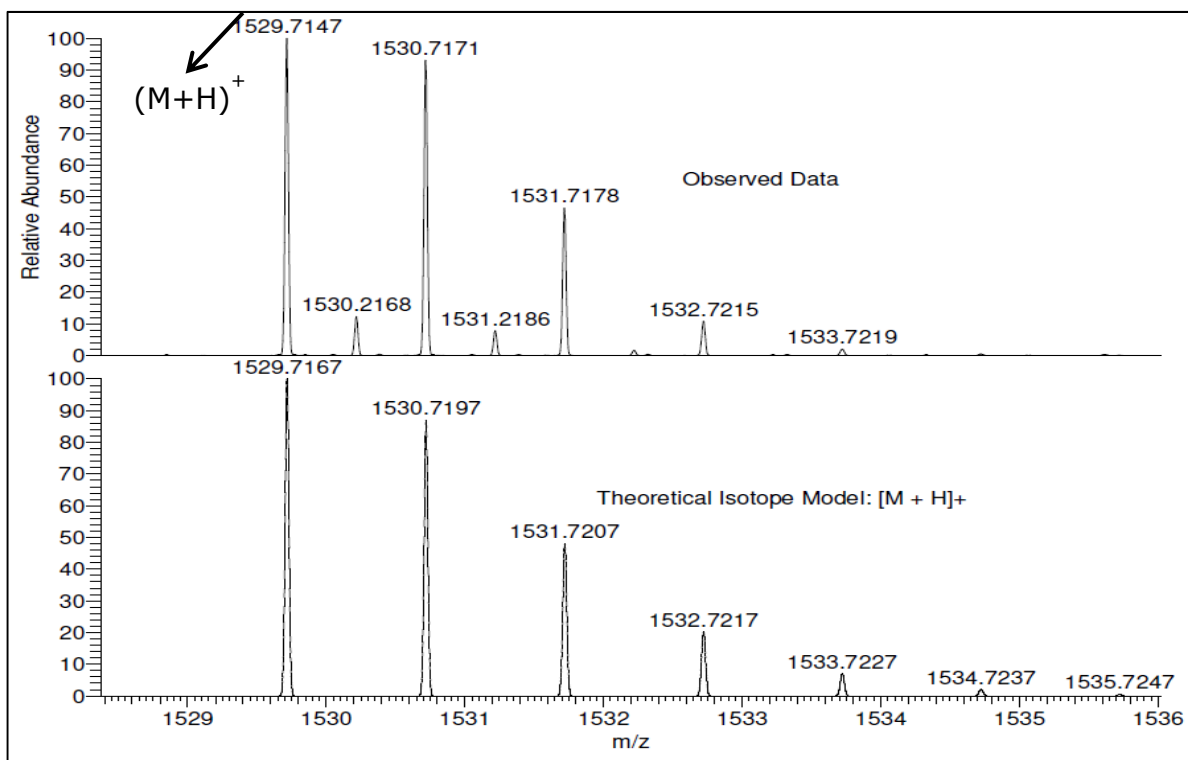
Molecular weight: 738.31

Compound 78

Molecular weight: 1345.40

Compound 80

Molecular weight: 1558.35

266

6. Supporting studies

- Presentation at the Robert Gordon University research student symposium, 2011.
- Poster presenation at the Univerisity of the Highlands and Islands (UHI) conference, 2012.
- Attainment of the Postgraduate certificate in Research methods, 2012.
- Poster presentation at the Institiute for Health and Welfare Research Showcase, 2012 and 2014.
- Poster presentation at the conference of the Academy of Pharmaceutical Scienctists of Great Britain (APSGB), 2013.
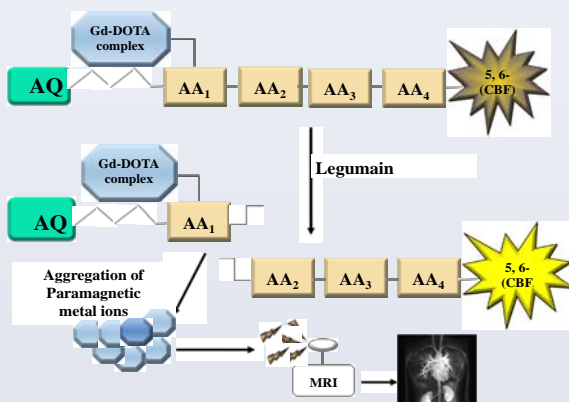
The design, synthesis and evaluation of selective, non-invasive imaging agents for atherosclerotic plaques.

