

**Synthesis and fluorescence properties of a novel  
legumain substrate probe**

**Deepthi Ravula**

**Thesis submitted in partial fulfilment of the requirements of Edinburgh Napier  
University for the degree of Master by Research**

**October 2020**

## Abstract

Despite great advancement in the diagnosis and treatment of many cancers, mortality rates remain high and the incidence of some forms of cancer is increasing.

Early diagnosis is recognised as an important factor to increase survival but there is currently a lack of reliable biomarkers of disease.

The focus of this research is the development of a new molecular probe for the putative cancer biomarker legumain (asparaginyl endoprotease) which is overexpressed in a wide variety of cancers and has been proposed as a viable biomarker of disease progression.

A novel FRET peptide substrate AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-piperazine-Rho (DR21) for legumain was designed with a peptide sequence containing a critical asparagine residue (cleavage site) at the P1 position for recognition and cleavage by the target enzyme. Uniquely, the N-terminus was capped by an aminoanthraquinone quencher (acceptor) group and a rhodamine B-derived fluorophore was incorporated at the C-terminus. DR21 was characterised by high resolution mass spectrometry and by NMR spectroscopy and efficient quenching of rhodamine fluorescence was demonstrated.

Proof of principle was demonstrated by fluorescence spectroscopy. DR21 was activated upon incubation with recombinant human legumain with concomitant release of intense rhodamine fluorescence consistent with legumain-induced activation of the fluorogenic probe. The ability to release a hydrophobic derivative of rhodamine has potential advantages over existing fluorogenic substrates due to possessing structural features to allow entry to cells to aid imaging for tumours that express legumain.

## CONTENTS

	<b>Acknowledgements</b>	<b>I</b>
	<b>Nomenclature</b>	<b>ii</b>
	<b>Abbreviations</b>	<b>iv</b>
	<b>List of figures</b>	<b>vi</b>
	<b>List of tables</b>	<b>viii</b>
<b>1</b>	<b>AIM</b>	<b>1</b>
<b>2</b>	<b>INTRODUCTION</b>	<b>2</b>
<b>2.1</b>	<b>CANCER</b>	<b>2</b>
<b>2.2</b>	<b>LEGUMAIN</b>	<b>5</b>
<b>2.2.1</b>	<b>Migratory and invasive properties of legumain.</b>	<b>6</b>
<b>2.3</b>	<b>EXAMPLES OF LEGUMAIN USED AS A BIOMARKER IN CANCER DIAGNOSIS</b>	<b>7</b>
<b>2.3.1</b>	<b>Example-1: Legumain as a biomarker for diagnosis and prognosis of human ovarian cancer.</b>	<b>7</b>
<b>2.3.2</b>	<b>Example-2: Legumain as a target in nanotherapeutic drug delivery system for inhibiting the tumour growth without systemic toxicity.</b>	<b>8</b>
<b>2.3.3</b>	<b>Example-3: In pancreatitis</b>	<b>8</b>
<b>2.3.4</b>	<b>Example-4: In human colorectal cancer</b>	<b>8</b>
<b>2.4</b>	<b>LEGUMAIN PRODRUGS</b>	<b>8</b>
<b>2.4.1</b>	<b>Example-1: Synthesis of a legumain substrate antitumour prodrug of etoposide.</b>	<b>8</b>
<b>2.4.2</b>	<b>Example-2: Legubicin - a legumain activated prodrug.</b>	<b>9</b>
<b>2.4.3</b>	<b>Example-3: EMC-AANL-DOX prodrug.</b>	<b>10</b>
<b>2.5</b>	<b>LEGUMAIN AS A DIAGNOSTIC PROBE</b>	<b>11</b>
<b>2.5.1</b>	<b>ACTIVITY BASED PROBES</b>	<b>11</b>
<b>2.5.2</b>	<b>SUBSTRATE BASED PROBES</b>	<b>12</b>
<b>2.5.2.1</b>	<b>Example-1: Substrate based probe targeting matrix metalloproteinases (MMPs).</b>	<b>13</b>
<b>2.5.2.2</b>	<b>Example-2: Substrate based probe targeting cathepsin-B</b>	<b>14</b>
<b>2.6</b>	<b>COMPARISON BETWEEN SUBSTRATE BASED PROBES AND ACTIVITY BASED PROBES</b>	<b>15</b>

<b>2.7</b>	<b>RHODAMINE-B</b>	<b>16</b>
<b>2.7.1</b>	<b>Example-1: Rhodamine-B based pH probe showing good photostability and membrane permeability</b>	<b>19</b>
<b>2.7.2</b>	<b>Rhodamine-B in FRET</b>	<b>20</b>
<b>2.8</b>	<b>FLUORESCENCE RESONANCE ENERGY TRANSFER (FRET)</b>	<b>20</b>
<b>2.8.1</b>	<b>Applications of FRET</b>	<b>21</b>
<b>3</b>	<b>RATIONAL DESIGN OF LEGUMAIN SUBSTRATE PROBE (DR-21)</b>	<b>23</b>
<b>4</b>	<b>RESULTS AND DISCUSSION</b>	<b>26</b>
<b>4.1</b>	<b>SYNTHESIS OF AMINOANTHRAQUINONE QUENCHER (DR-12)</b>	<b>26</b>
<b>4.1.1</b>	<b>Synthesis of AQ(4-OH)-<math>\beta</math>-Ala-O<sup>t</sup>Bu (DR-11)</b>	<b>26</b>
<b>4.1.2</b>	<b>Synthesis of AQ(4OH)-<math>\beta</math>-Ala-OH (DR-12)</b>	<b>29</b>
<b>4.2</b>	<b>SYNTHESIS OF RHODAMINE-B FLUOROPHORE (DR-16)</b>	<b>31</b>
<b>4.2.1</b>	<b>Synthesis of Rho-Pip-Boc (DR-13)</b>	<b>31</b>
<b>4.2.2</b>	<b>Synthesis of Rho-Pip-TFA (DR-14)</b>	<b>34</b>
<b>4.2.3</b>	<b>Synthesis of Boc-<math>\gamma</math>-Abu-Pip-Rho-TFA (DR-15)</b>	<b>36</b>
<b>4.2.4</b>	<b>Synthesis of <math>\gamma</math>-Abu-Pip-Rho-TFA (DR-16)</b>	<b>38</b>
<b>4.3</b>	<b>SYNTHESIS OF Asn(Trt)-<math>\gamma</math>-Abu-Pip-Rho-TFA FLUOROPHORE (DR-18)</b>	<b>40</b>
<b>4.3.1</b>	<b>Synthesis of Fmoc-Asn(Trt)-<math>\gamma</math>-Abu-Pip-Rho-TFA (DR-17)</b>	<b>40</b>
<b>4.3.2</b>	<b>Synthesis of Asn(Trt)-<math>\gamma</math>-Abu-Pip-Rho-TFA (DR-18)</b>	<b>43</b>
<b>4.4</b>	<b>Synthesis of AQ(4-OH)-<math>\beta</math>-Ala-Pro-Ala-Asn(Trt)-OH (DR-19)</b>	<b>46</b>
<b>4.5</b>	<b>SYNTHESIS OF FLUORESCENT PROBE (DR-21)</b>	<b>50</b>
<b>4.5.1</b>	<b>Synthesis of AQ(4-OH)-<math>\beta</math>-Ala-Pro-Ala-Asn(Trt)-<math>\gamma</math>-Abu-Pip-Rho (DR-20)</b>	<b>50</b>
<b>4.5.2</b>	<b>Synthesis of AQ(4-OH)-<math>\beta</math>-Ala-Pro-Ala-Asn-<math>\gamma</math>-Abu-Pip-Rho (DR-21)</b>	<b>52</b>
<b>4.6</b>	<b>DETERMINATION OF PARTITION COEFFICIENT OF DR-12, DR-16 AND DR-21</b>	<b>53</b>
<b>4.7</b>	<b>DETERMINATION OF MAXIMUM ABSORBANCE OF DR-12 AND DR-16</b>	<b>56</b>

4.8	DETERMINATION OF RELATIVE FLUORESCENCE INTENSITY OF DR-16 AND DR-21	57
4.9	FLUOROGENIC PROBE (DR-21) CLEAVAGE BY rh-LEGUMAIN	59
5	CONCLUSION	61
5.1	Suggestions for future work	62
6	EXPERIMENTAL	63
6.1	MATERIALS AND GENERAL PROCEDURES	63
6.2	SYNTHESIS METHODS	64
6.2.1	Synthesis of AQ(4-OH)- $\beta$ -Ala-O <sup>t</sup> Bu (DR-11)	64
6.2.2	Synthesis of AQ(4-OH)- $\beta$ -Ala-OH (DR-12)	64
6.2.3	Synthesis of Rho-Pip-Boc (DR-13)	65
6.2.4	Synthesis of Rho-Pip-TFA (DR-14)	66
6.2.5	Synthesis of Boc- $\gamma$ -Abu-Pip-Rho-TFA (DR-15)	66
6.2.6	Synthesis of $\gamma$ -Abu-Pip-Rho-TFA (DR-16)	67
6.2.7	Synthesis of Fmoc-Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-17)	68
6.2.8	Synthesis of Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-18)	68
6.2.9	SOLID PHASE PEPTIDE SYNTHESIS: Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)OH (DR-19):	69
6.2.9.1	Synthesis of H-Asn(Trt)-Resin	69
6.2.9.2.	General method for Fmoc amino acid coupling: Synthesis of H-Ala-Asn(Trt)-Resin	69
6.2.9.3	Synthesis of H-Pro-Ala-Asn(Trt)-Resin	70
6.2.9.4	Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)-Resin	70
6.2.9.5.	Cleavage of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)-OH (DR-19) from Resin	71
6.2.10	Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)- $\gamma$ -Abu-Pip-Rho (DR-20)	71
6.2.11	Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21)	72
6.3	DETERMINATION OF PARTITION COEFFICIENTS	72
6.4	UV-VIS ABSORPTION ASSAY	73
6.5	FLUORESCENCE ASSAY	73

<b>6.6</b>	<b>PROBE AQ(4-OH)-<math>\beta</math>-Ala-Pro-Ala-Asn-<math>\gamma</math>-Abu-Pip-Rho (DR-21) ACTIVATION USING RECOMBINANT HUMAN LEGUMAIN (rh- legumain)</b>	<b>74</b>
<b>7</b>	<b>STRUCTURE LIBRARY</b>	<b>76</b>
<b>8</b>	<b>REFERENCES</b>	<b>80</b>

## **Acknowledgements:**

I would like to express my sincere and heartfelt thanks to my supervisor Dr. Agnes Turnbull, Teaching student support tutor, Research fellow and my wonderful personal development tutor, who has helped, supported and encouraged me with her patience, politeness and valuable suggestions throughout this research project.

I would also like to express the deepest appreciation to my Director of Studies Dr. David J. Mincher, Reader in Experimental Chemotherapy Molecular Drug Design Laboratory, who was the genius and has the great attitude of understanding students and helping them. I would like to thank him for giving me this great opportunity to do this research project and for his continuous interest in teaching and soul of adventure in this research field. Without his guidance and persistent help this dissertation would not have been possible.

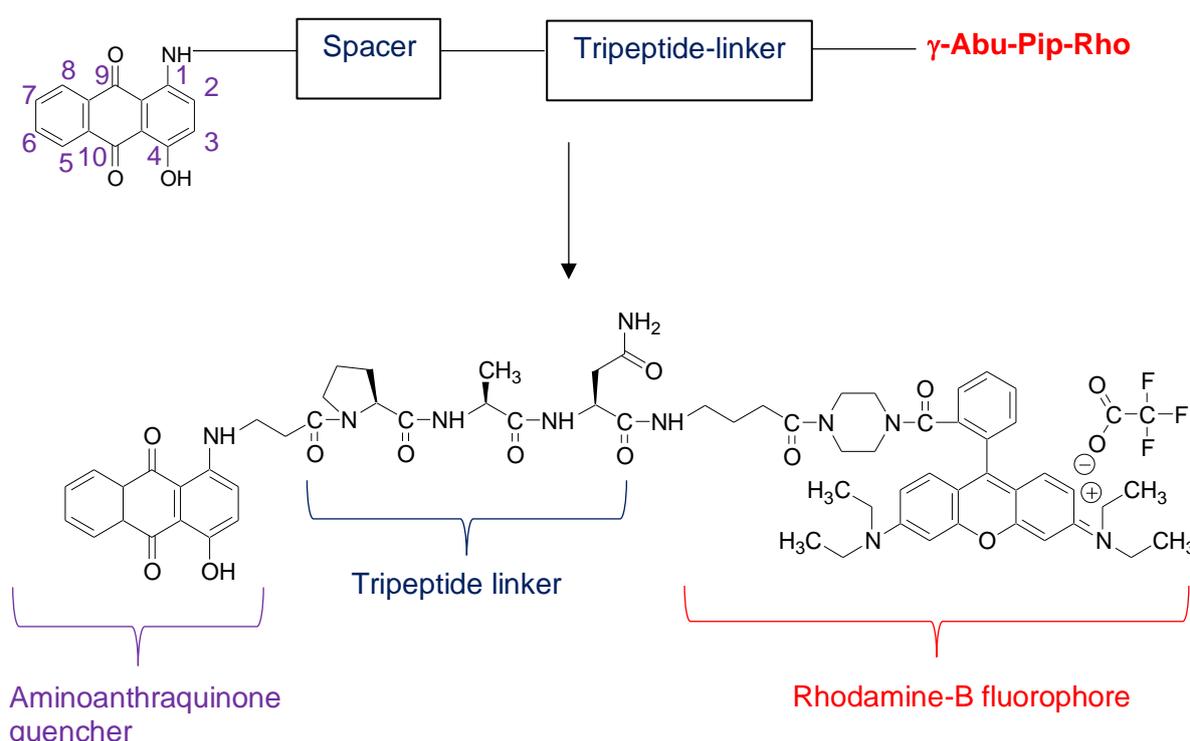
Actually, there are no words to express my thanks to Dr. David J. Mincher and Dr. Agnes Turnbull for their help in every aspect of my education and personal life. But, it is true that without their active guidance, help, cooperation and encouragement, I would not have made headway in the project.

I also acknowledge with a deep sense of reverence, my gratitude towards my husband, my kids and my family members who has always understood and helped me.

## Nomenclature:

In this research, the legumain substrate probe (DR-21) was synthesized, containing three main components; an aminoanthraquinone quencher, a legumain tripeptide substrate and a Rhodamine-B derivative fluorophore.

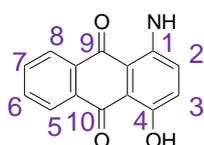
The numbering system for the aminoanthraquinone and rhodamine chromophores are described in this section along with examples of their simplified general structures and abbreviations.



Structural representation of the second generation of fluorogenic probe (DR-21).

## Aminoanthraquinone quencher:

In this research, the aminoanthraquinone was used as a quencher. The general structure and numbering system for the aminoanthraquinone (AQ) is shown below.



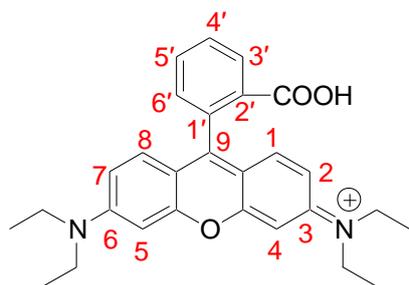
### Legumain tripeptide substrate:

This contains the amino acids alanine (Ala), proline (Pro) and asparagine (Asn).

Pro-Ala-Asn = (H-[Pro-Ala-Asn]-OH)

### Rhodamine-B fluorophore:

The structure of Rhodamine-B is given below, with the numbering system used for NMR signal assignments. The Rhodamine-B is represented by the three-letter abbreviation 'Rho' in this thesis.



**Abbreviations:**

ACN	Acetonitrile
AEP	Asparaginyl endopeptidase
Ala	Alanine
AQ	Anthraquinone
Asn	Asparagine
$\beta$ -Ala	$\beta$ -Alanine
DCM	Dichloromethane
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMF	<i>N,N'</i> -Dimethylformamide
DMSO	Dimethyl sulfoxide
Eq	Molar equivalent
ESI	Electrospray ionisation
Fmoc	Fluorenylmethoxycarbonyl
FRET	Fluorescence Resonance Energy Transfer
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxihexafluorophosphate
HOBt	Hydroxybenzotriazole
g	gramme
h	hour
M	Molar
mM	Millimolar
mg	Milligram(s)
min	Minutes
MS	Mass spectrum

MRI	Magnetic Resonance Image
NMR	Nuclear Magnetic Resonance
Pip	Piperazine (residue)
ppm	Parts per million
Pro	Proline
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
Rho	Rhodamine-B
RT	Room temperature
R <sub>f</sub>	Retention factor
<sup>t</sup> Boc	Tertiarybutoxycarbonyl
<sup>t</sup> Bu	Tertiarybutyl
TAM	Tumour associated macrophage
TFA	Trifluoroacetic acid
TLC	Thin layer chromatography
Trt	Trityl
UV	Ultraviolet

## List of figures

<b>Figure-1:</b>	An outline of legumain substrate probe DR-21 design and its concept of cleavage by legumain at the P1 site.	1
<b>Figure-2:</b>	Structure of second generation fluorogenic probe (DR-21) with rhodamine fluorophore (right) and aminoanthraquinone quencher (left) linked by tripeptide Pro-Ala-Asn.	1
<b>Figure-3:</b>	Diagrammatic representation of role of legumain in invasive and migratory properties of tumour cells	7
<b>Figure-4:</b>	Chemical structure of antitumour prodrug CBZ-Ala-Ala-Asn-ethylenediamine-etoposide	9
<b>Figure-5:</b>	Structure of prodrug Legubicin	10
<b>Figure-6:</b>	Structure of legumain activated doxorubicin prodrug.	10
<b>Figure-7:</b>	Schematic diagram of Activity based probe	12
<b>Figure-8:</b>	Structure of LP-1	12
<b>Figure-9:</b>	Schematic representation of MMP-2 protease activity using a coumarin-dabcyl fluorogenic substrate probe	13
<b>Figure-10:</b>	Diagrammatic representation of substrate based probe activity targeted to cathepsin	14
<b>Figure-11:</b>	Schematic representation of Cathepsin-B protease activity probed by Cbz-Lys-Lys-PABC-AMC	15
<b>Figure-12:</b>	Chemical structures of Rhodamine-B	17
<b>Figure-13:</b>	Cyclization of rhodamine amides	18
<b>Figure-14:</b>	Synthesis of rhodamine tertiary amide derivatives	18
<b>Figure-15:</b>	Chemical reaction involved in spirolactam ring closure and opening (in acidic condition) in Rhodamine-B probe (RCE)	19
<b>Figure-16:</b>	Schematic diagram illustrating the mechanism of energy transfer in FRET	21
<b>Figure-17:</b>	Schematic diagram of activation of SM9	22
<b>Figure-18:</b>	An outline of legumain substrate probe DR-21 design and its concept of cleavage by legumain at the P1 site	25
<b>Figure-19:</b>	<sup>1</sup> H NMR assignment of signals in DR-11	28
<b>Figure-20:</b>	<sup>1</sup> H NMR assignment of signals in DR-12	30

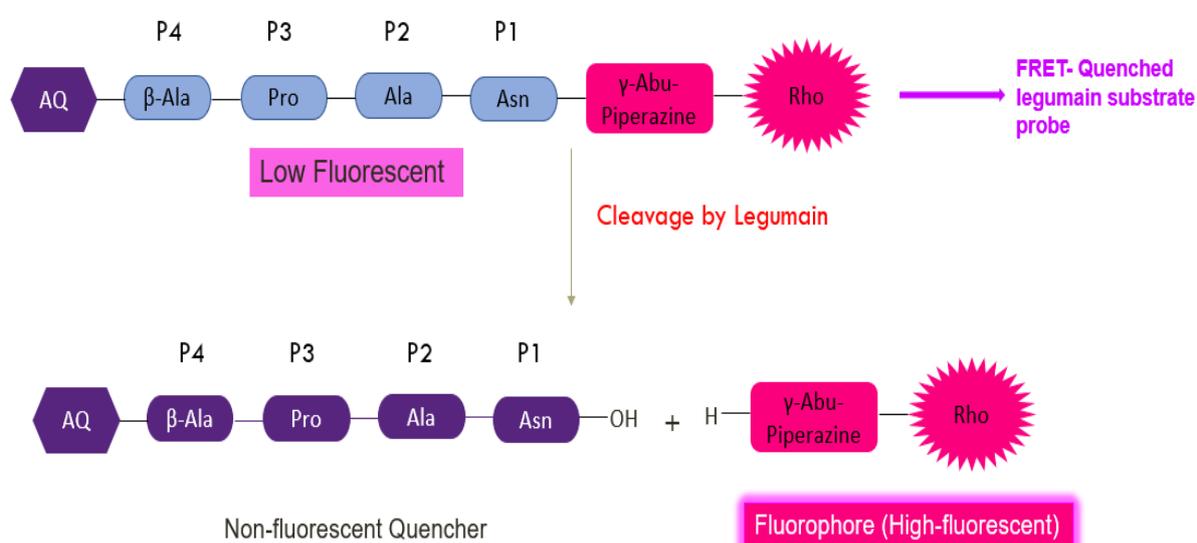
<b>Figure-21:</b>	1H NMR assignment of signals in DR-13	33
<b>Figure-22:</b>	1H NMR assignment of signals in DR-14	35
<b>Figure-23:</b>	1H NMR assignment of signals in DR-15	37
<b>Figure-24:</b>	1H NMR assignment of signals in DR-16	39
<b>Figure-25:</b>	1H NMR assignment of signals in DR-17	42
<b>Figure-26:</b>	A schematic representation of steps involved in synthesis of aminoanthraquinone-labelled tetrapeptide DR-19 by solid phase peptide synthesis method	46
<b>Figure-27:</b>	Structure of HZ22 reagent	47
<b>Figure-28:</b>	1H NMR assignment of signals in DR-19	49
<b>Figure-29:</b>	Relative lipophilicity/hydrophilicity of DR-12, DR-16 and DR-21	55
<b>Figure-30:</b>	Absorption spectra of AQ(4OH)- $\beta$ -Ala-OH (DR-12) and $\gamma$ -Abu-Pip-Rho-TFA (DR-16) (absorbance vs wavelength) in assay buffer at pH 5.0	57
<b>Figure-31:</b>	Emission spectra of $\gamma$ -Abu-Pip-Rho-TFA (DR-16) and AQ(4OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) (relative fluorescence intensity vs wavelength) in assay buffer at pH 5.0	58
<b>Figure-32:</b>	Incubation of DR-21 with rh-legumain in MES assay buffer (pH 5.0) at 37°C, $\lambda_{ex}$ 480 nm	59

## List of Tables

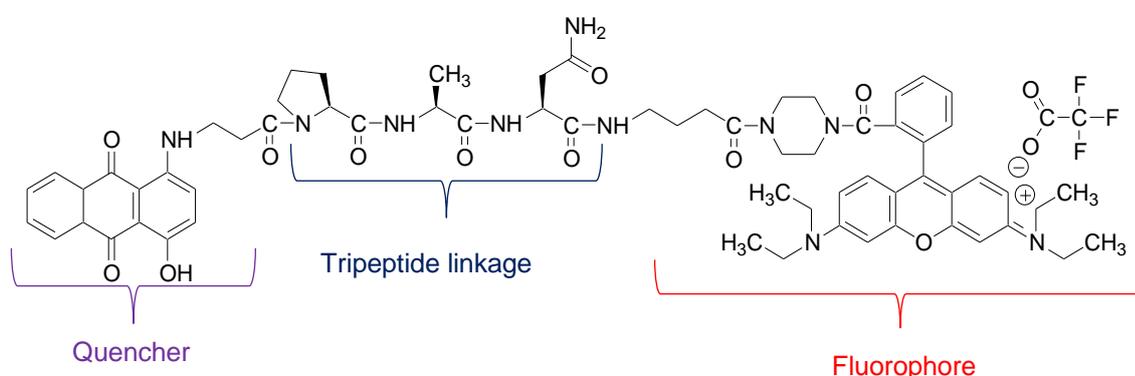
<b>Table-1:</b>	SwissADME predicted partition coefficient values	54
<b>Table-2:</b>	Determination of maximum absorbance of DR-12 and DR-16	56
<b>Table-3:</b>	Determination of Relative fluorescence intensity of DR-16 and DR-21	58

## 1. AIM

The aim of this research project is to synthesise a new chemical probe for cancer diagnosis and as a potential aid for planning the proper and efficient treatment method for the cancer patient (patient monitoring). This FRET probe was designed with the fluorophore, quencher and a Pro-Ala-Asn tripeptide linkage, to target a potential biomarker, legumain (which is overexpressed by cancer cells) for its proteolytic activity with the indication of probe activation via emission of fluorescence by the liberated fluorophore (Figures 1 and 2).



**Figure-1:** An outline of legumain substrate probe DR-21 design and its concept of cleavage by legumain at the P1 site.



**Figure-2:** Structure of second generation fluorogenic probe (DR-21) with rhodamine fluorophore (right) and aminoanthraquinone quencher (left) linked by tripeptide Pro-Ala-Asn.

## 2. INTRODUCTION:

### 2.1. CANCER:

The uncontrollable division and growth of cells is the major cause for formation of a tumour (Kim *et al.*, 2016). Cancer is characterised by cell modifications that include uncontrolled cell growth, morphological and cellular transformation, angiogenesis, dysregulation of apoptosis and metastasis. Cancer is the second most death-causing disease after cardiovascular diseases worldwide (Yousafzai *et al.*, 2018).

Tumours are divided into two types (Cooper & Sunderland, 2000): 1. Benign tumours: these tumours are restricted to one part of the body and can usually be removed by surgery. 2. Malignant tumours: these tumours spread to normal healthy tissues and circulate throughout the body by the circulatory system and are aggressively cancerous. This feature describes the process of metastasis.

**Treatment of cancer:** The cancer can be treated by different methods based on the type and stage of cancer. Such cancer treatment methods include: chemotherapy, radiotherapy, hormone therapy, immune therapy, stem cell transplant (in leukaemia), surgery and targeted therapy (examples: small-molecule drugs and monoclonal antibodies).

There emerged the development of chemotherapy in cancer treatment in the 1940s (Kim *et al.*, 2016). Chemotherapy is found as the one of the most important and efficient method of cancer treatments (Yousafzai *et al.*, 2018). Amongst the early compounds, N-mustard (which acts by binding to the DNA and blocks the gene replication necessary for the cell division thereby inhibits the cell division) was found to be a first chemical substance used against hematologic cancers by Goodmann and Gillan (Kim *et al.*, 2016). Later, the anticancer drugs like the antifolate aminopterin (discovered by Farber used for leukaemia) acts by blocking the folic acid metabolism thereby inhibits the cell division, 5-fluorouracil (discovered by Heidelberger in 1957) is an antimetabolite incorporated into RNA and DNA; vincristine (a vinca plant-derivative alkaloid developed in 1958) acts by suppressing the formation of cytoskeleton in the cell thereby blocking the cell division (Kim *et al.*, 2016) were developed. All of these agents are extremely systemically toxic and without selectivity for tumour cells.

Further investigation revealed the use of single types of anticancer drugs resulted in development of severe drug resistance and recurrence of cancer (Kim *et al.*, 2016). Development of chemotherapeutic drug resistance is found to be a common phenomenon during the treatment of primary and secondary tumours (Yousafzai *et al.*, 2018) and is the major factor in the failure of chemotherapy to effect cure.

Multidrug resistance (MDR) is known as a phenomenon in which the cancer cells are found to be sensitive at the initial stage of treatment with the anticancer drugs but later these cancer cells change their phenotype to become resistant towards multiple unrelated drugs which are structurally and functionally different and also may have different molecular targets (Kumar & Jaitak, 2019). The multidrug resistance is of two types: 1. Intrinsic MDR: occurs due to various characteristics of cells, 2. Acquired MDR: occurs due to genetic alterations of cells during chemotherapy (Kumar & Jaitak, 2019).

Some of the mechanisms that are responsible for the cause of MDR include: 1. Evolving adaptation of cancer cells to the microenvironment, 2. Mutations of oncogenes that become resistant to previous treatments, 3. ATP-binding cassette of transmembrane glycoproteins that expel drugs from the cell (overexpressed efflux pumps), 4. Survived cancer stem cells that escape from conventional therapies and 5. Activated cell growth factors and defence systems. Examples of some of the drugs susceptible to the development of drug resistance are doxorubicin, paclitaxel, colchicine etc (Ye *et al.*, 2019).

The problem of MDR with the older drugs continues to more recent treatment regimens. Xiang and his colleagues (2020) explained the drawbacks of chemotherapy in pancreatic cancer in elderly patients with the monotherapy of gemcitabine and combination therapy of gemcitabine with abraxane (Folfirinox) (Li *et al.*, 2020). This study described the efficacy of drugs gemcitabine and folfirinox towards advanced prostate cancer along with most common adverse effects like neutropenia, fatigue, anaemia, thrombocytopenia, vomiting, diarrhoea, sensory neuropathy, thromboembolism (Li *et al.*, 2020).

Therefore, the lack of selectivity in the chemotherapeutic treatment of cancer made the cancer a life threatening disease. Improvement in survival rates for cancer is known to associated with improved early diagnosis. There has evolved the development of

many diagnostic techniques (Ahdoot *et al.*, 2020; Feng *et al.*, 2017) and the use of biomarkers (Chen *et al.*, 2017; Merchant *et al.*, 2017) has also been important in diagnosis of cancer in early stages. This research project is focussed on this objective.

**Diagnosis of Cancer:** Cancer can be diagnosed by the use of various techniques such as: 1. Imaging procedures: MRI (magnetic resonance imaging) for example, for breast cancer (Ahdoot *et al.*, 2020), nuclear scanning, PET-CT scan (positron emission tomography) for lung cancer (Feng *et al.*, 2017), prostate cancer (Kallur *et al.*, 2017) where most cancers are diagnosed by biopsy. 2. Laboratory tests: high or low levels of certain substances (for example, enzymes or vitamin receptors) expressed by the cells can indicate a sign of cancer.

Some of the vitamins are used as radioimmunoimaging agents whereas some of the enzymes are used in the diagnosis of cancer by their overexpression in the cancer cells. Biotin is used in pre-targeted radioimmunoimaging with a biotinylated D-D<sub>3</sub> construct in lung cancer (Hong *et al.*, 2020), and riboflavin was used in the prognosis of esophageal squamous cell carcinoma (Li *et al.*, 2017).

Amongst the enzymes expressed in the tumour microenvironment, legumain was overexpressed in pancreatitis (Edgington-Mitchell *et al.*, 2016), ovarian cancer (Wang *et al.*, 2012) and prostate cancer.

Cathepsins, such as cathepsin-B, D, E and L are known to be overexpressed by most of the digestive cancer cells except liver cancer cells (Chen *et al.*, 2017); matrix metalloproteinases in lung cancer (Merchant *et al.*, 2017) and in gastric cancer (Yang *et al.*, 2017).

## 2.2. LEGUMAIN:

Even though there are major advances in cancer diagnosis and treatment, there remains a lack of biomarkers for early detection of cancer. Now-a-days legumain is found to be a potential biomarker for cancer and also a molecular target for imaging and drug targeting in cancer diagnosis and treatment.

Legumain, a cysteine endopeptidase named by Kembhavi *et al.* (1993) was found in many leguminous and other seeds, after they had isolated and characterized from *Vigna aconitifolia* (moth bean) (Chen *et al.*, 1997). The amino acid sequence of legumain from *Ricinus communis* (castor bean) and that of an enzyme derived from the fluke *Schistosoma mansoni* was found to be homologous, but the specificity of fluke enzyme was unknown initially whereas later on it was also found to be an asparaginyl endopeptidase on a test substrate which was introduced by Kembhavi (Chen *et al.*, 1997).

According to the MEROPS database peptidase classification (Rawlings *et al.*, 2018), legumain, a cysteine asparaginyl endopeptidase, belongs to the C13 family members, whereas all other lysosomal cysteine proteases identified to date, e.g. the cathepsins, are grouped in the C1 family. All lysosomal endopeptidases which were discovered prior to mammalian legumain are found to show a wide-ranging action on proteins, so the strict specificity of action of legumain towards asparagine bonds is found to be prominent (Gawenda *et al.*, 2007) with a remarkably restricted specificity, as it cleaves only substrate sequences having an asparagine (Asn) at the P1 site (Gawenda *et al.*, 2007).

