Synthesis of Activity-Based Protein Profiling Probes for Malaria and Hypertension Disease Models & Potential Novel ACE Inhibitors with an Attenuated Zinc Binding Group

Henry Kunda Kambafwile Jr

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

in the Department of Chemistry

Faculty of Science

UNIVERSITY OF CAPE TOWN

Supervisor: Prof. Kelly Chibale

Co-supervisor: Prof. Edward D. Sturrock

January 2012

Acknowledgments

My special thanks to:

My supervisor, Prof. Kelly Chibale and co-supervisor, Prof. Edward D. Sturrock for their wonderful guidance, support, encouragement throughout the course of this work.

Members of the UCT Medicinal Chemistry research group, Centre for Drug Discovery and the IIDMM Zinc Metalloprotease group for their positive contribution to this work.

Noel Hendricks, Pete Roberts, Pierro Benincasa, Dr. Wendy Krogger and Ms. Sylva Schwagger for various analyses.

My friends and colleagues in the pursuit of academic excellence and razzmatazz Chitalu Musonda P*h*D, Kambidima Wotela P*h*D, Sepo Hachigonta P*h*D , Kapambwe Lumbwe P*h*D for "*the energy of positivity and possibility, Docotele wama Docotele".*

The UCT, Department of Chemistry and NRF for the financial and facilities support.

My wife, best friend and partner, Dr. Judith Mwansa-Kambafwile BSc-Hb, MBChB, MPH for the love, complete support and encouragement in continuously challenging apparently impossible frontiers and without whom the strength to reach this point would not have been a possibility for me "it is here! Feel it!

My children and angels Kalaba, Chabu and Songwe "Soka 2010", this is for you.

My parents, Henry Sr. and Agnes Songwe, for their mountain moving support, prayers and seeing me "1982 Luanshya Trust School-Grade R through to 2011 University of Cape Town, PhD" I love you!

My brothers Aggrey, Frank (late), Timothy, Mwape, Michael and sisters Anne-Sandra(late), Ng'andwe and the greater Kambafwile/Luswili family, you have been an inspiration. Aunty Matilda, thank you!

Overall, The Alpha and Omega, the prince of peace, the mighty God *El Shaddai* you are indeed the "Great I am".

The pursuit of a PhD, is a journey that has been described as "one through which a (wo)man must walk alone"……..not on a single day have I been ALONE! – Henry K. Kambafwile

ABSTRACT

Angiotensin-converting enzyme (ACE) and P. *falciparum* A M1 (P*f*A-M1) are both zinc metalloproteases implicated in hypertension and malaria, respectively. Hypertension affects approximately 26 % of the world's population while each year over 300 million cases of malaria occur worldwide resulting in between 1.5 and 2.7 million deaths annually. Hypertension treatment with current ACE inhibitors is marred by unpleasant side effects, such as cough and angioedema. In malaria, the parasites continuously develop resistance to anti-malarial drugs where the disease is endemic. There is therefore a need for continuous research into the application of new techniques as well as the design of new small molecule chemical entities as probes or chemotherapeutic agents.

The work presented in this thesis is two-fold:

(i)The design and synthesis of a new series of hydroxamate, sulfhydryl and carboxylic acid-based inhibitor activity-based protein profiling probes for use in hypertension and malaria disease models. The hydroxamate probes were prepared by sequential Knoevanagel condensation of various aldehydes with Meldrums acid as starting material followed by conjugate reduction, alkylation with allyl bromide, decarboxylation and Ruthenium catalytic cleavage of a terminal alkene. Peptide coupling was then used to assemble the various sub-units, namely, the reactive groups, the probe linker, photocross-linker and the biotin reporter tag.

(ii)Design and synthesis of a new series of α-amino-phosphinic acid-based ACE inhibitors with an attenuated Zinc Binding Group (ZBG) using a one pot, 3-component synthesis involving equimolar quantities of amino acid, aldehyde and phosphonic acid. These were tested for ACE inhibition and the effect of the ZBG on domain-selectivity was investigated.

The sulfhydryl and carboxylate ABP's, hydroxamates as well as the phosphinic acids were evaluated with enzymatically active purified C- and N-domain constructs. Although the carboxylic acid probe **3.32a** showed no inhibition, the sulfhydryl probe **3.32b** exhibited an apparent IC_{50} of 25 µM.

For the hydroxamate inhibitors, compound **4.7d** was found to inhibit ACE by 86% at 5 µM. Amongst the phosphinic acid compounds with the attenuated zinc-binding group, **4.2i-7** inhibited the C-domain construct by 53% whereas **4.2i-2**, **4.2i-14**, and **4.2i-16** inhibited by 40, 39 and 36% respectively at 250nM. The different groups adjacent to the ZBG's had varying degrees of effect on the activity of the compounds with the most active **4.2i-7** having a pseudo-phenylalanine moiety in the α- and a propyl group in the α'-position relative to the phosphinic acid. On the other hand, compared to the 99% inhibition by Captopril in both domains, Enalaprilat analogue **4.4e** inhibits by 30 and 20% in the C- and N domains where Captopril analogue **4.5d** inhibit both domains by 20% respectively at 250 nM. None of the compounds showed significant selectivity for the C- or N-domains.

ABBREVIATIONS

ACE	Angiotensin-converting enzyme
Ang I	Angiotensin I
Ang II	Angiotensin II
AcSDKP	*N*-acetyl-seryl-aspartyl-lysyl-proline
BP	Blood Pressure
BK	Bradykinin
BBr_3	Boron Tribromide
CPA	Carboxypeptidase A
^{13}C NMR	Carbon 13 Nuclear Magnetic Resonance
CbzCl	Benzyl chloroformate
DMF	Dimethylformamide
DCC	Dicyclocarbodiimide
DMSO	Dimethyl sulphoxide
DCM	Dichloromethane
3-D	Three Dimensional
EDC-HCl	*N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide Hydrochloride
EOPB	ethyl-2-oxo-4-phenyl butyrate
FGI	Functional Group Interconversion
GnRH	Gonadotropin-releasing hormone
h	hours
HHL	hippuryl-histidyl-leucine
HOBt	1-Hydroxybenzotriazole
^{1}H NMR	Proton Nuclear Magnetic Resonance
HPLC	High Pressure Liquid Chromatography
IC_{50}	Inhibitor concentration for 50% activity reduction
KKS	Kallikrein-Kinin System
LHRH	Luteinising hormone-releasing hormone
LRMS	low resolution mass spectroscopy
MeOH	methanol

Mol	Mole
$NaHCO_3$	sodium bicarbonate
o-p	o-phthaldialdehyde
Pd/C	Palladium on Carbon
iPr_2Net	diisopropylethylamine
RAS	Renin-angiotensin system
$SOCl_2$	Thionyl chloride
sACE	Somatic ACE
tACE	Testis ACE
Z-FHL	*Z*-L-Phenalalyl-L-Histidyl-L-Leucine

The following abbreviations are used in the Experimental chapter:

s	singlet
d	doublet
dd	doublet of doublets
ddd	doublet of doublets of doublets
t	triplets
td	triplet of doublets
q	quartet
m	multiplet
mp	melting point
J	coupling constant
Hz	hertz
m/z	mass to charge ratio
δ	chemical shift
ppm	parts per million
R_f	retention factor
nM	nanomolar

CONTENTS

Chapter One

Introduction and Literature Review

1.0 Introduction

The first part of this thesis describes the design and synthesis of novel activity-based protein profiling probes based on the selective interaction of active metallo-proteases with synthesized inhibitor molecules. The general introduction in **Chapter 1** gives an overview of the current literature on the biological and patho-physiological role of the gluzincins with Angiotensin-Converting Enzyme (ACE) and P. *falciparum* A – M1 (P*f*A-M1) as examples. Excessive metalloprotease activity is a possible underlying cause for the development of many diseases ranging from inflammatory conditions to pathological tissue remodelling and cancer. The profiling of metalloprotease activity could thus provide a valuable diagnostic tool and shed light on the complicated role of these metalloproteases in health and disease. As metalloprotease activity is highly regulated *in vivo*, determination or profiling of the actual active isoforms that are present could provide a better picture of the (patho)-physiological situation. Several analytical techniques for distinguishing active from non-active, either in pro-protease, or zymogen form or inactivated by endogenous inhibitors from the literature have been outlined. **Chapter 1** will also gives a broad overview of the traditional methods for activity determination based on monitoring of substrate conversion and expands to the recently developed field of activity-based proteomics. This methodology uses small molecular weight inhibitors as

ligands labelled with biotin for tagging and visualisation of active proteases.

C**hapter 2** will discuss the rationale behind the design and synthesis of activity based probes targeting metalloproteases in the malaria and hypertension disease models. The hydroxamate-based inhibitor probe for the former as well as the sulfhydryl and carboxylic acid inhibitor probes for the latter include a photoreactive benzophenone moiety for cross-linking of the probe to the active metalloprotease and a biotin group for future visualisation and pull-down using biotin-streptavidin interaction. This chapter will also discuss the rationale as well as an overview of methodological prospects. **Chapter 3** encompasses a detailed retrosynthetic strategy followed by a synthetic discussion of the preparation of the described ABPP's is presented.

The importance of ACE in pathophysiology as well as its mechanism has been briefly described in **Chapter 1.** This enzyme is particularly interesting because it is a two-domain enzyme, an N- and a C-domain each with its own active site. It has been implicated in the control of hypertension and is an important therapeutic target. Since the 1950's ACE inhibitors have been used for its control. However this treatment has been associated with unpleasant side effects attributed to the lack of selectivity of these inhibitors for the hypertension controlling C-domain. Current generation ACE inhibitors inhibit both domains resulting in the disturbance of other homeostatic functions perfomed by the N-domain resulting in various side effects precipitated by the accumulation of substrates and

metabolites. From the foregoing, it is clear that the next generation of ACE inhibitors should be C-domain specific.

So apart from the synthesis of probes for future study of the behaviour of ACE, one other objective of this research was to design, synthesise and biologically evaluate a new series of potential novel C-domain specific inhibitors. Previous research has focused mainly on variations in the peptidomimetic carbon chain backbone whilst maintaining a strong zinc-binding group(ZBG) presence. This research will consider the effect on domain selectivity of an attenuated ZBG with a varied backbone. Our hypothesis is that the lack of domain selectivity exhibited by inhibitors, both existing and new ones synthesized in research, is because of the presence of a ZBG. The latter suppresses the influence of the more subtle but more specific interactions of the inhibitor backbone which have higher chances of achieving domain selectivity. Accordingly, **Chapter 4** elaborates on the structural ACE featuresas well as the need to achieve domain selective inhibition of this enzyme. It then proceeds to describe the design and synthesis of a small exploratory number of potential novel C-domain selective phosphinic acid based ACE inhibitors with an attenuated zinc-binding group. This part of the research will seek to investigate the effect attenuation would have on ACE-domain selectivity, utilizing structural features of the most C-domain selective inhibitor known, **RXPA380.** Current knowledge suggests that the domain selectivity of **RXPA380** is attributed to the tryptophan residue in the P2'position. This chapter will therefore also explore briefly the C-domain specificity-imparting effect of tryptophan by incorporating the residue into the P_2'-

position of two of the most commonly prescribed ACE-inhibitors, namely, captopril and enalaprilat. Chapter 4 will further describe the synthesis of "serendipitous" compounds synthesized from synthetic intermediates from Chapter 1 coupled to tryptophan.

In **Chapter 5** the biological results of the assays are summarized. In **Chapter 6** a conclusion of the results obtained as well as gives future perspectives are presented. Finally, **Chapter 7** gives detailed experimental procedures used in the synthesis and biological evaluation of the probes and inhibitors reported herein.

1.1 Literature Review

1.1.1 Metalloproteases

Proteases can be classified into four major classes: serine-, cysteine-, aspartic- and metalloproteases based on their residue or co-factor crucial in catalysis. Metalloproteases (MP's)are proteases where the water molecule that is used for hydrolysis is complexed to a metal ion in the catalytic centre of the enzyme. The metal ion is itself held in position by several amino acid residues. Proteases can be categorized into 15 different clans based on the structural diversity of their sub-sites. Ten of these are given in Table 1 below where one of the major families is the zinc metalloprotease superfamily. The latter are the zinc containing metalloproteases most of which exhibit a characteristic Histidine-Glutamic Acid-X-X-Histidine (HEXXH, single letter amino acid code, with X being any residue) consensus sequence. The two histidine (H) residues serve as zinc ligands and the glutamic acid (E) residue polarizes a water molecule involved in

nucleophilic attack at the scissile peptide bond of the substrate.

Clan/Sub-Clan	MEROPS Family Name	General Family name	Typical human Family members
MA (E) "Gluzincins"	M1	Aminopeptidase	Aminopeptidase N, Aminopeptidase A, Insulin-regulated aminopeptidase, Pyroglutamylaminopeptidase II
	M2	Angiotensin-Converting Enzyme (ACE) family	ACE, ACE2
	M3	ThimetOligopeptidase family	Thymetoligopeptidase (EP24.15), Neurolysin (EP24.16), Mitochondrial intermediate peptidase
	M13	Neprilysin family	Neprilysin, NEP2, Endothelin-converting enzyme (ECE-1), ECE-2, Damage-Induced neuronal peptidase (DINE).
	M41	FtsHendopeptidase family	i-AAA protease
MA (M) "Metzincins"	M8	Leishmanolysin family	Leishmanolysin-2
	M10A	Matrix metalloproteases	Collagenases. Gellatinases, stromelysins, matrilysin, membrane-type MMPs
	M12A	Astacin Subfamily	Meprin-α, β, Mammalian tolloid-like 1 protein
	M12B	Adamalysin Subfamily	ADAM proteases, AMADTS proteases
MC	M14A	Carboxypeptidase A subfamily	Carboxypeptidase A & B
	M14B	Carboxypeptidase E subfamily	Carboxypeptidase E & N

ME	M16A "Inverzincins"	Pitrilysin Subfamily	Insulysin, nardilysin, (NRD convertase)
	M16B		Mitochondrial processing peptidase-β subunit
	M16C		Eupitrilysin
MF	M17	Leucylaminopeptidase	Leucylaminopeptidase
MG	M24A	Methionylaminopeptidase subfamily	Methionyl aminopeptidase-1, -2
	M24B	Aminopeptidase P subfamily	Aminopeptidase P1 (solube), Aminopeptiase P2 (membrane bound), χ-Pro-dipeptidase
MH	M18		Aspartylaminopeptidase
	M20A		Ytosolic non-specific dipeptidase
	M28B		Glutamate carboxypeptidase !!
	M28D		Plasma Glutamate carboxypeptidase
MJ	M19	Membrane Dipeptidase family	Membrane Dipeptidase
MM	M50A		S2P protease
M-	M48A	Ste24 endopeptidase family	Farnesylated-protein-converting enzyme 1
	M49	Dipeptidyl-peptidase III family	Dipeptidyl-peptidase III
	M67		Pohl peptidase

Table 1.1: Classification of metallopeptidases. [1]

1.1.2 Zinc Metalloproteases

Zinc metalloproteases are a large and diverse class of enzymes involved and implicated in numerous and important biological phenomena both physiological and pathological. These are cell proliferation, signalisation, multiplication and migration, hormonal processing, angiogenesis, pathogen growth, tissue remodelling, hypertension and cancer.Their classification is based on the catalytic mechanism and the critical amino acid residues involved. In metalloproteases, an activated water molecule complexed to the divalent cation, usually Zn^{2+}, serves as the nucleophile, attacking the carbonyl group of the targeted peptide bond. Ten years ago, Rawlings and Barrett[2] proposed a classification system based not only on the catalytic mechanism, but also the evolutionary links between enzymes. According to this scheme, which has been updated in light of recent genomic and structural information and is now maintained on-line in the MEROPS peptidase database[3] most of the mammalian metalloproteases fall into one of 4 clans or sub-clans (Table 1.1).

By far the largest group, Clan MA, is distinguished by the active site motif HEXXH, where the two histidine residues, together with a third distant residue, bind the zinc atom. This clan is divided into two major sub-clans: MA(E), the "gluzincins," where the third ligand is a glutamate, and MA(M), "metzincins," which have a third histidine or an aspartate residue in this position.. This latter group is typified by the matrix metalloproteases and their relatives, including the "*a d*isintegrin*a*nd *m*etalloprotease" family (ADAMs). The gluzincins are further divided into families

based on homology. The families most relevant to mammalian signalling systems are M1 (aminopeptidases), M2 (angiotensin converting enzyme family), M3 (EP24.15/thimetoligopeptidase and EP24.16/neurolysin), and M13 (neutral endopeptidase family). The metallo-carboxyproteases belong to the clan MC, while a smaller group of peptidases, in which the active site motif is reversed (HXXEH; the so-called "inverzincins") and typified by insulin-degrading enzyme (insulysin), belong to the M16 family within the clan ME.

1.1.3 Angiotensin Converting Enzyme (ACE) as an example of a Zinc metalloprotease implicated in Hypertension

Gluzincins characterized to date include thermolysin-like enzymes in gram-positive and -negative bacteria (M4 family), mycolysin from streptomyces (M5), botulinum neurotoxin (M27), a family of aminopeptidases (M1), neprilysin-like enzymes (M13), and ACE (M2).Angiotensin-converting enzyme (ACE[4], EC 3.4.15.1) is a well-characterized membrane-bound zinc metalloprotease of the "gluzincins" family known to play a crucial role in cardiovascular homeostasis. It is an exopeptidase that cleaves dipeptides from the carboxyl-terminal end of various peptide substrates.[5] Like the similar carboxypeptidases, it is a zinc containing enzyme. Two reactions catalyzed by angiotensin-converting enzyme play a key role in blood pressure regulation. Firstly, it is the conversion of the inactive decapeptide angiotensin I (Ang I) to the potent vasopressor octapeptide angiotensin II[6] (Ang II) and secondly, inactivation of the vasodepressor nonapeptide bradykinin (BK).[7] The central role that ACE plays in the Renin-Angiotensin System (RAS) renders it a major drug target for cardiovascular disease and blood

pressure control. Apart from Ang I and BK, ACE is also involved with other substrates such as peptide *N*-acetyl-seryl-aspartyl-lysyl-proline (AcSDKP), a peptide known for its haemoregulatory roles. Even in this case, ACE functions as a dipeptidyl carboxypeptidase, releasing the KP fragment in the plasma.[8] This indicates that ACE plays a role in several other physiological systems apart from the RAS.