[1]J.R. Jesudoss, [1]G. Kay, [2]D.J. Mincher, [2]A. Turnbull, [1]A. Di Salvo, [1]C.L. Wainwright.

[1]Institute for Health & Welfare Research, Robert Gordon University, Aberdeen.
[2]Faculty of Health, Life and Social Sciences, Edinburgh Napier University, Edinburgh.

INTRODUCTION

Atherosclerosis is a serious medical condition wherein the arteries become blocked by fatty substances such as cholesterol, leading to hardened or narrow blood vessels. Early diagnosis and treatment of these atheromas is necessary as they may result in either the damage of organs or they may rupture, releasing a blood clot instigating a heart attack or stroke [1]. Recent studies have implicated the role of legumain- a cysteine protease, in the progression and activation of plaque rupture [2]. In addition, over expression of active legumain is found in unstable regions of atheromas. These findings suggest that legumain may have the potential to be exploited for targeted delivery of drugs/imaging agents.

AIM AND OBJECTIVES

The aim of this research project was the design, synthesis and evaluation of selective non-invasive imaging agents for atherosclerotic plaques. To achieve this, the study was split into three objectives.

• The synthesis, purification and characterisation of a small discrete library of peptide motifs, designed to be substrates for legumain.

• The covalent attachment of imaging agent (e.g., Gd complexes).

• Evaluation of *in vitro* cleavage and release of imaging agent.

METHODOLOGY

SOLID PHASE PEPTIDE SYNTHESIS

The peptide was synthesized by sequential addition of amino acids using PyBOP and DIPEA. Reaction completion was determined by a Kaiser test. The tetra-peptide sequence was then cleaved from the resin using 5% TFA in DCM. Coupling of the anthraquinone-spacer moiety was achieved using similar coupling conditions.

SOLUTION PHASE PEPTIDE SYNTHESIS

An anthraquinone moiety was coupled to the peptide using HOBT, TBTU, DIPEA in solution. After each coupling and deprotection the samples were purified by column chromatography and characterized by mass spectrometry. On substrate optimization, the final step of coupling an imaging agent will occur on a ε-lysine residue.
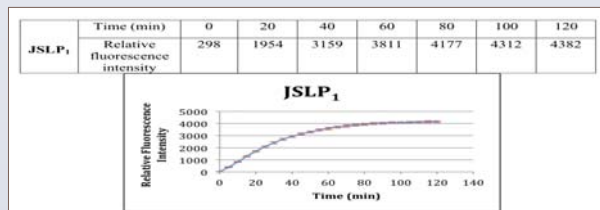
ENZYME ASSAY

These peptide sequences were then incubated with rhLegumain/Aparaginyl Endopeptidase at 37ºC for 2hours.

RESULTS AND DISCUSSION

Legumains are found to exhibit restricted specificity for peptide hydrolysis on the carboxyl side of asparagine residues [3]. By exploiting this condition six peptide substrates containing asparagine at P_1 position of the tetra peptide sequence were synthesized [4]. These were characterized by high resolution mass spectrometry.
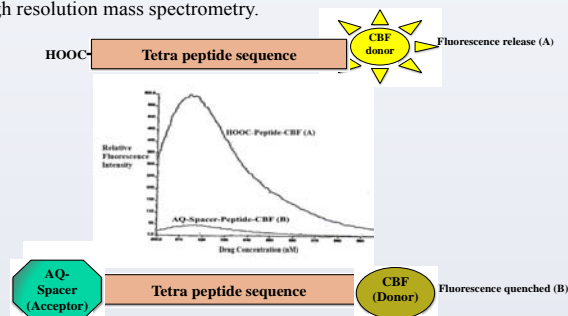
Each sequence contained a fluorophore (5(6) carboxyfluorescein) at the N-terminal. The peptide's fluorescence was measured prior to and after anthraquinone coupling. The results indicated that the anthraquinone system effectively quenched the peptide-fluorophore. It is hoped that this phenomenon can now be used to further elucidate peptide substrate specificity for legumain.

	Time (min)	0	20	40	60	80	100	120
$JSLP_1$	Relative fluorescence intensity	298	1954	3159	3811	4177	4312	4382

Moreover enzyme studies on these substrates clearly showed their specificity to legumain by the relative increase in fluorescence intensity when hydrolyzed on the carboxyl side of asparagine for e.g., $JSLP_1$.