According to the study of Liao (2011), legumain is also used as an ideal targeted ligand because of its following characteristics (Liao *et al.*, 2011):

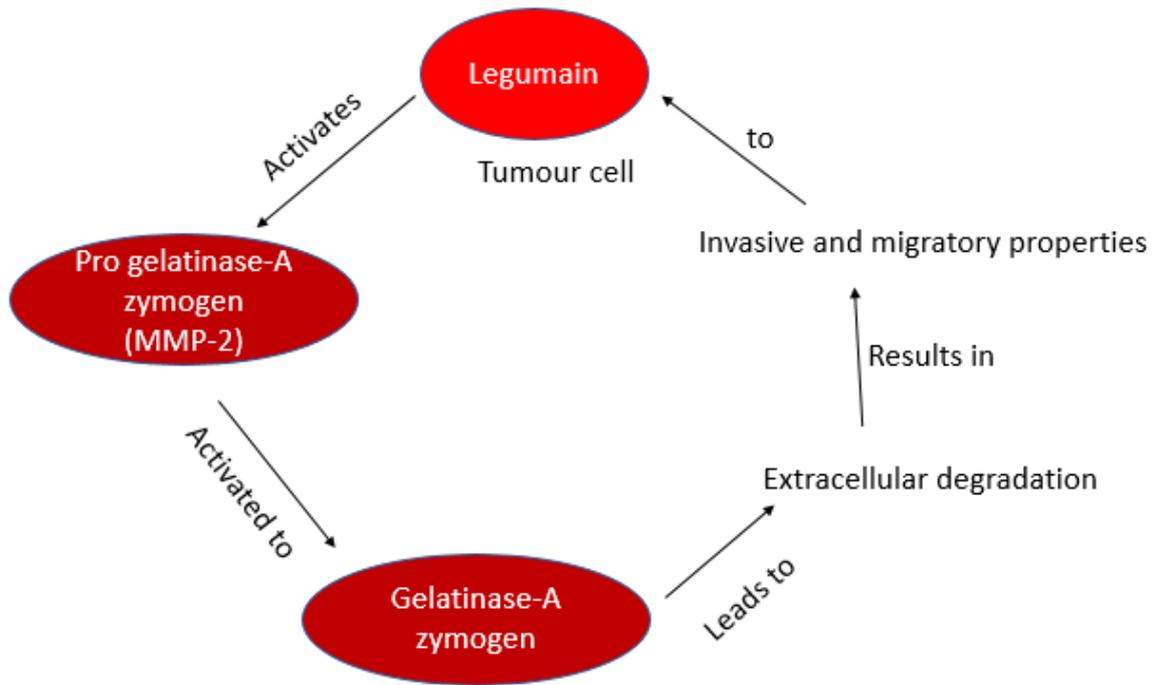
- i) overexpression by a majority of human solid tumours, including breast, colon and prostate tumours,
- ii) sparsely expressed in normal tissues,
- iii) overexpression in vivo by both tumour cells and proliferating endothelial cells in the tumour microenvironment (TME).

Legumain, was found in a number of mammalian tissues, such as the kidney, placenta, spleen, liver, and testis. Among these organs, legumain shows a higher

specific activity in the kidney (Morita *et al.*, 2007). Moreover, legumain was also found to be highly expressed in several types of tumours such as colon, prostate (Spiciarich *et al.*, 2017), breast (Murthy *et al.*, 2005), gastric cancers (Zhang *et al.*, 2016) and also *in vivo* in the murine colon carcinoma models. The legumain expression was found to be up-regulated during tumour development *in vivo*, signifying an environmental response, and which resembles a stress-responsive gene, being evidently elevated in especially cells that are subjected to environmental stress (Murthy *et al.*, 2005).

### **2.2.1. MIGRATORY AND INVASIVE PROPERTIES OF LEGUMAIN:**

Legumain is found to be expressed both intracellularly and on the cell surface of the tumour cells and also in the tumour associated endothelial cells (Murthy *et al.*, 2005; Guo *et al.*, 2013). Legumain is mainly found at the invadopodia, especially in membrane-associated vesicles of tumour cells. It is known that the cells expressing legumain show the migratory and invasive properties in the metastatic tumour cell surface. These migratory and invasive properties are mediated by increased extracellular matrix degradation, which occurs due to the activation of zymogens (zymogen-A) such as progelatinase-A (pro-MMP-2) (Murthy *et al.*, 2005; Guo *et al.*, 2013) by legumain in metastatic tumour cells (Figure-3). Animal tumour models generated with cells overexpressing legumain showed an *in vivo* behaviour of increased invasive growth and metastasis that is due to the phenotype which is a result from the proteolytic function of legumain. Hence, legumain acts as a target for inhibition of tumour growth and metastasis based on its enhancement of tumour growth and its unique restricted specificity (Murthy *et al.*, 2005; Guo *et al.*, 2013). The auto-activation of legumain leads to the dramatic change in its pH-stability profile such as the pro-legumain was found to be stable at near neutral pH, the auto-catalytically activated legumain was found to be stable at acidic pH, as found in the endo-lysosomal system, and then it is found to be destabilized at pH > 6.0 (Dall & Brandstetter, 2016).



**Figure-3:** Diagrammatic representation of role of legumain in invasive and migratory properties of tumour cells.

### 2.3. EXAMPLES OF LEGUMAIN USED AS A BIOMARKER IN CANCER DIAGNOSIS:

#### 2.3.1. Example-1: Legumain as a biomarker for diagnosis and prognosis of human ovarian cancer:

Although there are several biomarkers such as CEA (carcinoembryonic antigen) and CA19-9 (carbohydrate antigen) which are used in the diagnosis of ovarian cancer in clinical practice, some of the drawbacks in these detection methods, such as poor sensitivity and specificity which are more important aspects in early detection of potentially curable lesions, are not suitable for large-scale screening of the population.

Therefore, there is a pressing requirement to discover new or supplementary biomarkers to improve diagnosis and treatment of ovarian cancer.

In one study, the expression levels of proteins were compared between ovarian cancer and normal ovarian tissue, where the results indicated that legumain is over-expressed in ovarian cancer and upregulated expression of legumain increases ovarian cancer

cell migration and invasion in vitro. These results suggested that legumain is a potential novel biomarker for detection and prognosis of ovarian cancer (Wang *et al.*, 2012).

**2.3.2. Example-2: Legumain as a target in nanotherapeutic drug delivery system for inhibiting the tumour growth without systemic toxicity:** In the study by Li (2013) and his colleagues, they stated that the overexpression of legumain in gastric cancer compared to normal healthy cells resulted in increased in vivo metastasis and in vitro cell migration. They concluded that legumain was a potential novel biomarker for diagnosis and prognosis of gastric cancer (Li *et al.*, 2013).

**2.3.3. Example-3: In pancreatitis:** According to this study, legumain was found to be highly expressed especially in CD68 cells (cluster of differentiation 68 cells) of macrophages of the chronic pancreatitis compared to non-inflamed pancreatic cells (Edgington-Mitchell *et al.*, 2016). Hence, legumain acts as a potential biomarker in diagnosis of chronic pancreatitis.

**2.3.4. Example-4: In human colorectal cancer:** The study of Haugen (2015) and colleagues also examined the higher levels of legumain in tumour and stromal cells of colorectal cancer patients (Haugen *et al.*, 2015).

## **2.4. LEGUMAIN PRODRUGS:**

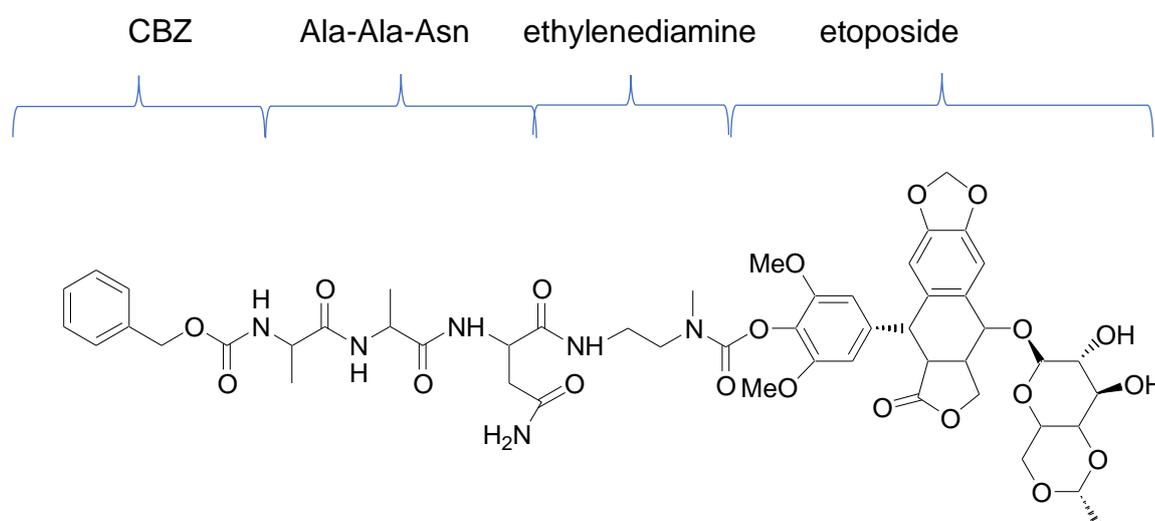
Legumain prodrugs are the inactive form of drugs which are conjugated with the substrate peptide (containing legumain cleavable amino acids) that upon hydrolysis by legumain results in the release of the active form of the drug. Thus, this mechanism of drug release from its prodrug results in reduced side effects and increased potency of the anticancer drugs, which are desired actions in cancer treatment. Examples of legumain prodrugs are discussed below:

### **2.4.1. Example-1:**

#### **Synthesis of a legumain substrate antitumour prodrug of etoposide:**

This study explained the synthesis of a prodrug which covalently binds a short peptide sequence etoposide (a chemotherapeutic agent that is still widely used in the clinic) to form a legumain substrate, CBZ-Ala-Ala-Asn-ethylenediamine-etoposide to show its

antitumour activity (Stern *et al.*, 2009). In this study, the prodrug was designed in such a way that it is cleaved by legumain in acidic conditions in cell lysosomes, at the C-terminus of the asparagine residue which is bound to ethylenediamine through a carbamate bond, creating an ethylenediamine–etoposide moiety. According to the study of Perry *et al.*, (2007), this cleavage occurs due to the cyclization of the ethylenediamine linker at physiological pH, and finally releases free etoposide in the cell cytoplasm (Figure-4) (Stern *et al.*, 2009).

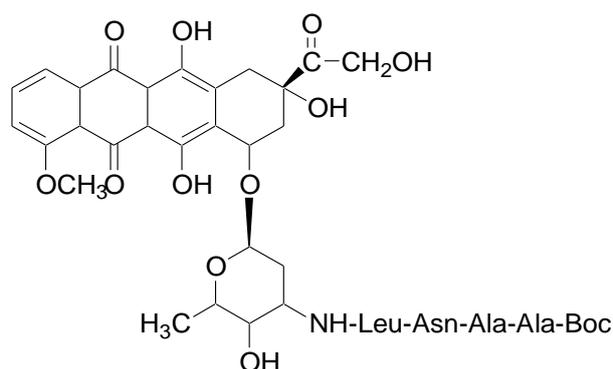


**Figure-4:** Chemical structure of antitumour prodrug CBZ-Ala-Ala-Asn-ethylenediamine-etoposide.

#### 2.4.2. Example-2:

##### Legubicin - a legumain activated prodrug:

A prodrug analogue was synthesised with doxorubicin where an asparaginyl endopeptidase substrate peptide Boc-Ala-Ala-Asn-Leu was added to the amino group of doxorubicin at the COOH terminal of leucine by means of a peptide bond to give a conjugated compound called (N-(*t*-butoxycarbonyl-L-alanyl-L-alanyl-L-asparaginyl-L-leucyl) doxorubicin) (Figure-5) which upon hydrolysis by legumain results in the conversion of prodrug into an active molecule, leucine-dox (leu-dox) which is responsible for an efficient cytotoxic activity. Moreover, the Boc group which is present at the NH<sub>2</sub> terminal inhibited the hydrolysis of the aminopeptidase peptidyl component (Liu *et al.*, 2003).

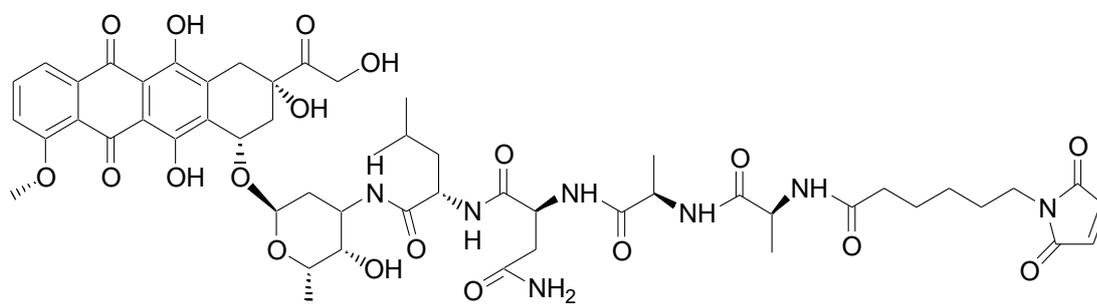


**Figure-5:** Structure of prodrug Legubicin.

### 2.4.3. Example-3:

#### EMC-AANL- DOX prodrug:

EMC-AANL-DOX(6-maleimidocaproyl-alanyl-alanyl-asparaginyl-leucine-Doxorubicin) (Figure-6) is a novel legumain based prodrug which showed a high selectivity and specificity towards paediatric neuroblastoma with less side effects. This prodrug is activated by the enzymatic hydrolysis of the amide bond between the leucine and asparagine by legumain thereby resulting in the release of the active molecule, leucine-doxorubicin into to the tumour cells to produce its cytotoxic action. And after cleavage the separated EMC (maleimide) group not only acts as a protecting group but also binds to the human serum albumin which is a major source of energy for the metabolism of the tumour cells (Zhang *et al.*, 2018).



**Figure-6:** Structure of legumain activated doxorubicin prodrug.

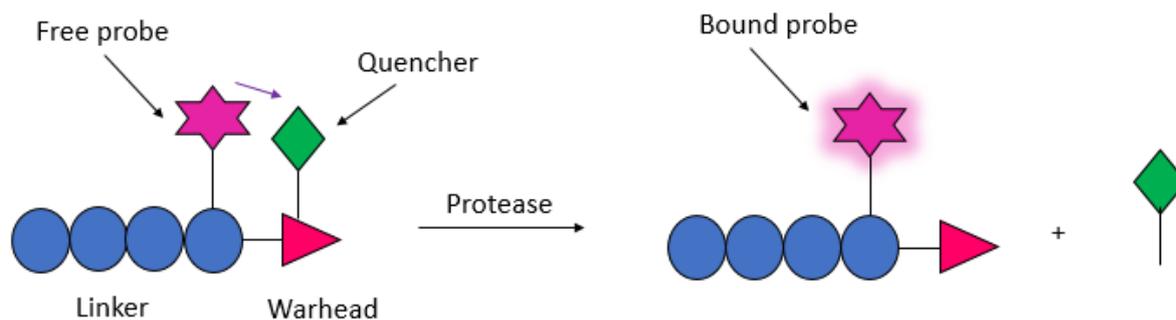
## 2.5. LEGUMAIN AS A DIAGNOSTIC PROBE:

The overexpression of legumain in tumour cells and its restricted specificity for the enzymatic cleavage at the P1 position of the asparagine made legumain a potential biomarker for designing probes for cancer diagnosis and treatment. The use of fluorescent probes to image the proteases in living cells helped to design desired specific targeted fluorescent probes. There are two types of probes:

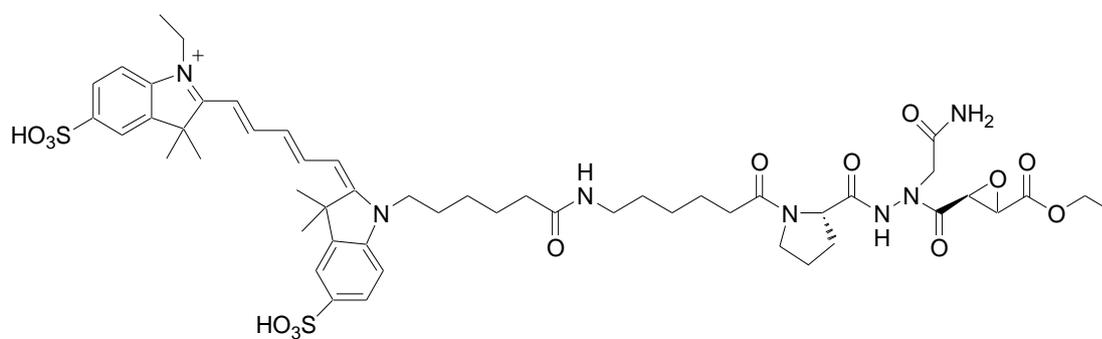
1. Activity based probe
2. Substrate based probe

### 2.5.1. ACTIVITY BASED PROBES:

Activity based probes are small molecules which upon cleavage by the protease, irreversibly bind to the active protease by a covalent bond and change the active site of the enzyme and produces the inhibitory action (Lee & Bogyo, 2010) (Figure-7). Irreversible legumain-specific inhibitors have been synthesised containing a Cbz Ala-Ala-Asn peptide scaffold (based on the sequence of a known substrate of legumain) and reactive electrophilic functional groups such as aza-Asn halomethylketones, aza-Asn epoxides, and aza-Asn Michael acceptors. These inhibitors are known to be highly potent *in vitro* but their potency and selectivity *in vivo* was unknown. Hence, a novel class of aza-Asn epoxide ABPs with the fluorophore Cy5 and a sequence of cell-permeable groups to target the legumain (Lee & Bogyo, 2010) was developed. The aza group of aza-Asn epoxide (LP-1) (Figure-8) ABPs acts as a unique scaffold which permits the addition of P1 Asn residue without interfering the stability of the compound. These aza-Asn epoxide ABPs are also found to be an efficient legumain inhibitors with very low cross-reactivity towards the cathepsins and caspases. Hence all these features made the aza-Asn epoxide ABPs for use as imaging probes (Lee & Bogyo, 2010).



**Figure-7:** Schematic diagram of Activity based probe. [Adapted and modified from Lee & Bogoy, 2010].



**Figure-8:** Structure of LP-1

### 2.5.2. SUBSTRATE BASED PROBES

Substrate-based probes are probes that produce a fluorescence signal after their cleavage by the protease enzyme. The proteases used for cleavage of the substrate probes includes the cysteine proteases (eg: legumain, cathepsin (Chowdhury *et al.*, 2014) matrix metalloproteinases (He *et al.*, 2017), serine proteases. Due to the overexpression of these proteases in tumours, this led to the design of a large number of substrate probes with a quencher and fluorophore which are used in the imaging of cancer cells.

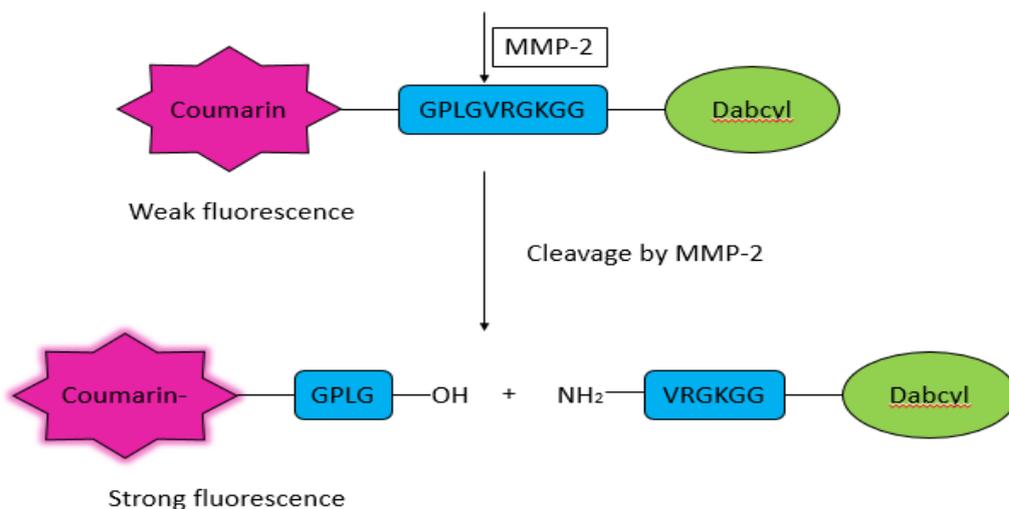
Selected examples of substrate probes that were designed based on matrix metalloproteinases and cathepsin proteases to image the cancer cells are given below.

### 2.5.2.1. Example-1:

#### Substrate based probe targeting matrix metalloproteinases (MMPs):

The expression of matrix metalloproteinases was found to be highly increased in various tumour cells and MMP-2 was found to be expressed in inactive zymogen form and converted to active MMP-2 by MT1-MMP which finally results in the tumour invasion, metastasis and angiogenesis. Hence the proteolytic activity of MMP-2 and MT1-MMP has led to the design of a molecular imaging probe for determining the tumour malignancy (He *et al.*, 2017).

A novel coumarin based fluorophore substrate probe was synthesised for screening, diagnosis and prognosis of cervical cancer. This coumarin based probe was prepared by coupling the fluorophore (Coumarin) with the quencher (Dabcyl-OSu) through a polypeptide chain (Coumarin-Gly-Pro-Leu-Gly-Val-Arg-Gly-Lys-Gly-Gly-Dabcyl) which is cleaved by MMP-2. Here, the emission spectrum of Coumarin and the absorption spectrum of Dabcyl were overlapped which formed the basic concept for the synthesis of this coumarin based fluorescence FRET probe (He *et al.*, 2017). In this probe MMP-2 protease cleaves the core MMP-2 substrate polypeptide chain between Glycine (G) and Valine (V) results in the separation of coumarin fluorophore and the quencher in the cancer cells (Figure-9) where the coumarin emitted increased fluorescence at 480 nm after cleavage from quencher (He *et al.*, 2017).



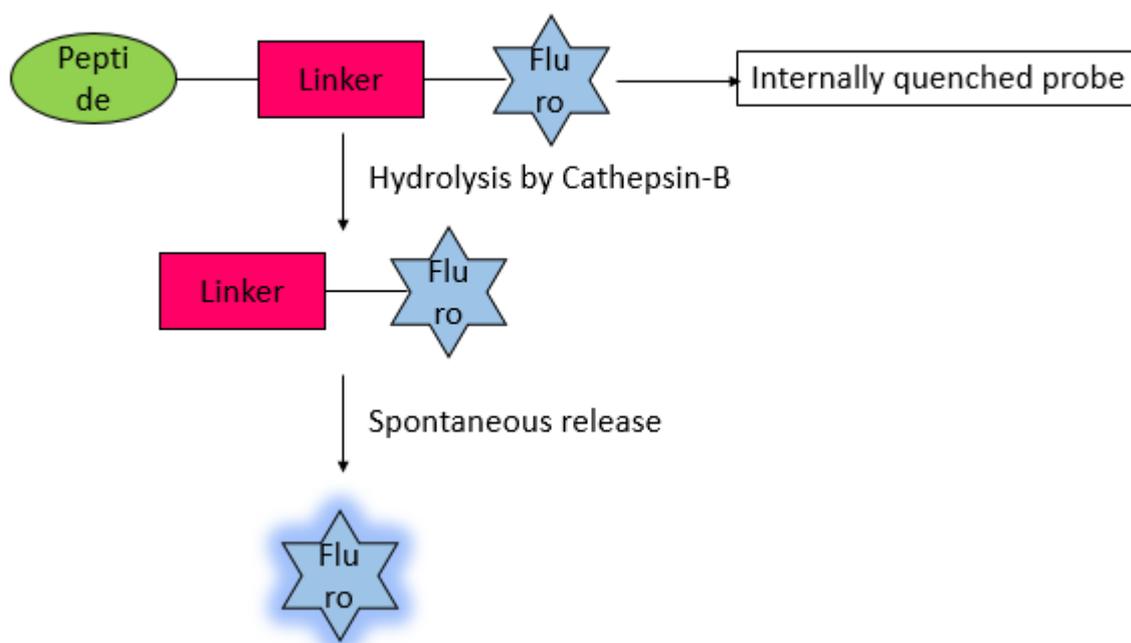
**Figure-9:** Schematic representation of MMP-2 protease activity using a coumarin-dabcyl fluorogenic substrate probe.

### 2.5.2.2. Example-2:

#### Substrate based probe targeting cathepsin-B:

Cathepsin B protease is found to be one of the most notable biomarkers used in the diagnosis and prognosis of various cancers because of its proteolytic activity (Chowdhury *et al.*, 2014). This literature describes a three component model prodrug-inspired strategy for the synthesis of a profluorophore where the peptide portion showed a greater affinity towards CTB (Cathepsin-B) when conjugated with the linker and a fluorophore (Figure-10).

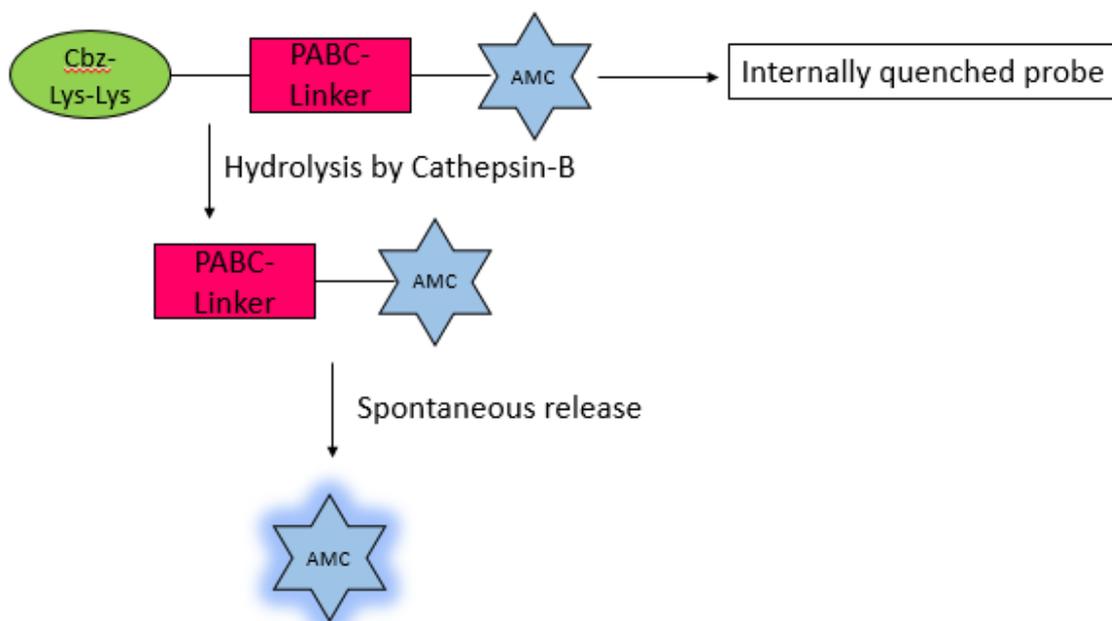
The three component model is designed in such a way that the CTB (which shows specificity towards the hydrophobic amino acid at P2 and a positively charged substituent at the P1 positions) cleaves the substrate molecule and results in the release of the fluorophore (Figure-12) (Chowdhury *et al.*, 2014).



**Figure-10:** Diagrammatic representation of substrate based probe activity targeted to cathepsin [Adapted and modified from Chowdhury *et al.*, 2014].

A further example of such probes includes **Cbz-Lys-Lys-PABC-AMC**. In this probe, the AMC acts as the fluorophore. PABC (p-aminocarbamate) is a self-destructive linker and benzyloxycarbonyl (Cbz)-Lys-Lys acts as a substrate which targets the CTB

(Chowdhury *et al.*, 2014). The CTB proteolytically cleaves the substrate molecule and thereby releases the fluorophore with a self-destructive linker which results in the restoration of fluorescence emitted by the highly fluorescent fluorophore (Figure-11) at the longer wavelength (480 nm).



**Figure-11:** Schematic representation of Cathepsin-B protease activity probed by Cbz-Lys-Lys-PABC-AMC.

## 2.6. COMPARISON BETWEEN SUBSTRATE BASED PROBES AND ACTIVITY BASED PROBES:

In the case of substrate-based probes, the specificity towards the targeted protease is monitored by the peptide recognition sequence. For example, the caspases show specificity towards the P1 of aspartic acid, legumain shows the specificity towards the P1 of asparagine (Blum *et al.*, 2009). One of the most important advantages with the use of substrate-based probes is the generation of signal amplification over time because of the condition where many substrates were processed by the single active protease and this amplification is found to be a most desired property.

Whereas, the activity based probes are known to have a value for non-invasive optical imaging applications. These activity based probes were controlled by both the peptide selectivity sequence and the type of reactive functional group or warhead used in the

probe (Blum *et al.*, 2009). The most important aspect of activity-based probes is that they form a covalent bond to the active protease of interest, which made them easy for biomedical analysis of targets once their *in vivo* imaging was performed. Moreover, because the ABP 's are small in size with a very short half-life *in-vivo*, these probes are quickly circulated and cleared in order to give high contrast images. These probes do not produce any signal amplifications (Blum *et al.*, 2009).

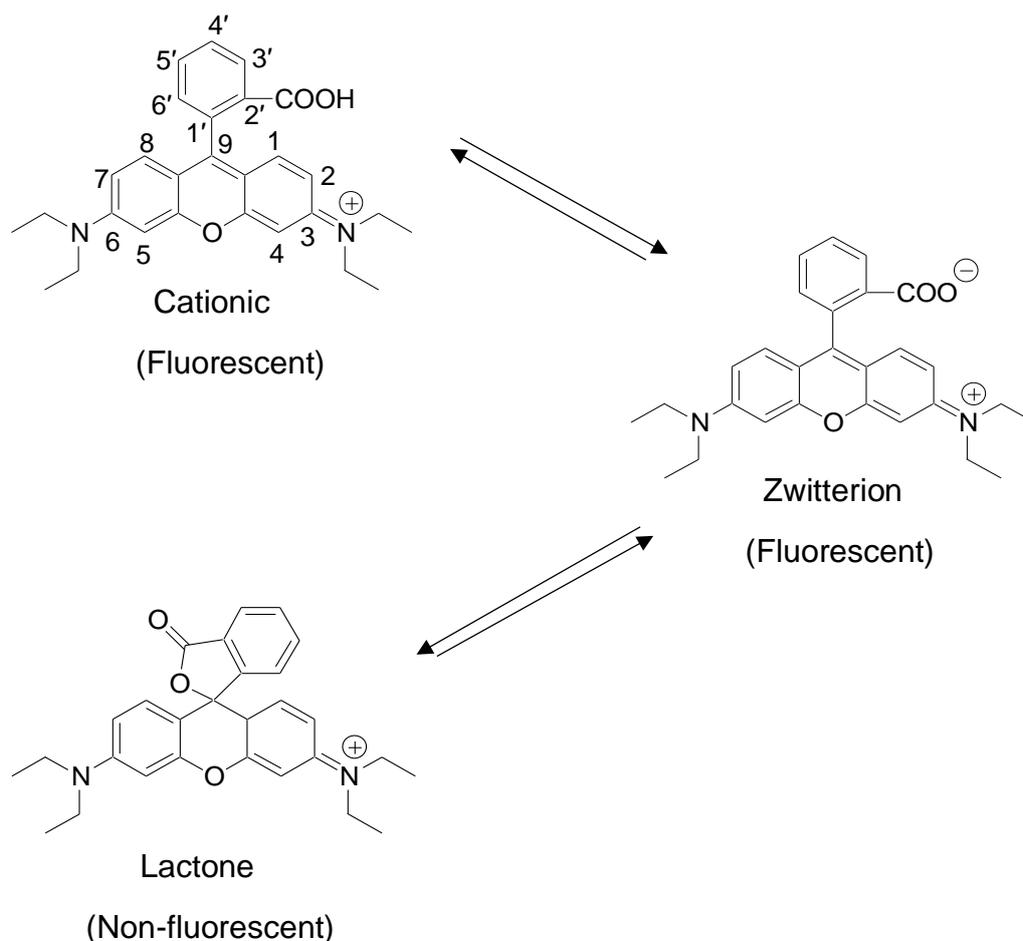
## **2.7. RHODAMINE-B:**

Rhodamine dyes are fluorophores, belonging to the xanthene family along with the fluorescein and eosin dyes. The rhodamine dyes are used as laser dyes, pigments, polymer-bioconjugate detectors and also in designing fluorescent probes because of their photophysical properties and resistance towards the radiant energy of light (Beija *et al.*, 2009). Rhodamine-B was found to show high quantum yield, is commercially available and less expensive than other dyes.