ACE is interesting not only because of its physiological and pathological characteristics, but also because it is the only known metalloprotease with two active-sites, attributed undoubtedly to gene duplication.[9] The two active-sites are subtly different, with each site displaying distinct substrate and inhibitor affinities, a phenomenon successfully exploited in the design of site-specific inhibitors which limit the formation of the vasoconstrictor angiotensin II with minimal effects on BK degradation.[10] In addition to its membrane-bound form, ACE also exists as two distinct soluble forms; a testis-specific, single active site ACE, arising from an alternate splice variant expressed in germ cells and required for male fertility. This isoform was recently used to solve the structure of the enzyme and is comparative to the C-domain of sACE[11] and a secreted form of the somatic enzyme, released from the extracellular membrane by a specific proteolytic event. The possible ADAM proteases involved in this "shedding" event, as well as the physiological consequences of ACE secretion have been described.[12,13]

1.1.4 P. *falciparum* A – M1 as an example of a Zinc Metalloprotease implicated in Malaria

Neutral aminopeptidases are ubiquitous in mammalian cells and are generally involved in protein processing by clipping *N*-terminal amino acids, such as the trimming of peptides to achieve optimal binding to major histo-compatibility complex class I molecules,[14] or in the final stages of protein degradation. Malaria parasites express two such enzymes, a 122 kDa M1- family alanyl aminopeptidase and a 67.8 kDa M17-family leucyl aminopeptidase, termed P*f*A-M1 (MAL13P1.56) and P*f*A-M17 (Pf14_0439), respectively).[15,16, 17,18,19, 20] P*f*A-M1 functions in the blood stage of malaria infection and is responsible for regulating the intracellular amino acid pool within malaria cells, most likely by releasing amino acids from host-derived haemoglobin. In addition to providing essential nutrients for parasite growth and development, digestion of haemoglobin is important for maintaining the osmotic integrity of infected red blood cells.[21] P*f*A-M1 and P*f*A-M17 are chemically validated therapeutic targets for a unique class of antimalarial drugs. Blocking the activity of these aminopeptidases by dipeptide analogs prevents growth of P. *falciparum* in culture and kills P. *Chabaudi* malaria in vivo.[22] P*f*A-M1 is able to efficiently cleave *N*-terminal amino acids from peptide substrates; it exhibits substrate specificity and rapidly hydrolyses substrates containing Leu, Ala, Arg, and Lys but can also cleave substrates containing Phe, Tyr, Ser, and Asn (but not Glu or Asp).[23,24] P*f*A-M1 binds a single, tightly-bound Zn^{2+} metal ion.[25] Its active site is buried deep in the structure, providing a solution to the problem of controlling appropriate substrate entry, as well as preventing hydrolysis of unwanted substrates.[26]

1.2 Mechanism of Proteolysis by Zinc Metalloproteases

The mechanism of action of both ACE and P*f*A–M1 is thought to mimic that of the prototypical zinc metalloprotease thermolysin, whose mechanism of action was elucidated in several studies and summarised by Matthews[27] (Fig. 1.1). Briefly, the reaction proceeds as a fairly typical general-base catalysis. A water molecule associated with the zinc attacks the carbonyl carbon of the scissile bond of the substrate. The glutamic acid (E384) of the metalloprotease HEXXH motif acts as the acceptor of a proton from the water. The negatively charged oxygen resulting from the binding of –OH is stabilized by the zinc, as well as tyrosine (Y523) and water in this intermediate (in thermolysin these correspond to a histidine and tyrosine, respectively). As the reaction proceeds, the E384 releases its proton to the amine nitrogen, thus destabilizing the scissile bond. This bond's nitrogen atom is stabilized by electrostatic interactions with the backbone carbonyls of A354 and S355. The consequent charge repulsion between the new amino terminal and zinc, as well as the stabilization of the C-terminal by the same, leads to separation of the products.

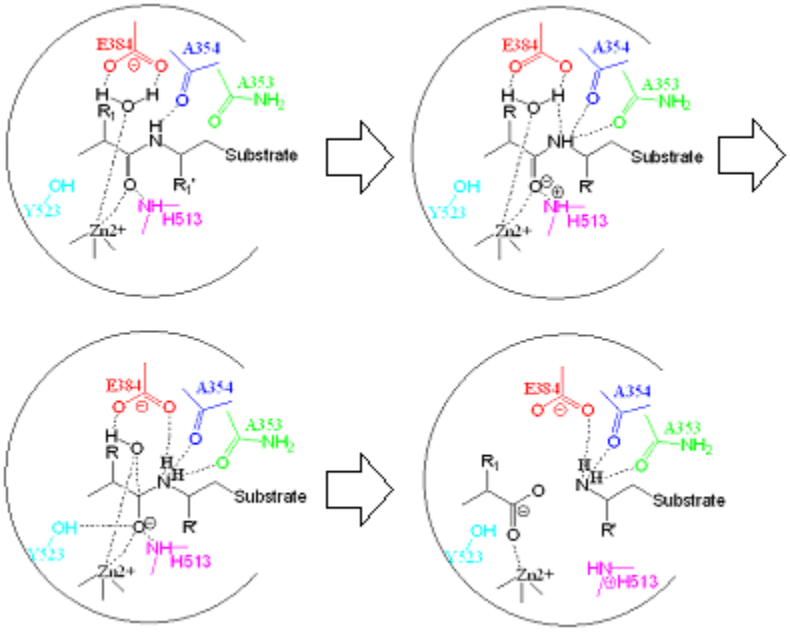


Figure 1.1: Catalytic mechanism of thermolysin.[27]

1.3 The Need to Study Enzymes utilizing new techniques

The majority of hypotheses associating gluzincins to disease states have been focussed on disregulation of the catalytic activity *in vivo*. This poses a challenge to the analysis of these enzymes, since the strict regulation of proteolytic activity means that traditional protein analysis techniques may not provide the required information. Since any given sample can contain three different forms of the metalloproteinase, the inactive zymogen, the inactivated TIMP-proteinase complex and the activated mature enzyme, a mere determination of gene expression by quantitative polymerase chain reaction (PCR) or protein amount by immunochemical techniques or conventional proteomics may be of limited diagnostic value.

Gluzincins are tightly regulated by multiple post-translational mechanisms *in vivo*, hindering their functional analysis by conventional genomic and proteomic methods. This thesis will describe a general strategy for creating activity-based proteomic probes for MPs by coupling a zinc-chelating reactive group to a benzophenone photo-crosslinker *via* a PEG linker. It was envisaged that this will promote selective binding and modification of MP active sites, in the malaria and hypertension disease models respectively.

1.4 Traditional Methods used to determine metalloprotease activity

Several analytical techniques to determine metalloprotease activity have been under development since the late 1980's whilst the earliest methods were described as early as the 1950's. These were based quite simply on conversion of a known (protein) substrate of the metalloprotease and are described below.

1.4.1 Substrate degradation

This is a technique used to monitor substrate conversion. A radiolabeled endogenous protein or substrate of interest is incubated with the protease and degradation is measured by gel electrophoresis, HPLC or Mass spectrometry. This is a simple technique that allows for continuous monitoring of enzyme activity. However, it depends on the availability of an appropriate substrate. It is further limited by the overlapping substrate specificity of metzincin subfamilies rendering quantification of individual proteases impossible.[28]

1.4.2 Gelatine or Collagen Zymography

This technique is a variation of substrate degradation but includes zymography. A substrate such as gelatine is co-polymerized into a polyacrylamide gel (SDS-PAG). The enzymes are then renatured and the incorporated subtrate is degraded. Visualization is then achieved using Coomassie staining yielding bright blue gel with transparent bands indicating gelatinase or collaginase activity. This is able to discriminate the zymogen and the active enzyme form by molecular size resolution through electrophoresis. However, gelatine zymography shows very few gelatinases and thus limited profiling. Furthermore, it does not give a confirmation of proteinase identity, or discriminate between free and tissue inhibitorof metalloproteases (TIMP) inhibited mature enzymes and although quantification is possible with densitometry, it has been described as not very accurate.[29]

1.4.3 Conversion of Biotinylated Gelatine

For improved quantification several reporter molecules have been incorporated into substrates, being either whole protein or small synthetic peptide sequences with internally quenched fluorophores. A biotinylated gelatine yields singly biotinylated fragments after proteolytic degradation that can be distinguished from the multiply biotinylated original substrate containing multiple biotin moieties by capture on immobilized streptavidin and incubated with horse radish peroxidise incubated streptavidin. The resulting signal is inversely proportional to the enzymatic activity. [30] This is a quantitative technique with good linearity and determines

the actual active enzyme in a sample. It is however limited by substrate specificity and gives only total gelatinase activity in the sample.

1.4.4 Fluorogenic Substrate Conversion

This technique utilizes an internally quenched approach which yields a fluorescent product after degradation by the active protease. The substrate is usually a small peptide containing the recognition sequence of the protease (or protease family) of interest. It is also a quantitative technique with good linearity giving a read out of the amount of active enzyme in a sample. However, it is reported to have limited substrate specificity and does not give information on the identity of individual metalloproteases.[31]

1.4.5 Activity-Based Enzyme Linked Immunosorbent Assay (ELISA)

The technique involves the activation of pro-urokinase which is modified for cleavage by a protease of interest. The latter is captured by an immobilized antibody and incubated with the modified pro-urokinase which is then activated. This activation is measured by monitoring conversion of a chromogenic substrate. This is a quantitative technique with good linearity and determines the actual active enzyme in the sample. The antibody-capture step is related to the protease of interest. However, it has been reported to have poor determination of single proteases rendering it limited with respect to profiling possibilities.[32,33]

1.4.6 Enzyme and Tissue Inhibitor of Metalloproteinase ratio (TIMP)

Activity-based analysis of metalloproteases may be based on the hypothesis that protease-related diseases are caused by an imbalance between proteolytic activity and the TIMP-based inhibition system. An excess of activated protease may over time deplete the pool of available TIMP and cause unwanted degradation of tissue since the protease is no longer controlled. This technique, however, may yield results of limited value since it depends on accurate quantification of the amount of active, mature protease that is present in a complex biological sample. A second difficulty is that the enzyme activities play a role within a tissue therefore quantities of enzyme and TIMPs measured in the soluble biological sample may not reflect the situation in the tissue. Often therefore, parallel studies are performed using tissue samples and histological staining techniques. This problem is particularly important in case of ADAMs, since many of the ADAMs are strong cell surface bound molecules that are not easily released into biological fluids. Another problem is the lack of specificity of TIMPs for individual MMPs, which may lead to incorrect conclusions. [34,35]

1.5 Proteomics

Since the analytical methods described above each have their shortcomings when applied to the family-wide profiling of metalloprotease activity, investigators have searched for alternatives. Proteomics is the study of proteins, their structure, localizations, functions, interactions and post-

translational mechanisms. Critical to modern drug discovery is the need to map protein structure-function as this is keyfor better understanding of cellular functions under both normal and diseased states.

The completion of the human genome project has no doubt been a major achievement in scientific research, opening up new and exciting possibilities to fully map and characterize all proteins expressed in cells.[36] However, the study of human proteome presents scientists with a task much more daunting than the human genome project. Functional assignment to these immense numbers of novel genes and gene products has become essential. In fact, the estimated > 100,000 different proteins expressed from 30000-40000 human genes make it extremely challenging, if not impossible with existing protein analysis techniques, to map the entire cellular functions at the translational level. Consequently, there have been rapid advances in the techniques and methods capable of large-scale proteomic studies. Among them, the recently developed high-throughput screening methods have enabled scientists to analyze proteins quickly and efficiently at an organism-wide scale. In the last couple of years various researchers have focused on the development or fine-tuning of the various methods for protein profiling,[37,38,39] a handful of which will be described in the coming sections of this thesis.

1.5.1 Conventional techniques for protein profiling

The conventional two-dimensional gel electrophoresis (2D-GE), in combination with advanced mass spectrometric techniques, has facilitated the rapid characterization of

thousands of proteins in a single polyacrylamide gel. The technique has the ability to separate thousands of proteins in a specific cell or tissue, including their post-translational modified forms. Thus the method is well suited for the global analysis of protein expression in an organism. This, together with the newly developed activity-based profiling approaches that target different classes of proteins, has now allowed the study of protein functions based on their intrinsic enzymatic activities.[40, 41] However, 2D-GE suffers from a number of long-standing problems, including low throughput, a limited dynamic detection range, poor reproducibility, low sensitivity, as well as its difficulties in analyzing hydrophobic, small, and very basic or acidic proteins. Incremental improvements in the 2D-GE technology, including the use of sensitive staining methods and higher-resolving gels, and sample fractionation prior to 2D-GE, have alleviated some of these problems.[42,43] Differential gel electrophoresis (DIGE)[44] and multiplexed proteomics approach[45] are some other recent developments related to 2D-GE. Yates and co-workers introduced a multidimensional protein identification technology (MudPIT).[46] In this technique, multiple types of columns are coupled together to separate proteins using different physiochemical properties in addition to molecular weights and charges, thereby extending the analytical range to proteins having high or low molecular weights, as well as low-abundance and insoluble proteins. MudPIT also allows the flexibility of incorporating a suitable affinity column to enable selective analysis of protein complexes and protein-protein interactions.

Isotope-coded affinity tagging (ICAT) is another chemical method developed for quantitative proteomics. It utilizes stable-isotope labelling to perform quantitative analysis of paired protein samples, followed by separation and identification of proteins within the complex mixtures with liquid chromatography and mass spectrometry.[47] The method relies on the use of an affinity based approach which permits the quantitative comparison of protein abundances between complex proteomes by MS analysis. It employs a chemical probe composed of a reactive group capable of covalently binding to a defined subset of amino acid side chains, an isotopically coded linker and an affinity tag for the isolation of reactive peptides. More recently, the gel-based ICAT has also been reported.[48] However, the strategy is limited to probing known proteins containing cysteine residues and depends on non-specific binding to ICAT reagents.

From the very beginning of the proteomic era, mass spectrometry (MS) has played a major role in the high-throughput identification of proteins following separation techniques such as 2D-GE and ICAT to name a few. With the advent of the 'soft' ionization techniques such as ESI and MALDI coupled with various mass analyzers including the ion trap, time-of-flight (TOF), quadrupole and Fourier transform ion cyclotron (FT-MS), the analysis of large intact protein complexes is now possible[49] thus providing complementary data to that obtained by well-established methods in structural biology such as electron microscopy, X-ray crystallography and NMR. Another MS-based technology for quantitative analysis of protein mixtures is known as the surface-enhanced laser desorption ionization-time of flight

(SELDI-TOF).[50] This technique utilizes stainless steel or aluminum-based supports, or chips, engineered with chemical or biological bait surfaces that allow differential capture of proteins. After the removal of non-specifically adhered proteins, the bound proteins are laser-desorbed and ionized for MS analysis. This field is currently being developed as a prominent technique when combined with advances in protein chip technology.

1.6 Activity Based Protein Profiling

As far as most enzymes are concerned, the overall protein expression levels do not strictly compare to their activity profiles. Different enzymes have intrinsically different catalytic activities. Moreover, various post translational modifications such as phosphorylation, glycosylation, acetylation, action of endogeneous activators/inhibitors, factors like pH and other native conditions significantly influence catalytic activity of enzymes. Most of the aforementioned methods, like 2DE-MS, still focus on measuring the changes in protein abundance and hence provide only an indirect estimate of dynamics in protein function. Undeniably, several important forms of post-translational regulation, as well as protein-protein and protein-small-molecule interactions,[51] may escape detection by these proteomic methods. In order to accomplish the functional analysis of proteins, a number of proteomic methods have been introduced in the past few years aiming to characterize the activity of proteins on a global scale. Large-scale yeast two-hybrid screens[52], [53] and epitope-tagging immuno-precipitation[54], [55] are some of the methods which are widely used to study protein-protein interactions. Although these

methods do have the benefit of assigning specific molecular functions to individual protein products, they normally depend on recombinant expression of proteins in non-natural environments and, therefore, do not directly evaluate the functional state of these bio-molecules in their native settings.

Recently developed activity-based protein profiling (ABPP) has opened up new ways of answering some of the challenges associated with proteomic research. The activity/affinity-based methods make use of small molecule/peptide based chemical probes to profile the functional state of enzyme families directly within a complex proteome. These probes are developed using tools of synthetic organic chemistry and have become a major attraction in the field of high-throughput functional proteomics. Originally developed by Cravatt*et al*., ABPP allows the proteases present in a crude proteome to be studied on the basis of their enzymatic activities rather than their relative abundance.[56, 57, 58, 59, 60, 61] The major advantage of the protein activity-based chemical approaches for profiling proteins includes the ability to target low abundance and membrane associated proteins within samples of high complexity. The strategy is able to bridge the gap between technologies such as the protein microarray and 2D-GE based techniques which study endogenous proteins by their expression, and combine the high throughput feature of 2D-GE with the ability of function-based protein studies. The general strategy in activity-based profiling typically involves a small molecule-based, active site-directed probe which targets a specific class of enzymes based on their enzymatic activity. Thus, the activity-based approach can potentially filter out

proteins that are not of interest and focus on certain classes of proteins. For example, **Fig. 1.2** below shows a general approach to activity-based enzyme profiling using a fluorescent activity-based probe. Among the different enzymes present in a complex proteome, the probe selectively labels a particular class of active enzymes. The labelled enzymes can be separated by SDS-PAGE and visualized by fluorescent imaging, which could be further characterized by mass spectrometry.