CONCLUSION AND FUTURE WORK

Six anthraquinone-peptide-fluorescein FRET systems, designed to be specific to legumain, have been synthesized and characterized. These peptide substrates may be exploited to not only deliver site specifically, contrast agents for magnetic resonance imaging but also pharmacological actives in a range of disease states.

ACKNOWLEDGMENTS

REFERENCES

[1] Lopez-Avila, V; and Spencer, JV. "Methods for Detection of Matrix Metalloproteinases as Biomarkers in Cardiovascular Disease." *Clinical Medicine: Cardiology*, 2008, 2:1–14.

[2] Clerin V, Shih HH, Deng N, Hebert G, Resmini C, Shields KM, Feldman JL, Winkler A, Albert L, Maganti V, Wong A, Paulsen JE, Keith JC Jr, Vlasuk GP, Pittman DD. Expression of the cysteine protease legumain in vascular lesions and functional implications in atherogenesis. Atherosclerosis. 2008, (1):53-66.

[3] Lee J, Bogyo M. Synthesis and evaluation of aza-peptidyl inhibitors of the lysosomal asparaginyl endopeptidase, legumain. Bioorg Med Chem Lett. 2012, (3):1340-3.

[4] Mathieu MA, Bogyo M, Caffrey CR, Choe Y, Lee J, Chapman H, Sajid M, Craik CS, McKerrow JH. Substrate specificity of schistosome versus human legumain determined by P1-P3 peptide libraries. Mol Biochem Parasitol. 2002,121(1):99-105.
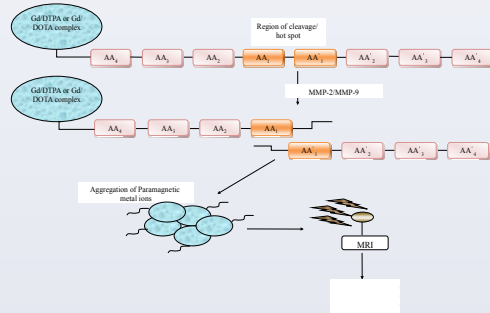
The design, synthesis and evaluation of selective , non-invasive imaging agents for atherosclerotic plaques

J. Jesudoss, G. Kay, A. Di Salvo, D. Mincher, C. Wainwright

School of Pharmacy and Life Sciences, Robert Gordon University, Aberdeen

Introduction, aim and objectives

Atherosclerosis is a slow process, where plaque develops at a very early stage of life and ultimately leads to narrowing of the blood vessel, impairing the flow of blood in susceptible patients (1). Research on MMPs has indicated that they play a vital role in the development of atherosclerotic plaques. Their expression is not only part of atherosclerosis pathogenesis, but may also be predictive of plaque instability. With characterization of peptide substrates specific to certain MMPs, their role has been exploited to not only deliver site specifically, pharmacological actives in a range of disease states both *in vivo* and *in vitro* but also as a device to localize imaging agents (3,4) .
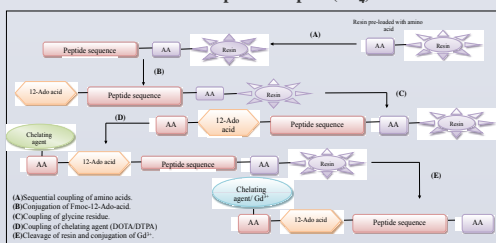

Aim and Objectives

The aim of this research project was the design, synthesis and evaluation of selective non-invasive imaging agents for atherosclerotic plaques. To achieve this, the study was split into three objectives.
• The synthesis, purification and characterisation of a small discrete library of peptide motifs, designed to be substrates for MMP-2/MMP-9.
• The covalent attachment of imaging agents (e.g., Gd complexes).
• Evaluation of *in vitro* cleavage and release of imaging agent.

Methodology

Synthesis of Gadolinium-DOTA-Peptide complex (JD₄)

(A)Sequential coupling of amino acids.
(B)Conjugation of Fmoc-12-Ado-acid.
(C)Coupling of glycine residue.
(D)Coupling of chelating agent (DOTA/DTPA).
(E)Cleavage of resin and conjugation of Gd³⁺.

Synthesis of JD₁

The peptide sequence was synthesised by the sequential conjugation of the amino acids, Leu-Nva-Ala-Leu-Ala-Ala using coupling agents HOBT, TBTU, DIPEA. After each coupling and deprotection the resin was washed with DMF (X1), DCM (X1), DCM/*i*-Pr-OH (X3), *i*-Pr-OH (X3), DCM (X3). Reaction completion was confirmed by Kaiser test.