Rhodamine derivatives, because of their extremely favourable photophysical properties and biocompatibility were found to be most widely used dyes for biomolecular detection and cellular imaging and were also used for sensing metal ions and biologically relevant species (Shen *et al.*, 2018).

Rhodamine-B undergoes intramolecular conversion in response to change in pH and high extinction coefficient which results in the formation of ring opened form (fluorescent) and ring closed form (spirolactam - non-fluorescent).

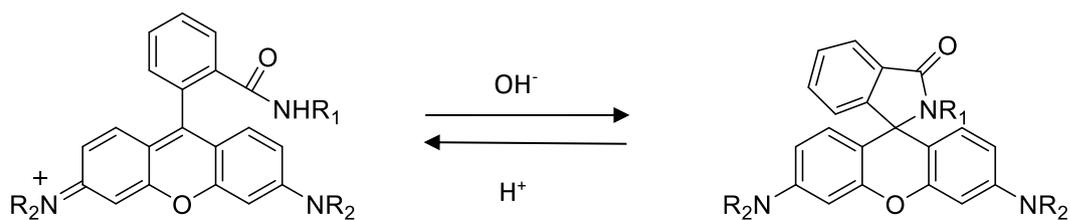
Beija (2009), explained that rhodamine-B was found in its cationic form because of protonation of carboxylic group in acidic solutions, whereas in basic solutions the rhodamine-B dissociates to convert into zwitterion form which shares the same chromophore as that of cation formed in acidic solution. However, the presence of the negative charge has an inductive effect on the central carbon atom of xanthene chromophore resulting in the hypsochromic shift of absorption and fluorescence maxima with decrease in extinction coefficient (Beija *et al.*, 2009). In the case of non-polar organic solvents, the zwitterionic form of rhodamine-B was interrupted by  $\pi$ -conjugation of the xanthene ring which results in irreversible formation of lactone ring (non-fluorescent) with very low fluorescent quantum yield compared to cationic and zwitterionic forms (Beija *et al.*, 2009) (Figure-12).



**Figure-12:** Chemical structures of Rhodamine-B.

Therefore, the Beija studies suggested the modification of the carboxylic acid group at position-2' formed a secondary amide (Figure-13) which is converted to the fluorescent ring-opened form at acidic pH and also in the presence of metal cations.

It has been noticed that modification at 2' position of rhodamine B (Figure-12) results in the formation of secondary amides of rhodamines such as **1** (Figure-13) that can undergo cyclization rapidly under all pH conditions especially in alkaline environments and convert the fluorescent to non-fluorescent lactams **2** (Figure-13). Thus, this reversible cyclization reaction made these compounds to restrict their use in most of the biological labelling experiments (Nguyen & Francis, 2003).



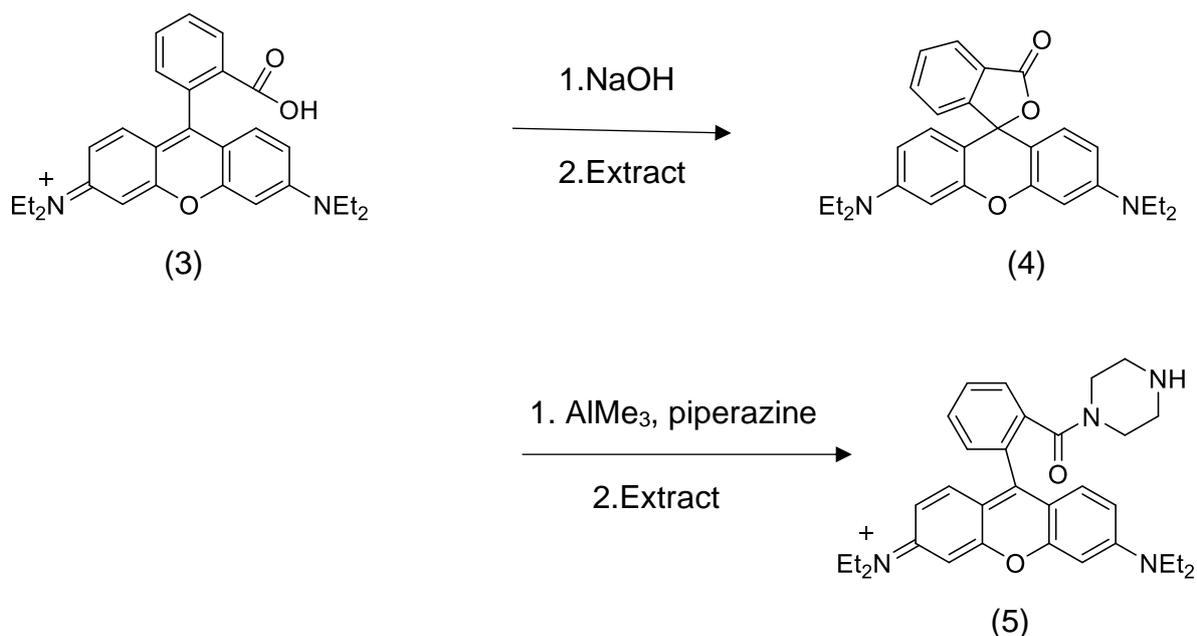
(1) Fluorescent

(2) Non-fluorescent

**Figure-13:** Cyclization of rhodamine amides.

Therefore, Nguyen and Francis developed a method to prevent this cyclization reaction and compared these transformation reaction conditions with other procedures that used Lewis Acids and carbodiimide coupling agents which produced unsatisfactory results and thereby declared this method as a most favourable one to produce a desired compound in appreciable yield.

This most advisable method involved the preparation of a tertiary amide **5** from the dye rhodamine-B **3**. The conversion of rhodamine-B **3** to tertiary amide **5** occurs by exposure of lactone **4** (available commercially or prepared from **3**) to 4 equivalents of piperazine and 2 equivalents of  $\text{AlMe}_3$  in refluxing  $\text{CH}_2\text{Cl}_2$  which thereby resulted in the formation of tertiary amide **5**, (Figure-14) (Nguyen & Francis, 2003).



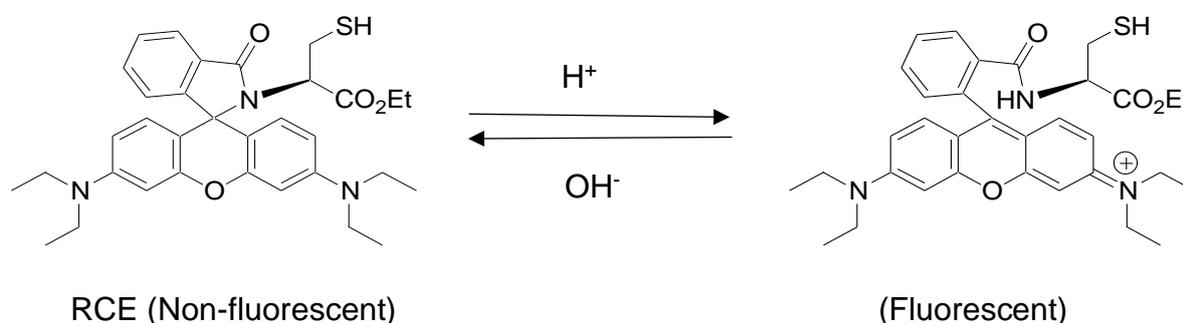
**Figure-14:** Synthesis of rhodamine tertiary amide derivatives.

### 2.7.1. Example-1:

#### Rhodamine-B based pH probe showing good photostability and membrane permeability

The sensing mechanism of rhodamine dyes is based on fluorescence enhancement which is caused by spirolactam ring-opening of rhodamine derivatives (Lv *et al.*, 2013). The opening and closure of the spirolactam ring of rhodamine derivative depends on the pH condition, the spirolactam ring remains closed in neutral or basic conditions, whereas the rhodamine derivatives are found to be non-fluorescent. While in acidic conditions,  $H^+$  leads to spirolactam ring-opening, by which the rhodamine derivatives exhibit strong fluorescence emission (Figure-14). Thus, rhodamine derivatives are found to be suitable to monitor acidic pH changes with enhanced fluorescence signals (Shen *et al.*, 2018; Lv *et al.*, 2013).

One of the novel rhodamine-B based pH probes developed in a study by Lv was RCE [Rhodamine B and cysteine ethyl ester (a lysosome-specific pH fluorescent indicator)], which was synthesized from rhodamine-B and cysteine ethyl ester with an excellent photostability and membrane permeable properties (Figure-15) (Lv *et al.*, 2013) and gave an enhanced fluorescence signal at acidic pH. It also showed longer absorption and emission wavelengths. The RCE probe also found to be suitable for studying acidic organelles with  $pK_a$  value 4.7. Because of the above properties, this probe has been successfully used for fluorescence imaging in living cells and used in staining the lysosomes selectively (Lv *et al.*, 2013).



**Figure-15:** Chemical reaction involved in spirolactam ring closure and opening (in acidic condition) in Rhodamine-B probe (RCE).

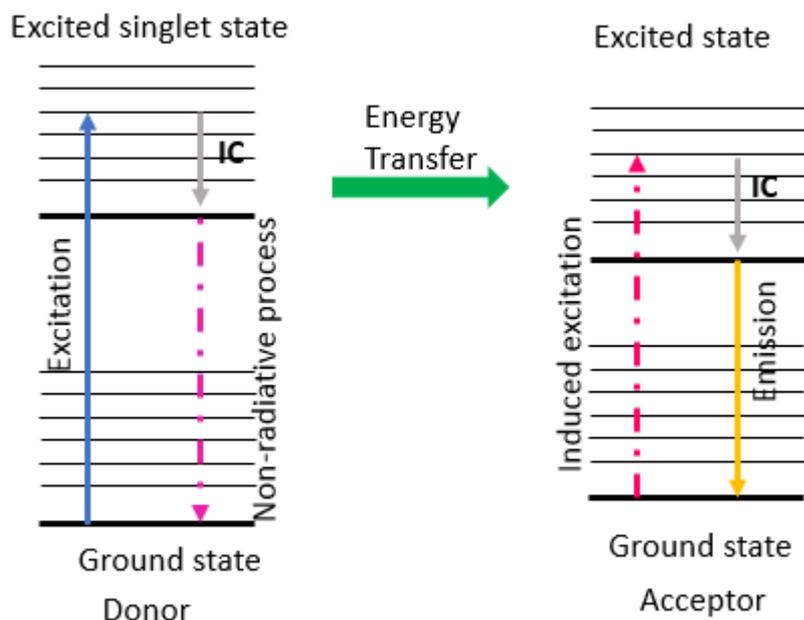
### 2.7.2. Rhodamine-B in FRET:

The properties such as high brightness, excellent photostability, and the capability to modulate the properties of the dye through substitution made the rhodamine dyes to be used more popularly as fluorescent and fluorogenic molecules. One of the most important properties of these dyes includes maintaining the equilibrium between an open, coloured, fluorescent quinoid form or a closed, colourless, nonfluorescent lactone, which can be controlled by appropriate substitution (Wysocki & Lavis, 2011).

### 2.8. FLUORESCENCE RESONANCE ENERGY TRANSFER (FRET):

**Fluorescence resonance energy transfer (FRET)** is a distance-dependent physical process in which energy is transferred from an excited donor molecular fluorophore (*D*) to acceptor fluorophore (*A*) non-radiatively by means of intermolecular long-range dipole–dipole interaction (Sekar & Periasamy, 2003). FRET can be measured accurately if the donor and acceptor molecule are at angstrom distances of 10–100 Å (Clegg, 1995) and it is found to be a highly efficient method if there is a Forster radius distance maintained between the donor and acceptor molecules.

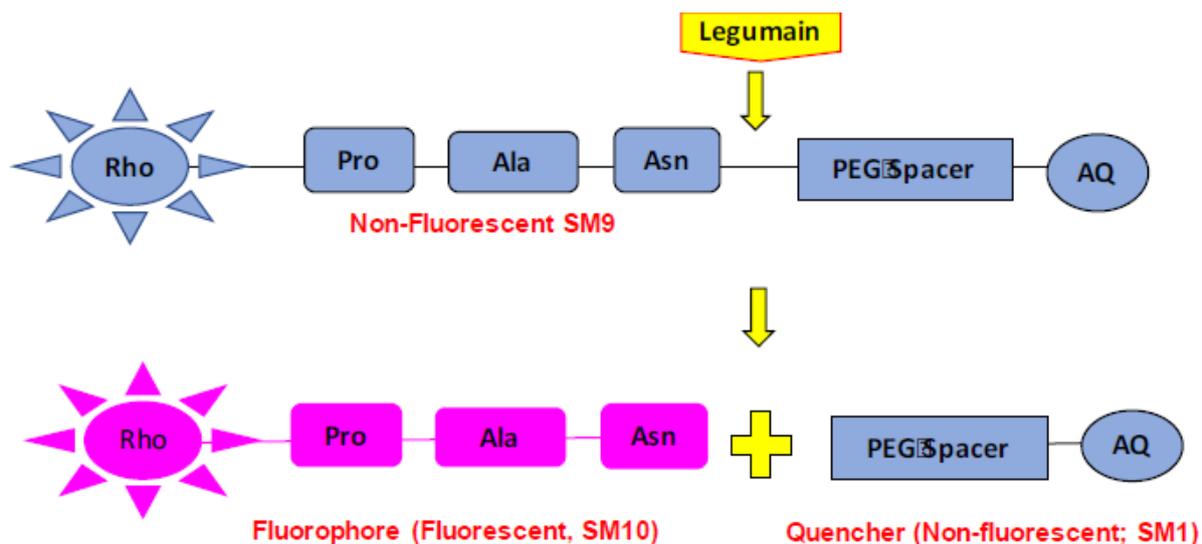
The mechanism of energy transfer in FRET is shown schematically in Figure-16. The fluorophore in the ground state is excited by the photon and reaches to the excited singlet state where it undergoes the internal conversion and vibrational relaxation (IC) and emits fluorescence non-radiatively by transferring the energy to the acceptor. After excitation by the induced energy transfer, the excited acceptor emits a photon and returns back to the ground state, if there is no quenching taking place.



**Figure-16:** Schematic diagram illustrating the mechanism of energy transfer in FRET.

### 2.8.1. Applications of FRET:

In previous research from this laboratory by Mathur *et al.*, (2016) a first generation fluorogenic FRET probe, SM9 [Rho-Pro-Ala-Asn-PEG-AQ(4-OH)], was developed based on the proteolytic activity of legumain in cancer because of its sensitive detection (based on rhodamine fluorescence). In SM9, the amino terminus of the peptide was capped with a Rhodamine-B fluorophore and the acid terminus of the peptide was coupled to an aminoanthraquinone which acts as a quencher for rhodamine B fluorescence (Figure-17). Cleavage of the probe by legumain at the carboxyl terminus of asparagine resulted in release of a fluorescent tripeptide with strong restoration of rhodamine fluorescence (Mathur *et al.*, 2016). However, one major problem with the fluorescent tripeptide (reporter fluorophore) released from SM9 is that it is not taken up by cells and for future *ex vivo* and *in vivo* applications, cell permeability of the fluorophore would be essential. This limitation provided the starting point for the research presented in this thesis. A re-designed fluorogenic probe was sought that could release a cell-permeating reporter fluorophore.



**Figure-17:** Schematic diagram of activation of SM9 [Adapted from: Mathur *et al.*, 2016].

This lack of cell permeability provided the motivation for this research and the synthesis of a second generation legumain substrate probe in which the positions of quencher and fluorophore were switched around giving the [AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-piperazine-Rho] DR-21 probe (Figure-17). In the present research, it was reasoned that reversing the positions of the quencher and fluorophore may provide a solution to the lack of permeability problem.

### 3. RATIONAL DESIGN OF LEGUMAIN SUBSTRATE PROBE (DR-21):

The rational design of this second generation fluorogenic legumain substrate probe (DR-21) contains three components - a quencher (aminoanthraquinone derivative), a fluorophore (Rhodamine-B-piperazine) and a tripeptide linker (Pro-Ala-Asn). These components are arranged in such a way that the quencher and the fluorophore are attached to either side of the tripeptide linker where the quencher is attached to the N-terminus of the tripeptide and the fluorophore is capping the carboxylic terminus of the tripeptide where the Asn (asparagine) is placed at the P1 position, to give AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho as shown in Figure-18. This arrangement of the components was made to direct legumain to cleave the substrate probe at the carboxylic end of the P1 position of Asn, given legumain shows restricted enzymatic cleavage at P1 position of Asparagine at acidic pH, (Gawenda *et al.*, 2007) resulting in the separation of the fluorophore from the quencher. In this case, the fluorophore (unlike the cleavage product of SM9) does not have a peptide residue or notably a carboxylic acid group that hinder take up by cells. In contrast, the piperazinyl-rhodamine should permeate cell membranes.

#### **Quencher:**

The aminoanthraquinone quencher in this probe is non-fluorescent which internally quenches the fluorescence of the fluorophore, so that the entire probe becomes non-fluorescent or substantially less fluorescent. Moreover, the aminoanthraquinone derivative, AQ(4-OH)- $\beta$ -Ala-OH was used as a quencher because of its wide range of absorbance wavelengths ( $\lambda_{abs}$  570-650 nm) which can overlap the emitted energy of the rhodamine fluorophore ( $\lambda_{em}$  560-585 nm) which results in efficient quenching where the fluorescence resonance energy was transferred (FRET) from the donor fluorophore to the acceptor quencher (Kodama *et al.*, 2006). This process of a FRET mechanism is one of the important steps in the activation of the probe to enhance the fluorescence of the fluorophore. Hence the quencher and the fluorophore are found to be important components to act as a pair in the design of the target fluorogenic substrate probe (DR21).

The literature of May and co-workers (2005), explained the previous development a non-fluorescent quencher based on the 1,4-diaminoanthraquinone molecule which can be capable of eliminating the background signals that are produced by the

quenching effect shown by using two fluorophores as a FRET pair. This non-fluorescent molecule was able to quench wide range of fluorophores, mostly which can emit at longer wavelengths. This specific quenching activity was because of its broad absorption wavelength which ranges from 550-750 nm of visible region (May *et al.*, 2005). The 1-amino-4-hydroxyanthraquinone chromophore was found in this project to be an equally efficient, non-fluorescent, sometimes called 'black-hole' quencher in an analogous way.

### **Peptide linkage:**

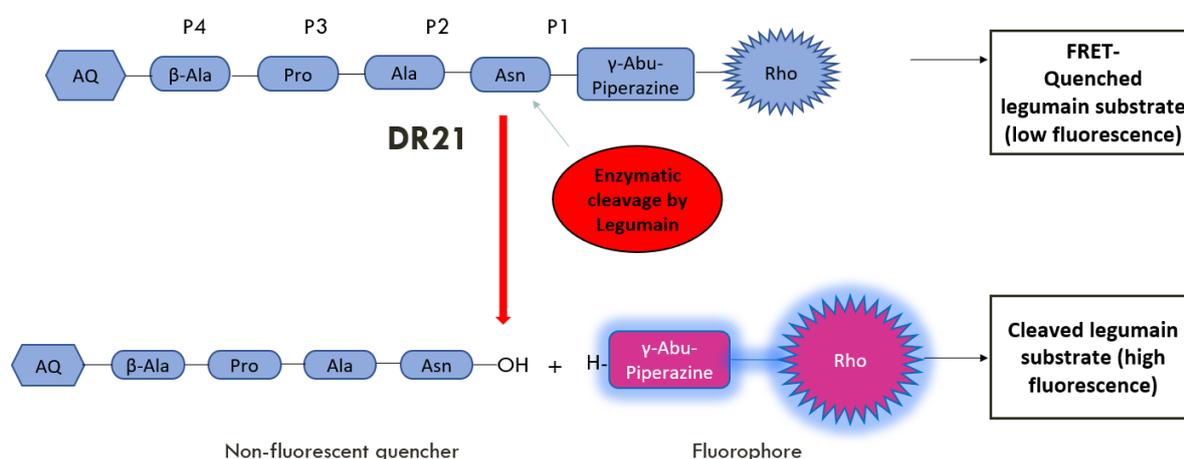
One further important step in substrate probe design is the alignment of the amino acid residues in the tripeptide and their COOH-terminal and N-terminal coupling with the fluorophore and the quencher respectively to be accommodated into the legumain active site. According to the Schechter and Berger (1967) nomenclature, the amino acids are numbered from P1, (non-primed) towards the N-terminus and numbered from P1' (primed) towards the carboxylic terminus of cleavage site (Schechter & Berger, 2012). Because of the enzymatic cleavage by restricted specificity of the legumain towards the Asn at the P1 position, the Asn was positioned at P1 in the tripeptide substrate molecule where the cleavage of the substrate occurs at the scissile bond between P1 and P1' (Choe *et al.*, 2006).

The Alanine, Proline and  $\beta$ -Alanine are positioned at the P2, P3 and P4 sites respectively because of their hydrophobicity and their efficacy towards the legumain. Stern and colleagues (2009), synthesised a chemotherapeutic agent that acts as a legumain substrate antitumour prodrug of etoposide which contains Asn at the P1 position and the Alanine is located in the P2 position in the tripeptide linkage chain to develop a final probe as CBZ-Ala-Ala-Asn-ethylenediamine-etoposide to show its antitumour activity (Stern *et al.*, 2009).

### **Fluorophore:**

The main aim of the synthesis is to determine the fluorogenic properties of the substrate probe to diagnose the tumour associated cell and this can be possible by the use of Rhodamine-B as a fluorophore. Due to the instability (ring closure) of the rhodamine-B, it was converted into a tertiary amide form by the piperazine residue where the rhodamine-B remains in open form in acidic conditions - which is the most favourable environment for the enzymatic activity of the legumain (lysosomes being

typically at pH 5.5). Moreover, the ring-opened rhodamine-B is known to show an emission between the 560-585 nm (Beija *et al.*, 2009). Therefore, the emission spectrum of the fluorophore can overlap the absorption spectrum of the quencher (570-650 nm) and confirms that the aminoanthraquinone quencher and rhodamine-B fluorophore pair as the best donor/quencher pair suitable for the synthesis of an efficient fluorogenic target substrate probe (DR-21). Because of the high sensitivity and selectivity towards a specific protease overexpressed by the tumour cells, the substrate based fluorogenic probes became most popular approach in molecular imaging of the tumour cells. Hence, the designed DR-21 substrate probe was predicted to be an efficient tumour diagnostic tool with a desired cell permeability property.



**Figure-18:** An outline of the target legumain substrate probe DR-21 design and its concept of cleavage by legumain at the P1 site.

## 4. RESULTS AND DISCUSSION

This section discusses the synthesis and fluorescence properties of the novel second-generation fluorogenic legumain substrate probe AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21). This section is followed by the Experimental sections describing the detailed chemical syntheses, purification and characterization of DR-21 and its intermediates; incubation of DR-21 with recombinant human legumain, physicochemical properties (partition coefficient), UV/Visible absorption and fluorescence studies.

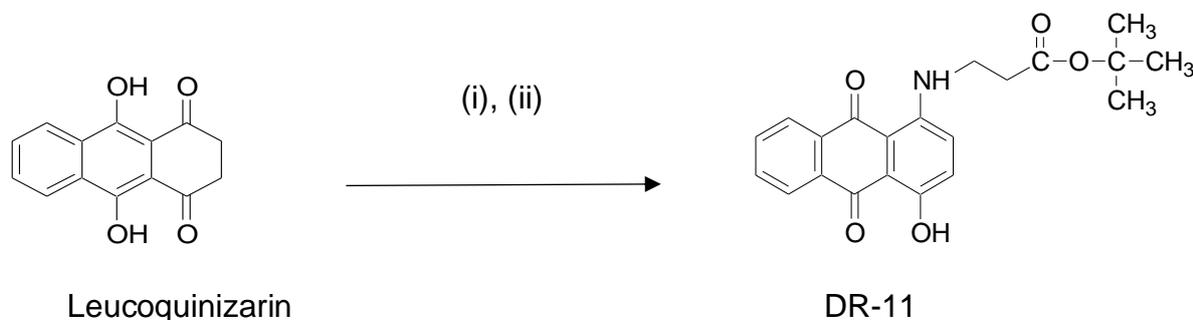
### 4.1. SYNTHESIS OF AMINOANTHRAQUINONE QUENCHER (DR-12)

The synthesis of the desired quencher moiety DR-12 involved the synthesis of DR-11 from leucoquinizarin followed by deprotection of DR-11 (<sup>t</sup>Bu removal) by treating with TFA.

#### 4.1.1. Synthesis of AQ(4-OH)- $\beta$ -Ala-O<sup>t</sup>Bu (DR-11):

Synthesis of DR-11 began with the controlled addition of  $\beta$ -Alanine-*t*-butyl ester hydrochloride to leucoquinizarin to afford the mono-aminated product.

#### Scheme-1: Reaction of leucoquinizarin with $\beta$ -Alanine-*t*-butyl ester HCl



**Reagents and conditions:** (i)  $\beta$ -Alanine-<sup>t</sup>butyl ester HCl, (ii)  $K_2CO_3$ , DMF, 95°C, 1 h, then  $Et_3N$  and  $O_2$ .

The procedure required anhydrous  $K_2CO_3$  as a base in DMF followed by heating on a water bath at 95°C for an hour [**Scheme-1**]. In this reaction,  $K_2CO_3$  was found to be effective for removal of HCl group from  $\beta$ -Alanine-<sup>t</sup>butyl ester HCl and this base can also be removed by addition of water at the solvent extraction stage. Once the reaction was cooled, the leuco-form reaction product was oxidised by triethylamine and aeration which resulted in the formation of a mono-substituted dark purple 4-

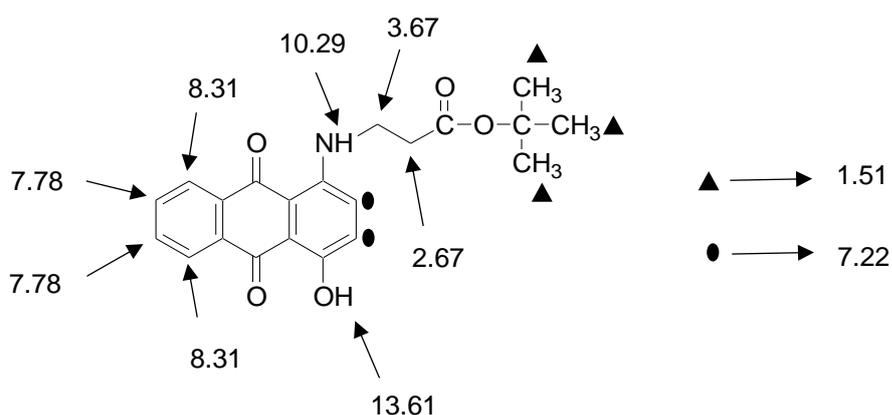
hydroxylated aminoanthraquinone compound along with two by-products. Leucoquinizarin upon amination results in formation of by-products such as (1) quinizarin – the oxidised product of unreacted leucoquinizarin, orange in colour, (2) di-substituted by-product because of reaction of  $\beta$ -Alanine-<sup>t</sup>butyl ester with leucoquinizarin at each of the 1, and 4-positions, blue in colour, and (3) desired mono-substituted compound from the nucleophilic substitution reaction, purple in colour. According to Kikuchi and co-workers (1982), study based on the kinetic and leuco compounds structures, leucoquinizarin gives mono- and di-substituted compounds upon amination. Amination of leucoquinizarin with butylamine resulted in formation of 1-butylamino-4-hydroxyanthraquinone(mono-substituted) and 1,4-bis(butylamino) anthraquinone derivative (di-substituted) by oxidation, but the same reaction resulted in formation of mono and di-substituted derivatives in their leuco forms in the presence of nitrogen (Kikuchi *et al.*, 1982). Moreover, Kikuchi and co-workers (1982) found the oxidation of leuco compounds in nitrobenzene resulted in formation of dealkylated compound such as 1-amino-4-hydroxy anthraquinone. The literature of Barasch and colleagues (1999) demonstrated the amination of leucoquinizarin with two molar excess of amine in the presence of nitrogen and followed by oxidation resulted in formation of a bi-substituted symmetrical analogue with a good yield (Barasch *et al.*, 1999). According to Nor and co-workers (2013), a good yield of di-substituted aminoanthraquinone derivatives were synthesised by the amination with butylamine using the catalyst  $\text{PhI}(\text{OAc})_2$  (Nor *et al.*, 2013).

In the synthesis of DR-11, amination was controlled to minimise bis-substitution. The purple coloured compound was extracted in chloroform and dried by adding anhydrous sodium sulfate which absorbs water from the organic extract. The crude compound was purified by column chromatography using silica gel as a stationary phase. The pure fractions of compound were eluted using dichloromethane as a mobile phase and with close monitoring by TLC. The pure (by TLC) compound was filtered from silica and dried to give the title compound (DR-11).

The purified product was characterised by its HRMS electrospray (+) mass spectrum which displayed a signal at  $m/z$  368.1497 Da  $[\text{M}+\text{H}]^+$  corresponding to the relative molecular mass of calculated for this species 368.1492 Da for DR-11.

Further evidence was provided by observing the signal at  $m/z$  757.2729 Da which is a result of a sodium adduct dimer species  $[2M+Na]^+$ . There was a good correlation between the observed data and theoretical isotope model data which confirmed the structure of expected compound DR-11.

The structure of the DR-11 compound was confirmed by its  $^1\text{H}$  NMR spectrum [Figure-19]. A nine-proton singlet signal at 1.51 ppm was assigned to the methyl protons of the  $\text{O}^t\text{Bu}$  group, two proton triplet at 2.67 ppm and two proton quartet at 3.67 ppm were assigned to  $\beta$  and  $\alpha$  positions of L-alanine respectively. The two proton doublet at 7.22 ppm were assigned to aromatic H2 and H3. A two proton doublet at 7.78 ppm were assigned to aromatic H2 and H3. A two proton multiplet at 7.78 ppm was assigned to aromatic H6 and H7; two proton multiplet at 8.31 ppm was assigned to the aromatic H5 and H8. A one proton triplet at 10.29 ppm was assigned to the aromatic NH and a singlet proton at 13.61 ppm was allocated to the phenolic OH group.

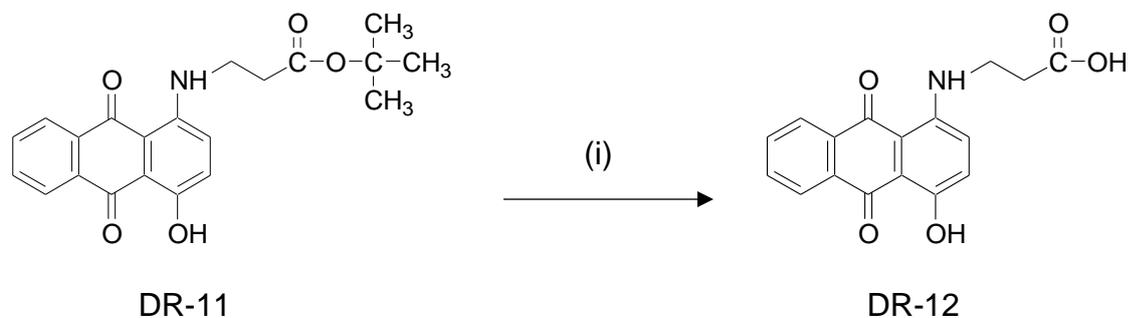


**Figure-19:**  $^1\text{H}$  NMR assignment of signals in DR-11.

#### 4.1.2. Synthesis of AQ(4OH)- $\beta$ -Ala-OH (DR-12):

The synthesis of DR-12 involved the standard removal of the O<sup>t</sup>Bu group from DR-11 by the addition of TFA.