Figure 1.2: Schematic representation of labelling of active proteases in a complex proteome by activity based probes. The probe selectively recognizes the active site of the active enzyme. Incorporation of a reporter molecule such as a fluorescent dye or biotin, the labelled protease can be visualized after gel electrophoresis,or enriched on affinity beads for further analysis.[57]

1.7 Activity-based probes

Activity-based probes are designed based on their potential application in proteomics. The design of these probes depends largely on the nature/complexity of the system under consideration as well as the type of information intended to be gathered. Highly specific probes offer invaluable insights into the topology of the enzyme active site and the catalytic mechanism of a particular enzyme or sub-class of enzymes. Hence they are very useful in the design of specific inhibitors. On the contrary, broad-spectrum probes

can label most proteins having the same kind(s) of activity and hence find more application in high-throughput activity based profiling of different classes of enzymes on a global scale. They typically are activity-, mechanism or affinity-based chemical probes which get attached to particular classes of enzymes, thus allowing for the large-scale protease identification, characterization, as well as fingerprinting experiments[62,63,64] Although affinity based probes are not strictly activity-based, their function in enzyme profiling is however intimately associated with the catalytic activity. The design template for activity-based probes generally comprises of a reactive unit, a linker unit, and a reporter unit. (Figure 1.3 below)

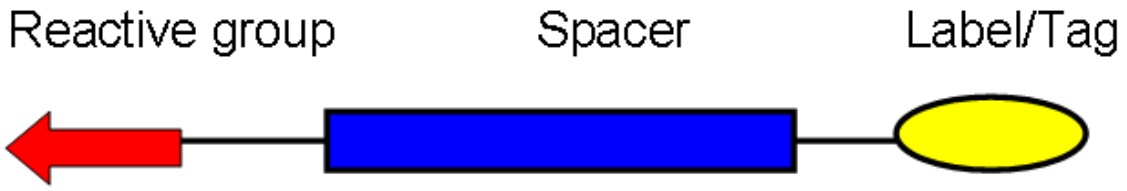


Figure 1.3: General Structure of an Activity-based protein profiling probe

ABPs are generally classified into two types: *directed ABP* and *non-directed ABP*. Fordirected ABP, the reactive group usually derives from a well-characterizedirreversible enzyme inhibitor. This mechanism-based inhibitor holds an electrophilewhich can selectively and covalently react with the active-site nucleophilic residue of a family of enzymes. Based on the inhibitor, it is conceptually straight forward todesign a directed ABP by incorporating the inhibitor moiety with a linker and areporter tag. This type of ABPs

includes fluorophosphonate probes derived fromserine protease inhibitor diisopropyl fluorophosphonate (DFP) for serine hydrolases,and epoxide probes derived from papain-like cysteine protease (PLCP) inhibitor E-64for PLCPs[65]. All directed ABP's haveexcellent target selectivity within a family of enzymes sharing cognate mechanismand function, and minimal interfamily cross-reactivity. For example, epoxide probeDCG-04 derived from E-64 only targets papain subfamily but not other cysteine protease subfamilies, or serine proteases or metalloproteases.[66] Due to a limited number of dedicated irreversible inhibitors, directed ABPs only covera small fraction of the active proteomes. Many enzymes including proteases, esterases and phosphatases catalyze hydrolytic reactions by acting on the substrates through a nucleophilic attack mechanism. Therefore, theoretically an ABP carrying amoderately reactive electrophile can covalently react with the nucleophile of the enzymes. Carbon electrophiles, including those present in reactive natural products, become the resource of choice as reactive groups of non-directed ABPs. Examples are sulfonate ester probes, and the spiro epoxide probes derived from the bioactive natural compounds fumagillin and luminacin D.[67]

1.7.1 Reactive unit

The reactive unit can be an activity or mechanism-based inhibitor of one particular class of enzymes. By reacting with the target enzymes in an activity-dependent manner, the reactive unit serves as a "guiding warhead" attaching the whole probe to the enzyme rendering the resulting probe-enzyme adducts easily distinguishable from other unmodified proteins. The chemical groups on reactive units may be fine

tuned to target different classes of enzymes based on their intrinsic activity even in the complex cellular environment. The reactive units used in ABPP are mostly derived from enzyme inhibitors. For cysteine proteases, a number of probes containing reactive units such as vinyl sulfones[68,69,70] epoxides [71, 72,73], α–halo or (acyloxy) methyl ketone substituents[74, 75] have been reported. Probes with sulfonate esters as reactive units can target different classes of enzymes such as thiolases, aldehyde dehydrogenases, epoxide hydrolases and so on. Probes conjugated to *p*–hydroxymandelic acid specifically label protein phosphatases[76,77] and hydroxamate based probes target metalloproteases.[78] Fluorophosphonate/fluorophosphate derivatives have been developed for selectively profiling serine hydrolases.[79,80] Typical reactive groups as described above are given in fig. 1.4 below.

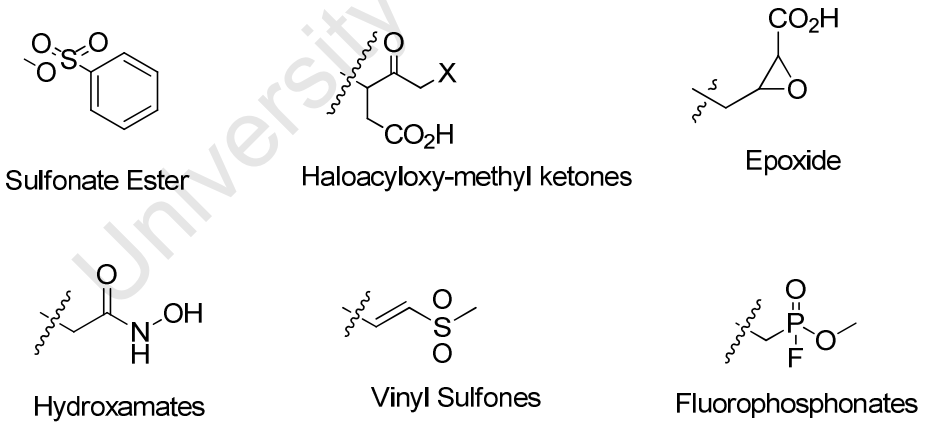


Figure 1.4: Examples of reactive groups used in Activity based protein profiling

1.7.2 The linker unit

The linker unit (Fig. 1.5 below) forms a bridge between the reactive group and the reporter tag unit. The main purpose of having a linker is to minimize the binding interference

of the tag to the enzyme. In other words, the linker unit functions as a spacer of appropriate length which allows the reactive unit to access the active site freely without considerable steric hindrance from the big reporter unit. Different types of linkers can be used based on the polarity of the active site of the target enzyme. Enzymes having hydrophobic residues in their active site are easily accessed by probes with a long alkyl chain as linker. Similarly a PEG linker may be used when polar residues are present in the active site. Thus the linker can actually enhance reactive unit–enzyme binding. The linker can also convey specificity elements to target the probe for a particular family of enzymes. A linker may function as a recognition unit when it is a peptide fragment which could define differential specificity of the probe towards different enzymes of the same class. Since an alkyl chain does not impart this kind of specificity, it is often found to be useful in the design of activity based probes for general profiling experiments such as those targeting all the enzymes in the same class. The non-specific linkers are more commonly used in large-scale protein profiling experiments whereby many different enzymes in the same class are analyzed simultaneously. Photo-cleavable linkers have been reported which facilitate quantitative proteome analysis.[81] Upon irradiation, the linker gets cleaved, releasing the labelled proteins, which could be easily isolated and purified by standard solid-phase methods.

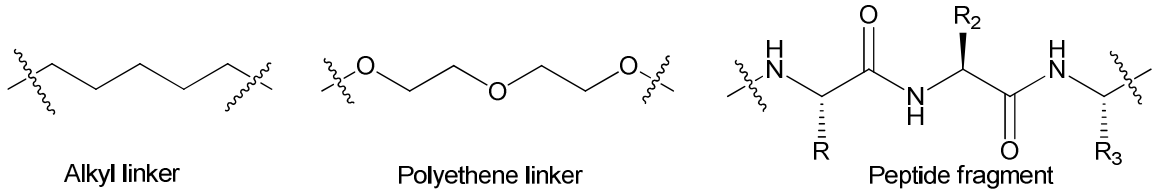


Figure 1.5: Typical linkers utilized in probes

1.7.3 The reporter unit

The reporter unit in the probe is either a fluorescent tag for sensitive and quantitative detection of labelled enzymes or an affinity tag which facilitates further protein enrichment, purification and identification. The tags must be compatible with standard SDS-PAGE techniques which allow for the simplest and inexpensive methods for protein separation and analysis. Biotin and its derivatives are the most commonly used affinity tags. Biotin facilitates detection by simple Western blot approaches using a reporter avidin molecule in place of a standard antibody. After biotin labelling, the protein can also be isolated from the crude cell lysate, enriched and purified by streptavidin agarose beads and subsequently identified by Mass Spectrometry. This strategy is well suited for the isolation of active enzymes that are present only in low abundance. When compared to affinity tags, fluorescent tags and radioactive tags offer high sensitivity for the detection experiments and are also much faster. They can be visualized by direct scanning of gels under a fluorescent scanner or phosphor-imager. The commonly used fluorescent tags are cyanine dyes such as Cy3 and Cy5, fluorescein and rhodamine examples of which are given in figure 1.5 below. These tags have a distinct colour which makes them suitable for potential multi-colour labelling experiments where a mixture of probes with different fluorescent units can be used to target different classes of enzymes in the same proteome, or the same class of enzymes from different proteome samples.

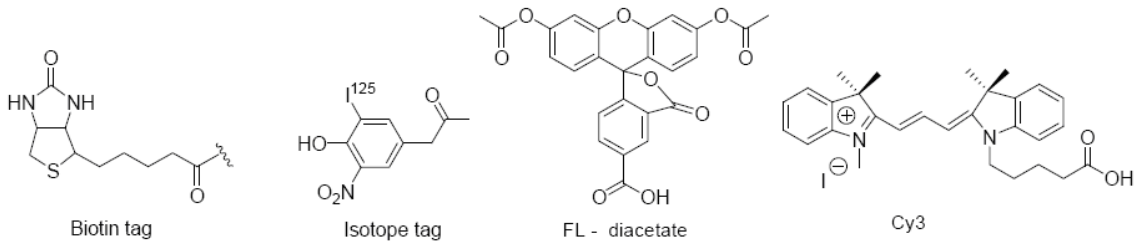


Figure 1.6: Typical reporter tags used in ABPP

1.7.3.1 Trifunctional Probe

ABPs may have a fluorescent or affinity tag. Another reported option is the combination of both fluorophore and biotin into one reporter tag. The resulting ABP is called a trifunctional ABP: a reactive group for activity-based labeling, and a fluorophore and a biotin for simultaneous in-gel detection and affinity purification.[82] Fluorescence scanning is a lot more sensitive than protein staining to locate and select enriched targets for MS analysis, and excludes co-isolated endogenously biotinylated proteins as background. However, trifunctional ABPs can have reduced probe reactivity due to a bulkier tag causing steric hindrance, and theseprobes are more challenging to synthesize.

1.7.3.2 Click-chemistry probe

The function of an ABP principally relies on the inhibitor moiety. However, thereporter tag inevitably modifies the structure and physicochemical features of theinhibitor, leading to changes in e.g. size and hydrophobicity. This may cause reducedcellular uptake or biased sub-cellular distribution of the ABP, particularly hamperingits application for *in vivo* profiling. The click-chemistry probe

is an untagged ABP,with a small chemical adapter for the attachment of a reporter tag after labeling. This “tag-free” ABP maximally resembles the prototype inhibitor and minimally affects itstarget binding, especially *in vivo*. Alkyne and azide are two such adapters. Following covalent labeling of protein targets, the probe is ligated to the reporter tag *in vitro*, through Cu(I)-catalyzed cycloaddition reaction (“click chemistry”).[83] In aqueous solution of biological samples, either adapter molecule is inertwhile their coupling is quick, efficient and specific under mild conditions. This two-step labeling strategy also allows one probe to accept diverse reporter tags, andsimplifies probe design and synthesis. Click-chemistry ABPs facilitate profiling protein activities in living cells and even in whole animals, and have a deep impact on ABPP techniques.[84]

1.8 Applications of ABPP

Directed activity-based probes have been developed for the profiling of various enzyme classesincluding cysteine proteases, serine hydrolases, catalytic β-subunits of proteasome, metalloproteases, histone deacetylases, kinases and nucleotide-binding proteins, phosphotases, ubiquitin-specific proteases, glycosidases and cytochrome P450s, and non-directed ABPs for several other enzymes classes, many of which are implicatedin important biological processes from normal metabolism to patho-physiology.[85] The rapidly-expanding ABP toolbox has been applied to both basic research on function and regulation of individual proteins, and medical research from pharmacological studies to clinical practices.

1.8.1 Applications in mammalian studies

1.8.1.2 Activity-based proteomics and enzyme mechanism studies

Genome-sequencing projects successfully deciphered genetic codes on an overwhelming scale, but unfortunately, it gives no information with regards to structure, function, regulation and interaction of proteins. Proteomics and ABPP promises to be the answer to the latter challenges. ABPs can not only be used to detect and measure activities at individual protein or on a proteomic scale. ABPP gives an activity readout of enzymes only in active forms for functional analysis. [86] ABPs are undoubtedly a powerful tool to study enzyme mechanisms,including the catalytic mode, substrate selectivity and molecular function.

1.8.1.3 Comparative ABPP and biomarkers/drug targets discovery

By comparing protein activity profiles of two or more proteomes under normal and diseased conditions or at different pathological stages, comparative ABPP identifies proteins whose activity level differs. These differential activities may be caused bydifferent protein accumulation levels, or by a higher order of regulation, e.g. presence of inhibitors. Differential protein activities can serve as biomarkers for diagnosis and/or drug targets for treatment. Cravatt *et al* reported the profiling of 33 human specimen proteomes from 28 breast tumours and 5 healthy breast tissues using APBs and identified 7 membrane-associated serine hydrolases with differential activities common to

breast cancer samples.[87] Notably, a cDNA microarray on the same samples provided a direct comparison of protein transcript and activity levels, and showed a negative correlation for thrombin and KIAA1363. This reveals that KIAA1363 is a potential breast cancer biomarker and drug target that would not have been recognized by molecular profiling methods other than ABPP. Another example is when profiling of the life cycle of *P. falciprum*, the causative agent of human malaria, with DCG-04 identified that a pathogen PLCP falcipain 1, but not falcipain 2 or 3, is up-regulated in invasive merozoites. [88] This implies that falcipan-1 plays a role in host cell invasion and that it is an anti-malaria drug target.

1.8.1.4 Competitive ABPP and inhibitor screening

A competitive version of ABPP allows screening of small molecule libraries and evaluation of inhibitors, since most inhibitors compete with the ABP for the same active site of a target protein. In a typical experiment, the model proteome is pre-treated with candidate inhibitors, and then labeled with ABP for the remaining active, not inhibited target proteins. Inhibitory efficacy is indirectly interpreted by the decrease in ABP labelling compared to inhibitor-untreated control, and is usually quantified and clustered into a heat-map for assessment. Screening of reversible inhibitors requires kinetically controlled conditions, because the end-point readout of ABPP is determined by both the affinity of the inhibitor and the rate of the probe reactivity. For this reason, labelling time must be optimized and restricted to the period before labelling with ABP reaches saturation.[89]

Competitive ABPP facilitates the selection of potent and specific inhibitors of known drug targets to generate novel drugs, and for proteins with unknown function to carry out chemical knockout assays. It offers several advantages over conventional substrate-based inhibitor screening. First, it takes place directly in physiologically relevant complex proteomes with no requisite for purified recombinant target protein. Second, it is "substrate-free", which makes the inhibitor selection for the target protein with unknown substrates possible. Third, the candidate inhibitors are examined against not only the target protein, but also many other family members with similar enzymatic mechanism, therefore, the selectivity of the inhibitors is evaluated simultaneously.[90]

1.9 Research objective

One of the major advantages of the ABPP method is that it allows the selective detection of targeted proteins from a group of unknown proteins spotted as arrays when incubated with potential ligands. Thus the whole process of separating and identifying proteins in a complex mixture can be simplified, by focusing on specific targeted proteins. Several types of probes have been reported that target different classes of proteins on the basis of their enzymatic activities.[91,92,93, 94]

The objective of this part of the project is to introduce ABPs in hypertension research by designing probes that can successfully bind and robustly label ACE as well as synthesise new probes for future use in malaria research. As an example, a summary of the rationale for the target

structure of the probes in the malaria model is given in figure 1.7 below.

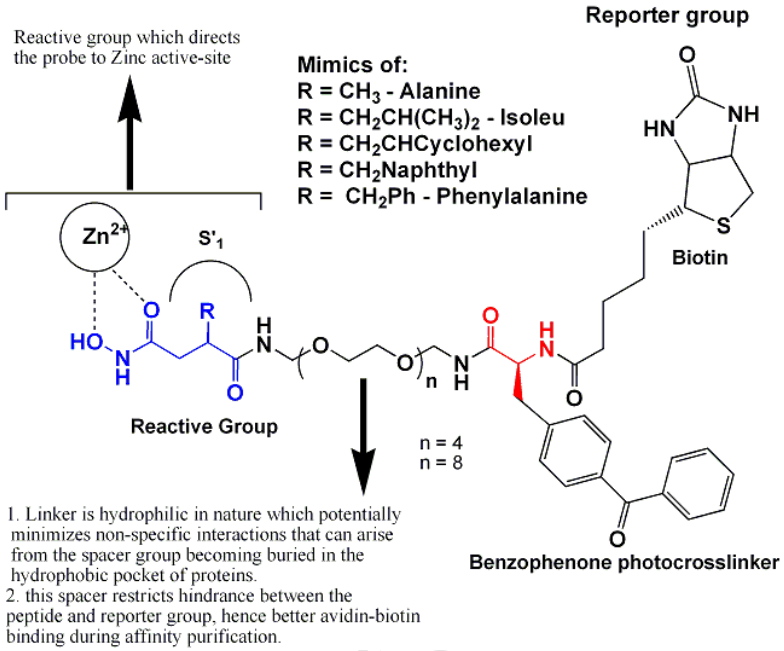


Figure 1.7: General structure of the target probes.

References

[1]Lew, R. A. *Protein and Peptide Letters* **2004**; 11(5):407-414.

[2]Rawlings, N.D.;Barrett, A.J. *Biochemical Journal* **1993**; 290:205.

[3]http://merops.sanger.ac.uk/

[4] Wong, J.; Patel, R.A.;Kowey, P.R. *Progress in Cardiovascular Diseases* **2004**;47: 116-130.

[5]Piquilloud, Y.;Reinharz, A.; Roth, M. *Biochimica et Biophysica Acta* **1970**; 206:136-142.

[6]Skeggs, L. T.; Jr., Lentz, K. E.; Kahn, J. R.; Shumway, N. P.; Woods, K. R., *Journal of Experimental Medicine* **1956**; 104:193-197.

[7] Yang, H.Y.;Erdos, E.G.; Levin, Y. *Journal of Pharmacology and Experimental Therapeutics* **1971**;177 (1): 291-300.

[8]Rieger, K. J.;Saez-Servent, N.;Papet, M. P.;Wdzieczak-Bakala, J.;Morgat, J. L.; Thierry, J.;Voelter, W.; Lenfant, M. *Biochemical Journal* **1993**; 296 (Pt 2): 373-378.

[9]Ehlers, M. R. W.; Fox, E. A.;Strydom, D. J.;Riordan, J. F. *Proceedings of the National Academy of Sciences U.S.A.* **1989**; 86: 7741-7745.

[10]Georgiadis, D.; Beau, F.;Czarny, B.; Cotton, J.; Yiotakis, A.; Dive. V. *Circulation Research* **2003**;93:148.

[11]Natesh, R.;Schwager, S.L.U.;Sturrock, E.D.; Acharaya, K.R. *Nature* **2003**;421:551-554.

[12]Tipnis, S.R.; Hooper, N.M.; Hyde, R., Karran, E.; Christie, G.; Turner, A.J. *Journal of Biological Chemistry* **2000**; 43:332-38.

[13]Donoghue, M.; Hsieh, F.;Baronas, E.; Godbout, K.; Gosselin, M.; Stagliano, N.; Donovan, M.; Woolf, B.;

Robison, K.;Jeyaseelan, R.;Breitbart, R.E.;Acton, S. *Circulation Research* **2000**; 87:e1.

[14]Serwold, T.; Gonzalez, F.; Kim, J.; Jacob, R.;Shastri, N. *Nature*, **2002**; 419(6906):480-483.

[15] www. plasmodb.org

[16]McGowan, S.; Porter, C.J.; Lowther, J.; Stack, C.M.; Golding, S.J.; Skinner-Adams, T.S.; Trenholme, K.R.; Teuscher, F.; Donnelly, S.M.; Grembecka, J. *Proceedings of the National Academy of Sciences USA* **2009**; **106** : 2537-2542.