Synthesis of JD₂

Fmoc-12-aminododecanoic acid was coupled to the peptide sequence using coupling agents HOBT, TBTU, DIPEA. This followed an adapted procedure by Espelt and co-workers (2).
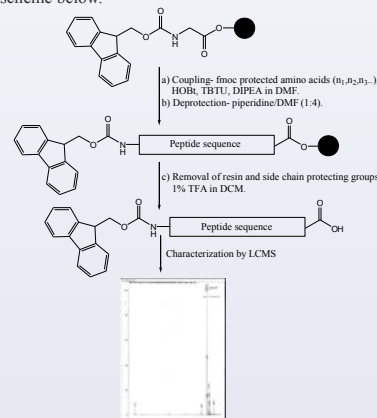
Synthesis of JD₃

A glycine residue was added to enhance the yield and facilitate efficient coupling of DOTA to the peptide sequence. DOTA-tris(*t*-Bu) was coupled to the sequence using coupling agents following the same parameters as JD₁ and JD₂. The resulting sequence (JD₃) was cleaved using TFA:H₂O:EDT:TA (87.5:2.5:4:6, v/v/v/v) for 4hrs at room temperature.

Synthesis of JD₄

The complexation of gadolinium chloride hexahydrate to the sequence will be undertaken at neutral pH.

Results and Discussion

A library of six peptides designed to be MMP-2/MMP-9 substrates have been synthesized and characterized by mass spectrometry. For example ESMS (+) which showed a signal at m/z 850.4 (M+1)⁺ corresponding to molecular mass of 849.97 as shown in the scheme below.

a) Coupling- fmoc protected amino acids (n₁,n₂,n₃..), HOBt, TBTU, DIPEA in DMF.
b) Deprotection- piperidine/DMF (1:4).

c) Removal of resin and side chain protecting groups 1% TFA in DCM.

Characterization by LCMS

An attempted synthesis of Gadolinium-DOTA-Peptide (JD₄) was also undertaken. Furthermore, the successful coupling of a hydrophobic moiety (Ado) was achieved. It is hoped that the addition of Ado to the peptide substrate will decrease the solubility of the cleaved contrast agent *in vitro* resulting in deposition of the imaging agent (5). The addition of a chelating agent and subsequent complexation reaction is now underway.

Conclusion and future work

Conclusion

The overall aim of this study is the development of a non-invasive imaging agent that can be deposited site specifically, via MMP cleavage, in vulnerable plaques. Six peptides designed to be MMP-2/MMP-9 substrates were synthesised. The structure of these peptide substrates was confirmed by mass spectrum analysis. Synthesis of a peptide based contrast agent was also undertaken. However, the complexation of the metal ion is yet to be done.

Future work

• Optimisation of conditions for mono substitution of DOTA/DTPA.
• To develop a library of oligopeptide MMP-2/MMP-9 substrates.
• Evaluate susceptibility of these peptide substrates to MMP-2/MMP-9 cleavage.

References

1)Avila VL and Spencer JV (2008). Methods for detection of matrix metalloproteinases as biomarkers in cardiovascular disease. *Clinical medicine*. **2**, 75-87.

2)Espelt L, Parella T, Bujons J, Solans C, Joglar J, Delgado A, Clapes P (2003). Stereoselective aldol additions catalyzed by dihydroxyacetone phosphate-dependent aldolases in emulsion systems: preparation and structural characterization of linear and cyclic iminopolyols from aminoaldehydes. *Chem. Euro. J.* **9**, 4887-4889.

3)Loftus IM, Naylor AR, Goodall S, Crowther M, Jones L, Bell PRF, Thompson MM (2000). Increased matrix metalloproteinase-9 activity in unstable carotid plaques: a potential role in acute plaque disruption. *Stroke*. **31**, 40–47.

4)Palazzuoli A, Iovine F, Scali C,Nuti R (2006). Acute Coronary Syndromes: From the Laboratory Markers to the Coronary Vessels. *Biomarkers Insight*. **2**, 123-30.

5)Jastrzebska B, Lebel R, Therriault H, McIntyre JO, Escher E, Guérin B, Paquette B, Neugebauer WA, Lepage M (2009). New enzyme-activated solubility-switchable contrast agent for magnetic resonance imaging: from synthesis to in vivo imaging. *Journal of Medicinal Chemistry*. **52**, 6: 1576-81.

Acknowledgement

269