##### Scheme-2: Reaction of DR-11 with TFA

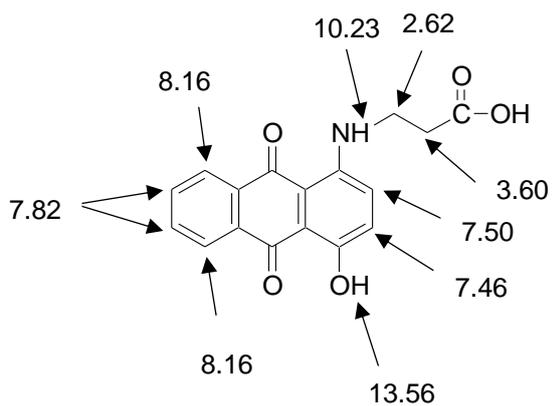


**Reagents and conditions:** (i) TFA, 3 h.

DR-11 protected as the tert-butyl ester was deprotected by treating with TFA and allowed to react for 3 hours [Scheme-2]. Once the completion of reaction was confirmed by TLC, then the solvent was evaporated and the crude product was precipitated by adding ether with ice cooling. The precipitate of DR-12 was filtered, collected to give title compound of DR-12 (homogeneous on TLC) which was deemed suitable for further reactions.

The pure DR-12 product was characterised by its HRMS ESI (-) mass spectrum which displayed a signal at  $m/z$  310.0721 Da for the species  $[M-H]^-$  which corresponded to the relative molecular mass of DR-12.

The structure of the anthraquinone-beta alanine adduct DR-12 was confirmed by its  $^1H$  NMR spectrum [Figure-20]. Two 2-proton triplets at 2.62 ppm and 3.60 ppm were assigned to the  $\alpha$  and  $\beta$  methylene groups of the beta-alanine respectively. Two coupled doublet protons at 7.50 ppm and 7.46 ppm were assigned to aromatic H2 and H3 respectively. A two proton multiplet at 7.82 ppm was assigned to the aromatic H6 and H7. A two proton multiplet at 8.16 ppm was assigned to aromatic H5 and H8. A singlet proton at 10.23 ppm was assigned to the aromatic NH and a singlet proton at 13.56 ppm was assigned to the aromatic OH.



**Figure-20:**  $^1\text{H}$  NMR assignment of signals in DR-12.

Moreover, The structure of DR-11 was also confirmed by  $^{13}\text{C}$  NMR spectroscopy in which all the 17 carbon signals were determined and accounted for. Two signals at 34.04 and 38.04 ppm were assigned to methylene groups. Total of 6 signals at 107.37, 112.76, 131.73, 134.52, 146.79 and 156.00 ppm accounted for 6 quaternary carbons. The 6 signals of aromatic carbons were found at 125.01, 125.78, 126.17, 128.55, 132.75 and 134.48 ppm. The three signals at 172.74, 180.63 and 186.56 ppm were assigned to amide carbonyl carbons.

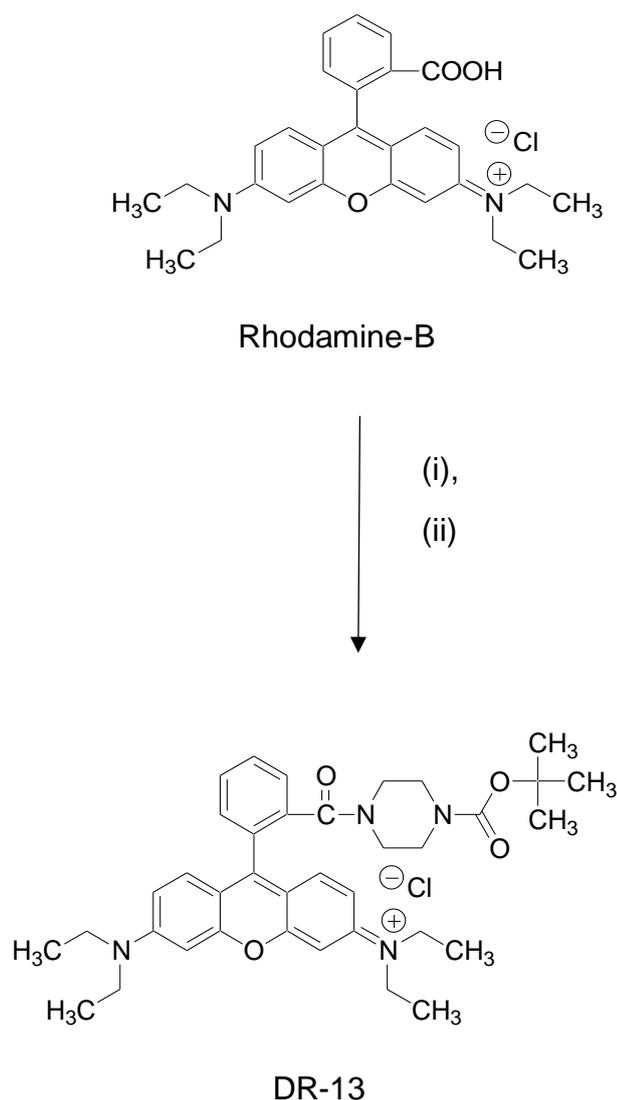
## 4.2. SYNTHESIS OF RHODAMINE-B FLUOROPHORE (DR-16):

DR-16 was synthesised by three following steps:

### 4.2.1. Synthesis of Rho-Pip-Boc (DR-13):

The synthesis of DR-13 involved the reaction of rhodamine-B with Boc-piperazine.

#### Scheme-3: Reaction of Rhodamine-B with Boc-Piperazine



**Reagents and conditions:** (i) HATU, ACN in DIPEA stir for 15 min, (ii) Boc-Piperazine

In this reaction, the rhodamine-B was added to the coupling reagent HATU and then dissolved in ACN which is used as an alternative solvent for DMF. The studies of Jad and co-workers found the alternative solvents for DCM (dichloromethane), DMF (N,N-

dimethylformamide) and NMP (N-methyl-2-pyrrolidone), moreover they stated DMF and NMP as undesirable solvents. Even though, the DMF solubilizes the amino acid derivatives and coupling reagents, it can decompose to formaldehyde and dimethylamine which finally threatens peptide synthesis. Therefore by focusing on the parameters of racemisation and coupling performance, ACN (acetonitrile) and THF (tetrahydrofuran) solvents were found to be excellent solvents in peptide synthesis both in solid and solution phase peptide synthesis compared to DMF and NMP (Jad *et al.*, 2015). The studies of MacMillan and co-workers suggested DMC, EtOAc and 2-Methyl-THF as alternative solvents to DMF and DCM in amide coupling processes which was demonstrated conducting amidation reactions using different carboxylic acids and amines with COMU (1-cyano-2-ethoxy-2-oxoethylideneaminoxy) dimethylamino-morpholino-carbenium hexafluorophosphate) as a coupling agent (MacMillan *et al.*, 2013).

For the preparation of the rhodamine-piperazine fluorophore required in this research, it was necessary to activate the carboxylic acid group of the rhodamine in the manner used in peptide synthesis. The activation reaction of rhodamine B with powerful coupling agent HATU was started by addition of DIPEA. The activated reaction mixture was allowed to react with Boc-piperazine. The reaction mixture was poured into chloroform and washed with water then dried. The solvent was removed by evaporation under reduced pressure and the crude product was purified by column chromatography. The fractions of pure compound were confirmed by TLC, combined, filtered and dried to give the desired compound DR-13 [**Scheme-3**].

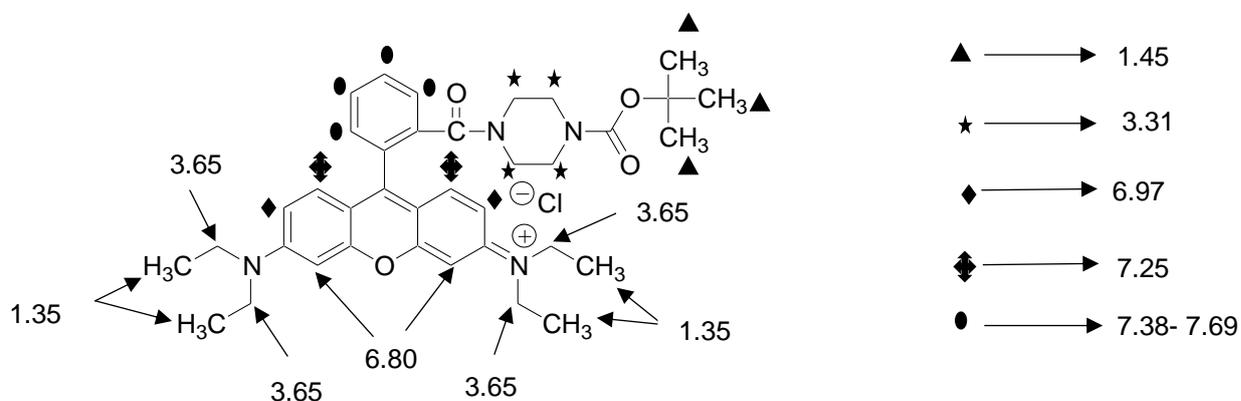
Given the spirolactone ring of rhodamine-B undergoes cyclization at its 2' position and converts the fluorescent compound to its non-fluorescent form in acidic pH, the rhodamine-B was coupled to piperazine to prevent this reversible cyclisation through tertiary amide formation thus making the rhodamine-Pip conjugate stable across a wide pH range to show its fluorescent activity. The literature of Nguyen and Francis (2003), suggested that the exposure of the lactone ring of rhodamine-B to 4 equivalents of piperazine and 2 equivalents of AlMe<sub>3</sub> along with refluxing in chloroform showed the successful formation of a tertiary amide of rhodamine-B (Nguyen & Francis, 2003). But Pierangelo Gobbo and his colleagues (2015), reported the synthesis of amides of rhodamine-B with piperazine at basic pH. This literature described the reaction of rhodamine-B with mono-BOC-protected piperazine by using

HBTU as the coupling agent which resulted in the successful synthesis of a rhodamine-B piperazine adduct with higher yield and purity without using  $\text{AlMe}_3$  (Gobbo *et al.*, 2015). According to Basavaraj Padmashali and his colleagues (2017) the HATU and DIPEA were found suitable in coupling of acids and amines for the successful synthesis of indole piperazine methanone derivatives 3A-J where the indole piperazines were synthesised with substituted aromatic and heterocyclic acids (Padmashali *et al.*, 2017).

The purified rhodamine-piperazine conjugate here, DR-13 was characterised by its HRMS ESI (+) mass spectrum which displayed a signal at  $m/z$  611.3584 Da  $[\text{M}-\text{Cl}]^+$  corresponding to the mass of 611.3584 Da for the DR-13 cation.

Further evidence of similarity between observed data and theoretical isotope model confirmed the structure of desired cation of compound DR-13.

The structure of the DR-13 compound was confirmed by its  $^1\text{H}$  NMR spectrum [Figure-21]. Twelve protons of triplet signal at 1.35 ppm were assigned to methyl groups of rhodamine-B, nine protons of singlet at 1.45 ppm were assigned to methyl protons of the Boc group. The eight protons of multiplet at 3.31 ppm were assigned to methylene groups of piperazine. An eight proton quartet at 3.65 ppm was assigned to methylene group of rhodamine-B. The two proton doublet at 6.80 ppm was assigned to aromatic H4 and H5; two proton doublet at 6.97 ppm was assigned to aromatic H2 and H7, the two proton doublet at 7.25 ppm was assigned to aromatic H1 and H8. A one proton multiplet at 7.38 ppm, a one proton multiplet at 7.55 ppm and two proton multiplet at 7.69 ppm accounted for aromatic protons of rhodamine-B, from the integration. All aromatic protons were accounted for by integration.

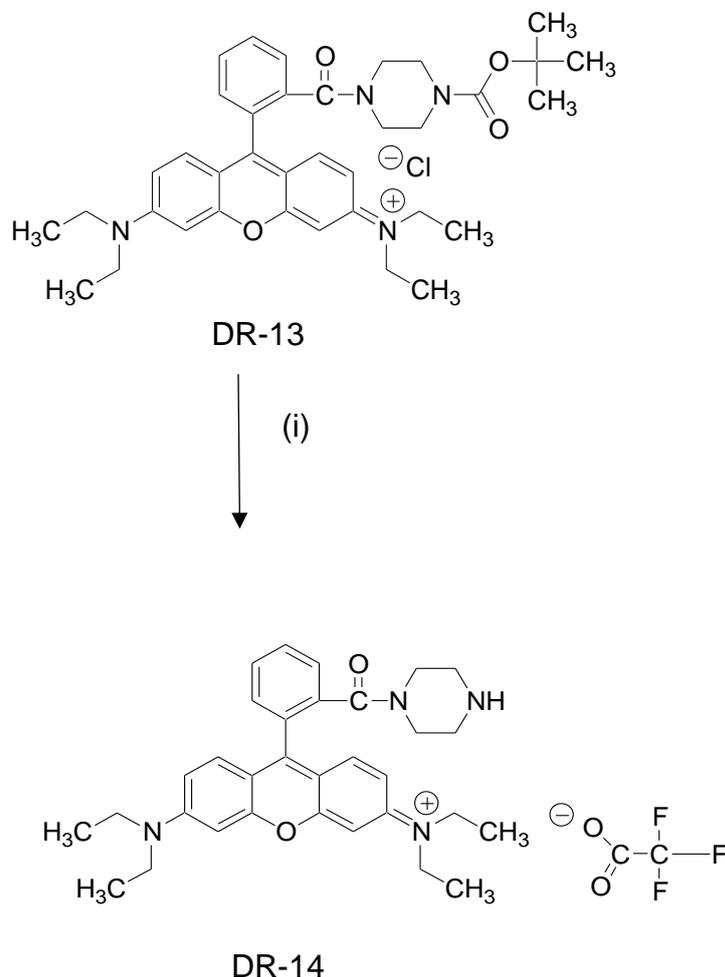


**Figure-21:**  $^1\text{H}$  NMR assignment of signals in DR-13

#### 4.2.2. Synthesis of Rho-Pip-TFA (DR-14):

Synthesis of DR-14 involved the standard deprotection of DR-13.

#### Scheme-4: Reaction of Rho-Pip-Boc (DR-13) with TFA:



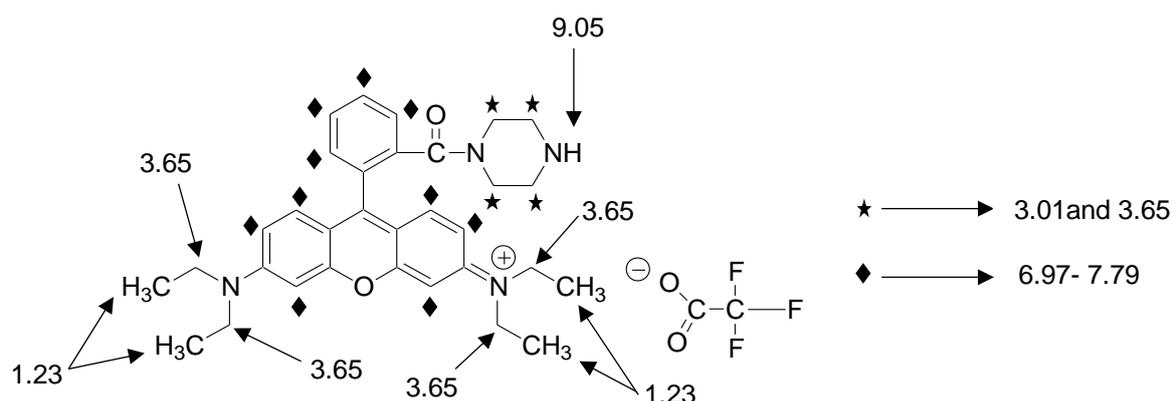
**Reagents and conditions:** (i) TFA, 1 h.

DR-13 was deprotected by treating with trifluoroacetic acid for an hour at room temperature. The formation of DR-14 trifluoroacetate salt was confirmed by TLC and was obtained in solid form by treating with diethyl ether on ice. The salt precipitate of DR-14 was filtered from ether and dried to give a pure salt of DR-14 [**Scheme-4**]. Wilhelmsen and his co-workers (2018) in their experimental investigation on amination and Boc-group deprotection, found different solvents and additives can be used for deprotection of a Boc-group. From which, TFA was found as a best deprotecting agent for the deprotection of tert-butyl carbamates and also known to facilitate the reductive amination reactions of weakly basic amines. Moreover, their experiments also found that Lewis acids can also deprotect the Boc group but showed poor results in reductive

amination reactions. Wilhelmsen and his co-workers (2018) also found applications of aluminium trichloride, titanium (IV) chloride and boron trifluoride diethyl etherate which are capable of deprotection of a Boc group in a only a few minutes (Wilhelmsen *et al.*, 2018). TFA was also known to deprotect the protected aromatic hydroxyl groups (Monteiro *et al.*, 2017).

The purified dried DR-14 was characterised by its HRMS mass spectrum which displayed a signal at  $m/z$  511.3054 Da  $[M-2TFA-H]^+$  corresponding to the mass of 511.3068 Da for DR-14 cation. There was observed a signal at  $m/z$  256.1570 which is because of the two cationic charges, by which it confirms the mass of DR-14 was divided by two charges as per the  $m/z$  ratio (appeared at half mass;  $z = 2$ ).

The structure of the DR-14 compound was further confirmed by its  $^1H$  NMR [Figure-22]. Twelve proton triplet signals at 1.23 ppm were assigned to methyl groups of rhodamine-B; four protons of multiplet at 3.01 ppm were assigned to methylene groups of piperazine. From twelve protons of quartet at 3.65 ppm, eight protons were assigned to methylene group of rhodamine-B and four protons were assigned to methylene group of piperazine. The signals at 6.97, 7.15, 7.54 and 7.79 ppm were assigned to the aromatic protons of rhodamine-B. A singlet proton at 9.05 ppm was allocated to the amide proton in the spectrum. Thus, all the protons were accounted for by chemical shift and integration.

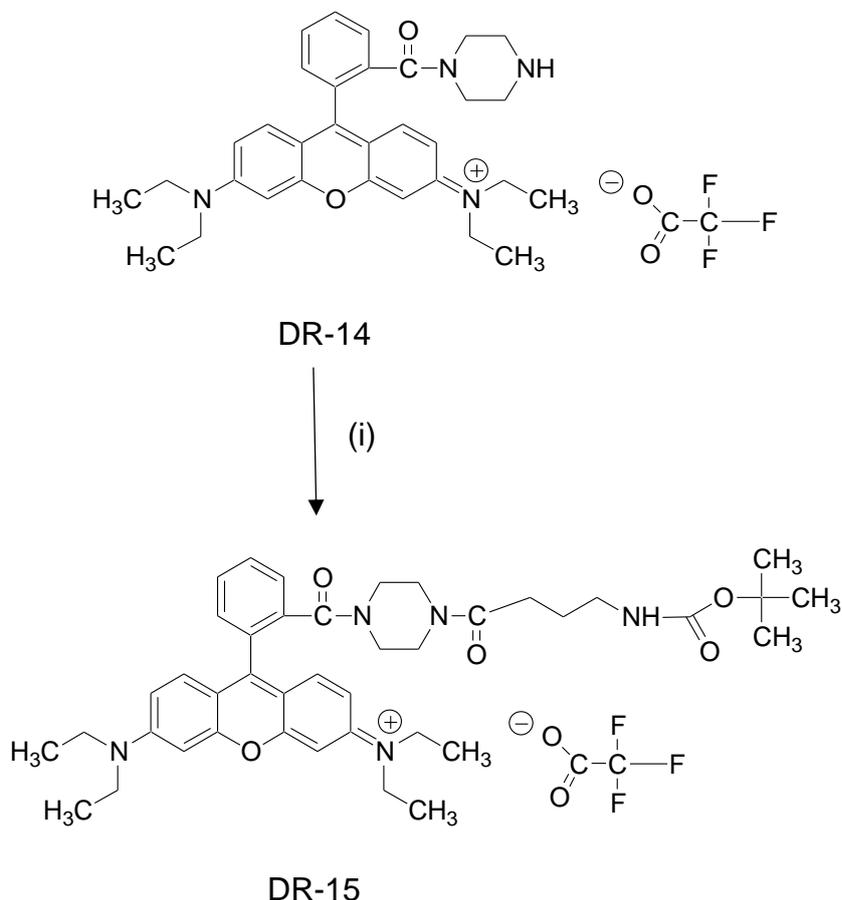


**Figure-22:**  $^1H$  NMR assignment of signals in DR-14.

#### 4.2.3. Synthesis of Boc- $\gamma$ -Abu-Pip-Rho-TFA (DR-15):

Reaction of DR-14 with Boc- $\gamma$ -Abu-OH by using coupling agent HATU with the base DIPEA resulted in the formation of DR-15.

#### Scheme-5: Reaction of Rho-Pip-TFA (DR-14) with BOC- $\gamma$ -Abu-OH:



**Reagents and conditions:** (i) HATU, Acetonitrile, DIPEA 1 h.

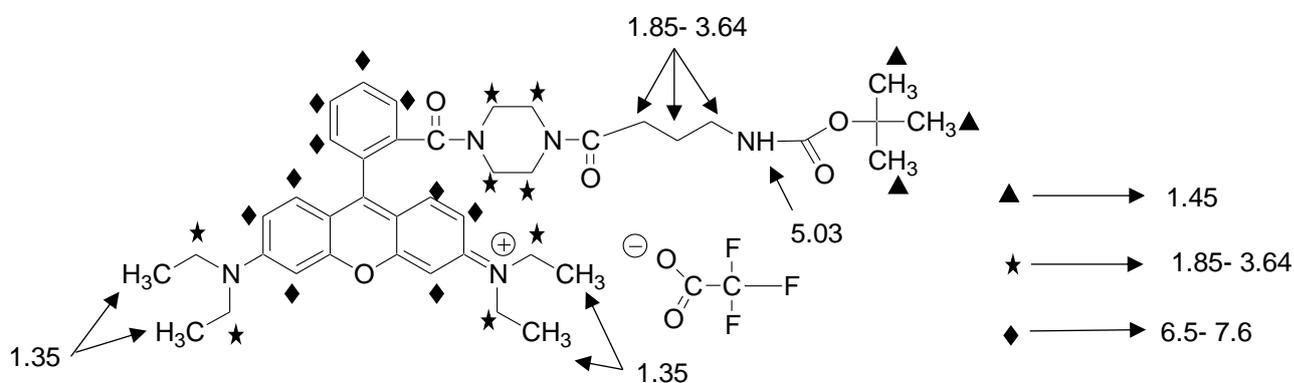
Boc- $\gamma$ -Abu-OH and HATU in ACN were added to DIPEA and allowed to react for activation. DR-14 in ACN was added to DIPEA and then added to the activated reaction mixture followed by stirring for an hour. After stirring, the completion of reaction was confirmed by the TLC spots of the synthesised desired compound.

1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxide hexafluorophosphate (HATU), an immonium salt, which is mostly used as an effective coupling reagent in the synthesis of amides and peptides because of its advantages like fast reaction rate, availability and low cost. The literature by Leas (2018) described the mechanism of HATU as a coupling reagent involves the carbon transfer in the

formation of a cyclic compound (Leas *et al.*, 2018). The compound was partially purified by solvent extraction followed by collection of pure compound by column chromatography through TLC monitoring. The pure compound was dried and solidified by adding diethyl ether in cool condition. The precipitate of compound was collected, filtered from ether and dried under vacuum to give pure dried of title compound DR-15 [Scheme-5].

The purified dried DR-15 was characterised by its HRMS mass spectrum which displayed a signal at  $m/z$  696.4108 Da  $[M-TFA]^+$  corresponding to the mass of 696.4119 Da for DR-15 cation, while the anion was trifluoroacetate. There was found to be an excellent agreement between experimentally observed data and theoretical isotope pattern for the desired cation  $[DR-15 \text{ Cation}]^+$ .

The structure of the DR-15 compound was also confirmed by its  $^1\text{H}$  NMR spectrum [Figure-23]. Twelve protons of triplet signal at 1.35 ppm were assigned to methyl groups of rhodamine-B, the nine proton singlet at 1.45 ppm were assigned to methyl protons of the Boc group. From 22 protons of multiplet signals from 1.85- 3.64 ppm, eight protons were assigned to methylene groups of Abu, six protons were assigned to methylene groups of piperazine and remaining eight protons were assigned to methylene groups of rhodamine-B. A one-proton singlet signal at 5.03 ppm was assigned to amide group proton. Ten protons at 6.5- 7.0, 7.05, 7.4 and 7.6 accounted for aromatic protons of rhodamine-B in the spectrum, from the integration. All aromatic protons were accounted for by integration.

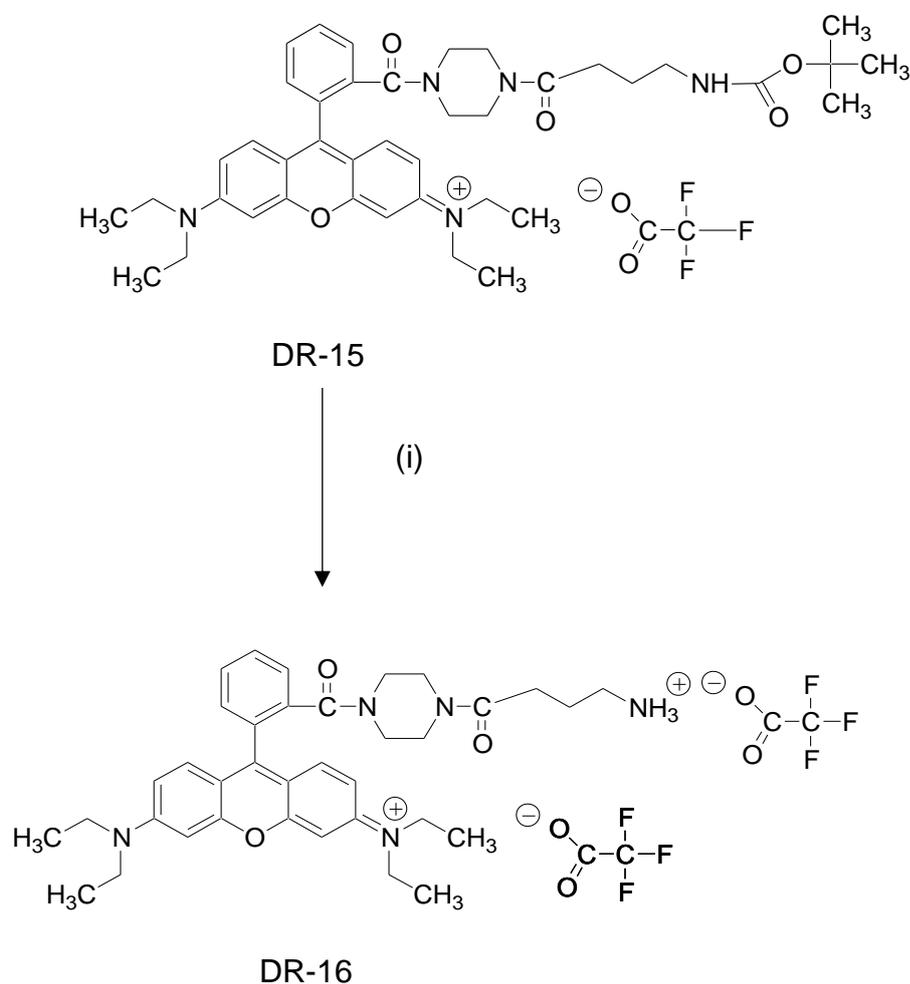


**Figure-23:**  $^1\text{H}$  NMR assignment of signals in DR-15.

#### 4.2.4. Synthesis of $\gamma$ -Abu-Pip-Rho-TFA (DR-16):

Synthesis of DR-16 involves the deprotection of DR-15 by treating with TFA.

##### Scheme-6: Reaction of Boc- $\gamma$ -Abu-Pip-Rho-Boc with TFA:



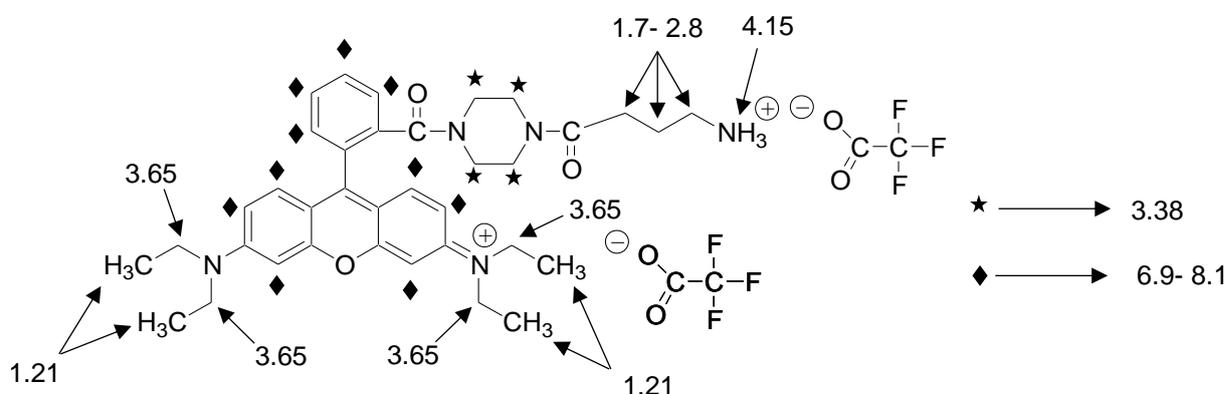
**Reagents and conditions:** (i) TFA, 30 min.

We successfully synthesised DR-16 by deprotecting DR-15 with TFA [**Scheme-6**]. The DR-15 was allowed to react with TFA for 30 minutes and the completion of reaction was confirmed by TLC. Boc groups are more easily removed than tert-butyl esters which require longer reaction times.

Boc-aminoadipic acid tert-butyl ester protecting group is removed by treating with excess amount of dry HCl in EtOAc about 3 hrs which results in production of 97% yield of a desired deprotected compound (Goodacre *et al.*, 1994). The studies of Isidro-Llobet and co-workers (2009) demonstrated that Boc-groups are used in both solid and solution phase synthesis and this Boc-group can be removed by treating with

25-50% of TFA in DCM (Isidro-Llobet *et al.*, 2009).

The structure of the DR-16 compound was confirmed by its  $^1\text{H}$  NMR spectrum [Figure-24]. Twelve protons of triplet signals at 1.21 ppm were assigned to methyl groups of rhodamine-B. Six protons of Abu were assigned to multiplet signals from 1.75- 2.8 ppm. Eight protons of doublet at 3.38 ppm signal were assigned to methylene groups of piperazine and the eight protons of a quartet at 3.65 ppm were assigned to methylene groups of rhodamine-B. Three protons of the amino group were found at 4.15 ppm as a broad singlet signal. All ten aromatic protons of rhodamine-B were accounted at 6.95, 7.05, 7.5 and 7.6 - 8.1 ppm.