[17]Skinner-Adams, T.S.; Lowther, J.; Teuscher, F.; Stack, C.M.; Grembecka, J.; Mucha, A.; Kafarski, P.; Trenholme, K.R.; Dalton, J.P.; Gardiner, D. L.*Journal of Medicinal Chemistry* **2007**; 50: 6024-6031.

[18]Skinner-Adams, T.S.; Stack, C. M.; Trenholme, K.R.; Brown, C.L.; Grembecka, J.; Lowther, J.; Mucha, A.; Drag, M.; Kafarski, P.; McGowan, S.; Whisstock, J.C.; Gardiner, D. L.; Dalton, J. P. *Trends in Biochemical Sciences* **2010**;, 35(1):53-61.

[19]Stack CM, Lowther J.; Cunningham E.; Donnelly S; Gardiner D L; Trenholme K R; Skinner-Adams T S; Teuscher F.;Grembecka J; Mucha A; Kafarski P; Lua L; Bell A;Dalton J P. *Journal of Biological Chemistry* **2007** 282(3):2069-2080.

[20]Dalal, S.;Klemba,M.;*Journal of Biological Chemistry* **2007**; 282(49):35978-35987.

[21]Liu, J.;Istvan, E.S.;Gluzman, I.Y.; Gross, J.; Goldberg, D.E.;*Proceedings of the National Academy of Sciences USA* **2006**; 103(23):8840-8845.

[22]Flipo, M.; Beghyn, T.; Leroux, V.; Florent, I.; Deprez, B P.; Deprez-Poulain, R F. *Journal of Medicinal Chemistry* **2007**; 50(6):1322-1334.

[23]Allary, M.; Schrevel, J.; Florent, I. *Parasitology* **2002**;

125(Pt 1):1–10.

[24]Florent, I.; Derhy, Z.; Allary, M.; Monsigny, M.; Mayer, R.; Schrevel J.; *Mol Biochemical Parasitology* **1998**; 97(1–2):149–160.

[25]Maric, S.; Donnelly, S. M.; Robinson, M. W.; Skinner–Adams T; Trenholme, K. R.; Gardiner, D. L.; Dalton, J. P.; Stack, C. M.; Lowther J., *Biochemistry* **2009**; 48(23):5435–5439.

[26]McGowan, S; Porter, C J.; Lowther, J; Stack, C M.; Golding, S J.; Skinner–Adams, T S.; Trenholme, K R.; Teuscher, F; Donnelly, SM.; Grembecka, J; Mucha, A; Kafarski, P; DeGori, R; Buckle, A M.; Gardiner, D L.; Whisstock, J C.; Dalton, JP.*Proceedings of the National Academy of Sciences USA* **2009**; 106(8), 2537–2542.

[27] Matthews, B.W. *Accounts of Chemical Research* **1988**; 21: 333–340.

[28]Tint,H. *Arch. Biochemical and Biophysical Research Communications* **1961**;92: 154–158.

[29]Hume,M.E.; Siegel,M.S.; Polakoski,K.L. *Journal of Andrology* **1987***;* 8: 221–224.

[30]Ratnikov,B.; Deryugina, E.; Leng, J.; Marchenko, G.; Dembrow, D.; Strongin A. *Analytical Biochemistry* **2000**; 286: 149–155.

[31]Stack,M.S.;Gray,R.D. *Journal of Biological Chemistry* **1989***;*264, 4277–4281.

[32]Verheijen,J.H.; Nieuwenbroek, N. M.; Beekman, B.; Hanemaaijer, R.; Verspaget, H. W.; Ronday, H.K.; Bakker, A. H. *Biochemical Journal* **1997**;323 (Pt 3):603–609.

[33]Hanemaaijer,R.;Visser,H.;Konttinen,Y.T.;Koolwijk,P.; Verheijen,J.H. *Matrix Biology* **1998**;17: 657–665.

[34]Tamarina,N.A.;McMillan,W.D.;Shively,V.P.;Pearce,W.H. *Surgery* **1997***;*122:264–271.

[35]Jung,K.; Lein, M.; Ulbrich, N.; Rudolph, B.; Henke, W.;

Schnorr, D.; Loening, S. A. *Prostate* **1998;**34: 130–136.

[36]Ezzell, E. *Scientific American.* **2000;***283*: 64.

[37]Yeo, D.S.Y.;Panicker, R.C.; Tan, L.P.; Yao, S.Q. *Combinatorial Chemistry and High Throughput Screening* **2004;***7:* 213.

[38]Uttamchandani, M.; Huang, X.; Chen, G.Y.J.; Tan, L.P.; Yao, Y.Q. *Mol. Biotechnology* **2004;**28: 227.

[39]Sun, H.;Panicker, R.C.; Yao, S.Q. *Biopolymers* **2007;**88: 141.

[40]Hu, Y.;Huan, X.; Chen, G.Y.J.; Yao, S.Q. *Molecular Biotechnology* **2004;**28: 63.

[41]Huang, X.; Tan, E.L.P.; Chen, G.Y.J.; Yao, S.Q. *Application of Genomics and Proteomics* **2003;**2:225.

[42]30. Rabilloud, T. *Proteomics* **2002;**2: 3.

[43]Gauss, C.;Kalkum, M.; Lowe, M.;Lehrach, M.H.;Klose, J. *Electrophoresis* **1999;**20: 575.

[44]Unlu, M.; Morgan, M.E.; Minden, J.S. *Electrophoresis* **1997;**18:2071.

[45]Steinberg, T.H.; Pretty, K.; Berggren, K.N.; Kemper, C.; Jones, L.;Diwu, Z.;Haugland, R.P.;Pattonet, W.F. *Proteomics* **2001;**1: 841.

[46]Washburn, M.P., Wolters, D., Yates, J.R. *Nature Biotechnology* **2001**, *19*, 242.

[47]Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., Aebersold, R. *Nature Biotechnology* **1999**, *17*, 994.

[48]Smolka, M., Zhou, H., Aebersold, R. *Molecular & Cellular Proteomics* **2002**, *1*, 19.

[49]Daniel, J.M., Friess, S.D., Rajagopalan, S., Wendt, S., Zenobi, R. *International Journal of Mass Spectrometry* **2002**, *216*, 1.

[50]Issaq, H.J., Conrads, T.P., Prieto, T.A., Tirumalai, R., Veenstra, T.D. *Analytical Chemistry* **2003**, *75*, 148A.

[51]Kobe, B., Kemp, B.E., *Nature* **1999**, *402*, 373.

[52]Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R.,Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P. *Nature* **2000**, *403*, 623.

[53]Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., Sakaki, Y. *Proceedings of the National Academy of Sciences USA.* **2001**, *98*, 4569.

[54]Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz,J., Rick, J.M., Michon, A.M., Cruciat, C.M. *Nature* **2002**, *415,* 141.

[55]Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A,Taylor, P., Bennett, K., Boutilier, K. *Nature* **2002**, *415*, 180.

[56]Speers, A. E., Cravatt, B. F. *ChemBiochem* **2004**, *5*, 41.

[57]Adam, G. C., Sorensen, E. J., Cravatt B. F. *Molecular and Cellular Proteomics* **2002**, *1*, 781.

[58]Adam, G. C., Sorensen, E. J., Cravatt, B. F. *Nature Biotechnology* **2002**, *20*, 805.

[59]Liu, Y., Patricelli, M. P., Cravatt, B. F. *Proceedings of the National Academy of Sciences U.S.A.* **1999**, *96*,14694.

[60]Kidd, D., Liu,Y., Cravatt, B. F. *Biochemistry* **2001**, *40*, 4005.

[61]Cravatt, B.F., Sorensen, E.J., *Current Opinion in Chemical Biology* **2000**, *4*, 663.

[62]Chan, E.W.S., Chattopadhaya, S., Panicker, R.C., Huang, X., Yao, S.Q. *Journal of the American Chemical Society* **2004**, *126*, 14435.

[63]Sun, H., Panicker, R.C., Yao, S.Q. *Biopolymers.* **2007**, *88*, 141.

[64]Liau, M.L., Panicker, R.C., Yao, S.Q. *Tetrahedron Letters* **2003**, *44*, 1043.

[65]Liu, Y., Patricelli, M.P. and Cravatt, B.F.*Proc. National*

Academy of Science.USA, **1999** 96(26), 14694-14699.

[66]Greenbaum, D.C., Arnold, W. D., Lu, F., Hayrapetian, L., Baruch, A., Krumrine, J., Toba, S., Chehade, K., Brömme, D., Kuntz, I. D. and Bogyo, M. (2002a) *Chemical Biology* **2002a** 9(10), 1085-1094.

[67]Evans, M.J., Saghatelian, A., Sorensen, E.J. and Cravatt, B.F. *Nat. Biotech.,* **2005** 23(10), 1303-1307.

[68]Uttamchandani, M., Liu, K., Panicker, R.C., Yao, S.Q. *Chemical Communications* **2007**,1518.

[69]Wang, G., Uttamchandani, M., Chen, G.Y.J., Yao, S.Q. *Organic Letters* **2003**, *5*, 737.

[70]Wang, G., Yao, S.Q. *Organic Letters***2003**, *5*, 4437.

[71]Greenbaum, D., Medzihradszky, K. F., Burlingame, A., Bogyo, M. *Chemical Biology* **2000**, 7, 569.

[72]Hashida, S., Towatari, T., Kominami, E., Katunuma, N. *Journal of Biochemistry* **1980**, *88*, 1805.

[73]Bossard, M. J., Tomaszek, T. A., Thompson, S. K., Amegadzie, B. Y., Hanning, C. R., Jones, C., Kurdyla, J. T., McNulty, D. E., Drake, F. H., Gowen, M., Levy, M. A. *Journal of Biological Chemistry* **1996**, *271*,12517.

[74]Schoellmann, G., Shaw, E. *Biochemistry* **1963**, *2*, 252.

[75]Shaw, E., Mares-Guia, M., Cohen, W. *Biochemistry* **1965**, *4*, 2219.

[76]Lo, L.C., Pang, T.L., Kuo, C.H., Chiang, Y.L., Wang, H.Y., Lin, J.J. *Journal of Proteome Research* **2002**, *1*, 35.

[77]Srinivasan, R., Uttamchandani, M., Yao, S.Q. *Organic Letters* **2006**, *8*, 713.

[78]Chan, E.W.S., Chattopadhaya, S., Panicker, R.C., Huang, X., Yao, S.Q. *Journal of the American Chemical Society* **2004**, *126*, 14435.

[79]Liu, Y., Patricelli, M. P., Cravatt, B. F. *Proceedings of the National Academy of Sciences, U.S.A.* **1999**, *96*,14694.

[80]Kidd, D., Liu,Y., Cravatt, B. F. *Biochemistry* **2001**, *40*, 4005.

[81]Zhou, H., Ranish, J.A., Watts, J.D., Aebersold, R. *Nature Biotechnology* **2002**, 19, 512.

[82]Adam, G.C., Sorensen, E.J. and Cravatt, B.F. *Molecular and Cellular Proteomics,* **2002b** 1(10), 828–835.

[83]Speers, A.E., Adam, G.C.; Cravatt, B.F. *Journal of the American.Chemical Society* **2003** 125(16),4686–4687.

[84]Speers, A.E. and Cravatt, B.F. *ChemBioChem,* **2004** 5(1), 41–47.

[85]Cravatt, B.F., Wright, A.T., Kozarich, J.W. *Annu. Rev. Biochem.,***2008** 77(1), 383–414.

[86]Phillips, C.I. and Bogyo, M. *Cell. Microbiol.,***2005** 7(8), 1061–1076.

[87]Jessani, N., Niessen, S., Wei, B.Q., Nicolau, M.,Humphrey, M., Ji, Y., Han, W., Noh, D.,Yates, J.R., Jeffrey, S.S., Cravatt, B.F. *Nat. Meth.,* **2005a** 2(9), 691–697.

[88]Greenbaum, D.C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradszky, K.F., Engel, J., DeRisi, J., Holder, A.A. Bogyo, M. *Science,* **2002c** 298(5600), 2002–2006.

[89]Jessani, N.;Cravatt, B.F.,*Current Opinion in Chemical Biology,* **2004** 8(1), 54–59.

[90]Speers, A.E.; Cravatt, B.F. *ChemBioChem,* **2004** 5(1), 41–47.

[91]Zhu, Q., Girish, A., Chattopadhaya, S., Yao, S.Q. *Chemical Communications* **2004**, 1512.

[92]Wang, G., Uttamchandani, M., Chen, G.Y.J., Yao, S.Q. *Organic Letters* **2003**, *5*, 737.

[93]Liau, M.L., Panicker, R.C., Yao, S.Q. *Tetrahedron Letters* **2003**, *44*, 1043.

[94]Chen, G.Y.J., Uttamchandani, M., Zhu, Q., Wang, G., Yao, S.Q. *ChemBioChem*.**2003**, *4*, 336.

Chapter 2

Rationale and Design of Activity Based Probes

2.1 Introduction

This chapter illustrates efforts towards the design of hydroxamate, carboxylate and sufhydryl based probes to potentially target the metalloproteases involved in hypertension and malaria such as Angiotensin Converting Enzyme (ACE) and P. *faliciparum* A – M1(P*f*A–M1).

2.2 Affinity tag approach to develop a metalloprotease probe for an *in vitro* proteomic experiment in Hypertension and Malaria

As described in detail in **Chapter 1** a number of methods have been developed for identification of new metalloproteases and their interacting proteins (e.g. metalloprotease substrates)[1,2,3] including cDNA pool expression strategies, two-hybrid approaches and two-dimensional gel electrophoresis (2D-GE), just to name a few. Among them, the 2-D approach, in which global proteomic profiles of normal and diseased cells were compared, often generates a few hundreds of candidate protein spots, the majority of which are proteins unrelated to disease. This thus makes it difficult, as well as impractical, to be used in large-scale studies of proteins interacting/associating with proteases during the diseased state. Already researchers are developing new methods to

simplify protein analysis and allow high-throughput screenings. The activity-based protein profiling tag approach can potentially filter out proteins that are not of interest and focus on certain classes of proteins (e.g. based on activity)[4, 5, 6, 7, 8] while protein chips and microarrays offer the attractive prospect of screening thousands of proteins simultaneously.[9, 10] Unknown proteins could be spotted on arrays and incubated with potential ligands, which bind and allow the detection of the target proteins. The whole process of separating and identifying proteins in a complex mixture hence may be simplified, by zooming in only onto specific target proteins. Our metalloprotease probes are intended and hypothesized to selectively target all zinc metalloproteases in a single proteomic experiment for the targeted hypertension and malaria disease models.

2.3 Design of Malaria ABPP probes for use in malaria disease model

Generally, as described in **Chapter 1**, our probe is made up of four parts. Firstly, it has a hydroxamate, carboxylate or sulfhydril reactive group, secondly a simple polyethylene glycol linker as spacer, thirdly a benzophenone photocrosslinker and finally a fluorescent tag for easy visualization and purification of target proteins.

2.3.1 Reactive group

Gluzincins such as ACE and Aminopeptidases are potentially highly interesting targets for therapeutic intervention.[11, 12] Successful production of synthetic inhibitors with therapeutic value has been achieved, especially in the

field of cardiovascular disease with several inhibitors and chemical entities currently on the market. However, many of these, especially inhibitors of ACE have been associated with undesirable side-effects due to their lack of domain selectivity.[13, 14, 15, 16] Most ACE inhibitors synthesized to date are based on a Zn^{2+}-chelating group fitted onto a peptide-like backbone mimicking the endogenous protein substrate of the protease. Following the Schechter and Berger nomenclature[17] three distinct inhibitor types can be distinguished: inhibitors with a peptide sequence N-terminal (P_1-P_2) or C-terminal (P_1'-P_2') of the scissile peptide bond (i.e. the Zn^{2+} ion in the catalytic site) and compounds with a zinc binding group (ZBG) located in the middle of the inhibitor with a backbone protruding on either side.

Several functionalities such as carboxylic acid derivatives, thiol groups, phosphinates, ketones can be used as effective ZBGs but the most effective ZBG for producing high-affinity inhibitors has been the hydroxamate moiety. The latter have been found immensely suitable in elucidating the binding requirements and chemically reactive residues of the active sites of zinc metalloproteases and are a popular approach as zinc chelators.[18] However, their failure in clinical trials has been, among other reasons, their lack of selectivity. This failure has been mainly due to the strong bidentate zinc(II)- binding capacity of hydroxamic acid group (Figure 2.1). This strong ligation of the Zn^{2+} ion, which is the main determinant of the overall binding energy,[19] effectively replaces the water molecule and inhibits the catalytic mechanism. However, it suppresses the overall

effect of weaker but more specific non-covalent interactions of the inhibitor backbone with the substrate specific regions of enzymes[18, 20] and as a result has led only to the development of a large number of highly efficient but broad spectrum zinc metalloproteases inhibitors with low nanomolar affinity.[21]

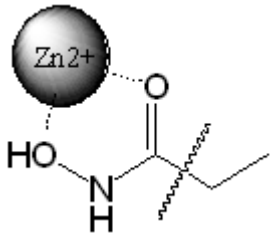


Figure 2.1: Bidentate binding of hydroxamates with zinc metalloproteases.

It is this lack of selectivity that has been exploited in the design of the probe reactive group structure. Hydroxamates as a class have been reported to inhibit P*f*A-M1 a zinc aminopeptidase found in the malaria parasite Plasmodium *falciparum*[22] as well as exhibit *in vitro* schizontocidal activity.[23] They are also reported as inhibitors of ACE in hypertension.[24] There also exists a plethora of literature on their success in closely related matrix-metalloproteases.[25] Generally, therefore, their promiscuity renders them as an optimal binding scaffold which accommodates a broad range of active-site structures found in metalloproteases and were thus proposed as the zinc-chelator or "war head" in this research.

2.3.2 Polyethylene Glycol (PEG) Linker

PEGylation is a procedure of growing interest for enhancing the therapeutic and biotechnological potential of peptides and non-peptides.[26] When a PEG is properly linked to a polypeptide, it modifies many of its features while the main biological functions, such as enzymatic activity or receptor recognition, may be maintained. The potential of the conjugation of PEG's to proteins emerged in the 1970's.[27] PEGylation produces alterations in the physiochemical properties including changes in conformation, electrostatic binding and hydrophobicity improving solubility. Several authors have reviewed different aspects of PEGylation with emphasis on the polypeptide and protein conjugation and also on the modification of peptides and non-peptide moieties.[28, 29, 30, 31, 32, 33, 34]

The use of this PEG linker was proposed due to its hydrophilicity which would potentially minimize non-specific interactions that can arise from the spacer group becoming buried in the hydrophobic pocket of proteins. Secondly, this spacer restricts hindrance between the peptide and reporter group, hence better avidin-biotin binding during affinity purification. The PEGylation process would involve the preparation of suitable homo-bifunctionalization of the polymer at the terminals.

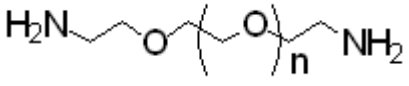


Figure 2.2: Homo-bifunctional linker for PEGylation to probe

2.3.3 Benzophenone Photocrosslinker

Taking into account the fact that in the catalytic mechanism of metalloproteases the nucleophile is not part of the protein itself but rather an active-site bound water molecule, the development of covalent inhibitor probes for these enzymes has to rely on a different principle. Described probes for MMPs[35, 36] and for ADAM proteases[37] all rely on inhibitor probes that have an interaction with the catalytic zinc ion (usually a hydroxamate zinc-chelating group) combined with a reactive photo-crosslinking group that forms a covalent bond with the enzyme after *uv* irradiation. Given that the active site is located at least 32 Å deep inside the protein, a probe with sufficient length spanning the active site and outside regions of the protein is ideal to ensure the photocrosslinker and biotin remain on the outside for photocrosslinking, visualization and affinity purification.

2.3.4 Reporter Group

Biotin **2.3a** (Fig. 2.3) is an entity utilized in cell growth, fatty acid production, fat and amino acid metabolism as well as participates in the citric acid cycle. It is also important in the maintenance of steady blood sugar levels as well as a co-factor responsible for carbon dioxide transfer in several carboxylase enzymes.[38] The attachment of biotin to various chemical sites, a process called biotinylation, is used as an important laboratory technique to study various processes including protein localization, protein interactions, DNA transcription and replication. Biotin binds very tightly to

the tetrameric protein avidin, streptavidin and neutravidin.[39, 40] Biotin may be covalently linked, or tagged, to a molecule for biochemical assays in biotinylation. Since avidins bind preferentially to biotin, biotin-tagged molecules can be extracted from a sample by mixing them with beads with covalently-attached avidin, and washing away anything unbound to the beads.

Biotin was thus the choice for the activity based probes in this research to allow for future avidin-based purification of target structures. The method would enable the enrichment and purification of labelled proteins by affinity capture with streptavidin beads. This biotin-streptavidin interaction can be broken using various methodologies, freeing the biotinylated structure for further analytical purposes.[41]

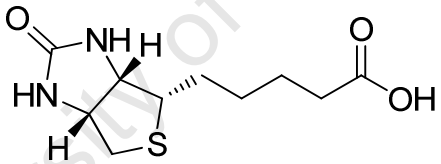


Figure 2.3: Structure of Biotin

2.4 Methodological prospects

The analytical use of small molecule inhibitors of metalloproteases has the potential to greatly improve our understanding of the roles of these proteases in malaria and hypertension disease conditions. By employing this functional proteomics technique described schematically in steps 1 to 6 (figures 2.5 to 2.10) below, the analysis may move from correlating the well known and well-described ACE and aminopeptidases in disease towards profiling of all active metalloproteases. This may lead to new insights, and

potentially to new targets for pharmaceutical intervention and the opportunity for a better follow-up of therapy. On the other hand, this methodology may also be used to identify unknown protein targets that interact with newly developed as well as existing inhibitors thus allowing better optimization of inhibitors prior to entering clinical trials.

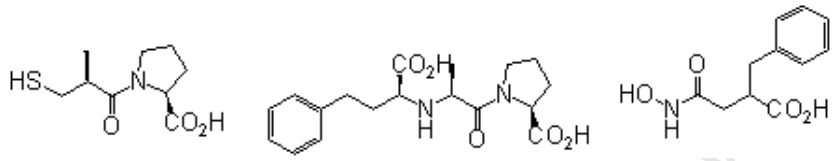


Figure 2. 4: Step 1. **Selection of a suitable available inhibitor, or synthesis of novel inhibitors**

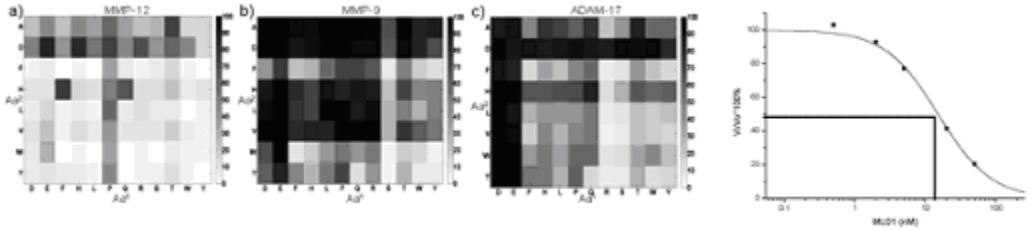


Figure 2. 5: Step 2. Screening of the inhibitory profile using a relevant metalloproteases

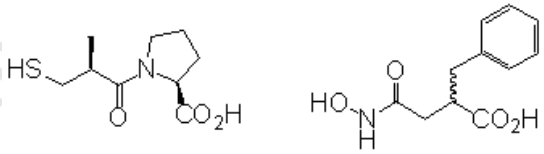


Figure 2.6: Step 3. Selection of best-suited inhibitor for profiling.

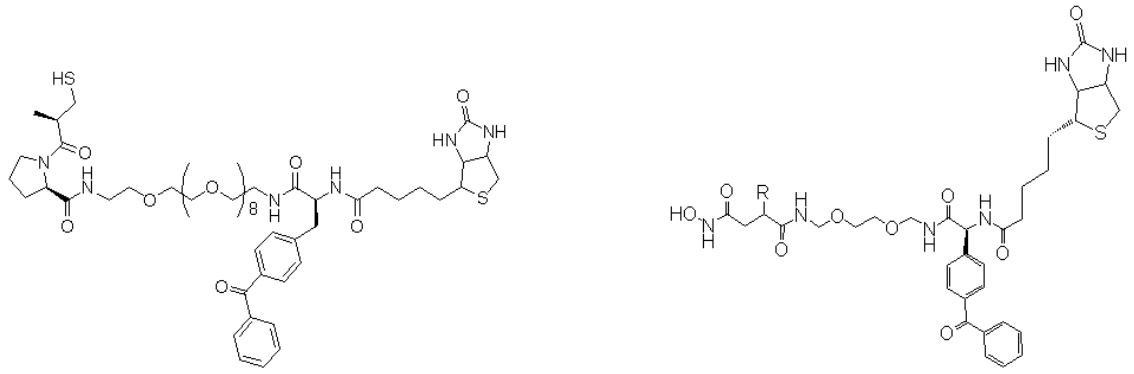


Figure 2.7: Step 4. Conversion of inhibitor into suitable structure for future activity-based proteomics via synthesis of activity-based probe by introducing spacer, photo-crosslinker and a reporter molecule

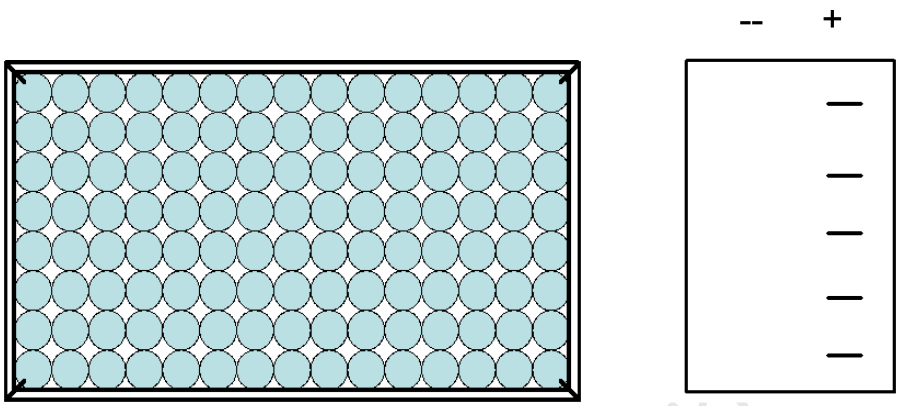


Figure 2. 8: Step 5. *In-vitro* testing of newly synthesized probes for activity-based proteomics with proteases by labelling of samples of interest with the activity based probe and visualisation of active metalloproteases after gel electrophoresis

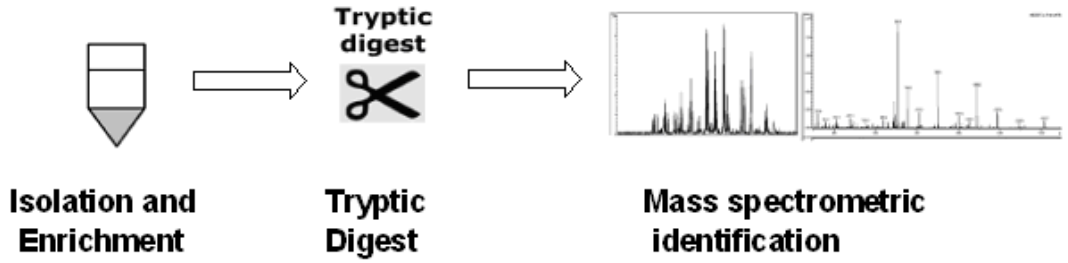


Figure 2.9: Step 6. Isolation and enrichment of active proteases from biological samples by enrichment of photo-labelled proteins with immobilized streptavidin. This is then followed by a tryptic digestion of enriched proteins and mass spectrometric identification.

Similar methodology have been reported by several authors and applied extensively in MMPs which are related to Zinc metalloproteases. However, this thesis describes that synthesis of novel probes with PEG linkers which have incorporated the use of a benzophenone photocrosslinker. This is the first time that such probes targeting future

analysis of protease activity in the malaria and hypertension disease models specifically targeting ACE and P*f*A-M1 have been synthesized. It is also the first time that the known ACE inhibitors namely suflhydril captopril as well as the carboxylate enalaprilat have been employed as reactive groups in such probes. Successful binding of the probe molecule would lead to insight into the interaction of these known inhibitors with other active enzymes in proteomes. From the X-ray crystal structure, the active-site of ACE is known to be located 32Å deep inside the 32kda enzyme.[42] Application of the PEG_8 probe which is sufficiently long is expected to access this active-site. However, we envisaged and hypothesized that if the shorter PEG_3 probes were able to access the active-site and label active ACE, we would be able to have further insight into the conformation and positioning of the active-site when ACE is in action.

References

[1] Cryns, V.L., Byun, Y., Rana, A., Mellor, H., Lustig, K.D., Ghanem, L., Parker, P.J., Kirschner, M.W., Yuan, Y.J., *Journal of Biological Chemistry* **1997**, *272*, 29449.

[2] Wu, Y.H., Shih, S.F., Lin, Y.J. *Journal of Biological Chemistry* **2004**, *279*, 19264.

[3] Wan, J.H., Wang, J.L., Cheng, H.P., Yu, Y.T., Xing, G.C., Qiu, Z.Y., Qian, X.H.,He, F.C. *Electrophoresis*, **2001**, *22*, 3026.

[4] Adam, G. C., Sorensen, E. J., Cravatt B. F. *Molecular and Cellular Proteomics* **2002**, *1*, 781.

[5] Adam, G. C., Sorensen, E. J., Cravatt, B. F. *Nature Biotechnology* **2002**, *20*, 805.

[6] Liu, Y., Patricelli, M. P., Cravatt, B. F. *Proceeding of the National Academy of Sciences U.S.A.* **1999**, *96*,14694.

[7] Kidd, D., Liu,Y., Cravatt, B. F. *Biochemistry* **2001**, *40*, 4005.

[8] Cravatt, B.F., Sorensen, E.J., *Current Opinion in Chemical Biology* **2000**, *4*, 663.

[9] MacBeath, G., Schreiber, S. L. *Science* **2000**, *289*, 1760.

[10] Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N.,Jansen, R., Dinlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., Snyder, M. *Science* **2001**, *293*, 2101.

[11] Smith, A.I., Turner, A.J., *Cellular and Molecular Life Sciences*, **2004** 61 2675–2676.

[12] Moore, HE., Davenport, EL., Smith, EM., Muralikrishnan, S., Dunlop, AS., Walker, B.A., Krige, D., Drummond, A.H., Hooftman, L., Morgan, G.J., Davies, F. E. *Molecular*

Cancer Therapeutics **2009** 8; 762.

[13] Okumura H, Nishimura E, Kariya S, et al. *Yakugaku Zasshi* **2001**;121(3):253-7.

[14] Rang H.P., Dale, M.M, Ritter, J.M.. *Pharmacology*. Churchill Livingstone, London. 4th ed. **1999** 290-296.

[15] Lacourciere Y, Brunner H, Irwin R, Karlberg BE, Ramsay LE, Snavely DB, Dobbins TW, Faison EP, Nelson EB. *Journal of Hypertension* **1994**;12: 1387-1393.

[16] Messerli, F. H. and Nussberger, J. *Lancet* **2000** 356, 608-609.

[17] Schecter, I.; Berger, A. *Biochemical and Biophysical Research Communications* **1967**, *27*, 157-162.

[18] Cuniasse, P.; Devel, L.; Makaritis, A.; Beau, F.; Georgiadis, D.; Matziari, M.; Yiotakis, A.; Dive, V. *Biochimie* **2005**, 87, 393.

[19] Babine,R.E. & Bender,S.L. *Chemical Reviews* **1997**; 97, 1359-1472.

[20] Rao, B. G. *Current Pharmaceutical Design* **2005**, 11, 295.

[21] Whittaker,M., Floyd,C.D., Brown,P. & Gearing,A.J. *Chemical Reviews* **1999**; 99, 2735-2776.

[22] Flipo, M.; Beghyn, T.; Leroux, V.; Florent, I.; Deprez, B.P..; Deprez-Poulain, R. F. *Journal of Medicinal Chemistry* **2007**; 50:1322-1334.

[23] Mishra, R. C.; Tripathi, Renu; Katiyar, Diksha; Tewari, Neetu; Singh, Deepti; Tripathi, R. P. *Bioorganic*

Medicinal Chemistry **2003** 11: 5363–74.

[24] Turbanti, T., Cerbai, G., Di Bugno, C., Giorgi, R., Garzelli, G., Criscuoli, M., Renzetti, AR., Subissi, A., Bramanti, G., *Journal of Medicinal Chemistry* **1993 36** 699–707.

[25] Fisher , JF., Mobashery, S. *Cancer Metastasis Review* **2006** 25: 115–136.

[26] Veronese, F.M. *Biomaterials* 22 (2001) 405–417.

[27] Abuchowski, A.; McCoy, J. R.; Palczuk, N. C.; van Es T.; Davis, F. F. *Journal of Biological Chemistry* **1977**; 252 (11): 3582–3586.

[28] Monfardini, C.; Veronese, F.M. *Bioconjugate Chemistry* **1998**;9:418–450.

[29] Katre, NV. *Advanced Drug Delivery Review* **1993**; 10:91–114.

[30] Hooftman G, Herman S, Schacht E. *Journal of Bioactive Compatible Polymers* **1996**;11:135–159.

[31] Nucci, M.L., Shorr, R., Abuchowski, A. *Advanced Drug Delivery Reviews* **1991**;6:133}51.

[32] Zalipsky, S. *Advanced Drug Delivery Reviews* **1995**;16:157}82.

[33] Delgado, C., Francis, G.E., Derek, F.F. *Critical Reviews in Therapeutic Drug Carrier Systems* **1992**;9:249–304.

[34] Dreborg, S, Akerblom, E,B. *Critical Reviews in Therapeutic Drug Carrier Systems* **1990**;6:315}65.

[35] David, A., Steer, D., Bregant, S., Devel, L., Makaritis, A., Beau, F., Yiotakis, A., Dive, V. *Angewandte Chemie International Edition in English* **2007**; 46 3275–3277.

[36] Chan,E.W.S., Chattopadhaya,S., Panicker,R.C., Huang,X.Yao,S.Q. *Journal of the American Chemical Society* **2004** 126, 14435-14446.

[37] Leeuwenburgh,M.A. Geurink, P., Klein, T., Kauffman, H., van der Marel, G., Bischoff, R., Overkleeft, H.,*Organic Letters* **2006***;* 8, 1705-1708.

[38] Zempleni J, Wijeratne SS, Hassan YI. *Biofactors* **2009;** 35 (1): 36-46.

[39] Bonjour, JP. *International Journal of Nutritional Research* **1977** ; 47 : 107-118.

[40] Green, N.M. *Advanced Protein Chemistry* **1975** 29: 85-133.

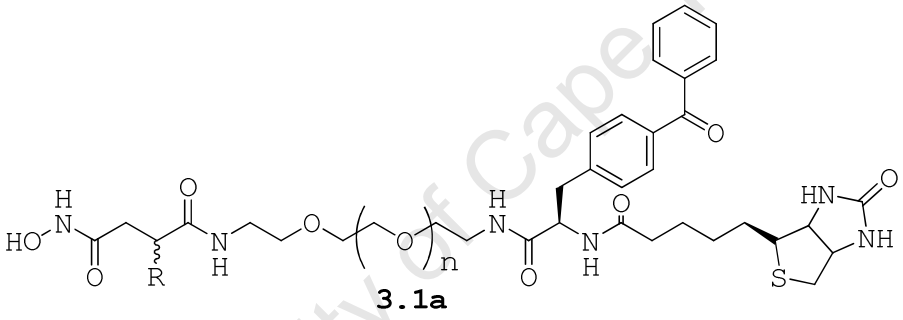
[41] Holmberg, A., Blomstergren, A., Nord, O. *Electrophoresis* **2005** 26 (3): 501-10.

[42] Acharya, K. R.; Sturrock, E. D.; Riordan, J. F.; Ehlers, M. R., *Nature Reviews Drug Discovery* **2003**; 2: 891-902.

Chapter 3

3.0 Introduction to Chemical synthesis

Chapter two described the design of the probes to be employed in future for both the hypertension and malaria disease models. Furthermore, it described the rationale behind the synthesis of such probes. It also looked briefly at the methodological prospects to be employed in the application of these probes. **Chapter 3** will give a retro-synthetic strategy followed by a detailed synthetic discussion which culminated into the synthesis of the targeted probes with general structure **3.1a** below(Figure 3.1).