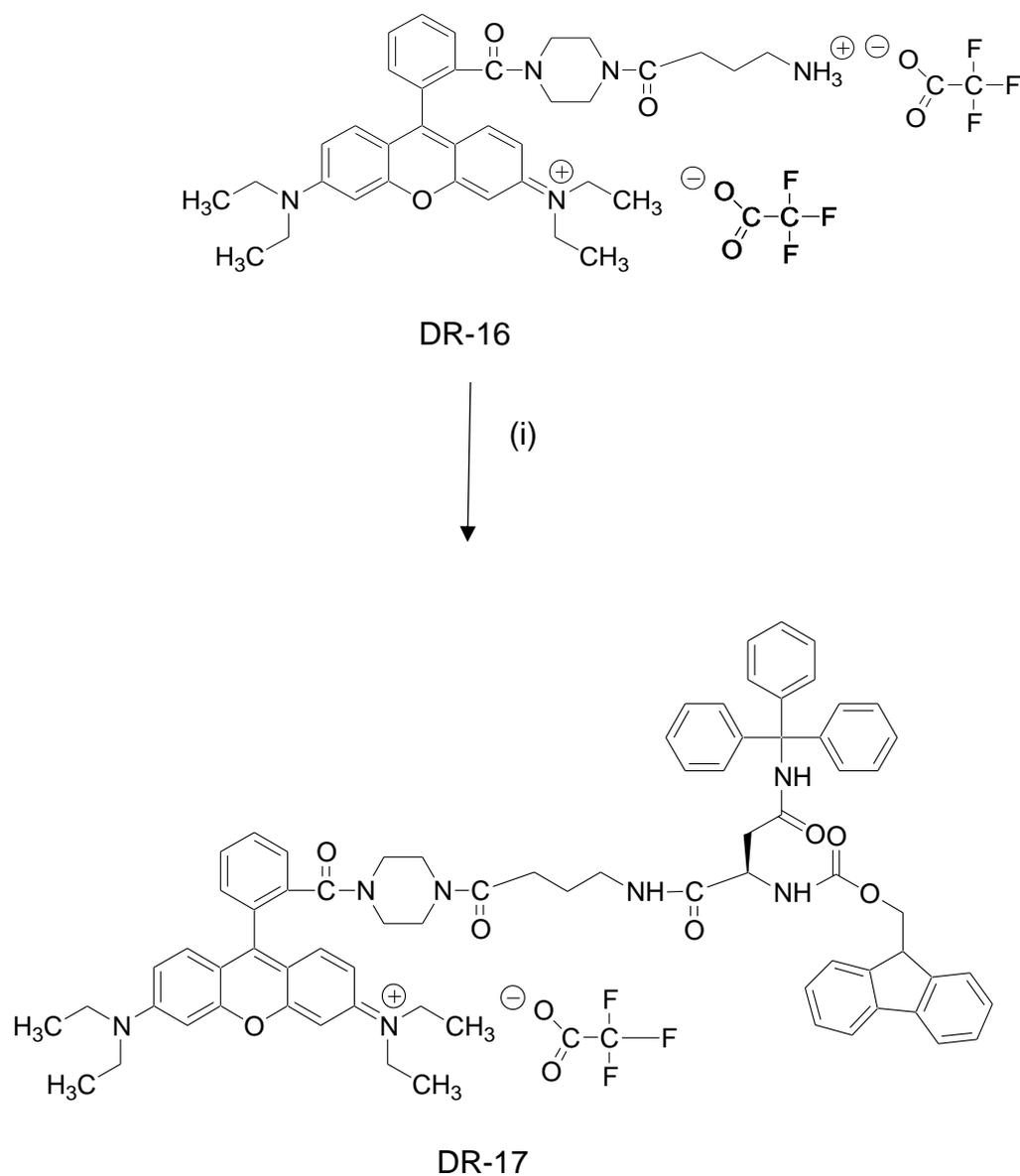
**Figure-24:**  $^1\text{H}$  NMR assignment of signals in DR-16.

### 4.3. SYNTHESIS OF Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA FLUOROPHORE (DR-18)

#### 4.3.1. Synthesis of Fmoc-Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-17):

DR-17 was synthesised by coupling DR-16 with Fmoc-Asn(Trt)OH by using HATU as a coupling agent in the presence of DIPEA base.

#### Scheme-7: Reaction of $\gamma$ -Abu-Pip-Rho-TFA (DR-16) with Fmoc-Asn(Trt)-OH



Reagents and conditions: (i) HATU, DMF, DIPEA, RT 24h.

HATU, Fmoc-Asn(Trt)-OH in DMF were allowed react in DIPEA for 15 minutes for its activation. The DR-16 in DMF was added to the activated reaction mixture and allowed to react for 24 hours. Once the reaction was completed the resultant desired compound formation was confirmed by TLC by using Butanol: Acetic acid: Water (4:5:1) and CHCl<sub>3</sub>: 8% Methanol as the solvent systems. The reaction mixture with desired target compound was dissolved in chloroform and washed with water, followed by collection of organic solvent, which was treated with anhydrous sodium sulphate, stirred well and allowed to stand for overnight for the complete removal of water molecules/ drops from the organic solvent. The crude compound was filtered from anhydrous sodium sulphate under vacuum and reduced its volume by using rota evaporator. The crude compound was purified by column chromatography, where the pure compound was collected in chloroform: 8% methanol by continuous monitoring with TLC. The fractions containing the pure compound were combined, filtered, and evaporated. The pure compound was treated with diethyl ether to form the precipitate on ice. The precipitate of DR-17 was filtered from ether and dried under vacuum to afford DR-17 homogeneous on TLC [**Scheme-8**].

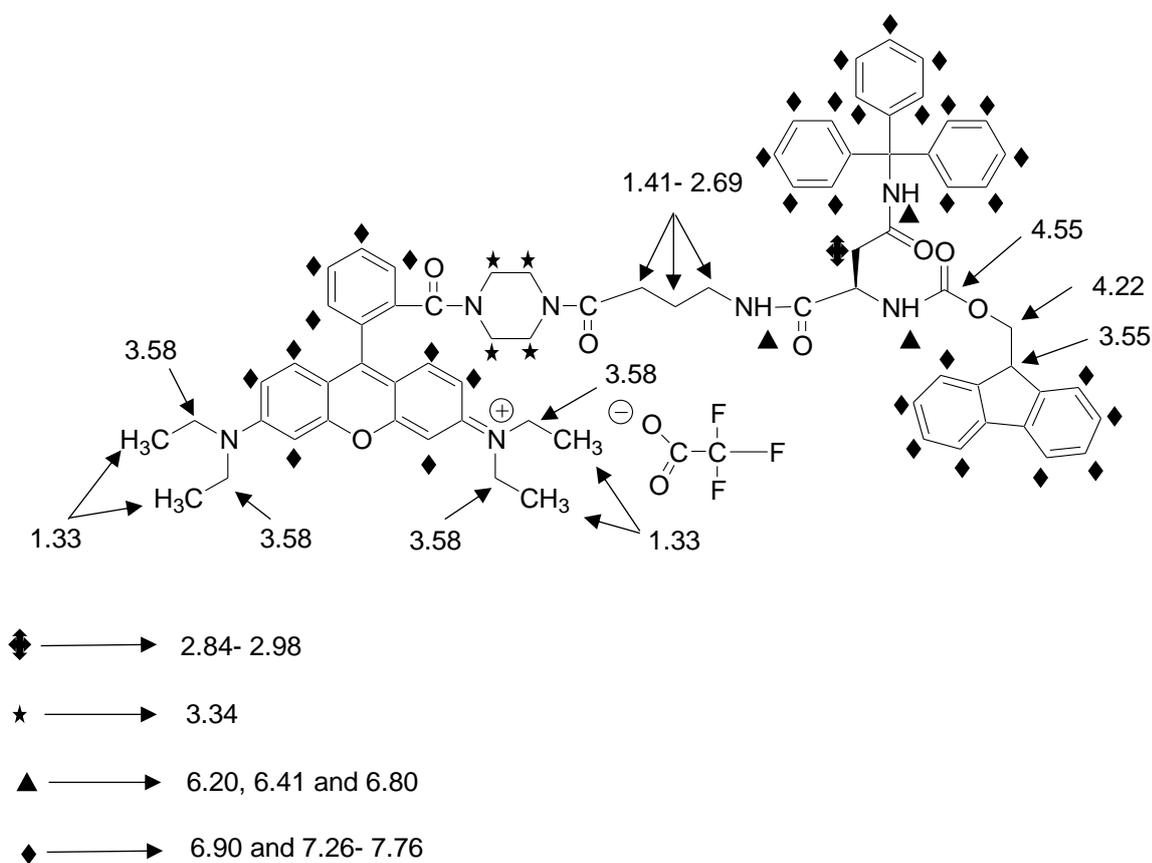
9-Fluorenylmethoxycarbonyl (Fmoc), is a protecting group initially introduced for solution phase reactions (solution chemistry) (Carpino & Han, 1972) but because of its instability caused by the acylation reaction of the amino groups of the components, hence it is proved as unsuitable (Bodanszky *et al.*, 1979).

Fmoc is found as one of the urethane protecting groups such as benzyl carbamate (benzyloxycarbonyl) and Boc groups which are most commonly used in activation and coupling reactions for the inhibition of racemisation process (Behrendt *et al.*, 2016).

The purified dried DR-17 was characterised by its HRMS ESI (+) mass spectrum which displayed a signal at  $m/z$  1174.5806 Da [M-TFA]<sup>+</sup> corresponding to the mass of 1174.5806 Da for DR-17 cation. There was an excellent agreement between the observed data and theoretical isotope model for the confirmation of structure of the expected DR-17 cation.

The structure of the DR-17 compound was confirmed by its <sup>1</sup>H NMR spectrum [**Figure-25**]. A twelve proton signal at 1.33 ppm was assigned to methyl groups of rhodamine-B. From six protons of Abu, two protons were assigned to signals from 1.41- 1.46 ppm, two were assigned at 3.34 ppm and the remaining two protons were assigned from

2.37 to 2.69. The two protons multiplet of methylene group of asparagine were assigned to signals from 2.84 to 2.98 ppm. Eight protons of doublet at 3.34 ppm signal were assigned to methylene groups of piperazine. The singlet signal at 3.55 ppm was assigned to the 9H proton of the Fmoc ring system. The eight protons of multiplet at 3.58 ppm were assigned to methylene groups of rhodamine-B. A multiplet of two protons of methylene group of Fmoc were found at 4.22 ppm. A singlet proton of  $\alpha$ -position of asparagine was found at 4.55 ppm and three protons of amide groups were found at signals of 6.20, 6.41 and 6.80 ppm with singlet peaks. All 33 aromatic protons of DR-17 (10H of Rhodamine-B, 15H of trityl group asparagine and 8H of the Fmoc group) were accounted for at 6.70 and from 7.26 to 7.76 ppm.

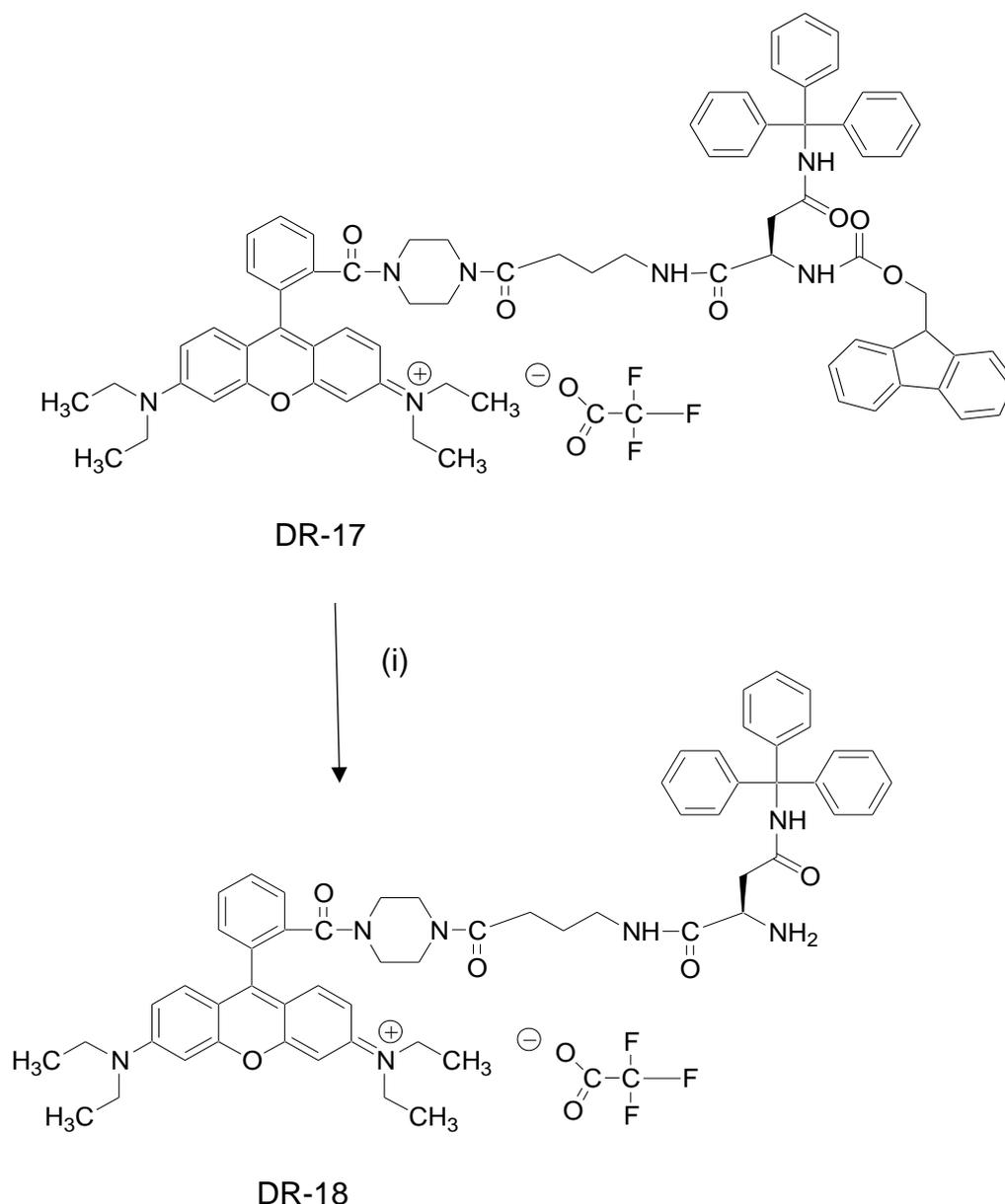


**Figure-25:**  $^1\text{H}$  NMR assignment of signals in DR-17

#### 4.3.2. Synthesis of Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-18):

Synthesis of DR-18 involved the removal of Fmoc-group from DR-17 by treating with 20% piperidine.

**Scheme-8: Reaction of Fmoc-Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-17) with 20% piperidine:**



**Reagents and conditions:** (i) 20% piperidine, DMF, RT, 15 min.

The synthesis of DR-18 involved the removal of Fmoc group from DR-17 by reacting with 20% piperidine in DMF **[Scheme-8]**. DR-17 dissolved in DMF was treated with 20% piperidine and allowed to react for 15 minutes. The reacted product was solvent

extracted by chloroform. The crude product in dried organic solvent was purified by column chromatography, where the desired product was collected by using chloroform: methanol (4:1) as a mobile phase. The purification was done twice for the complete removal of UV active Fmoc-piperidine adduct which was pale yellow in colour eluted in initial fractions of column chromatography purification. This step resulted in poor yield of the compound. The fractions containing the pure compound were combined and dried to give DR-18 compound.

The Fmoc, an amine protecting group is stable towards the acids and catalytic hydrogenation but found to be easily cleaved in mildly basic and non-hydrolytic conditions such as in liquid ammonia with piperidine, morpholine and ethanolamine at room temperature (Carpino & Han, 1972). Behrendt and his colleagues (2016) found the dibenzofulvene as a first cleaved product of Fmoc deprotection which can be re-attachable to the liberated amine making it more difficult for its further separation (Behrendt *et al.*, 2016). Therefore, removal of the Fmoc is an crucial step in synthesising the target compound; inefficient deprotection leads to poor yield of compound because of repeated purification which results in loss of compound (Luna *et al.*, 2016). Whereas, the initial product of Boc deprotection results in butylene formation which is volatile in nature but with limitations such as lability to secondary amine bases (Behrendt *et al.*, 2016). The Fmoc group is found to be mostly deprotected by primary amines and less commonly by tertiary amines but the deprotection with cyclic secondary amines was found to be best method because of their nucleophilicity (Luna *et al.*, 2016). Piperidine was found to be most popularly used secondary amine even though in spite of its disadvantage of forming aspartimide which can be minimised by use of some bases. Hence, the DMF is also most commonly used base in Fmoc removal for its interchain aggregation disruption action (Luna *et al.*, 2016). Therefore, removal of Fmoc is a crucial step in synthesising the targeted compound here, inefficient deprotection leads in poor yield of compound because of repeated purification which results in loss of compound (Luna *et al.*, 2016). Therefore, 20% piperidine along with DMF was used for the removal of Fmoc in the synthesis of DR-18.

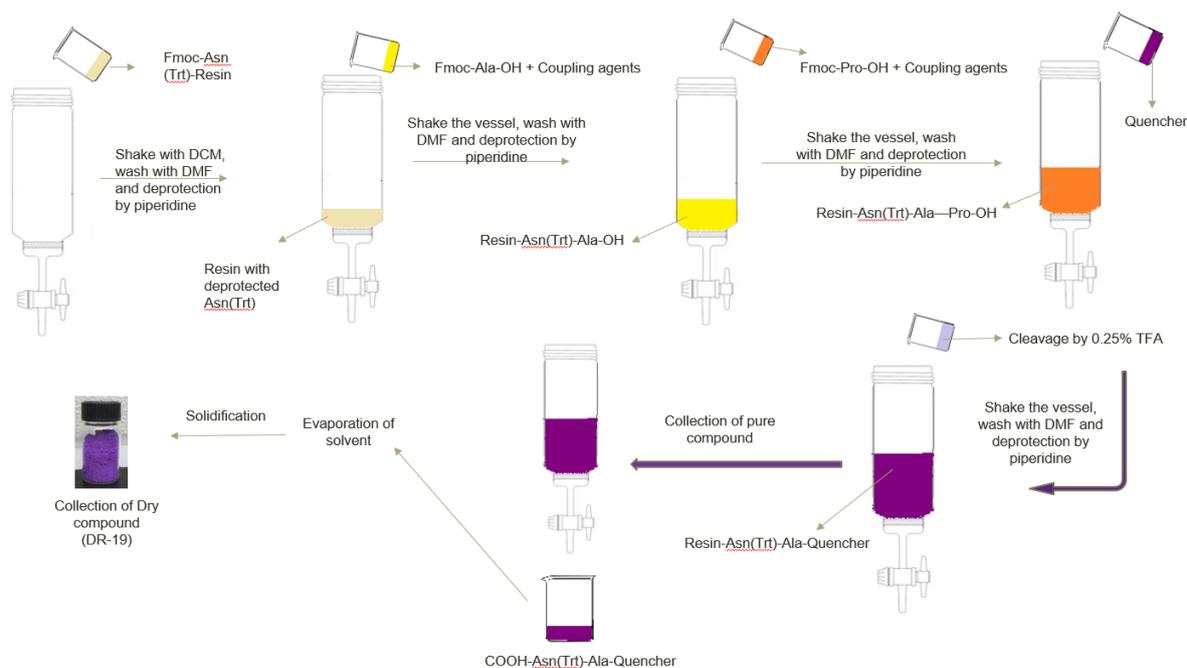
The purified dried DR-18 was characterised by its HRMS mass spectrum which displayed a signal at  $m/z$  952.5126 Da  $[M-TFA]^+$  corresponding to the mass of 952.5131 Da for DR-18 cation.

## **SOLID PHASE PEPTIDE SYNTHESIS:**

R.B. Merrifield described solid phase peptide synthesis in 1963 by stating “solid phase peptide synthesis method involves the stepwise addition of amino acids for a growing peptide by means of a covalent bond formation with a solid resin”. Solid phase peptide synthesis was developed in order to overcome some of the disadvantages of the solution phase peptide synthesis which includes difficulties with solubilities and purification procedures during the synthesis of polypeptide chain. This method was found to be more convenient, simple and speed technique compared to solution phase peptide synthesis (Merrifield, 1963). Swelling of resin beads is a foremost important step for the maximum permeation of activated N-protected amino acids within the crosslinked polystyrene beads which enhances yield of coupling. DCM was known to show maximum swelling properties when the resin was treated with DCM for 20- 30 minutes.

The Fmoc-group was found to be a best amine protecting group used in solid phase peptide synthesis method for more than a decade because of its orthogonal protecting strategy (which means deprotecting the specific protective group without effecting the other groups in multi-protected structures) (Fields & Noble, 2009).

The literature of Luna (2016) illustrated the mechanism of Fmoc-protecting group removal which includes two steps: (i) removal of acidic proton by mild base (mostly a secondary amine) at position-9 of the fluorene ring of the Fmoc group and (ii)  $\beta$ -elimination results in formation of a dibenzofulvene (DBF) intermediate which is surrounded by a secondary amine. DMF was found to be best solvent for the Fmoc-removal steps because of its polar nature and electron donor capability (Luna *et al.*, 2016). Moreover, PyBop and HOBt were found to enhance the coupling of hindered amino acids by reducing the racemisation. The DIPEA was known to be a best base for the activation of reaction mixture.



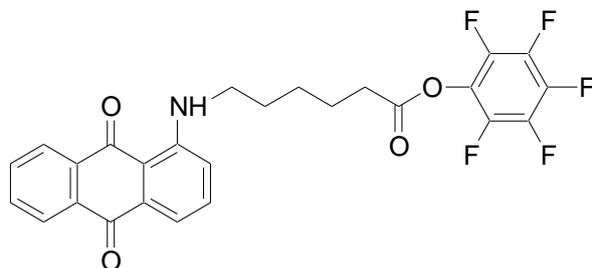
**Figure-26:** A schematic representation of steps involved in synthesis of aminoanthraquinone-labelled tetrapeptide DR-19 by solid phase peptide synthesis method.

#### 4.4. Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)-OH (DR-19):

Anthraquinone-quencher labelled tetrapeptide DR-19 was successfully synthesised by solid phase peptide synthesis which involves a series of sequential coupling and deprotection reactions followed by on-resin capping with the AQ quencher DR-12. The coupling reaction was performed by using the PyBop and HoBt as coupling agents in DIPEA base with the appropriate Fmoc protected amino acids, and the Fmoc-group was deprotected by 20% piperidine in DMF. These coupling and deprotection reactions were followed by subsequent washings with DMF and colour test by HZ22 reagent.

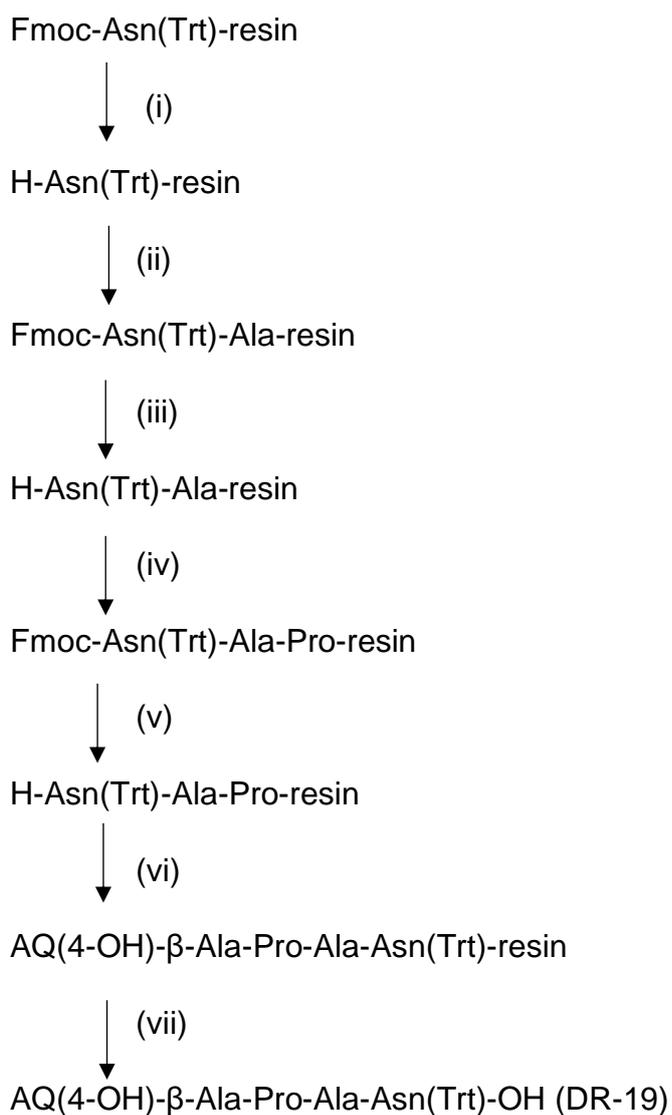
HZ22 is a pentafluorophenolate active ester which is a conjugate of anthraquinone and aminohexanoic acid. It is used as colour test reagent after every coupling and deprotection reaction steps in solid phase peptide synthesis, where it determines the presence or absence of primary and secondary amino groups on resin beads for respective reaction steps. HZ22 colour test was used as an alternative to the Kaiser test to overcome the disadvantages of Kaiser test such as no positive result for

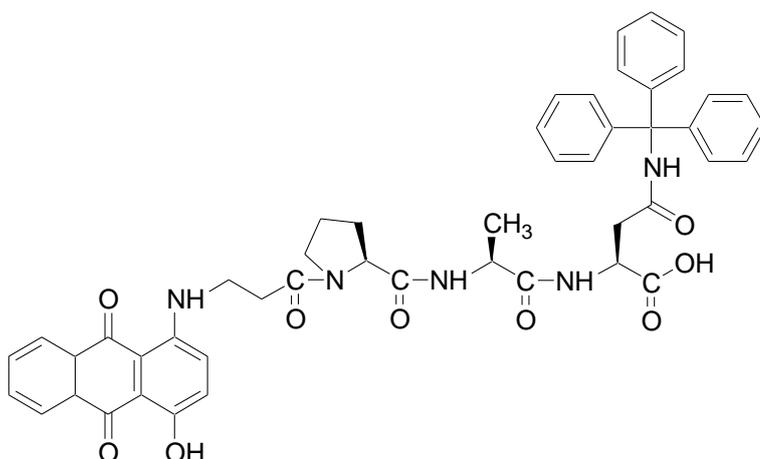
secondary amines, use of hazardous chemicals such as potassium cyanide etc (Zhang,H., 2017).



**Figure-27:** Structure of HZ22 reagent.

**Scheme-9: Reaction of tripeptide amino acids (Resin-Asn(Trt)-Ala-Pro-OH) with DR-12:**





DR-19

**Reagents and conditions:** (i) 20% piperidine, DMF, RT, 15 min, (ii) Fmoc-Ala-OH, PyBop, HOBT, DIPEA, RT, 30 min, (iii) 20% piperidine, DMF, RT, 15 min, (iv) Fmoc-Pro-OH, PyBop, HOBT, DIPEA, RT, 30 min, (v) 20% piperidine, DMF, RT, 15 min, (vi) DR-12, PyBop, HOBT, DIPEA, RT, 2 h, (vii) 0.25% TFA, RT, 15 min.

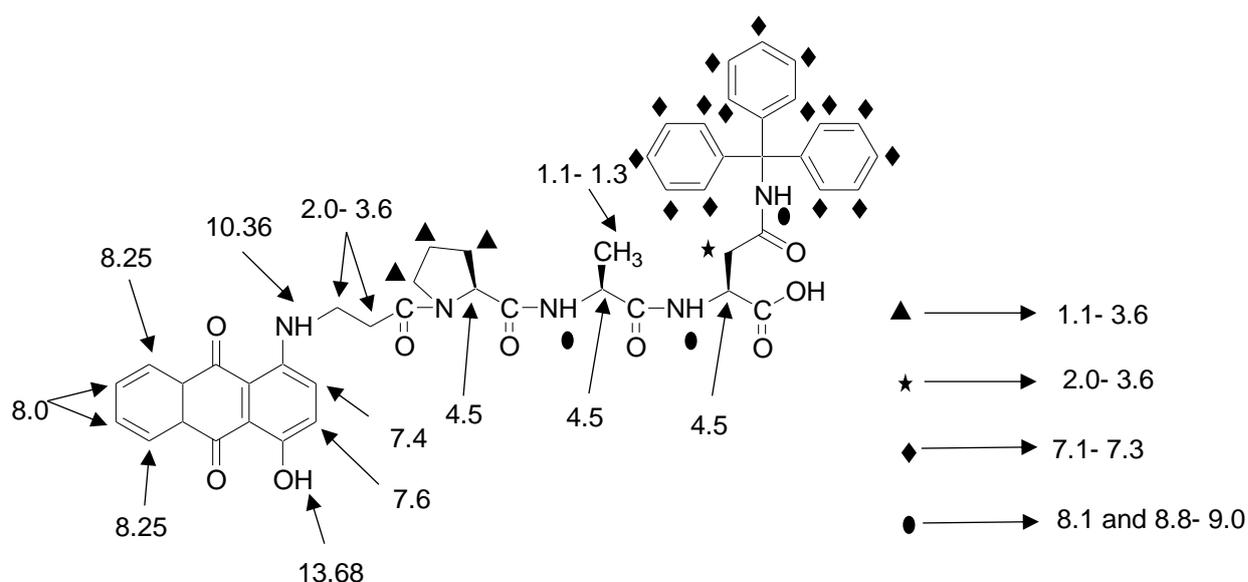
Synthesis of DR-19 by solid phase peptide synthesis was described in **section 6.2.9**. In this method the tripeptide was synthesised on resin with subsequent coupling of desired Fmoc protected amino acid with coupling agents PyBop and HOBT in DIPEA on to the resin followed by deprotection of Fmoc group by 20% piperidine in DMF. The activated DR-12 in DIPEA was coupled with deprotected tripeptide amino acids resin and allowed to react for 2 hours. The completion of coupling reaction was confirmed by observing the resin beads which convert into purple colour.

These purple colour resin beads indicated the formation of DR-19 which was collected from resin by treating with 0.25% TFA for 15 minutes or by washing DR-19 resin SPPS reaction vessel with 0.25% TFA **[Scheme-9]**. The collected filtrate solvent from reaction vessel was evaporated and treated with diethyl ether to form the solid precipitate on ice. The solid precipitate of DR-19 was collected from ether and dried to give title compound DR-19.

Therefore, the procedure of DR-19 synthesis by solid phase peptide synthesis method overcame the drawback of poor yield of the desired compound synthesised by solution phase peptide synthesis method as in **section 6.2.8**.

The purified dried DR-19 was characterised by its HRMS ESI(-) mass spectrum which displayed a signal at  $m/z$  834.3137 Da  $[M-H]^-$  corresponding to the relative molecular mass of 834.3145 Da for DR-19.

The structure of the DR-19 compound was confirmed by its  $^1H$  NMR spectrum [Figure-28]. From the five protons of multiplet from 1.1- 1.3 ppm, two protons were assigned to methylene group of proline and three protons were accounted for methyl group of alanine. All the twelve protons of the methylene groups of alanine, proline and asparagine were found at signals from 2.0 - 3.6 ppm by integration. All the ten aromatic protons of aminoanthraquinone were assigned to signals at 7.4, 7.6, 8.0 and 8.2 ppm. The three amide group protons of alanine, proline and asparagine were assigned to signals at 8.0, 8.8 and 9.0 ppm. The amino group proton of the anthraquinone and hydroxy group proton were assigned to a triplet and singlet at 10.36 and 13.68 ppm respectively.