3.1a

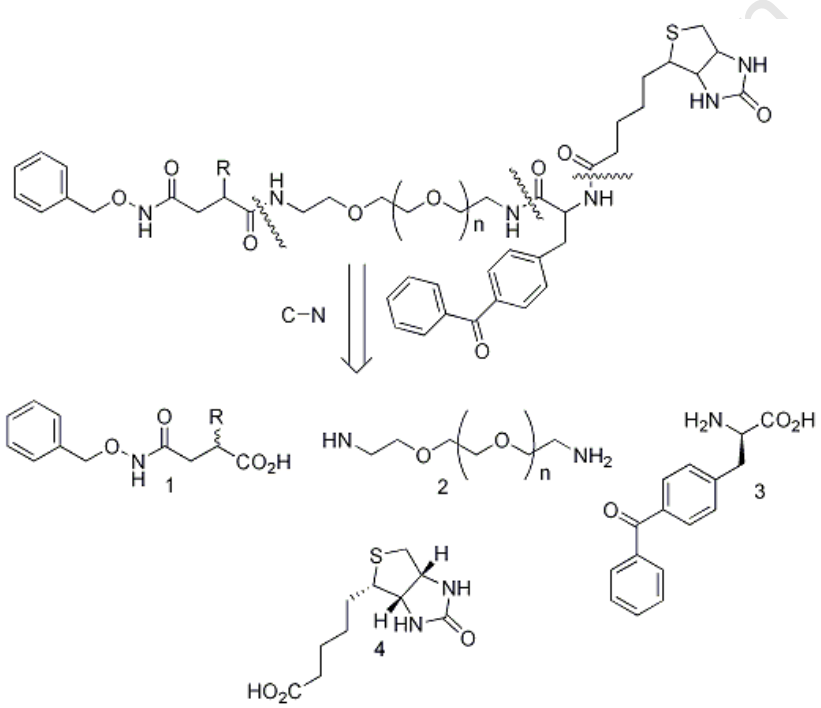
Figure 3.1: Probes general structure

3.1 Retrosynthetic strategy

3.2.1 Overall restrosynthetic strategy

To determine the best possible synthetic route to the targeted probes, a retrosynthetic strategy was conducted. Retrosynthesis involves working backwards from the target molecule in order to devise a suitable synthetic route. In the case of the present study, the retrosynthetic analysis revealed that a convergent route, involving several basic 'building blocks' joined together relatively late on in the synthesis to give the final product was the best strategy.

C–N peptide bond disconnection of the designed probe general structure **3.1a** leads to four sub-units.(Scheme 3.1 below) The four adducts being the o-benzyl protected reactive group, the PEG spacer,(S)-2-amino-3-(4-benzoylphenyl)propanoic acid (Benzophenone) photo-crosslinker and 5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoic acid (biotin) tag identified as **1**, **2**, **3** and **4** respectively. Sub-units **3** and **4** are commercially available whereas **1** and **2** had to be synthesized.

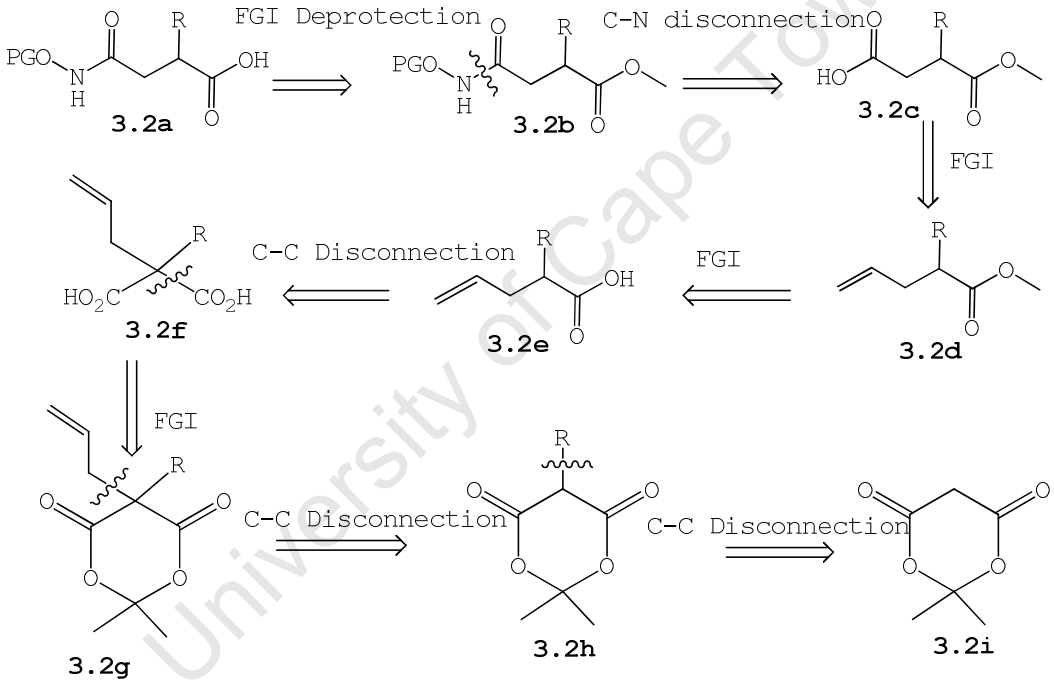


Scheme 3.1: Retrosynthetic strategy for the complete probe showing sub-units **1**,**2**, **3**, and**4**.

3.2.2 Retrosynthesis of the reactive group

The *O*-benzyl protected reactive group **1**(scheme 3.1) is a succinic acid derivative. The stereochemistry at the steroegenic centre could be either the *S*- or *R*- and resolution was thus not important. Succinate hydroxamates have been reported to have been synthesized utilizing several strategies including malonate esters as starting material.[1] Retrosynthetically, we envisaged an initial functional group

interconversion of the carboxylate **3.2a** to the ester **3.2b.**(Scheme 3.2)A *C–N* disconnection at the benzyl protected hydroxamateof **3.2b** would give the carboxylic acid **3.2c** which can be prepared from **3.2d** by a functional group interconversion (FGI).Intermediate **3.2d** itself may easily be formed from carboxylic intermediate **3.2e.** The latter is a product of 1,3-dicarboxylic acid **3.2f. 3.2g** is a dialkylated derivative of 2,2-dimethyl-1,3-dioxane-4,6-dione **3.2i** commonly known as Meldrum's acid and this was chosen as the starting material.



Scheme 3.2: Retrosynthetic strategy for the reactive group.

3.2.3 Synthetic Strategy

3.2.3.1 Introduction

Synthesis started with 2,2-dimethyl-1,3-dioxane-4,6-dione**3.2i** (Meldrum's Acid). The latter is an active methylene compound with a rigid cyclic structure, high acidity (pKa = 4.9) and can undergo hydrolysis very easily.[2] It is well known that Meldrum's acid can undergo Knoevenagel condensation.[3, 4]Reports

on alkylation studies reveal that unlike dialkyl malonates it has a strong tendency to undergo bis-alkylation even with one mole equivalent of alkyl halide[5] and only in exceptional cases does mono-alkylation result.[6,7] The interest in the synthesis of monoalkylated Meldrum's acid as well as mixed dialkylated Meldrum's acid continues to evolve due to their use in the synthesis of dialkylated malonates as well as barbiturates.

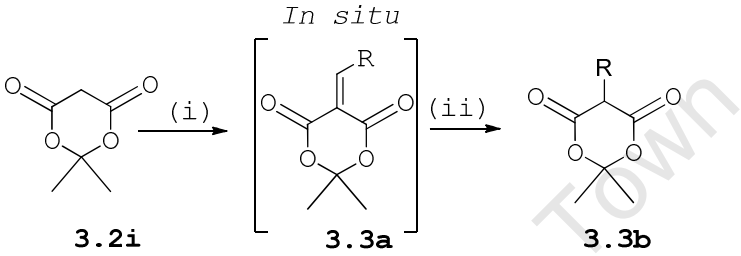
The difficulties in preparing dialkylated Meldrum's acid wherein the two alkyl substituents are different can be overcome by alkylation of mono-alkylated Meldrum's acid, which in turn can be prepared by several described methods. Firstly, they may be prepared by a Knoevenagel condensation of Meldrum's acid with carbonyl compounds to ylidene Meldrum's acid followed by conjugate reduction with sodium borohydride[8] or sodium tellurohydride[9] or by cuprous chloride catalyzed conjugate Grignard addition to ylidene Meldrum's acid.[10] The second reported option is by condensation of Meldrum's acid with aromatic aldehydes in the presence of formic acid and triethylamine,[11]or thirdly by the reduction of acylated Meldrum's acid with sodium cyanoborohydride.[12]A fourth option is the reductive alkylation of Meldrum's acid with carbonyl compounds using boranetriethylamine.[13]

All these methods, however, require a multi-step reaction sequence to be carried out. Therefore a simple, one pot method for the preparation of monoalkylated as well as mixed dialkylated Meldrum's acids was preferred.[14] It involved the Knoevenagel condensation of Meldrum's acid with various aldehydes utilizing K_3PO_4 as a catalytic base.

3.2.3.2 Preparation of the Mono-alkylated Meldrum's acid

The reaction sequence depicted in scheme 3.3 below was carried out in ethanol medium by K_3PO_4 as base catalyst giving the various Knoevenagel condensation products *in situ*. This

specific solvent–base combination not only effected Knoevenagel condensation of Meldrum's acid **3.2i** with aromatic and some aliphatic aldehydes to give arylidene Meldrum's acids **3.3a,** but also permited the reduction of **3.3a** with $NaBH_4$ in the same pot to furnish monoalkylated Meldrum's acid derivatives **3.3b.** The sequence of reactions was carried out at room temperature and were monitored easily by TLC.

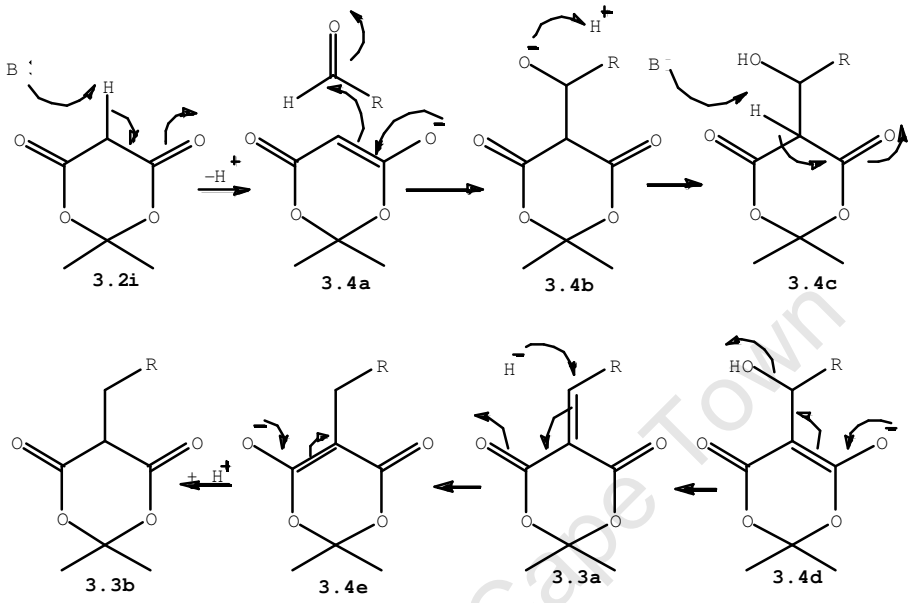


Compound	3.3b–3	3.3b–4	3.3b–5	3.3b–6
R – Group	$CH_2CH(CH_3)_2$	CH_2Cyclohexyl	CH_2Ph	CH_2Naphthyl
Yield	88	79	91	90

Scheme 3.3: Reagents and conditons. (i) RCHO, K_3PO_4EtOH, RT, 4 h (ii) $NaBH_4$, 3 h.

Mechanistically, the Knoevenagel condensation proceeds *via* the enolate ion generated by the catalytic base. K_3PO_4 is a strong enough base to form a significant concentration from the 1,3-dicarbonyl compound **3.2i** but not strong enough to form the enolate from the aldehyde. The formation of the enolate **3.4a** in scheme 3.4 below can be from either tautomer of **3.2i.** The enolate ion **3.4a** attacks the electrophilic carbonyl of the aldehyde and the resulting intermediate **3.4b** captures the initially abstracted proton which is in the reaction mixture resulting in the aldol **3.4c.** However, there is still one acidic proton between the two carbonyl groups in **3.4c** so enolate anion formation is again easy and dehydration of **3.4d** follows to give the unsaturated product **3.3a.** It is at this point that 1,4–conjugate reduction of the unsaturated

alkylidene product **3.3a**is conveniently achieved with $NaBH_4$, resulting in **3.4e.** Acidification with dilute HCl then gives way to the mono-substituted Meldrum's acid **3.3b.**



Scheme 3.4: Mechanism of Knoevenagel condensation and conjugate reduction.

A section of the typical ^1H NMR of the monosubstituted Meldrum's acid, namely 5-benzyl-2,2-dimethyl-1,3-dioxane-4,6-dione **3.3b–5.** (Fig. 3.2,Phenyl region not shown)

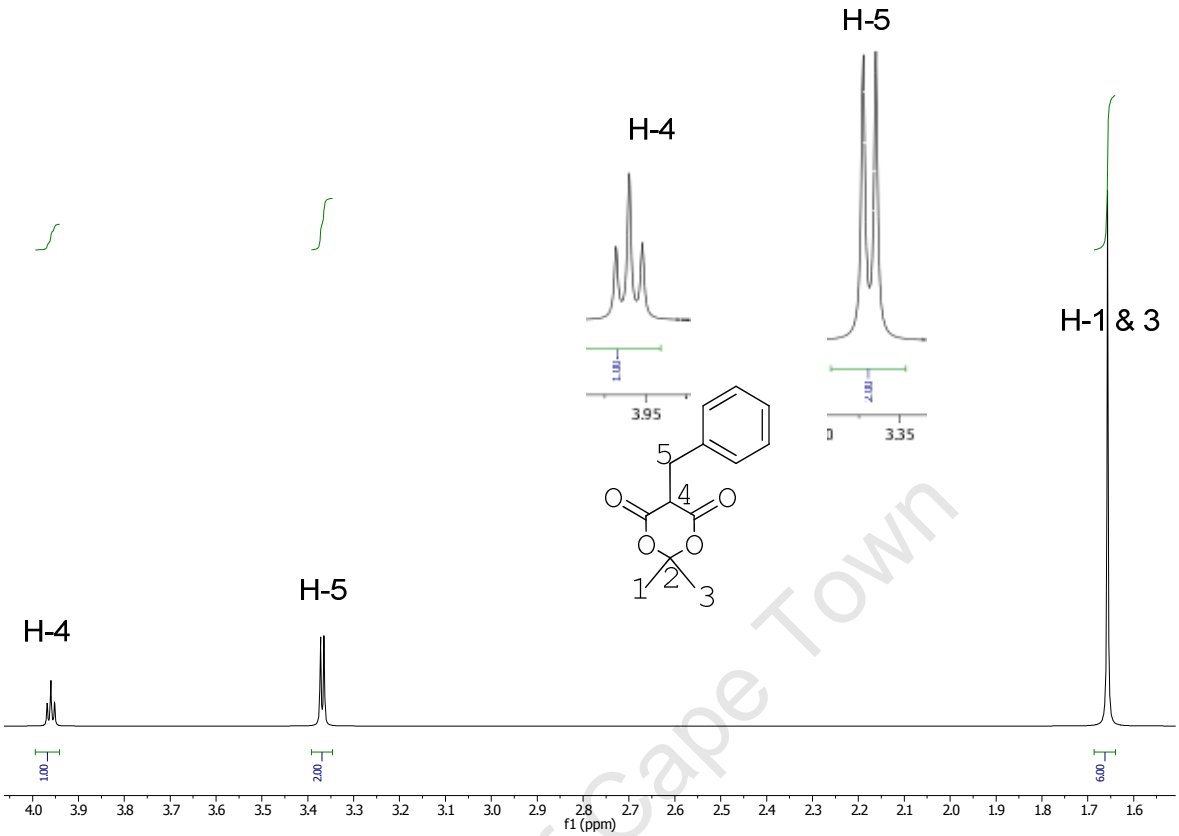
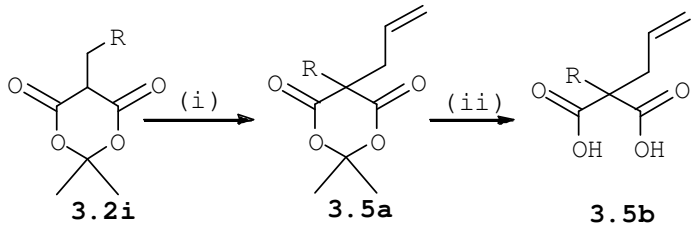


Figure 3.2: **A section of a typical ^{1}H NMR of mono-substituted Meldrum's acid** 3.3b–5 **in $CDCl_3$.**

From the foregoing, a singlet was observed in the region ~1.65 ppm integrating for six C–1 and C–3 protons. Secondly, the C–5 protons were each revealed as doublet around 3.35 ppm and finally the C–4 proton came as a triplet at ~3.95 ppm.

3.2.3.3 Preparation of the disubstituted Meldrum's acid

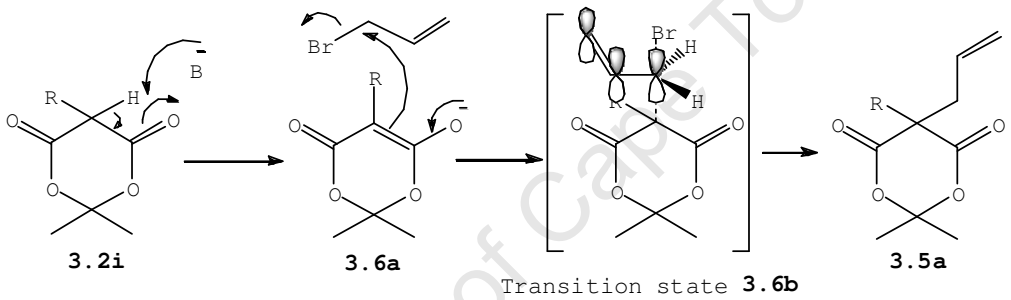
The desired unsymmetrical 5-benzyl-5-substituted Meldrum's acids **3.5a** were prepared in one-step by reacting the mono-substituted Meldrum's acid **3.3b** with 2 equivalents of allyl bromide using K_2CO_3 in DMF. (Scheme 3.5)



Compound	3.5a-3	3.5a-4	3.5a-5	3.5a-6
R – Group	$CH_2CH(CH_3)_2$	CH_2Cyclohexyl	CH_2Ph	CH_2Naphthyl
Yield	88	80	75	90

Scheme 3.5: Reagents and Conditions (i) Allyl Bromide, K_2CO_3, DMF, 0 °C – rt., overnight (ii) 4M NaOH, 1,4-Dioxane (1:1) reflux 100°C,5 h.

Mechanistic details of the di-substitution reaction are shown in scheme 3.6.It proceeds *via* a typical S_N2 reaction mechanism. Allyl compounds react rapidly by this mechanism because the double bond can stabilize the transition state **3.6b** by conjugation with the allylic π-bond.



Scheme 3.6: Mechanism of S_N2 reaction with allyl bromide.