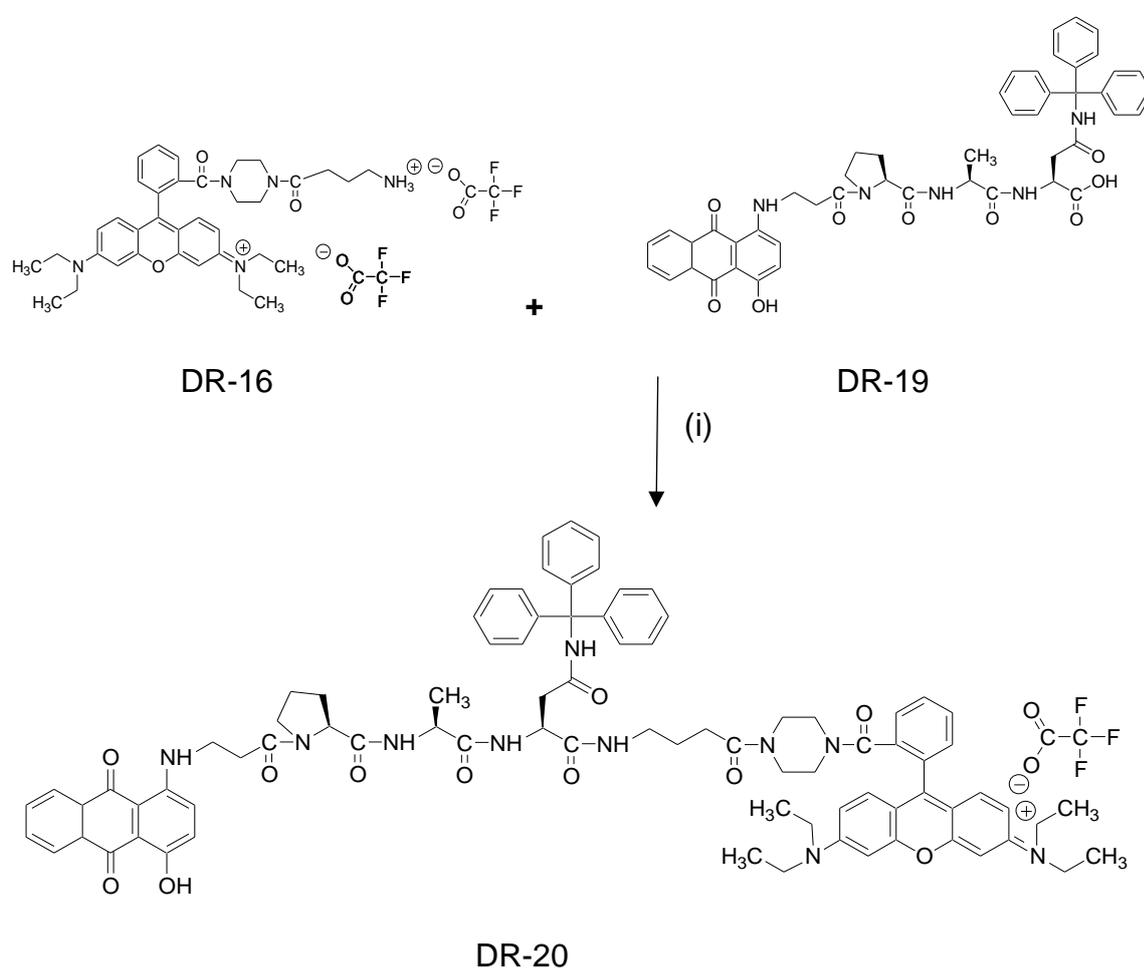
**Figure-28:**  $^1H$  NMR assignment of signals in DR-19

## 4.5. SYNTHESIS OF FLUOROGENIC PROBE (DR-21)

### 4.5.1. Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)- $\gamma$ -Abu-Pip-Rho (DR-20):

Synthesis of DR-20 involved the coupling of the two preassembled components in a convergent synthesis. DR-16 (the fluorophore) was coupled with DR-19 (tripeptide linked quencher) in DIPEA with the coupling reagents.

#### Scheme-10: Reaction of $\gamma$ -Abu-Pip-Rho-TFA (DR-16) with AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)-OH (DR-19):



**Reagents and conditions:** (i) PyBop, HOBt and DIPEA, RT, 2h.

PyBop, HOBt and DR-19 (quencher-labelled tripeptide) in DMF were added to DIPEA and allowed to react for 20 minutes for activation. After 20 minutes the DR-16 fluorophore component in DMF was added to the activated reaction mixture and allowed to react for 2 hours. The product in reaction mixture was solvent extracted and

purified by column chromatography followed by solidification by diethyl ether. The precipitate was dried to give the title compound DR-20 [**Scheme-10**].

The purified dried DR-20 was characterised by its HRMS mass spectrum which displayed a signal at  $m/z$  1413.6708 Da  $[M-TFA]^+$  corresponding to the mass of 1413.6707 Da for DR-20 cation.

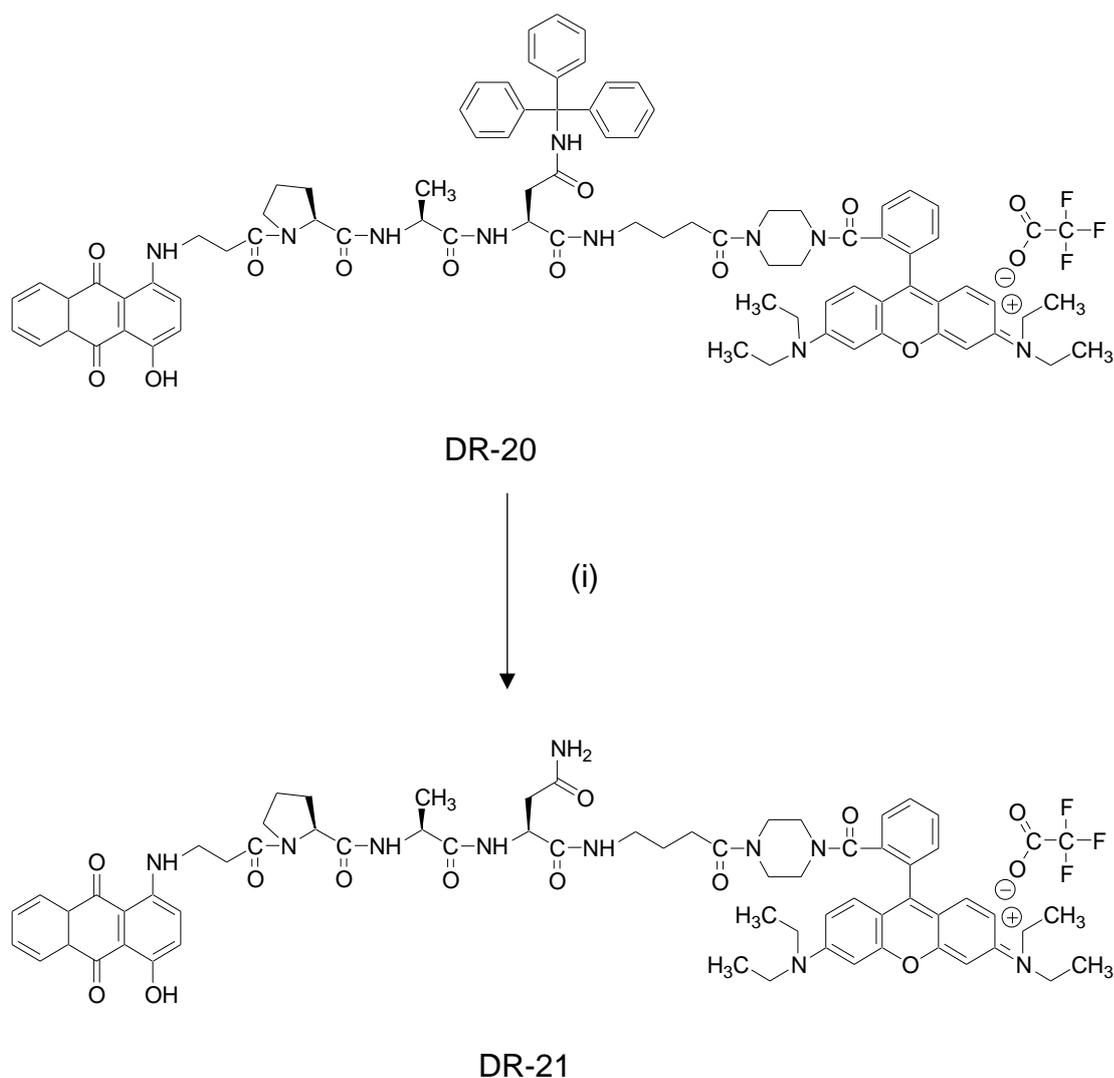
The spectrum displayed one more signal at  $m/z$  718.3300 because of two charges which was represented as  $[M-TFA+Na]^{2+}$  and appeared at half mass ( $z = 2$ ).

Moreover, the good correlation between the observed data and theoretical isotope model confirmed the expected structure of the DR-20 cation.

#### 4.5.2. Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21):

The target fluorogenic molecular probe DR-21 was synthesised by removal of the trityl protecting group from asparagine by treatment with TFA.

#### Scheme-11: Reaction of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)- $\gamma$ -Abu-Pip-Rho (DR-20) with TFA:



**Reagents and conditions:** (i) TFA, RT, 2 h.

DR-21 was successfully synthesised by adding the DR-20 with TFA [Scheme-11] and allowing to react for 2 hours with addition of a drop of water at the start of reaction. This reaction involved the removal of triphenylmethyl (trityl) group of asparagine by TFA. But there may be chance of formation a stable by-product ester which is formed because of reaction of the triphenylmethyl cation with TFA, hence the water was added for avoidance of this ester by-product affording triphenylmethanol, removed by

chromatography. The completion of reaction was confirmed by TLC by comparing the newly formed (DR-21) TLC spot with DR-20. The desired product was solvent extracted and purified by column chromatography.

The purified DR-21 was characterised by HRMS mass spectrum which displayed a signal at  $m/z$  1171.5607 Da  $[M-TFA]^+$  corresponding to the mass of 1171.5611 Da for DR-21 cation. The spectrum displayed a signal at  $m/z$  597.2745 Da (half mass) because of the doubly-charged species which was represented as  $[M-TFA-Na]^{2+}$ .

Further evidence of good correlation between observed data and theoretical isotope data confirmed the expected cationic structure of DR-21.

#### **4.6. DETERMINATION OF PARTITION COEFFICIENT OF DR-12, DR-16 AND DR-21:**

Drug discovery and development is a very time and resource consuming process in which different parameters must be evaluated in order to develop an effective drug with high biological activity and low toxicity (Daina *et al.*, 2017). Determination of physicochemical properties of the drug is essential in order to find clear information about its pharmacokinetics (absorption, distribution, metabolism and excretion) properties by which the cell permeability and duration of drug in its active form at the site of action can be determined. One of these parameters is the Partition Coefficient, which gives an indication of how lipophilic or hydrophilic a drug is. Lipophilicity describes the behaviour of the drug molecules when the drug molecules are in contact with lipid membranes and proteins (Daina *et al.*, 2014). Hence, lipophilicity plays an important role in determining the drug permeability through phospholipid membranes. In contrast, hydrophilic drugs are preferentially found in hydrophilic compartments such as blood serum. The partition coefficient is a measure of how a solute is distributed between two immiscible solvents. A polar will tend to favour the aqueous, whereas a non-polar compound will favour the organic phase (normally octanol). Compounds with a high partition coefficient value are lipophilic and those with low values are hydrophilic. As partition coefficient are expressed as a logarithm (Log P), lipophilic drugs have positive values and hydrophilic have negative values. Although the distribution coefficient can be measured experimentally (Leo *et al.*, 1971), in this

research project, SwissADME (Daina *et al.*, 2014) predictive software to estimate the partition coefficients for selected compounds

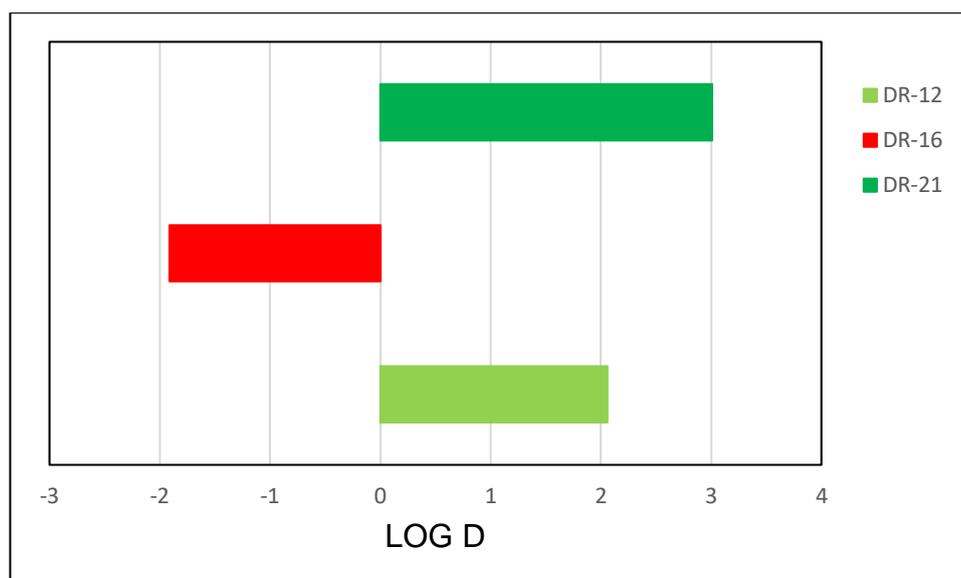
In this research, the partition coefficient of AQ(4OH)- $\beta$ -Ala-OH (DR-12),  $\gamma$ -Abu-Pip-Rho-TFA (DR-16) and AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) were predicted using SwissADME software. This was found to be a simple and easy tool for determining the physicochemical properties of the drug which can be used by the persons who cannot operate the CADD (computer aided drug design) (Daina *et al.*, 2017).

SwissADME uses an average of five predictive models to produce consensus log  $P_{o/w}$  estimation. The positive value of consensus log  $P_{o/w}$  indicates the lipophilic nature of the molecule, whereas negative value of consensus Log  $P_{o/w}$  indicates the hydrophilic nature of the molecule.

According to the SwissADME results (**Table-1**), the partition coefficient values of the quencher (DR-12), the fluorophore (DR-16) and the probe (DR-21) were 2.06, -1.91 and 3.01, respectively. Therefore, the final probe (DR-21) were predicted to be lipophilic and its cleaved products predicted to be lipophilic (DR-12) and hydrophilic (DR-16) in nature (Figure-29).

Compound	consensus Log $P_{o/w}$	Property
DR-12	2.06	Lipophilic
DR-16	-1.91	Hydrophilic
DR-21	3.01	Lipophilic

**Table-1:** SwissADME predicted partition coefficient values.



**Figure-29:** Relative lipophilicity/hydrophilicity of DR-12, DR-16 and DR-21.

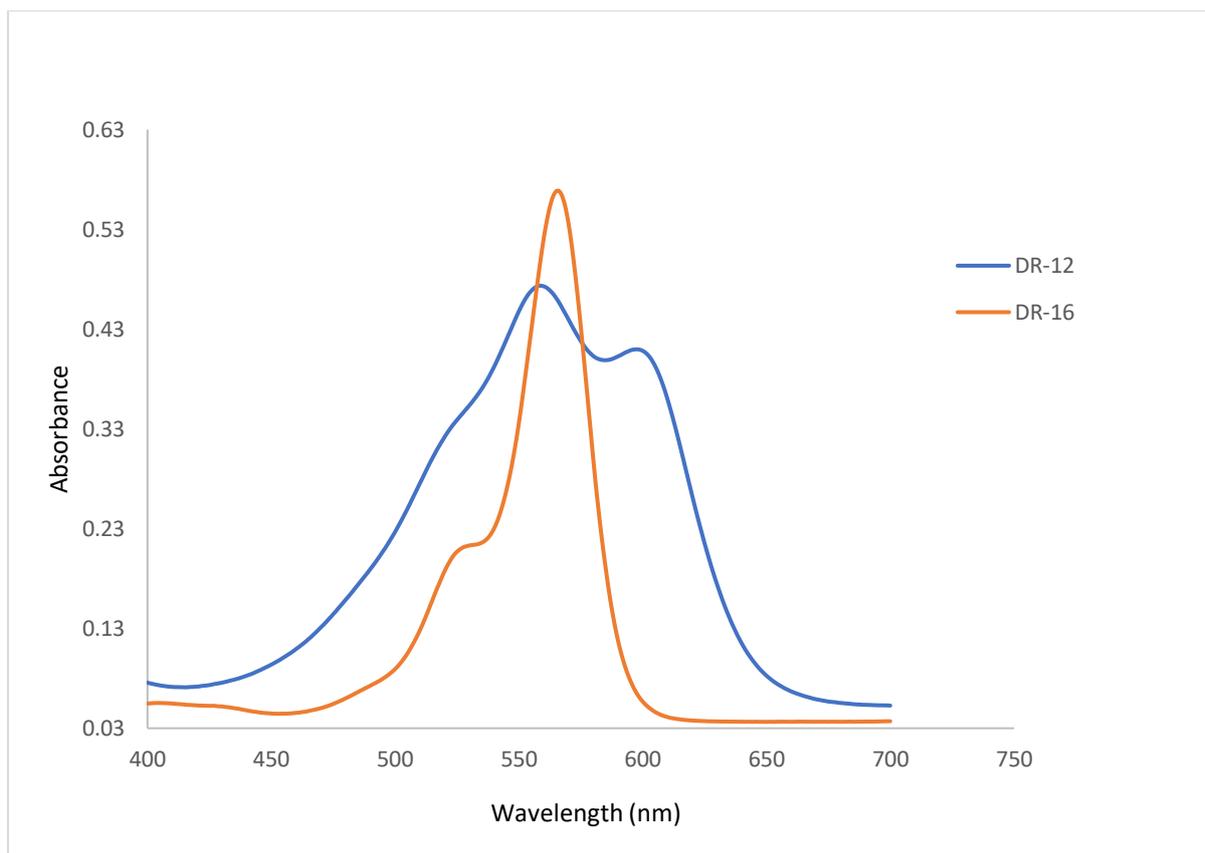
Based on the logP values, both the compounds DR-12 and DR-21 were found to be lipophilic and may be permeable to acidic organelles like lysosomes, recycling endosomes and Golgi complex and become protonated and sequestered. As DR-21 is predicted to be cell permeable, it may be cleaved by legumain to its derived products DR-12 and DR-16 in the lysosome as the legumain is overexpressed in acidic environment of lysosome (Lunde *et al.*, 2019). According to the literature of Lloyd (2000), lysosomes were found to accumulate organic amines at lower pH compared to surrounding cytoplasm (Lloyd, 2000). The lysosomes are also found to be involved in lysosomal trapping process with the ionizable amines of lipophilic and cationic amphiphilic drugs are accumulated. According to Kazmi (2013), the ionised form of lipophilic amines do not diffuse freely back across the lysosomal membrane (Kazmi *et al.*, 2013). Although the predictive software indicates that the fluorophore fragment  $\gamma$ -Abu-Pip-Rho-TFA (DR-16), that would be released upon legumain cleavage, is hydrophilic rather than lipophilic, it does have a primary amino group that could be protonated in and accumulate in lysosomes through this mechanism. Fluorescent  $\gamma$ -Abu-Pip-Rho-TFA (DR-16) could then be detected by fluorescence imaging, with higher fluorescence potentially correlating with greater levels of legumain; these experiments should be a priority in future studies on DR-21.

#### 4.7. DETERMINATION OF MAXIMUM ABSORBANCE OF DR-12 AND DR-16:

The maximum wavelength at which a substance absorbs maximally is called the maximum absorbance or lamda max, denoted by  $\lambda_{\max}$ . The absorbance spectra measure the maximum absorption of a substance which is determined by UV-visible spectroscopy. This was measured for the quencher DR-12 to determine  $\lambda_{\max}$  and also the wavelength range over which DR-12 could act as a quencher for Rodamine-B. The absorbance spectra was recorded for DR-16 because  $\lambda_{\max}$  was used as the excitation wavelength in fluorescence experiments. In this research, we measured the maximum absorbance of DR-12 (50  $\mu\text{M}$ ) and DR-16 (5  $\mu\text{M}$ ) in assay buffer by recording the absorbance spectra between 400-700nm wavelength. **Table-2** and **Figure-30** shows the data of absorption spectrum of DR-12 and DR-16, where the DR-12 shows the maximum absorbency of 0.47355 at 558 nm and DR-16 shows the maximum absorbance of 0.5689 at 565 nm. These results indicate the quencher absorbs between 400-650 nm, with a maximum around 558 nm, and can quench fluorophores that emit within this wavelength range.

Compound	Concentration ( $\mu\text{M}$ )	Wavelength ( $\lambda_{\max}$ ) (nm)	Absorbance
DR-12	50	558	0.4735
DR-16	5	565	0.5689

**Table-2:** Determination of maximum absorbance of DR-12 and DR-16.



**Figure-30:** Absorption spectra of AQ(4OH)- $\beta$ -Ala-OH (DR-12) and  $\gamma$ -Abu-Pip-Rho-TFA (DR-16) (absorbance vs wavelength) in assay buffer at pH 5.0.

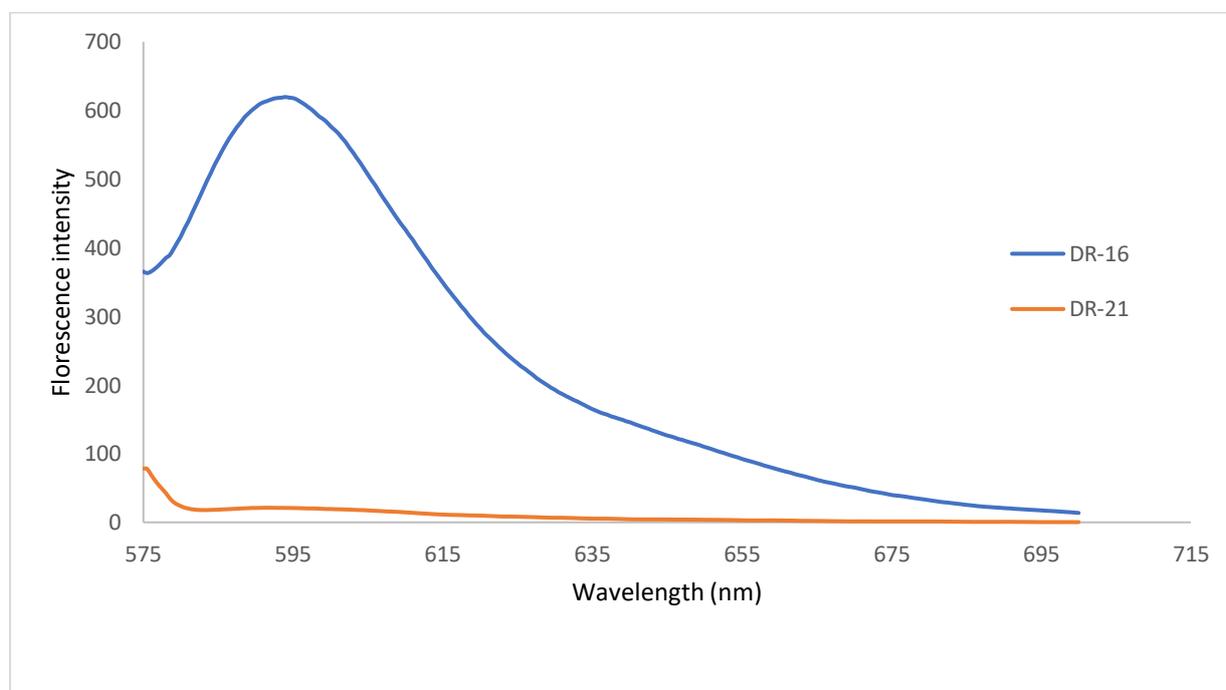
#### 4.8. DETERMINATION OF FLUORESCENCE INTENSITY OF DR-16 AND DR-21:

Emission of light from a fluorophore when it is excited by a shorter wavelength of light is known as fluorescence. The measurement of emitted fluorescence is called as fluorescent intensity. Emission spectrum studies determines the intensity of fluorescence released from the fluorophore and the quenched fluorophore. In this, research, the relative fluorescence intensity of 0.1  $\mu$ M concentration of DR-16 and DR-21 were determined at a given wavelength. The 0.1  $\mu$ M concentration of DR-16 and DR-21 compounds were prepared from their respective stock solutions and diluted by assay buffer, while the fluorescent intensity was measured at fluorescence emission wavelength between 575-700nm, scan speed of 500nm/min with excitation slit and emission slit widths of 4.5 nm and 5.5 nm respectively. **Table-3** and **Figure-31** give the fluorescence intensity of the fluorophore (DR-16) and the probe (DR-21), where the DR-16 had maximum emission at 594 nm with relative fluorescence intensity of

619.55 and the DR-21 had maximum emission at 594 nm with relative fluorescence intensity of 21.0104. These results show that the emission of the fluorophore at 594 nm overlaps with the absorbance wavelength of quencher, with an absorbance range of 400-650 nm. Moreover, the observations of relative fluorescence intensities of DR-16 and DR-21 clearly shows that this was DR-21 was 29 folds lower for DR-21 than DR-16. Therefore, DR-21 proved to be an efficient FRET substrate because of its lower fluorescence emission compared to DR-16 which indicated that the DR-21 was completely quenched.

Compound	Concentration ( $\mu\text{M}$ )	Maximum Emission Wavelength (nm)	Relative Fluorescence Intensity
DR-16	0.1	594	619.55
DR-21	0.1	594	21.01

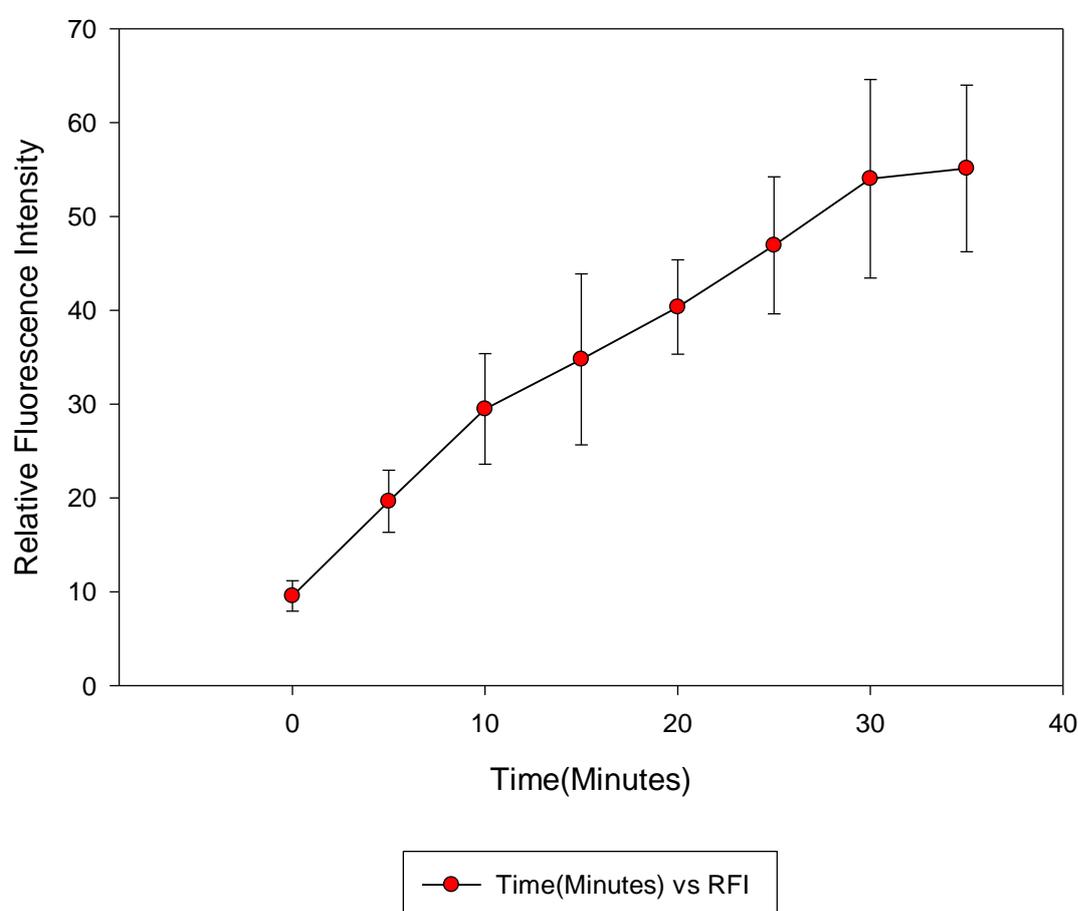
**Table-3:** Determination of Relative Fluorescence Intensity of DR-16 and DR-21.



**Figure-31:** Emission spectra of  $\gamma$ -Abu-Pip-Rho-TFA (DR-16) and AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) (relative fluorescence intensity vs wavelength) in assay buffer at pH 5.0.

#### 4.9. FLUOROGENIC PROBE (DR-21) CLEAVAGE BY rh-LEGUMAIN:

A fluorimetric assay was performed to test if the probe AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) was a substrate by incubating with recombinant human legumain at 37°C in MES assay buffer (pH 5.0) at 37°C. The fluorescence response was observed for every five minutes over two hours. The results showed an increase in fluorescence of over time because of the cleavage of probe by legumain, although further studies using HPLC would be required to show whether this occurred at P1 site of asparagine. **Figure-32** shows increase in fluorescence with time ( $n = 3 \pm SD$ ).



**Figure-32:** Incubation of DR-21 with rh-legumain in MES assay buffer (pH 5.0) at 37°C,  $\lambda_{ex}$  480 nm ( $n = 3$ ).

Therefore, this experiment concluded that the synthesized target fluorogenic molecular probe DR-21 showed increased fluorescence after its cleavage by rh-legumain and that it was a fluorogenic substrate for legumain.

All outcomes of these experiments confirmed the expected behaviour of the legumain substrate FRET probe wherein DR-21 was essentially completely quenched with 29-fold lower fluorescence intensity than DR-16 (reporter fluorophore) under the same conditions. Fluorescence intensity increased with time after activation by rh-legumain (assumed to be) at the P1 position of asparagine (at its C-terminus) of probe DR-21 resulting in release of the fluorophore from quencher.

## 5. CONCLUSION:

In this research project, the designed legumain substrate fluorogenic probe DR21 AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) was successfully synthesized and its cleavage by legumain in vitro was also realized. The quencher was synthesized in good yield and purity from commercially available, inexpensive leucoquinizarin by controlled monoamination and further functionalization. It was shown that the quencher could be successfully used in solid-phase synthesis reactions to cap the N-terminus of the peptide targeted to legumain. Due to its intense purple colour, monitoring the end-capping reaction using classic methods or through the use of HZ22 reagent were not possible. Nevertheless, the N-labelled peptide was successfully and cleanly removed from the resin. The fluorophore was synthesized from inexpensive rhodamine-B dye and stabilized by incorporating a tertiary amine group with piperazine to prevent reversible cyclization commonly observed with rhodamine and derivatives.

The synthesized compounds were characterized by mass spectrometry and NMR spectroscopy. The partition coefficient of DR-12, DR-16 and DR-21 were predicted using Swiss ADME software which indicated DR-16 as hydrophilic whereas DR-12 and DR-21 as lipophilic. The latter supports the proposal that the new molecular probe DR-21 should be cell permeable. Moreover, DR-21 upon incubation with rh-legumain revealed the increase in fluorescence intensity by the enzymatic cleavage of DR-21 (reasonably assumed to be at the C-terminus of Asn) and release of fluorophore which was likely to be taken up by cells based on its predicted partition coefficient value. Clearly, these indications need to be established by experimental determination of partition coefficients and by studying compound entry into cells by microscopy. It is the expectation that DR21 will have advantages over its prototype SM9 if the released fluorophore is able to enter cells and therefore be retained in a tumour mass in contrast to its analogue.

## **5.1. Suggestions for future work:**

In this project, the extent of cell permeability of DR-21 probe on live cancer cells was not measured. So, future work is required for confocal microscopy experiments to observe cellular uptake of the legumain probe and to compare with uptake and sub-cellular localization of its released fluorophore. This information will be useful to assess the potential of the probe as a tool in early cancer diagnosis in response to overexpressed legumain. Furthermore, HPLC methods should be applied to the in vitro incubation of DR21 with rh-legumain to decipher unambiguously the cleavage site as the expected C-terminus of the key Asn residue. The compounds (standards) required are already available from the synthesis procedures (intact probe and its two cleavage products).

## 6. EXPERIMENTAL

### 6.1. MATERIALS AND GENERAL PROCEDURES:

1. TLC (Thin Layer Chromatography): Kieselgel 60 F<sub>254</sub> preloaded aluminium sheets were used. The coloured solutions spotted on the TLC plate were visible to the naked eye and colourless spots were observed under ultra-violet light.
2. Column Chromatography: Silica gel 60 (VWR) was used to prepare the stationary phase of the columns.
3. Mini solvent extraction: Mini solvent extraction prior to TLC when high boiling solvents had been used in the reaction. It was performed by the addition of dichloromethane or chloroform (2 mL) and water (8 mL) to one drop of the reaction mixture. After the separation of aqueous layer and organic layers, the organic layer was used for TLC.
4. All the solvents used in the experiments were obtained from Fischer Scientific, all the amino acids and coupling agents were obtained from Merck; TFA was from Molekula.
5. The UV absorbance of compounds was analysed using a Perkin Elmer Lambda 25 UV Visible spectrophotometer.
6. High resolution electrospray mass spectra were recorded on an LTQ Orbitrap XL or Waters Xevo G2-S (EPSRC National Mass Spectrometry Facility, University of Swansea).
7. The proton <sup>1</sup>H and carbon <sup>13</sup>C Nuclear Magnetic Resonance (NMR) spectra were recorded on a Bruker AC300 NMR spectrometer at 25 °C at 300.1 MHz and 75.47 MHz, respectively with samples dissolved either in DMSO-<sub>d6</sub> or CDCl<sub>3</sub>.
8. Fluorescence was recorded using a Perkin Elmer LS 55 Fluorescence Spectrophotometer.

## 6.2. SYNTHESIS METHODS

### 6.2.1. Synthesis of AQ(4-OH)- $\beta$ -Ala-O<sup>t</sup>Bu (DR-11):

9,10-Dihydroxy-anthracene-1,4-dione (leucoquinizarin) (2 g, 8.26 mmol), L-alanine t-butyl ester hydrochloride (6.006 g, 33.06 mmol) and potassium carbonate (4.1 g, 29.8 mmol) were dissolved in DMF (30 ml) and heated over a water bath (at 95 °C) for 1h. Once cooled, triethylamine (1 mL) was added to the solution and the crude product was oxidised by bubbling through air for 1h then refrigerated for overnight. The crude product was collected by solvent extraction, where the organic layer was washed with water (x5) and dried using anhydrous sodium sulfate, which was filtered off the filtrate was evaporated to dryness.

TLC of crude product (chloroform: ethyl acetate, 19:1):- R<sub>f</sub>: 0.83 (purple product), R<sub>f</sub>: 0.88 (orange) (quinizarin), R<sub>f</sub>: 0.69 (Blue).

TLC of crude product (dichloromethane):- R<sub>f</sub>: 0.34 (purple product), R<sub>f</sub>: 0.39 (orange) quinizarin), R<sub>f</sub>: 0.06 (blue, 1,4-bis product).

The crude product in dichloromethane was purified by column chromatography using silica gel (20x4.5cm) and dichloromethane as a mobile phase solvent system. The fractions containing the pure product (DR-11) were collected, combined, filtered and evaporated using a rotary evaporator to give the title compound (DR-11). Yield (1.2438g, 43%).

HRMS m/z: 757.2729(100%) [2M+Na]<sup>+</sup>; 368.1497 (42%) [M+H]<sup>+</sup>.

Calcd for [C<sub>21</sub>H<sub>22</sub>NO<sub>5</sub>]<sup>+</sup> 368.1492; Found 368.1497.

<sup>1</sup>H NMR (CDCl<sub>3</sub>, 300MHz):  $\delta$  1.51 (s, 9H, AQ-NH-CH<sub>2</sub>-COO-C(CH<sub>3</sub>)<sub>3</sub>), 2.67 (t, 2H, -CH<sub>2</sub>-COO), 3.67 (q, 2H, AQ-NH-CH<sub>2</sub>), 7.22 (d, 2H, H-2 and 3), 7.78 (m, 2H, H-6 and 7), 8.31 (m, 2H, H-5 and 8), 10.29 (t, 1H, AQ-NH), 13.61 (s, 1H, AQ-OH).

### 6.2.2. Synthesis of AQ(4-OH)- $\beta$ -Ala-OH (DR-12):

DR11 (1.24 g, 3.3 mmol) was dissolved in trifluoroacetic acid (TFA) (5 ml) and kept at rt for 3 h; the solvent was evaporated and diethyl ether (25 ml) was added, then

refrigerated for 4h. The precipitate was collected, filtered and dried under vacuum to give the title compound (DR-12). Yield (0.7128 g, 67%).

TLC (dichloromethane: methanol, 9:1):- R<sub>f</sub> value of DR-11:0.9 (purple), R<sub>f</sub> value of DR-12: 0.43 (purple).

HRMS ESI (-) m/z: 266.0823 (100%) [(C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>)<sub>2</sub>]<sup>-</sup>; 310.0721(5%) [M-H]<sup>-</sup>.

Calcd for [C<sub>17</sub>H<sub>12</sub>NO<sub>5</sub>]<sup>-</sup> 310.0721; Found 310.0721.

<sup>1</sup>H NMR (DMSO-<sub>d6</sub>, 300MHz): δ 2.62 (t, 2H, -CH<sub>2</sub>-COOH), 3.60 (t, 2H, AQ-NH-CH<sub>2</sub>), 7.46 (d,1H, AQ-H3), 7.50 (d, 1H, H-2), 7.82 (m, 2H, H-6 and H-7), 8.16 (m, 2H, H-5 and H-8), 10.23 (s, 1H, AQ-NH), 13.56 (s,1H, AQ-OH).

<sup>13</sup>C NMR (DMSO-<sub>d6</sub>, 75.5 MHz): δ (ppm) 38.04 (-ve, CH<sub>2</sub>), 34.04 (-ve, CH<sub>2</sub>), 107.37 (ab, q, C<sub>AR</sub>-AQ), 112.76 (ab,q, C<sub>AR</sub>-AQ), 125.01 (+ve, C<sub>AR</sub>-AQ), 125.78 (+ve, C<sub>AR</sub>-AQ), 126.17 (+ve, C<sub>AR</sub>-AQ), 128.55 (+ve, C<sub>AR</sub>-AQ), 131.73 (ab, q, C<sub>AR</sub>-AQ), 132.75, (+ve, C<sub>AR</sub>-AQ), 134.48, (+ve, CAR-AQ), 134.52, (ab, q, C<sub>AR</sub>-AQ), 146.79 (ab, q, C<sub>AR</sub>-AQ), 156.00 (ab, C<sub>AR</sub>-AQ), 172.74 (ab, C=O), 180.63 (ab, C=O), 186.56 (ab, C=O).

### 6.2.3. Synthesis of Rho-Pip-Boc (DR-13):

A solution of rhodamine-B (1 g, 2.1 mmol) and HATU (1.19 g, 3.1 mmol) in acetonitrile (ACN) (5 ml) was added to DIPEA (1092 μl, 6.3 mmol) and allowed to react for 15 min. Boc-piperazine (0.43 g, 0.22 mmol) in acetonitrile was added to the activated reaction mixture and allowed to react for overnight at rt.

The reaction mixture was dissolved in chloroform (50 ml) and the organic layer was washed with water (5×100 mL), dried with anhydrous sodium sulfate, filtered and evaporated.

TLC of reaction mixture: (Chloroform: Methanol, 9:1):- R<sub>f</sub>: 0.37 (pink product).

The evaporated dried residue was dissolved in chloroform and purified by column chromatography using silica gel (20×7.3cm) where chloroform was the mobile phase. Purification of the product was carried out by increasing the polarity of mobile phase of the column to chloroform: methanol (19:1). The fractions containing the pure product were collected, combined, filtered and evaporated to give the title compound (DR-13).

HRMS ESI (+) m/z: 611.3584 (100%) [M-Cl]<sup>+</sup>.

Calcd for [C<sub>37</sub>H<sub>47</sub>N<sub>4</sub>O<sub>4</sub>]<sup>+</sup> 611.3584; Found 611.3584.

<sup>1</sup>H NMR (CDCl<sub>3</sub>, 300MHz): δ 1.35 (t, 12H, 4× CH<sub>3</sub>-Rho), 1.45 (s, 9H, C(CH<sub>3</sub>)<sub>3</sub>-Boc), 3.31 (m, 8H, 4×CH<sub>2</sub>-N-Pip), 3.65 (q, 8H, 4×CH<sub>2</sub>-N-Rho), 6.80 (d, 2H, Rho-H-4 and 5), 6.97 (d, 2H, Rho-H-2 and 7), 7.25 (d, 2H, Rho-H-1 and 8), 7.38 (m, 1H, Rho-CH<sub>AR</sub>), 7.55 (m, 1H, Rho-CH<sub>AR</sub>), 7.69 (m, 2H, 2×Rho-CH<sub>AR</sub>).

#### 6.2.4. Synthesis of Rho-Pip-TFA (DR-14):

DR-13 was dissolved in trifluoroacetic acid (TFA) (3 ml) and allowed to react for 1 h, followed by evaporation of solvent by rotary evaporator. Diethyl ether (25 ml) was added to the dried residue and refrigerated. The precipitate was collected, filtered and dried to get the title compound (DR-14). Yield (1.8582 g).

TLC (chloroform: methanol, 9:1):- R<sub>f</sub> value of DR-14 0.34 (pink product).

HRMS (+) m/z: 511.3054 (100%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]<sup>+</sup>; 256.1570 (10%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]<sup>2+</sup>.

Calcd for [C<sub>21</sub>H<sub>22</sub>N<sub>4</sub>O<sub>2</sub>]<sup>+</sup> 511.3068; Found 511.3054.

<sup>1</sup>H NMR (DMSO, 300MHz): δ 1.23 (t, 12H, 4× CH<sub>3</sub>-Rho), 3.01 (m, 4H, 2×CH<sub>2</sub>-N-Pip), 3.65 (q, 12H, 4×CH<sub>2</sub>-N-Rho and 2×CH<sub>2</sub>-N-Pip), 6.97 (d, 2H, 2×Rho-CH<sub>AR</sub>), 7.15 (m, 4H, 4×Rho-CH<sub>AR</sub>), 7.54 (m, 1H, Rho-CH<sub>AR</sub>), 7.79 (m, 3H, 3×Rho-CH<sub>AR</sub>), 9.05 (s, 3H, NH<sub>3</sub><sup>+</sup>).

#### 6.2.5. Synthesis of Boc-γ-Abu-Pip-Rho-TFA (DR-15):

Boc-γ-Abu-OH (0.21 g, 0.36 mmol) and HATU (0.39 g, 0.34 mmol) were dissolved in acetonitrile (15 ml) and then DIPEA (398 μL, 0.75 mmol) was added and allowed to react for 15 min for activation. DR-14 (0.6 g, 0.91 mmol) was dissolved in acetonitrile (30 ml) and added to DIPEA (240 μl, 0.45 mmol) and then allowed to react with activated reaction mixture for 1 h. After 1 h, the reaction mixture was dissolved in chloroform and the organic layer was washed with water (×5), dried over anhydrous sodium sulfate, filtered and evaporated.

TLC of reaction mixture: (chloroform: methanol, 9:1):- R<sub>f</sub>: 0.3 (pink product).

The dried residue was purified by silica gel column by increasing the polarity of mobile phase (dichloromethane: 2% methanol to dichloromethane: 8% methanol). Fractions containing the desired compound were collected, combined, filtered and allowed to evaporate overnight.

TLC of compound:(dichloromethane: methanol, 4:1):- R<sub>f</sub>: 0.73 (pink product)

Diethyl ether (6 ml) was added and the residue allowed to solidify in a refrigerator overnight. The solid precipitate was collected, filtered under vacuum and dried to give the title compound (DR15). Yield (0.5573 g, 83%).

HRMS ESI (+) m/z: 696.4108 (100%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]<sup>+</sup>. Calcd for [C<sub>41</sub>H<sub>55</sub>N<sub>5</sub>O<sub>5</sub>]<sup>+</sup> 696.4119; Found 696.4108.

<sup>1</sup>H NMR (CDCl<sub>3</sub>, 300MHz): δ 1.35 (t, 12H, 4× CH<sub>3</sub>-Rho), 1.45 (s, 9H, C(CH<sub>3</sub>)<sub>3</sub>-Boc), 1.85-3.64 (m, 4×CH<sub>2</sub>-Abu, 3×CH<sub>2</sub>-N-Pip and 4×CH<sub>2</sub>-N-Rho), 5.03 (s, 1H, Rho-Pip-γ-Abu-NH), 6.5-7.0 (m, 2H, Rho-CH<sub>AR</sub>), 7.05 (m, 4H, 2×Rho-CH<sub>AR</sub>), 7.4 (s, 1H, Rho-CH<sub>AR</sub>), 7.6 (s, 3H, 2×Rho-CH<sub>AR</sub>).

#### 6.2.6. Synthesis of γ-Abu-Pip-Rho-TFA (DR-16):

DR15 (0.5 g) was dissolved in TFA (3 ml) and allowed to react for 30 min, then evaporated. The diethyl ether was added to the dried residue and then cooled. The solid precipitate was collected from ether by filtration under reduced pressure and dried under vacuum to give the pure compound (DR-16). Yield (0.3837 g, 75%).

TLC of product: (dichloromethane: 8% methanol):- R<sub>f</sub>: 0.5 (pink product).

<sup>1</sup>H NMR (DMSO-<sub>d</sub>6, 300MHz): δ 1.21 (m, 12H, 4× CH<sub>3</sub>-Rho), 1.75 (m, 2H, CH<sub>2</sub>-Abu), 2.4 (t, 2H, CH<sub>2</sub>-Abu), 2.8 (m, 2H, CH<sub>2</sub>-Abu), 3.38 (d, 8H, 4×CH<sub>2</sub>-N-Pip), 3.65 (q, 8H, 4×CH<sub>2</sub>-N-Rho), 4.15 (s, 3H, NH<sub>3</sub>), 6.95 (s, 2H, 2×Rho-CH<sub>AR</sub>), 7.05 (s, 4H, 4×Rho-CH<sub>AR</sub>), 7.5 (t, 1H, Rho-CH<sub>AR</sub>), 7.6- 8.1 (m, 3H, 3×Rho-CH<sub>AR</sub>).

### 6.2.7. Synthesis of Fmoc-Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-17):

DR-16 (0.1 g, 0.134 mmol), Fmoc-Asn(Trt)-OH (0.1 g, 0.16 mmol), HATU (0.23 g, 0.16 mmol) and DIPEA (0.093  $\mu$ L, 0.536 mmol) were dissolved in DMF (10 ml) at rt for 12h. The reaction mixture was dissolved in chloroform and washed with water ( $\times$ 5) and dried over anhydrous sodium sulfate, filtered and evaporated. The dried compound was purified by silica gel column chromatography by increasing the polarity of mobile phase from chloroform to chloroform:8% methanol. Fractions containing the major product were collected, combined, filtered and evaporated. The dried residue was suspended in diethyl ether and cooled. The precipitate was collected, filtered under reduced pressure to give the title compound. Yield (0.0920 g, 58%).

TLC of compound: (chloroform: 8% methanol):- R<sub>f</sub>: 0.25 (pink product).

HRMS ESI (+) m/z: 347.1835 (100%); 1174.5803 (25%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]. Calcd for [C<sub>74</sub>H<sub>76</sub>N<sub>7</sub>O<sub>7</sub>]<sup>+</sup> 1174.5083; Found 1174.5806.

<sup>1</sup>H NMR (CDCl<sub>3</sub>, 300MHz):  $\delta$  1.33 (q, 12H, 4 $\times$  CH<sub>3</sub>-Rho), 1.41- 1.46 (m, 2H, CH<sub>2</sub>-Abu), 1.82 (s, 2H, CH<sub>2</sub>-Abu), 2.37- 2.69 (m, 2H, CH<sub>2</sub>-Abu), 2.84- 2.98 (m, 2H, CH<sub>2</sub>-Asn), 3.34 (d, 8H, 4 $\times$ CH<sub>2</sub>-N-Pip), 3.55 (s, 1H, CH-9-Fmoc), 3.58 (m, 8H, 4 $\times$ CH<sub>2</sub>-N-Rho), 4.22 (m, 2H, Fmoc-CH<sub>2</sub>-CO), 4.55 (s, 1H,  $\alpha$ -CH-Asn), 6.20 (s, 1H, NHCO), 6.41 (s, 1H, NHCO), 6.70 (s, 2H, 2 $\times$ Rho-CH<sub>AR</sub>), 6.80 (s, 1H, NHCO), 7.26- 7.39 (m, 24H, 24 $\times$ CH<sub>AR</sub>), 7.60 (m, 4H, 4 $\times$ CH<sub>AR</sub>), 7.76 (d, 3H, 3 $\times$ Rho-CH<sub>AR</sub>).

### 6.2.8. Synthesis of Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-18):

DR-17 (0.06 g) was dissolved in 20% piperidine in DMF (1 ml) and allowed to react for 15 min, the reaction mixture was dissolved in chloroform followed by washing the organic layer with water ( $\times$ 5 times) and evaporated.

TLC of reaction mixture: (chloroform: methanol, 4:1):- R<sub>f</sub>: 0.5 (pink product)

The evaporated compound was purified by silica gel chromatography using dichloromethane: methanol (4:1) as a mobile phase. Fractions containing the desired compound were collected, combined, filtered and evaporated.

The evaporated compound was purified again by silica gel column chromatography using chloroform: methanol (9:1) as a mobile phase. Fractions containing the desired

pure compound were collected, filtered and evaporated. The evaporated compound was suspended in diethyl ether and cooled to form a precipitate which was collected and dried under vacuum. Yield (0.01 g, 21%).

TLC of compound: (dichloromethane: methanol, 4:1):- R<sub>f</sub>: 0.6 (pink product).

HRMS (+) m/z: 710.4033 (100%); 952.5126 (50%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]. Calcd for [C<sub>59</sub>H<sub>66</sub>N<sub>7</sub>O<sub>5</sub>]<sup>+</sup> 952.5131; Found 952.5126.

### **6.2.9. SOLID PHASE PEPTIDE SYNTHESIS: Synthesis of AQ(4-OH)-β-Ala-Pro-Ala-Asn(Trt)OH (DR-19):**

#### **6.2.9.1. Synthesis of H-Asn(Trt)-Resin**

Fmoc-Asn(Trt)-Novasyn® TGT resin (1 g) was suspended in DCM and shaken in reaction vessel for 1h to swell, then washed with DMF (2×2 mL).

#### **General method for Fmoc Deprotection:**

Step-1: The resin in reaction vessel was treated with 20% piperidine in DMF (5 ml) and allowed to shake for 10 min.

Step-2: The 20% piperidine in DMF was drained off from reaction vessel.

Step-3: Steps-1 and 2 were repeated

Step-4: The deprotected resin was washed with DCM (3×4 mL)

#### **General method for colour test:**

A sample of the resin beads were added to 2-3 drops of HZ22 reagent (pentafluorophenolate active ester) and allowed to react for 5 min and washed with DCM, the formation of red coloured beads indicated the completion of deprotection reaction whereas the formation of colourless beads indicates the completion of coupling reactions with added Fmoc-protected amino acids.

#### **6.2.9.2. General method for Fmoc amino acid coupling**

##### **Synthesis of H-Ala-Asn(Trt)-Resin:**

Step-1: Fmoc-Ala-OH (0.09 g, 0.28 mmol), PyBOP (0.14 g, 0.27 mmol), HOBT (0.04 g, 0.27 mmol) were dissolved in DMF (10 ml) and then added to DIPEA (150 μL, 0.836 mmol) and allowed to activate for 15 min.

Step-2: Activated reaction mixture was added to resin in reaction vessel and shaken for 30 min then the reaction mixture was drained off from the reaction vessel.

Step-3: Steps-1 and 2 were repeated once again and resin was washed with DCM (3x 3mL).

#### **DEPROTECTION OF Fmoc-group:**

The amino acid coupled resin was deprotected by deprotection steps and washed with DCM (3x 3mL) to perform the colour test for the confirmation of deprotection reaction.

#### **6.2.9.3. Synthesis of H-Pro-Ala-Asn(Trt)-Resin:**

Step-1: Fmoc-Pro-OH (0.1 g, 0.28 mmol), PyBOP (0.14 g, 0.27 mmol), HOBt (0.04 g, 0.27 mmol) were dissolved in DMF (10 ml) and then added to DIPEA (150  $\mu$ L, 0.836 mmol) and allowed to activate for 15 min.

Step-2: Activated reaction mixture was added to resin in the reaction vessel and shaken for 30min on an electric shaker then the reaction mixture was drained off from the reaction vessel.

Step-3: Steps-1 and 2 were repeated once again and resin was washed with DCM 3x 3mL) to perform the colour test for confirmation of a coupling reaction.

#### **DEPROTECTION OF Fmoc-group:**

The coupled resin was deprotected by deprotection steps and washed with DCM 3x 3mL) to perform the colour test for the confirmation of deprotection reaction.

#### **6.2.9.4. Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)-Resin:**

Step-1: DR-12 (0.09 g, 0.28 mmol), PyBOP (0.14 g, 0.27 mmol), HOBt (0.04 g, 0.27 mmol) were dissolved in DMF (10 ml) and then added to DIPEA (150  $\mu$ L, 0.836 mmol) and allowed to react for 45 min for activation.

Step-2: Activated reaction mixture was added to resin in reaction vessel and shaken for 2 hrs on electric shaker then the reaction mixture was drained off from the reaction vessel.

Step-3: The resin was washed with DMF (1x 3mL) then followed by washing with DMF (3x 3mL) and DCM (3x 3mL) alternatively.

Step-4: Colour test for coupling reaction:

The formation of purple colour resin beads indicates the successful coupling reaction.

Step-5: After confirmation of coupling by colour test, the resin was washed for 3 times with DMF for 2<sup>nd</sup> time coupling.

Step-6: Steps from 1 to 3 were repeated and the resin was suspended in DCM in reaction vessel.

#### **6.2.9.5. Cleavage of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)-OH (DR19) from Resin:**

Resin-bound DR-19 was suspended in DCM was washed with 0.25% TFA in DCM until all the purple coloured compound was collected, then the resin was washed with DCM for 3 times. The filtrate containing the compound was evaporated and suspended in diethyl ether to form the solid precipitate of title compound. Yield (0.2304 g, 93%).

HRMS ESI (-) m/z: 834.3137 (100%) [M-H]<sup>-</sup>. Calcd for [C<sub>48</sub>H<sub>44</sub>N<sub>5</sub>O<sub>9</sub>]<sup>-</sup> 834.3145; Found 834.3137.

<sup>1</sup>H NMR (DMSO-d<sub>6</sub>, 300MHz):  $\delta$  1.1.-1.3 (m, 5H, -CH<sub>2</sub> and CH<sub>3</sub>-Ala), 2.0 (m, 2H, CH<sub>2</sub>), 2.8 (m, 2H, CH<sub>2</sub>), 3.0 (m, 4H, 2xCH<sub>2</sub>), 3.65 (m, 4H, 2xCH), 4.5 (m, 3H, 3x $\alpha$ -CH), 7.19-7.3 (m, 15H, Asn-CH-Trt), 7.4 (m, 1H, AQ-CH<sub>AR</sub>-H<sub>2</sub>), 7.6 (m, 1H, AQ-CH<sub>AR</sub>-H<sub>3</sub>), 8.0 (m, 2H, 2xAQ-CH<sub>AR</sub>-H<sub>6</sub> and H<sub>7</sub>), 8.1(d, 1H, NHCO), 8.25 (m, 2H, 2xAQ-CH<sub>AR</sub>), 8.85 (d, 1H, NHCO), 9.0 (d, 1H, NHCO), 9.0 (d, 1H, NHCO), 10.36 (t, 1H, AQ-NH), 13.68 (s, 1H, AQ-OH).

#### **6.2.10. Synthesis of AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn(Trt)- $\gamma$ -Abu-Pip-Rho (DR-20):**

DR-19 (1.6 g, 0.09 mmol), PyBOP (1.2 g, 0.11 mmol), and HOBt (0.04 g, 0.11 mmol) in DMF were added to DIPEA (120  $\mu$ L, 0.329 mmol), allowed to react for 20 mins for activation of the reaction mixture. After 20 mins, DR-16 (1.4 g, 0.09 mmol) in DMF was added and stirred for 1 h. After 1 h, the organic layer of reaction mixture was dissolved in DCM (10 mL) and washed with water (5x 100 mL) and dried by adding anhydrous sodium sulphate followed by filtration, evaporation and purification by silica gel column using DCM: 4% methanol as a mobile phase. Fractions containing the product were combined, filtered, evaporated and suspended in ether to form the solid precipitate. The solid precipitate was collected and dried under vacuum to give the title compound (DR20). Yield (0.0961 g, 33%).

TLC of compound: (dichloromethane: methanol, 9:1):- R<sub>f</sub>: 0.42 (purple product).

HRESMS (+) m/z: 718.3300 (100%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]<sup>2+</sup>; 1413.6708 (35%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]<sup>+</sup>.

Calcd for [C<sub>84</sub>H<sub>89</sub>N<sub>10</sub>O<sub>11</sub>]<sup>+</sup> 1413.6707; Found 1413.6708.

#### **6.2.11. Synthesis of AQ(4-OH)-β-Ala-Pro-Ala-Asn-γ-Abu-Pip-Rho (DR-21):**

TFA was added to DR-20 (0.08 g) and allowed to react for 3 h with addition of a drop of water. After 3 h, the reaction mixture was evaporated and the residue was purified by silica gel column by using chloroform: methanol (4:1) as a mobile phase. Fractions containing the compound were collected, combined, filtered, evaporated and again columned for purification by silica gel column and by increasing the polarity of mobile phase from chloroform: methanol, 9:1 to chloroform: methanol, 4:1. Fractions containing the pure product were collected, combined, filtered, evaporated and suspended in diethyl ether and refrigerated for overnight. The solid precipitated was collected from ether and dried to give the desired title compound (DR-21). (0.044 g, 66%).

TLC of compound: (chloroform: methanol, 4:1):- R<sub>f</sub>: 0.5 (purple product).

HRMS ESI (+) m/z: 597.2745 (100%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>+Na]<sup>2+</sup>; 1171.5607(25%) [M- C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>]<sup>+</sup>.

Calcd for [C<sub>65</sub>H<sub>75</sub>N<sub>10</sub>O<sub>11</sub>]<sup>+</sup> 1171.5611; Found 1171.5607.

Calcd for [C<sub>65</sub>H<sub>75</sub>N<sub>10</sub>O<sub>11</sub>Na]<sup>2+</sup> 597.2752; Found 597.2745.

### **6.3. DETERMINATION OF PARTITION COEFFICIENTS:**

The partition coefficient of compounds DR-12, DR-16 and DR-21 were estimated using SwissADME software where the consensus value indicates the logP. <http://www.swissadme.ch>

## 6.4. UV-VIS ABSORPTION ASSAY

### Materials:

Assay buffer [50 mM MES hydrate, 250 mM NaCl, pH 5.0]; quartz cuvette for UV-Vis spectrophotometer-3 mL; Compounds tested:  $\beta$ -Ala-AQ(4-OH) (DR-12) (1 mg/mL stock solution in DMSO, diluted to 50  $\mu$ M concentration by assay buffer) and  $\gamma$ -Abu-Pip-Rho-TFA (DR-16) (1 mg/mL stock solution in DMSO, diluted to 5  $\mu$ M concentration by assay buffer), Absorption wavelength range: 400-700 nm.

### Procedure:

- 1 mg/mL stock solution of DR-12 was prepared by dissolving 1.47 mg of DR-12 in 1.47 ml of DMSO.

For the determination of absorbance of DR-12, the 50  $\mu$ M concentration of DR-12 was prepared by diluting 47  $\mu$ L of DR-12 from 1 mg/ml of stock solution with 2953  $\mu$ L of Assay buffer in 3000  $\mu$ L quartz cuvette and measured the absorbance between 400-700 nm from the UV-Vis absorption spectrum.

- 1 mg/mL stock solution of DR-16 was prepared by dissolving 1.44 mg of DR-16 in 1.44 ml of DMSO.

For the determination of absorbance of DR-16, the 5  $\mu$ M concentration of DR-16 was prepared by diluting 16  $\mu$ L of DR-16 from 1 mg/ml of stock solution with 2984  $\mu$ L of Assay buffer in 3000  $\mu$ L quartz cuvette and measured the absorbance between 400-700 nm from the UV-Vis absorption spectrum.

## 6.5. FLUORESCENCE ASSAY

### Materials:

Assay buffer [50 mM MES hydrate, 250 mM NaCl, pH 5.0]; fluorescence cuvette-3 mL. Compounds tested:  $\gamma$ -Abu-Pip-Rho-TFA (DR-16) (1 mg/mL stock solution in DMSO, diluted to 0.1  $\mu$ M concentration by assay buffer) and AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) (1 mg/mL stock solution in DMSO, diluted to 0.1  $\mu$ M concentration by assay buffer).

**Procedure:**

- 1 mg/mL stock solution of DR-16 was prepared by dissolving 1.44 mg of DR-16 in 1.44 ml of DMSO.

For the determination of fluorescence intensity of DR-16, the 0.1  $\mu$ M concentration of DR-16 was prepared by diluting 0.322  $\mu$ L of DR-16 from 1 mg/ml of stock solution with 2999.67  $\mu$ L of Assay buffer in a 3000  $\mu$ L fluorescence cuvette and measured the fluorescence emission between 575-700 nm, A 1 mg/mL stock solution of DR-21 was prepared by dissolving 1.42 mg of DR-21 in 1.42 ml of DMSO.

For the determination of fluorescence intensity of DR-21, the 0.1  $\mu$ M concentration of DR-21 was prepared by diluting 0.386  $\mu$ L of DR-21 from 1 mg/ml of stock solution with 2999.61  $\mu$ L of Assay buffer in a 3000  $\mu$ L fluorescence cuvette and measured the fluorescence emission between 575-700 nm.

**6.6. PROBE AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21) ACTIVATION USING RECOMBINANT HUMAN LEGUMAIN (rh-legumain):****Materials:**

Activation buffer (50mM sodium acetate, 100mM NaCl, pH 4.0), assay buffer [50mM MES hydrate (Sigma) 250mM NaCl, pH 5.0], recombinant human legumain, AQ(4-OH)- $\beta$ -Ala-Pro-Ala-Asn- $\gamma$ -Abu-Pip-Rho (DR-21).

**Procedure:**

Preparation of 5 $\mu$ L recombinant human legumain aliquots:

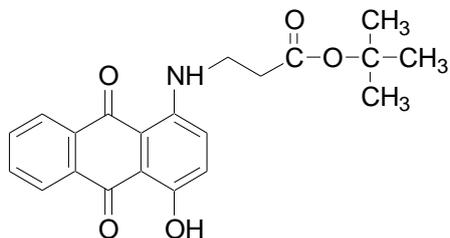
Recombinant human legumain (10  $\mu$ g) was diluted with activation buffer (100  $\mu$ L, pH 4.0) and divided into 20 aliquots of 5  $\mu$ L (each contains 500 ng of rh-legumain). These aliquots were frozen at -80°C.

**Procedure for Probe Fluorescence activation Assay:**

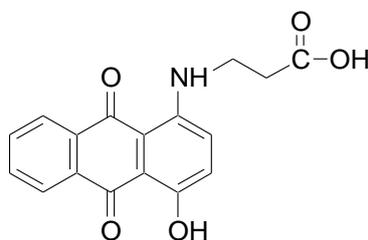
An aliquot of rh-legumain (5  $\mu$ L) was defrosted by incubating in a water bath for 2 h at 37°C, followed by dilution with assay buffer (495  $\mu$ L, pH 5.0) to make 1 ng/ $\mu$ L. A solution of DR-21 (1  $\mu$ M) was prepared by diluting with assay buffer from 1 mg/mL of stock solution of DR-21 in DMSO. The probe was activated by adding 3.85  $\mu$ L of DR-

21 with 176.15  $\mu\text{L}$  rh-legumain followed by diluting with assay buffer (120  $\mu\text{L}$ ) to make up the final volume of fluorescence cuvette (300  $\mu\text{L}$ ). The relative intensity of fluorescence of probe DR-21 with rh-legumain was observed for every 5 min time interval at 37°C.