The quaternized Meldrum's acid derivatives **3.5a** were easily isolated in an analytically pure form by extraction and further purified by recrystallization from MeOH. Substrate isolation was facilitated by the exceptional crystallinity of Meldrum's acid derivatives. A typical ^1H NMR for this disubstituted structure, 5-allyl-5-benzyl-2,2-dimethyl-1,3-dioxane-4,6-dione **3.5a-5.** (Fig. 3.3)

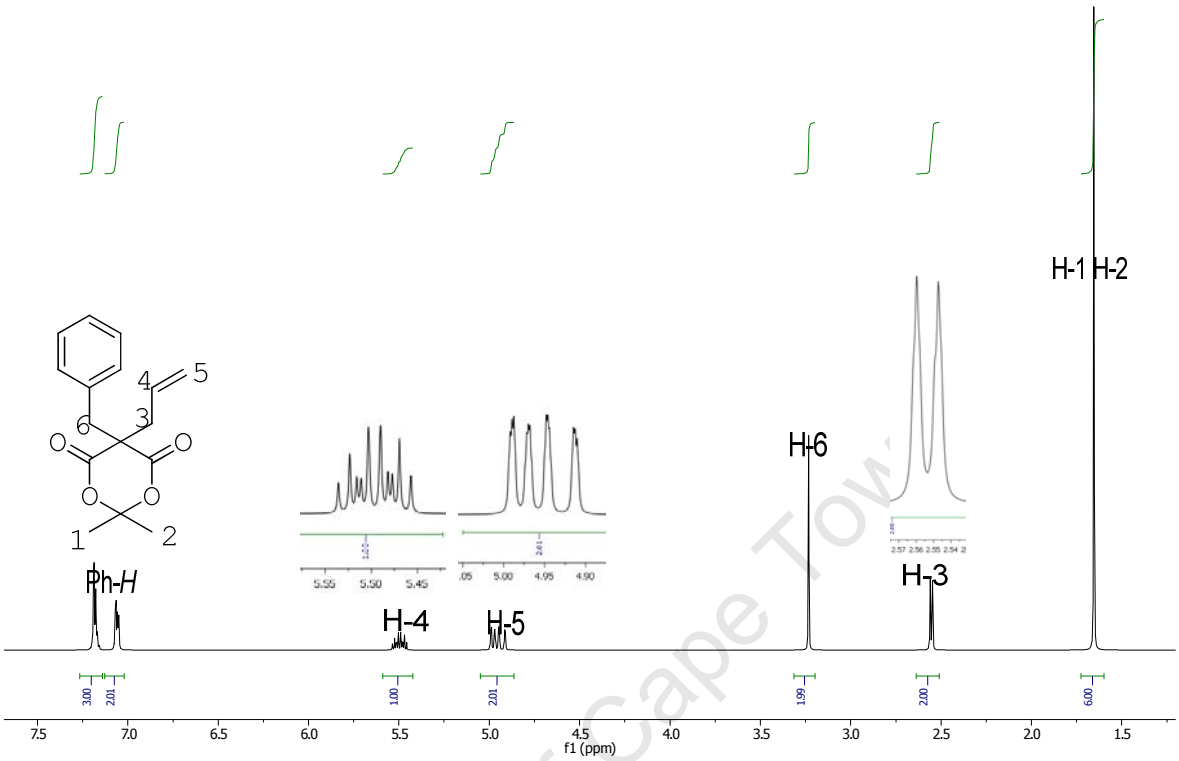


Figure 3.3: Section of typical ^{1}H NMR of 5-allyl-5-benzyl-2,2-dimethyl-1,3-dioxane-4,6-dione **3.5a-5**.

Amongst this group of intermediates, the common diagnostic features in the ^{1}H NMR were the appearance of the characteristic vinylic proton signals downfield doublet of doublets at ~4.95 ppm and a multiplet at ~5.50 ppm integrating for the two C-5 protons and one C-4 proton respectively. The H-3 protons typically came as a doublet around ~2.55 ppm.

3.2.3.4 Preparation of the benzyl protected hydroxamates

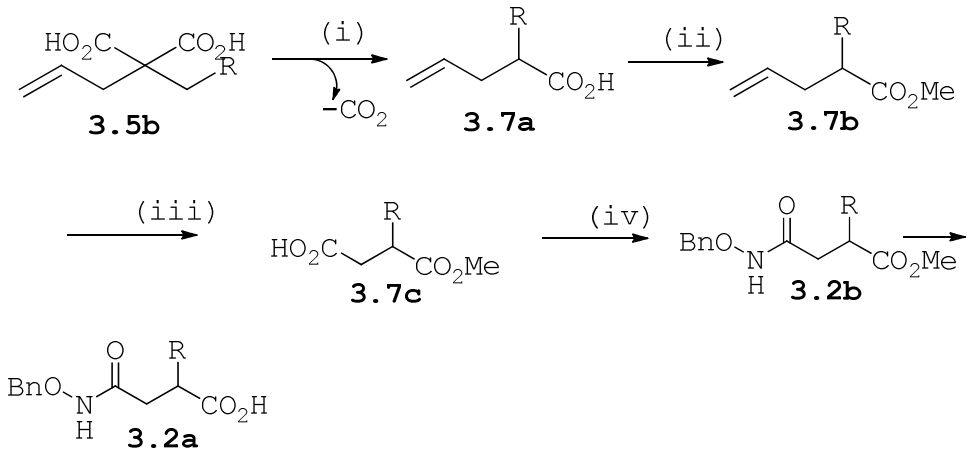
Hydrolysis of the acetonide ring of **3.5a** was achieved by treatment with sodium hydroxide in H_2O:1,4-dioxane (1:1)affording diacid **3.5b (**Scheme 3.5) in yields as per table 3.1.

Compound	3.5b-3	3.5b-4	3.5b-5	3.5b-6
R – Group	$CH_2CH(CH_3)_2$	CH_2Cyclohexyl	CH_2Ph	CH_2Naphthyl
Yield	77	70	88	80

Table 3.1: Yields for compound series 3.b.

Subsequent thermal decarboxylation in DMSO resulted in **3.7a** and condensation with methanol in the presence of $SOCl_2$ afforded the ester **3.7b** (Scheme 3.7). Ruthenium catalyzed oxidation of the terminal olefin gave the scission product carboxylic acid **3.7c.** RuO_4 was generated by vigorously stirring the required (stoichiometric) amount of $RuO_2.2H_20$ and a 6 molar excess of $NaIO_4$ in the tri-phasic system CH_3CN–CCl_4–H_20 (2:1:1) and the organic layer added portion-wise to the olefin **3.7b** dissolved in CCl_4.

Peptide coupling of the resultant carboxylic acid **3.7c** with commercially available benzyl hydroxylamine was achieved using molar equivalents of commercially available *N*-(3-dimethylamonipropyl)-*N'*-ethylcarbodiimide hydrochloride salt (EDC-HCl) in the presence of 1-hydroxybenzotriazole (1-HOBt) and *i*Pr_2NEt in DMF. The mixture was stirred at room temperature for 18h giving the fully protected precursor **3.2b.** Alkaline hydrolysis of the methoxy-ester isolated products yielded the desired benzyl protected succinic acid based reactive group compound series **3.2a-1** to **3.2a-6.**

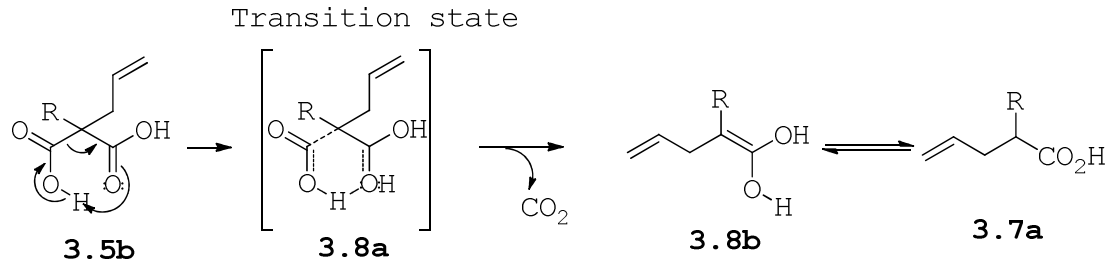


Scheme 3.7: Reagents and conditions. (i) DMSO, 130 °C, 3 h, (40 – 65%) (ii) $SOCl_2$, MeOH, 0 °C – rt, 3 h, quant. (iii) RuO_2(cat.), $NaIO_4$, Acetonitrile-CCl_4-water (2:1:1), rt, 48 h (45 – 65%); (iv) $BnONH_2$-HCl, EDC-HCl, DIEA, HOBt, DMF, RT, 18 h (60 – 84%).

Compound	3.2a-1	3.2a-2	3.2a-3	3.2a-4	3.2a-5	3.2a-6
R – Group	H	CH_3	$CH_2CH(CH_3)_2$	CH_2Cyclohexyl	CH_2Ph	CH_2Naphthyl
Yield	90	70	70	72	75	84

3.2.3.5 Mechanistic details of decarboxylation

Decarboxylation is an elimination type reaction resulting in the loss of a carbon dioxide molecule. Simple carboxylic acids rarely undergo decarboxylation. However, carboxylic acids with a carbonyl group at the 3- (or β-) position readily undergo thermal decarboxylation and a typical example are the derivatives of malonic acid. Structure **3.5b** is a malonic acid derivative. The reaction proceeds *via* a cyclic transition state. The mechanism begins with the protonation of the carbonyl, followed by breaking of the O-H bond then formation of the *pi*-bond as the carbon-carbon bond breaks all in a rearrangement of six electrons in a cyclic six-membered transition state **3.8a**(scheme 3.8). Tautomerisation of the enol **3.8b** leads to the carboxylic acid product **3.8c** (2-benzylpent-4-enoic acid **3.7a-5** ^{1}HNMR given in Fig.3.4**).**



Scheme 3.8: Mechanism of thermal decarboxylation of the malonic acid derivatives.

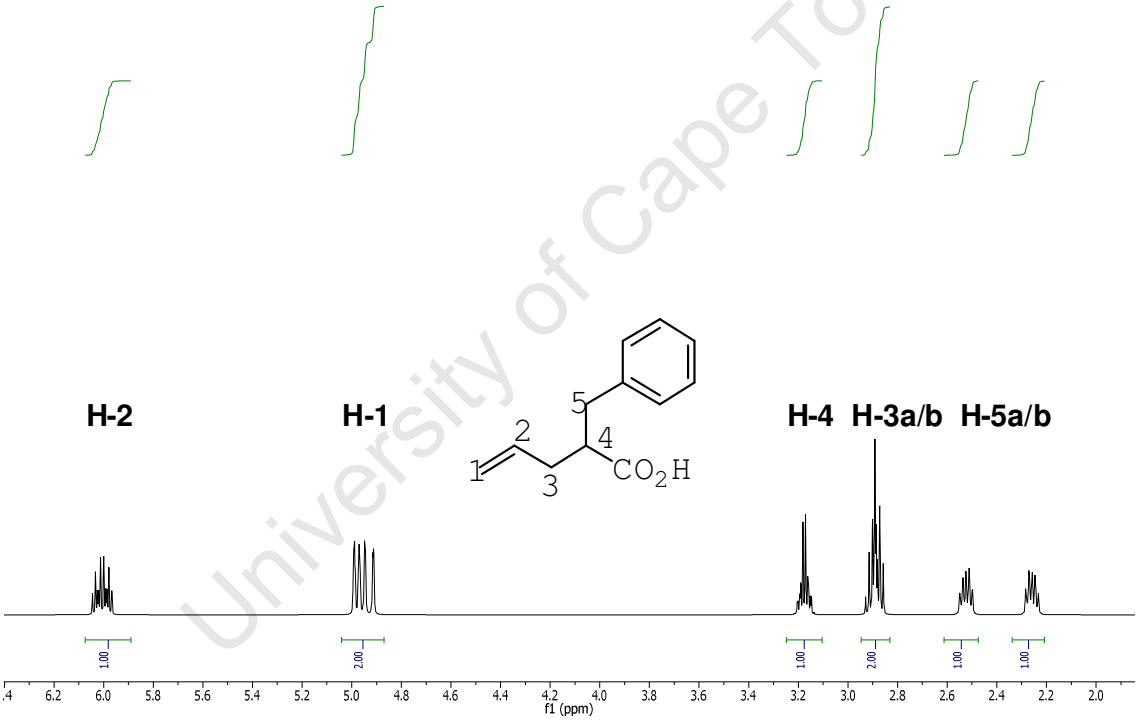
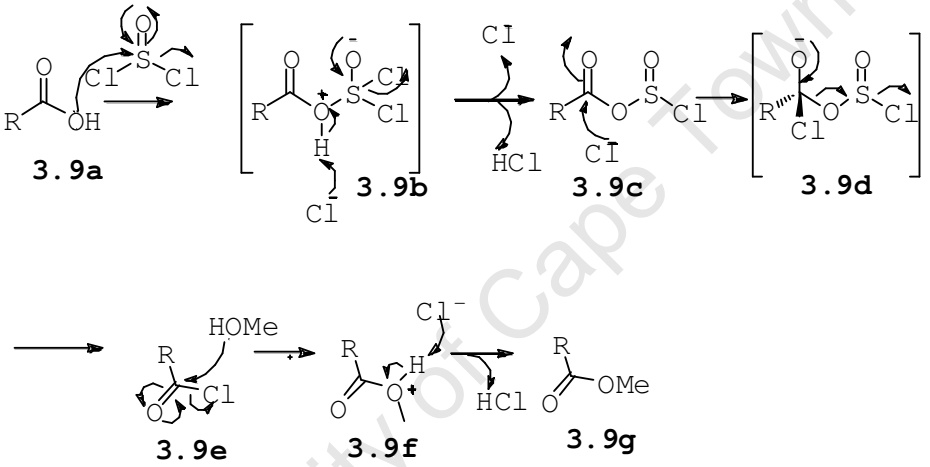


Figure 3.4: A section of ^{1}H NMR of 2-benzylpent-4-enoic acid **3.7a–5**in $CDCl_3$(phenyl region not shown).

3.2.3.6 Mechanism of methoxyester protection

The carboxylate group with general structure **3.9a** (Scheme 3.9) is the effective nucleophile. A substitution onto $SOCl_2$ takes place in a similar fashion as nucleophilic acyl substitution, with the loss of a chloride anion as a good leaving group resulting in transition intermediate **3.9b.** This is then

followed by the loss of a second chloride as well as abstraction of a proton by the first chloride leading to formation of HCl. The second chloride then serves as the nucleophile in the nucleophilic acyl substitution of **3.9c.** The next step **3.9d** driven primarily by the loss of sulfur dioxide then gives rise to a very reactive acid chloride intermediate **3.9e** that is sufficiently electrophilic to be attacked by the methanol oxygen lone pair to ultimately form the desired methyl ester **3.9g** after the departure of HCl in **3.9f.**

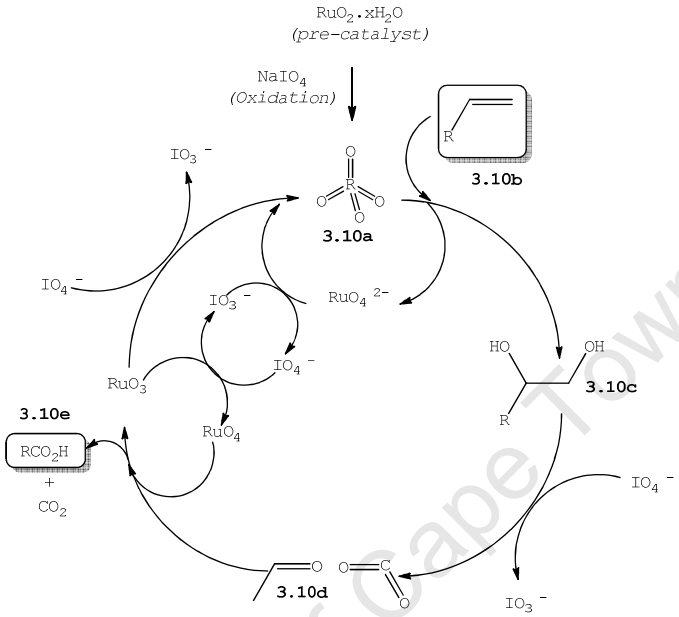


Scheme 3.9: Mechanism for methoxy-protection of the carboxylic group

3.2.3.7 Ruthenium tetraoxide catalyzed cleavage mechanism

Ruthenium tetroxide is a powerful oxidizing agent classically used for oxidative C–C-bond cleavage.[15,16] It is a diamagnetic tetrahedral ruthenium compound. It may be generated from $RuCl_3 \cdot xH_2O$ or RuO_2 as pre-catalyst with $NaIO_4$. A plausible mechanistic cycle for the cleavage of alkenes with ruthenium dioxide and sodium periodate[17] is given in scheme 3.10. The pre-catalyst $RuO_2.xH_2O$ is converted into the catalytically active RuO_4 **3.10a** *in situ*. Subsequently, the tetraoxide converts the alkene **3.10b** into a vicinal diol **3.10c**, which is oxidatively cleaved by periodate into an aldehyde **3.10d** with

the release of carbon dioxide. The aldehyde is subsequently oxidized with RuO_4 to afford the desired carboxylic acid **3.10e.** Ruthenium(VI) species that are produced during the oxidations, are oxidized back to RuO_4 through periodate.



Scheme 3.10: Proposed mechanism for the Ruthenium tetra-oxide catalyzed cleavage of terminal alkenes.[18]

A typical ^1H NMR of the hemi-succinate product **3.10e–5** is given below (Fig. 3.5).