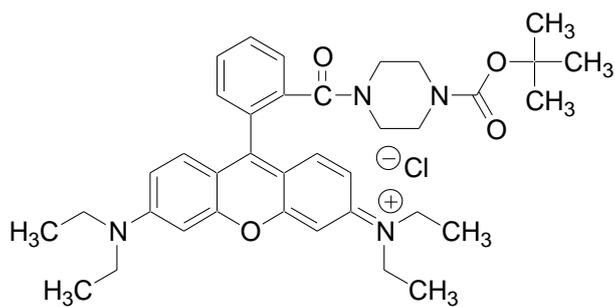
## 7. STRUCTURE LIBRARY



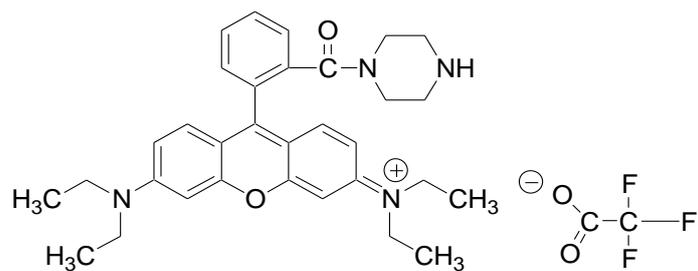
AQ(4-OH)-β-Ala-O<sup>t</sup>Bu (DR-11)



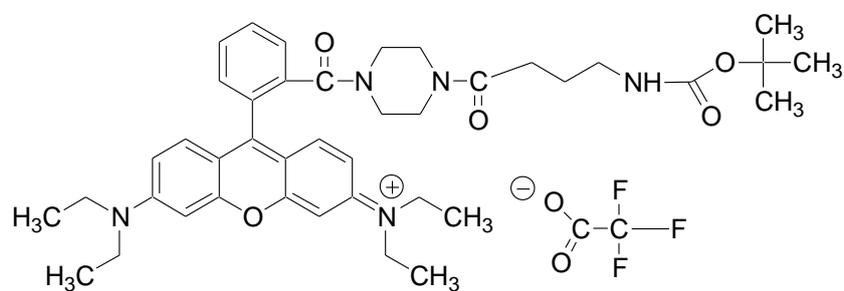
AQ(4OH)-β-Ala-OH (DR-12)



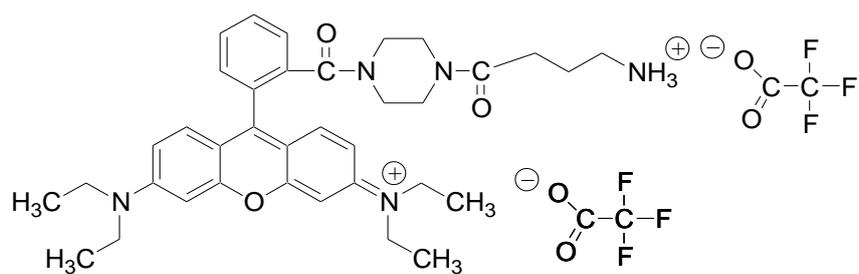
Rho-Pip-Boc (DR-13)



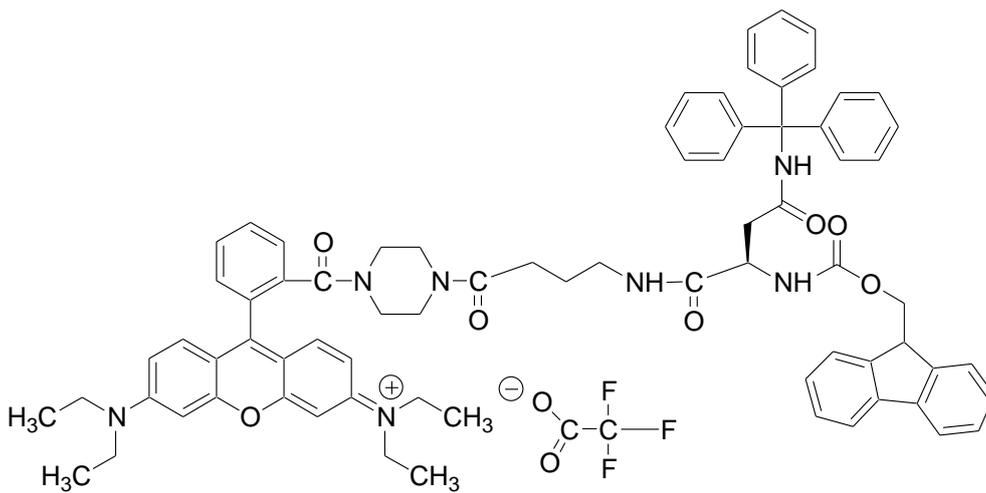
Rho-Pip-TFA (DR-14)



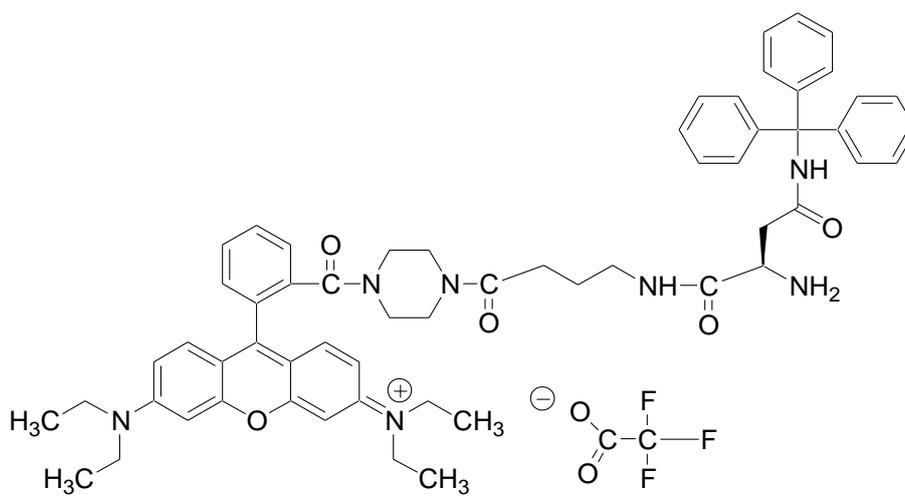
Boc- $\gamma$ -Abu-Pip-Rho-TFA (DR-15)



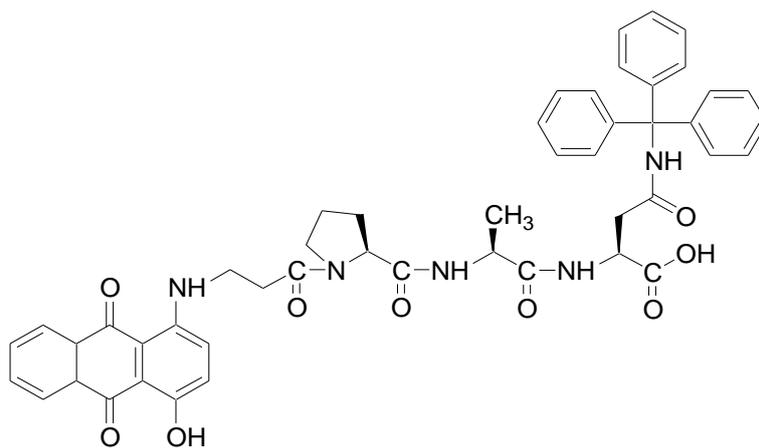
$\gamma$ -Abu-Pip-Rho-TFA (DR-16)



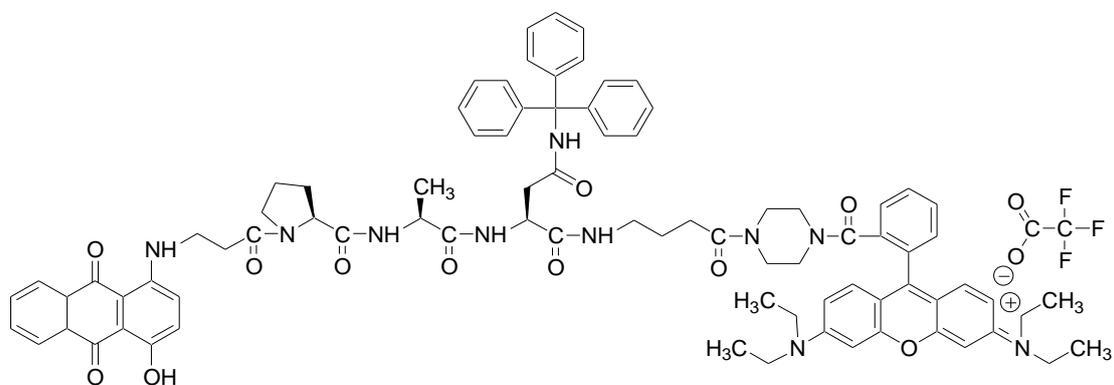
Fmoc-Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-17)



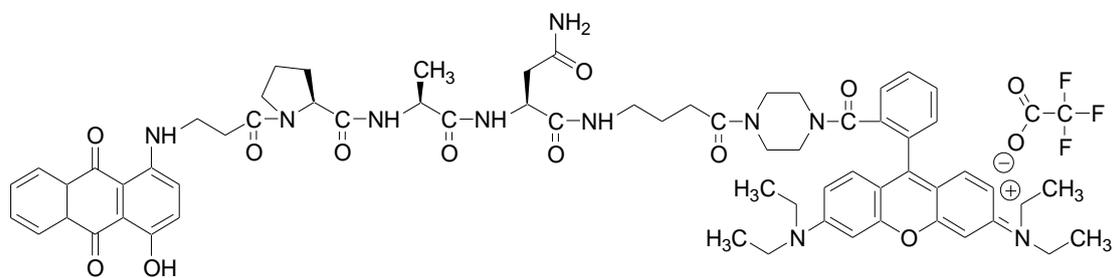
Asn(Trt)- $\gamma$ -Abu-Pip-Rho-TFA (DR-18)



AQ(4-OH)-β-Ala-Pro-Ala-Asn(Trt)OH (DR-19)



AQ(4-OH)-β-Ala-Pro-Ala-Asn(Trt)-γ-Abu-Pip-Rho (DR-20)



AQ(4-OH)-β-Ala-Pro-Ala-Asn-γ-Abu-Pip-Rho (DR-21)

## 8. REFERENCES

- Ahdoot, M., Wilbur, A. R., Reese, S. E., Lebastchi, A. H., Mehralivand, S., Gomella, P. T., Bloom, J., Gurram, S., Siddiqui, M., Pinsky, P., Parnes, H., Linehan, W. M., Merino, M., Choyke, P. L., Shih, J. H., Turkbey, B., Wood, B. J., & Pinto, P. A. (2020). MRI-Targeted, Systematic, and Combined Biopsy for Prostate Cancer Diagnosis. *New England Journal of Medicine*, *382*(10), 917–928. <https://doi.org/10.1056/NEJMoa1910038>
- Barasch, D., Zipori, O., Ringel, I., Ginsburg, I., Samuni, A., & Katzhendler, J. (1999). Novel anthraquinone derivatives with redox-active functional groups capable of producing free radicals by metabolism: Are free radicals essential for cytotoxicity? *European Journal of Medicinal Chemistry*, *34*(7–8), 597–615. [https://doi.org/10.1016/S0223-5234\(00\)80029-X](https://doi.org/10.1016/S0223-5234(00)80029-X)
- Behrendt, R., White, P., & Offer, J. (2016). Advances in Fmoc solid-phase peptide synthesis. In *Journal of Peptide Science* (Vol. 22, Issue 1, pp. 4–27). John Wiley and Sons Ltd. <https://doi.org/10.1002/psc.2836>
- Beija, M., Afonso, C. A. M., & Martinho, J. M. G. (2009). Synthesis and applications of Rhodamine derivatives as fluorescent probes. *Chemical Society Reviews*, *38*(8), 2410. <https://doi.org/10.1039/b901612k>
- Blum, G., Weimer, R. M., Edgington, L. E., Adams, W., & Bogoyo, M. (2009). Comparative Assessment of Substrates and Activity Based Probes as Tools for Non-Invasive Optical Imaging of Cysteine Protease Activity. *PLoS ONE*, *4*(7), 6374. <https://doi.org/10.1371/journal.pone.0006374>
- Bodanszky, M., Deshmane, S. S., & Martinez, J. (1979). Side Reactions in Peptide Synthesis. 11.1 Possible Removal of the 9-Fluorenylmethoxycarbonyl Group by the Amino Components during Coupling. *Journal of Organic Chemistry*, *44*(10), 1622–1625. <https://doi.org/10.1021/jo01324a008>
- Carpino, L. A., & Han, G. Y. (1972). The 9-Fluorenylmethoxycarbonyl Amino-Protecting Group. *Journal of Organic Chemistry*, *37*(22), 3404–3409. <https://doi.org/10.1021/jo00795a005>

- Chen, J. M., Dando, P. M., Rawlings, N. D., Brown, M. A., Young, N. E., Stevens, R. A., Hewitt, E., Watts, C., & Barrett, A. J. (1997). Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *The Journal of Biological Chemistry*, 272(12), 8090–8098. <https://doi.org/10.1074/JBC.272.12.8090>
- Chen, S., Dong, H., Yang, S., & Guo, H. (2017). Cathepsins in digestive cancers. In *Oncotarget* (Vol. 8, Issue 25, pp. 41690–41700). Impact Journals LLC. <https://doi.org/10.18632/oncotarget.16677>
- Choe, Y., Leonetti, F., Greenbaum, D. C., Lecaille, F., Bogoyo, M., Brömme, D., Ellman, J. A., & Craik, C. S. (2006). Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *The Journal of Biological Chemistry*, 281(18), 12824–12832. <https://doi.org/10.1074/jbc.M513331200>
- Chowdhury, M. A., Moya, I. A., Bhilocha, S., McMillan, C. C., Vigliarolo, B. G., Zehbe, I., & Phenix, C. P. (2014). Prodrug-inspired probes selective to cathepsin B over other cysteine cathepsins. *Journal of Medicinal Chemistry*, 57(14), 6092–6104. <https://doi.org/10.1021/jm500544p>
- Clegg, R. M. (1995). Fluorescence resonance energy transfer. In *Current Opinion in Biotechnology* (Vol. 6). [https://ac-els-cdn-com.ezproxy.napier.ac.uk/0958166995800166/1-s2.0-0958166995800166-main.pdf?\\_tid=09868820-27d2-4895-accd-161c3595f6db&acdnat=1550072010\\_bae31ca622730485fa16b76c1c304731](https://ac-els-cdn-com.ezproxy.napier.ac.uk/0958166995800166/1-s2.0-0958166995800166-main.pdf?_tid=09868820-27d2-4895-accd-161c3595f6db&acdnat=1550072010_bae31ca622730485fa16b76c1c304731)
- Cooper, G., & Sunderland, M. (2000). *The Cell: A Molecular Approach 2nd Edition*. In *Sinauer Associates* (pp. 1–820).
- Daina, A., Michielin, O., & Zoete, V. (2014). ILOGP: A simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *Journal of Chemical Information and Modeling*, 54(12), 3284–3301. <https://doi.org/10.1021/ci500467k>
- Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7. <https://doi.org/10.1038/srep42717>

- Dall, E., & Brandstetter, H. (2016). Structure and function of legumain in health and disease. *Biochimie*, 122, 126–150. <https://doi.org/10.1016/J.BIOCHI.2015.09.022>
- Edgington-Mitchell, L. E., Wartmann, T., Fleming, A. K., Gocheva, V., Van Der Linden, W. A., Withana, N. P., Verdoes, M., Aurelio, L., Edgington-Mitchell, D., Lieu, T., Parker, B. S., Graham, B., Reinheckel, T., Furness, J. B., Joyce, J. A., Storz, P., Halangk, W., Bogyo, M., & Bunnett, N. W. (2016). Legumain is activated in macrophages during pancreatitis. *Am J Physiol Gastrointest Liver Physiol*, 311, 548–560. <https://doi.org/10.1152/ajpgi.00047.2016.-Pancre>
- Feng, M., Yang, X., Ma, Q., & He, Y. (2017). Retrospective analysis for the false positive diagnosis of PET-CT scan in lung cancer patients. *Medicine*, 96(42), e7415. <https://doi.org/10.1097/MD.00000000000007415>
- Fields, G. B., & Noble, R. L. (2009). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *International Journal of Peptide and Protein Research*, 35(3), 161–214. <https://doi.org/10.1111/j.1399-3011.1990.tb00939.x>
- Gawenda, J., Traub, F., Lück, H. J., Kreipe, H., & von Wasielewski, R. (2007). Legumain expression as a prognostic factor in breast cancer patients. *Breast Cancer Research and Treatment*, 102(1), 1–6. <https://doi.org/10.1007/s10549-006-9311-z>
- Gobbo, P., Gunawardene, P., Luo, W., & Workentin, M. S. (2015). Synthesis of a toolbox of clickable rhodamine B derivatives. *Synlett*, 26(9), 1169–1174. <https://doi.org/10.1055/s-0034-1380191>
- Goodacre, J., Ponsford, R. J., Sterling, I., Muraki, M., & Mizoguchi, T. (1994). Selective Removal of an IV-BOC Protecting Group in the Presence of a ferri-Butyl Ester and Other Acid-Sensitive Groups. In *Chem. Pharm. Bull* (Vol. 59, Issue 2). <https://pubs.acs.org/sharingguidelines>
- Guo, P., Zhu, Z., Sun, Z., Wang, Z., Zheng, X., & Xu, H. (2013). Expression of Legumain Correlates with Prognosis and Metastasis in Gastric Carcinoma. *PLoS ONE*, 8(9), e73090. <https://doi.org/10.1371/journal.pone.0073090>

- Haugen, M. H., Boye, K., Nesland, J. M., Pettersen, S. J., Egeland, E. V., Tamhane, T., Brix, K., Maelandsmo, G. M., & Flatmark, K. (2015). High expression of the cysteine proteinase legumain in colorectal cancer - Implications for therapeutic targeting. *European Journal of Cancer*, *51*(1), 9–17. <https://doi.org/10.1016/j.ejca.2014.10.020>
- He, G., Yang, L., Qian, X., Li, J., Yuan, Z., Li, C., Li, J., Li, J., Yuan, Z., Yuan, Z., Li, C., & Li, C. (2017). A coumarin-based fluorescence resonance energy transfer probe targeting matrix metalloproteinase-2 for the detection of cervical cancer. *International Journal of Molecular Medicine*, *39*(6), 1571–1579. <https://doi.org/10.3892/ijmm.2017.2974>
- Hong, Z., Deng, S., Shi, Y., Xie, Y., You, J., Wang, W., Huang, H., & Liu, Z. (2020). Pretargeted radioimmunoimaging with a biotinylated D-D<sub>3</sub> construct and <sup>99m</sup>Tc-DTPA-biotin: strategies for early diagnosis of small cell lung cancer. *Journal of International Medical Research*, *48*(7), 030006052093716. <https://doi.org/10.1177/0300060520937162>
- Isidro-Llobet, A., Álvarez, M., & Albericio, F. (2009). Amino Acid-Protecting Groups. *Chemical Reviews*, *109*(6), 2455–2504. <https://doi.org/10.1021/cr800323s>
- Jad, Y. E., Acosta, G. A., Khattab, S. N., De La Torre, B. G., Govender, T., Kruger, H. G., El-Faham, A., & Albericio, F. (2015). Peptide synthesis beyond DMF: THF and ACN as excellent and friendlier alternatives. *Organic and Biomolecular Chemistry*, *13*(8), 2393–2398. <https://doi.org/10.1039/c4ob02046d>
- Kallur, K., Ramachandra, P., Rajkumar, K., Swamy, S., Desai, I., Rao, R., Patil, S., Sridhar, P., Madhusudhan, N., Krishnappa, R., Bhadrasetty, V., Kumara, H., Santhosh, S., & Ajaikumar, B. (2017). Clinical utility of gallium-68 PSMA PET/CT scan for prostate cancer. *Indian Journal of Nuclear Medicine*, *32*(2), 110–117. <https://doi.org/10.4103/0972-3919.202255>
- Kazmi, F., Hensley, T., Pope, C., Funk, R. S., Loewen, G. J., Buckley, D. B., & Parkinson, A. (2013). Lysosomal sequestration (trapping) of lipophilic amine (cationic amphiphilic) drugs in immortalized human hepatocytes (Fa2N-4 cells). *Drug Metabolism and Disposition*, *41*(4), 897–905. <https://doi.org/10.1124/dmd.112.050054>

- Kikuchi, M., Yamagishi, T., & Hida, M. (1982). Kinetic Studies on the Amination of Leucoquinizarin. In *Bulletin of the Chemical Society of Japan* (Vol. 55, Issue 4, pp. 1209–1212). <https://doi.org/10.1246/bcsj.55.1209>
- Kim, K.-W., Roh, J. K., Wee, H.-J., Kim, C., Kim, K.-W., Roh, J. K., Wee, H.-J., & Kim, C. (2016). Advancement of the Science and History of Cancer and Anticancer Drugs. In *Cancer Drug Discovery* (pp. 25–56). Springer Netherlands. [https://doi.org/10.1007/978-94-024-0844-7\\_2](https://doi.org/10.1007/978-94-024-0844-7_2)
- Kodama, S., Asano, S., Moriguchi, T., Sawai, H., & Shinozuka, K. (2006). Novel fluorescent oligoDNA probe bearing a multi-conjugated nucleoside with a fluorophore and a non-fluorescent intercalator as a quencher. *Bioorganic & Medicinal Chemistry Letters*, 16, 2685–2688. <https://doi.org/10.1016/j.bmcl.2006.02.016>
- Kumar, A., & Jaitak, V. (2019). Natural products as multidrug resistance modulators in cancer. In *European Journal of Medicinal Chemistry* (Vol. 176, pp. 268–291). Elsevier Masson SAS. <https://doi.org/10.1016/j.ejmech.2019.05.027>
- Leas, D. A., Wu, J., Ezell, E. L., Garrison, J. C., Vennerstrom, J. L., & Dong, Y. (2018). Formation of 2-Imino Benzo[e]-1,3-oxazin-4-ones from Reactions of Salicylic Acids and Anilines with HATU: Mechanistic and Synthetic Studies. *ACS Omega*, 3(1), 781–787. <https://doi.org/10.1021/acsomega.7b01824>
- Lee, J., & Bogyo, M. (2010). *Development of Near-Infrared Fluorophore (NIRF)-Labeled Activity-Based Probes for in Vivo Imaging of Legumain*. <https://doi.org/10.1021/cb900232a>
- Leo, A., Hansch, C., & Elkins, D. (1971). Partition coefficients and their uses. *Chemical Reviews*, 71(6), 525–616. <https://doi.org/10.1021/cr60274a001>
- Li, N., Liu, Q., Su, Q., Wei, C., Lan, B., Wang, J., Bao, G., Yan, F., Yu, Y., Peng, B., Qiu, J., Yan, X., Zhang, S., & Guo, F. (2013). Effects of legumain as a potential prognostic factor on gastric cancers. *Medical Oncology*, 30(3), 621. <https://doi.org/10.1007/s12032-013-0621-9>

- Li, S.-S., Xu, Y.-W., Wu, J.-Y., Tan, H.-Z., Wu, Z.-Y., Xue, Y.-J., Zhang, J.-J., Li, E.-M., & Xu, L.-Y. (2017). Plasma Riboflavin Level is Associated with Risk, Relapse, and Survival of Esophageal Squamous Cell Carcinoma. *Nutrition and Cancer*, 69(1), 21–28. <https://doi.org/10.1080/01635581.2017.1247890>
- Li, X., Huang, D. bing, Zhang, Q., Guo, C. xiang, Fu, Q. han, Zhang, X. chen, Tang, T. Y., Su, W., Chen, Y. W., Chen, W., Ma, T., Gao, S. L., Que, R. S., Bai, X. L., & Liang, T. B. (2020). The efficacy and toxicity of chemotherapy in the elderly with advanced pancreatic cancer. *Pancreatology*, 20(1), 95–100. <https://doi.org/10.1016/j.pan.2019.11.012>
- Liao, D., Liu, Z., Wrasidlo, W., Chen, T., Luo, Y., Xiang, R., & Reisfeld, R. A. (2011). Synthetic enzyme inhibitor: a novel targeting ligand for nanotherapeutic drug delivery inhibiting tumor growth without systemic toxicity. *Nanomedicine: Nanotechnology, Biology and Medicine*, 7(6), 665–673. <https://doi.org/10.1016/J.NANO.2011.03.001>
- 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 1. In *Cancer Research* (Vol. 63). <http://cancerres.aacrjournals.org/content/canres/63/11/2957.full.pdf>
- Lloyd, J. B. (2000). Lysosome membrane permeability: Implications for drug delivery. *Advanced Drug Delivery Reviews*, 41(2), 189–200. [https://doi.org/10.1016/S0169-409X\(99\)00065-4](https://doi.org/10.1016/S0169-409X(99)00065-4)
- Luna, O. F., Gomez, J., Cárdenas, C., Albericio, F., Marshall, S. H., & Guzmán, F. (2016). Deprotection reagents in Fmoc solid phase peptide synthesis: Moving away from piperidine? *Molecules*, 21(11). <https://doi.org/10.3390/molecules21111542>
- Lunde, N. N., Bosnjak, T., Solberg, R., & Johansen, H. T. (2019). Mammalian legumain – A lysosomal cysteine protease with extracellular functions? In *Biochimie* (Vol. 166, pp. 77–83). Elsevier B.V. <https://doi.org/10.1016/j.biochi.2019.06.002>

- Lv, H.-S., Huang, S.-Y., Zhao, B.-X., & Miao, J.-Y. (2013). A new rhodamine B-based lysosomal pH fluorescent indicator. *Analytica Chimica Acta*, 788, 177–182. <https://doi.org/10.1016/J.ACA.2013.06.038>
- MacMillan, D. S., Murray, J., Sneddon, H. F., Jamieson, C., & Watson, A. J. B. (2013). Evaluation of alternative solvents in common amide coupling reactions: Replacement of dichloromethane and N,N-dimethylformamide. *Green Chemistry*, 15(3), 596–600. <https://doi.org/10.1039/c2gc36900a>
- Mathur, S., Mincher, D., Turnbull, A., Stevens, C., & Poole, A. (2016). Design and evaluation of novel theranostic fluorogenic dual probe-prodrug in cancer. *European Journal of Cancer*. [https://doi.org/10.1016/S0959-8049\(16\)61501-0](https://doi.org/10.1016/S0959-8049(16)61501-0)
- May, J. P., Brown, L. J., 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 & Biomolecular Chemistry*, 3(14), 2534. <https://doi.org/10.1039/b504759e>
- Merchant, N., Nagaraju, G. P., Rajitha, B., Lammata, S., Jella, K. K., S.Buchwald, Z., S.Lakka, S., & Ali, A. N. (2017). Matrix metalloproteinases: Their functional role in lung cancer. In *Carcinogenesis* (Vol. 38, Issue 8, pp. 766–780). Oxford University Press. <https://doi.org/10.1093/carcin/bgx063>
- Merrifield, R. B. (1963). Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society*, 85(14), 2149–2154. <https://doi.org/10.1021/ja00897a025>
- Monteiro, L. S., Oliveira, S., Paiva-Martins, F., Ferreira, P. M. T., Pereira, D. M., Andrade, P. B., & Valentão, P. (2017). Synthesis and preliminary biological evaluation of new phenolic and catecholic dehydroamino acid derivatives. *Tetrahedron*, 73(43), 6199–6209. <https://doi.org/10.1016/j.tet.2017.09.012>
- 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(7), 1417–1424. <https://doi.org/10.1016/j.febslet.2007.02.064>

- Murthy, R. V., Arbman, G., Gao, J., Roodman, G. D., & Sun, X.-F. (2005). Legumain expression in relation to clinicopathologic and biological variables in colorectal cancer. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 11(6), 2293–2299. <https://doi.org/10.1158/1078-0432.CCR-04-1642>
- Nor, S. M. M., Sukari, M. A. H. M., Azziz, S. S. S. A., Fah, W. C., Alimon, H., & Juhan, S. F. (2013). Synthesis of new cytotoxic aminoanthraquinone derivatives via nucleophilic substitution reactions. *Molecules*, 18(7), 8046–8062. <https://doi.org/10.3390/molecules18078046>
- Nguyen T & Francis M.B. Francis. (2003). Practical Synthetic Route to Functionalized Rhodamine Dyes. <https://doi.org/10.1021/OL035135Z>
- Padmashali, B., Mallikarjuna, S. M., & Sandeep, C. (2017). Acid amine coupling of (1h-indole-6-yl) piperazin-1-yl) methanone with substituted acids using HATU coupling reagent and their antimicrobial and antioxidant activity Acid amine coupling of (1h-indole-6-yl)piperazin-1-yl)methanone with substituted acids using hatu coupling reagent and their antimicrobial and antioxidant activity. *Article in International Journal of Pharmaceutical Sciences and Research*, 8(7), 2879–2885. [https://doi.org/10.13040/IJPSR.0975-8232.8\(7\).2879-85](https://doi.org/10.13040/IJPSR.0975-8232.8(7).2879-85)
- Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., & Finn, R. D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research*, 46. <https://doi.org/10.1093/nar/gkx1134>
- Schechter, I., & Berger, A. (2012). *Biochemical and Biophysical Research communications on the size of the active site in proteases. i.. Papain* (Vol. 27, Issue 2). <https://doi.org/10.1016/j.bbrc.2012.08.015>
- Sekar, R. B., & Periasamy, A. (2003). Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *The Journal of Cell Biology*, 160(5), 629–633. <https://doi.org/10.1083/jcb.200210140>

- Shen, S.-L., Zhang, X.-F., Ge, Y.-Q., Zhu, Y., Lang, X.-Q., & Cao, X.-Q. (2018). A near-infrared lysosomal pH probe based on rhodamine derivative. *Sensors and Actuators B: Chemical*, 256, 261–267. <https://doi.org/10.1016/J.SNB.2017.10.103>
- Spiciarich, D. R., Nolley, R., Maund, S. L., Purcell, S. C., Herschel, J., Iavarone, A. T., Peehl, D. M., & Bertozzi, C. R. (2017). Bioorthogonal Labeling of Human Prostate Cancer Tissue Slice Cultures for Glycoproteomics. *Angewandte Chemie International Edition*, 56(31), 8992–8997. <https://doi.org/10.1002/anie.201701424>
- 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(3), 500–510. <https://doi.org/10.1021/bc800448u>
- Wang, L., Chen, S., Zhang, M., Li, N., Chen, Y., Su, W., Liu, Y., Lu, D., Li, S., Yang, Y., Li, Z., Stupack, D., Qu, P., Hu, H., & Xiang, R. (2012). Legumain: A biomarker for diagnosis and prognosis of human ovarian cancer. *Journal of Cellular Biochemistry*, 113(8), 2679–2686. <https://doi.org/10.1002/jcb.24143>
- Wilhelmsen, C. A., Dixon, A. D. C., Chisholm, J. D., & Clark, D. A. (2018). *Synthesis of N-Substituted 3-Amino-4-halopyridines: A Sequential Boc-Removal/Reductive Amination Mediated by Brønsted and Lewis Acids*. <https://doi.org/10.1021/acs.joc.7b02966>
- Wysocki, L. M., & Lavis, L. D. (2011). Advances in the chemistry of small molecule fluorescent probes. In *Current Opinion in Chemical Biology* (Vol. 15, Issue 6, pp. 752–759). Elsevier Current Trends. <https://doi.org/10.1016/j.cbpa.2011.10.013>
- Yang, M. D., Lin, K. C., Lu, M. C., Jeng, L. Bin, Hsiao, C. L., Yueh, T. C., Fu, C. K., Li, H. T., Yen, S. T., Lin, C. W., Wu, C. W., Pang, S. Y., Bau, D. T., & Tsai, F. J. (2017). Contribution of matrix metalloproteinases-1 genotypes to gastric cancer susceptibility in Taiwan. *BioMedicine (France)*, 7(2), 18–24. <https://doi.org/10.1051/bmdcn/2017070203>

- Ye, Q., Liu, K., Shen, Q., Li, Q., Hao, J., Han, F., & Jiang, R. W. (2019). Reversal of multidrug resistance in cancer by multi-functional flavonoids. In *Frontiers in Oncology* (Vol. 9, Issue JUN, p. 487). Frontiers Media S.A. <https://doi.org/10.3389/fonc.2019.00487>
- Yousafzai, N. A., Wang, H., Wang, Z., Zhu, Y., Zhu, L., Jin, H., & Wang, X. (2018). Exosome mediated multidrug resistance in cancer. *American Journal of Cancer Research*, 8(11), 2210–2226. <http://www.ncbi.nlm.nih.gov/pubmed/30555739>
- Zhang, H. (2017). *Design and synthesis of novel prodrugs to modulate GABA receptors in cancer*. PhD, Edinburgh Napier University.
- Zhang, M., Jiang, Z., Chen, S., Wu, Z., Chen, K., & Wu, Y. (2018). Legumain correlates with neuroblastoma differentiation and can be used in prodrug design. *Chemical Biology and Drug Design*, 91(2), 534–544. <https://doi.org/10.1111/cbdd.13116>
- Zhang, Y., Wu, Y.-Y., Jiang, J.-N., Liu, X.-S., Ji, F.-J., & Fang, X.-D. (2016). MiRNA-3978 regulates peritoneal gastric cancer metastasis by targeting legumain. *Oncotarget*, 7(50), 83223–83230. <https://doi.org/10.18632/oncotarget.12917>