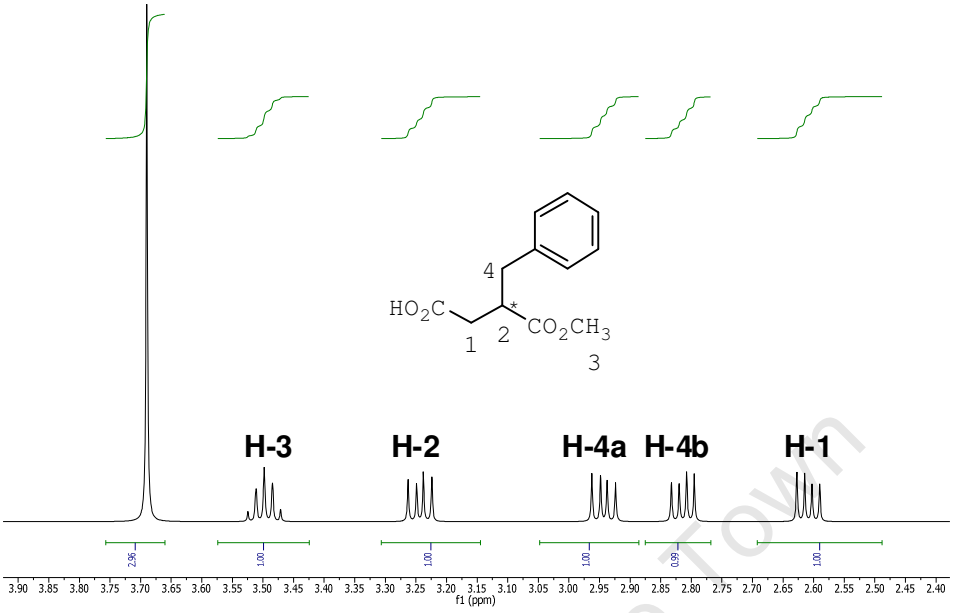


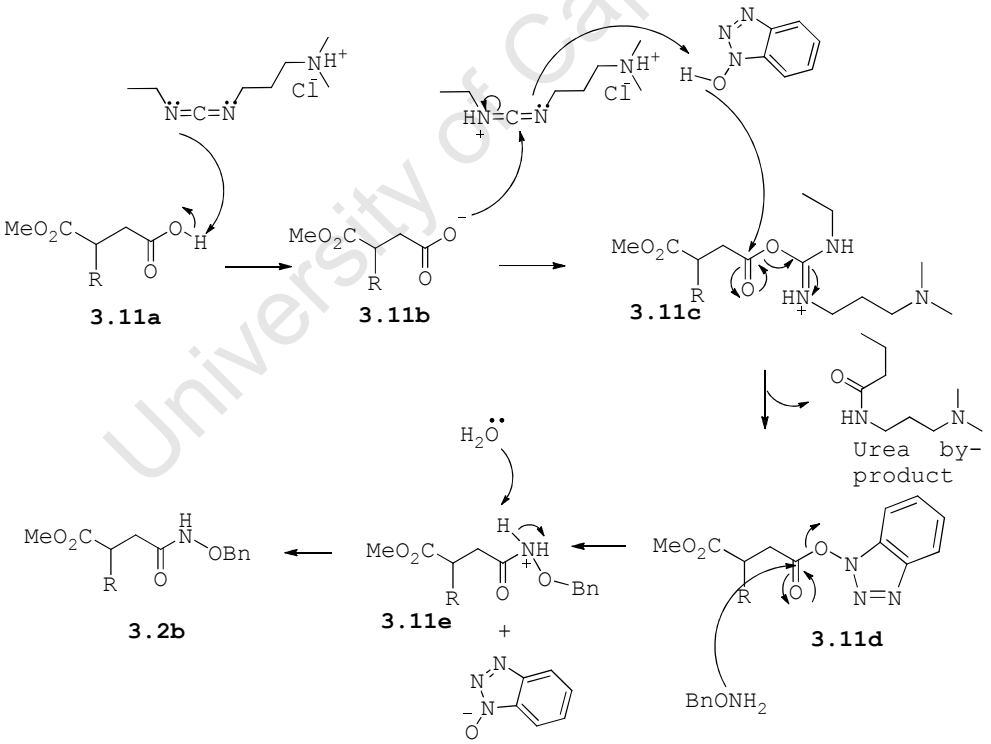
Figure 3.5: A section of ^{1}H NMR of hemisuccinate **3.10e–5** the RuO_4 catalyzed scission product in $CDCl_3$. (Phenyl region not shown).

All the ^{1}HNMR spectra of these intermediates displayed common diagnostic features. The most notable features, compared to the spectrum for the precursor showing successful cleavage of the terminal alkene and formation of **3.10e–5** included the disappearance of the signals for the vinylic protons in the ^{1}HNMR. This was corroborated by the loss of one carbon peak in the ^{13}C NMR spectra and the appearance of a second carboxylic acid peak at ~177.8ppm.

3.2.3.8 Mechanistic details of the peptide crosslinking

In the peptide cross-coupling mechanism, a lone electron pair from a nitrogen of the zero crosslinker EDC (scheme 3.11) abstracts the most acidic proton from the substrate **3.11a** forming the carboxylate anion **3.11b**. Subsequent attack by this anion on the *sp* carbon of the protonated EDC species results in the formation of an *O*-acylisourea intermediate **3.11c**. The latter rapidly undergoes a nucleophilic acyl substitution reaction releasing the urea by-product which is water soluble

and ends up in the aqueous layer during workup, resulting in **3.11d**. HOBt is an auxiliary which enhances the reaction. It is a strong nucleophile owing to the destabilizing effect of the nitrogen lone pair electrons on the negatively charged oxygen. This results in a very fast substitution reaction which reduces the concentration of *O*–acylisourea, and ultimately the extent to which racemisation can occur. The short lifespan of the activated *O*–acylisourea intermediate reduces the occurrence of *O*– to *N*–acyl migration leading to *N*–acyl ureas. HOBt's weakly acidic character prevents racemisation. A nucleophilic acyl substitution couples the benzyl hydroxylamine to the carboxylic acid moiety resulting in an amide bond formation and product **3.2b** *via* intermediate **3.11e** and regeneration of the auxiliary HOBt (scheme 3.11).



Scheme 3.11 : Mechanism of the peptide coupling reaction.

Fig. 3.6 below shows the typical ^{1}H NMR of 2-benzyl-4-((benzyloxy)amino)-4-oxobutanoic acid **3.2a–5** below which is representative of the protected reactive group series of compounds.

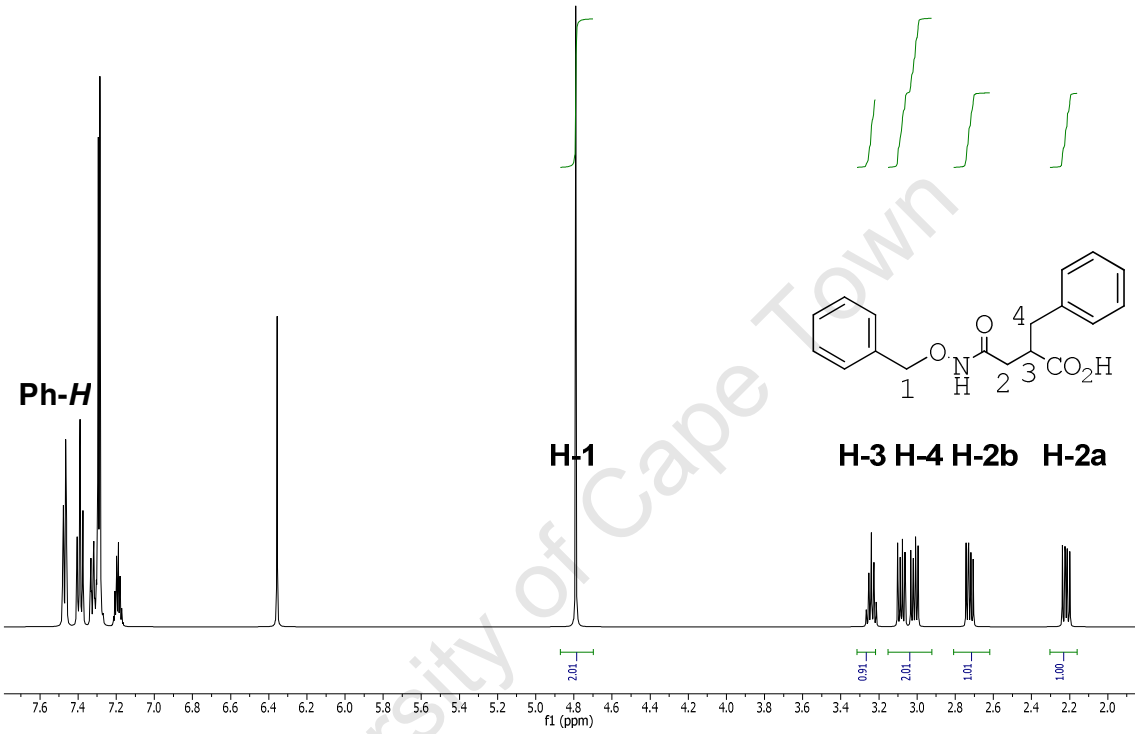


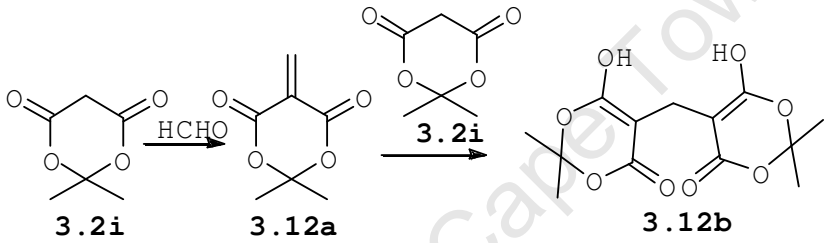
Figure 3.6: A section of a typical ^{1}HNMR of 2-benzyl-4-((benzyloxy)amino)-4-oxobutanoic acid **3.2a–5**.

Key diagnostic features for this spectrum were the doublet for C-2 proton pair at ~2.20 ppm and another in the region of 2.7 ppm, the pair of C-4 proton doublet of doublets at ~3.1 ppm and ~3.2 ppm respectively followed by a multiplet for C-3 protons at ~3.3ppm. The C-1 diastereotopic methylene proton pair were revealed as a singlet around ~4.8 ppm.

3.2.3.9 Preparation of the Succinate derivatives from anhydrides

Several methods have been reported for the preparation of monoalkylidene derivatives **3.3a** (scheme 3.3) of Meldrum's acid **3.2i** by reaction with aldehydes. The method of Cope[19] and that

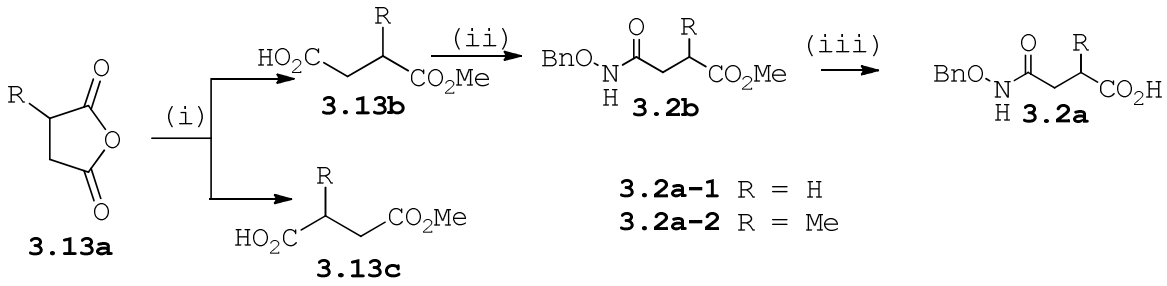
was used in this research gives fair-to-good yields with aromatic or hindered aliphatic aldehydes. However, with straight-chain aliphatic aldehydes, it is reported to give polymers. The major disadvantage of this normal base- or Lewis-acid catalysed[20] Knoevenagel condensation is that the rate of Michael addition of **3.2i** to the product **3.12a** is, unfortunately, often comparable to the rate of formation of **3.3a** itself, leading to extensive formation of the double condensation product such as 5,5'-methylenebis(6-hydroxy-2,2-dimethyl-4H-1,3-dioxin-4-one) **3.12b** (scheme 3.12 below).



Scheme 3.12: Scheme depicting formation of 5,5'-methylenebis(6-hydroxy-2,2-dimethyl-4H-1,3-dioxin-4-one)

To circumvent the problem of formation of the undesired polymer, a procedure employed by Gullotti et al[21] for the preparation of the glycine and alanine succinic acid derivatives **3.2a-1** and **3.2a-2** was utilized. This involved a simple alcoholysis of succinic anhydride **3.13a** with an equimolar amount of methanol. However, in the case of the 2-methyl succinate, as expected, the reaction proceeded with low regioselectivity and resulted in two products namely the desired **3.13b** and **3.13c** (35:65% respectively). These two products were easily isolated on silica gel with 20% EtOAc in hexane (rf 0.3 and 0.35). The desired regeoisomer **3.13b** was the minor product due to steric crowding around the carbonyl carbon adjacent to the methyl group restricting nucleophilic attack. The relative percentages of the regioisomer found were in agreement with the literature.[22] The usual coupling and

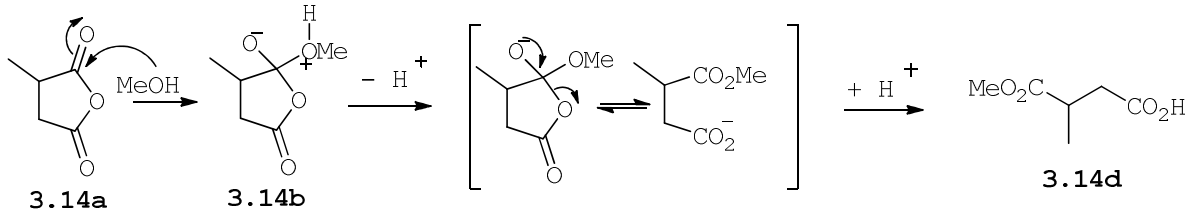
deprotection procedures were then employed resulting in **3.2b** and **3.2a** respectively.



Compound	3.2a-1	3.2a-2
R – Group	H	CH_3
Yield	90	60

Scheme 3.13: Reagents and Conditions. (i) MeOH, Reflux 110 °C, 5 h, (quant.) (ii) $BnONH_2$-HCl, EDC-HCl, DIEA, HOBt, DMF, RT, 18 h (60 – 90%).

The mechanism for the ring opening reaction of the anhydrides enroute to the succinyl derivatives is shown in scheme 3.14 below. The RCOO- moiety is not a very good leaving group (pk_a = about 5) and does not leave on its own and thus a *second-order* mechanism applies. Under these circumstances the kinetics are bimolecular (rate = k[$(RCO)_2O$][MeOH]) and the rate determining step is formation of a tetrahedral intermediate. The methanol oxygen lone pair attacks either one of the electrophilic carbonyls of **3.14a** resulting in **3.14b.** A subsequent proton transfer leads to unstable tetrahedral intermediate **3.14c.** The oxygen lone pair of the latter then attacks leading to ring opening giving way to the alkoxide species **3.14c.** Finally, a proton transfer leads to the formation of product 4-methoxy-3-methyl-4-oxobutanoic acid **3.14d.**(^{1}HNMR in Fig 3.7)



3.14a **3.14b** Unstable tetrahedral Intermediate **3.14c** **3.14d**

Scheme 3.14: Mechanism for the ring opening for the anhydride 3-methyldihydrofuran-2,5-dione **3.14a.**

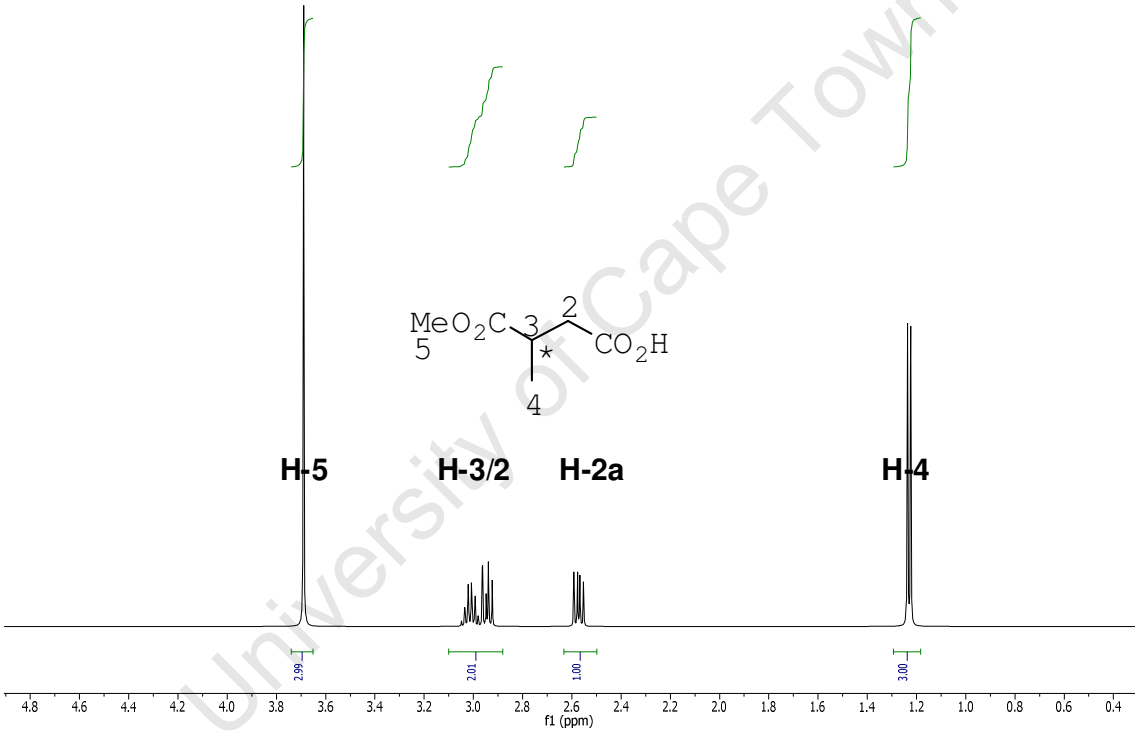
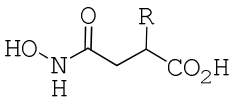


Figure 3.7: A section of a typical ^{1}H NMR in $CDCl_3$ of synthesized compound product 4-methoxy-3-methyl-4-oxobutanoic acid **3.13b-1.**

3.2.4 Preparation of hydroxamate reactive groups for screening with ACE

Before attempting the synthesis of probes for the hypertension model where ACE is one of the target enzymes, and according to our methodological prospects (Chapter 2) biological assays

were performed with the reactive group series **3.15-1** to **3.15-2**(Scheme 3.15).

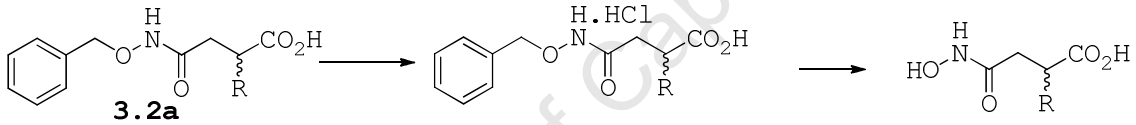


3.15

3.15-1 R = H
3.15-2 R = Me
3.15-3 R = Isopropyl
3.15-4 R = CH2Cyclohehyl
3.15-5 R = Benzyl
3.15-6 R = CH2Napthyl

Scheme 3.15: General structure of reactive group with unmasked hydroxamate.

This series were prepared by deprotection of the precursor series **3.2a** by hydrogenolysis in EtOH/Pd-C after conversion to the respective HCl salt. The general scheme is given in 3.16 below and a representative ^{1}H NMR of 2-benzyl-4-(hydroxyamino)-4-oxobutanoic acid **3.16f** is given in figure 3.8.



3.2a

41 - 60%

3.16

Compound	3.16b	3.16c	3.16d	3.16e	3.16f	3.16g
R – Group	H	CH_3	$CH_2CH(CH_3)_2$	CH_2Cyclohexyl	CH_2Ph	CH_2Naphthyl
Yield	55	60	58	60	41	60

Scheme 3.16:Preparation of hydroxamate reactive groups for biological assays. Reagents and conditions.: (i) NaOH, H_2O (ii) HCl(aq) (iii) H_2/Pd-C, MeOH, rt, 3atm.

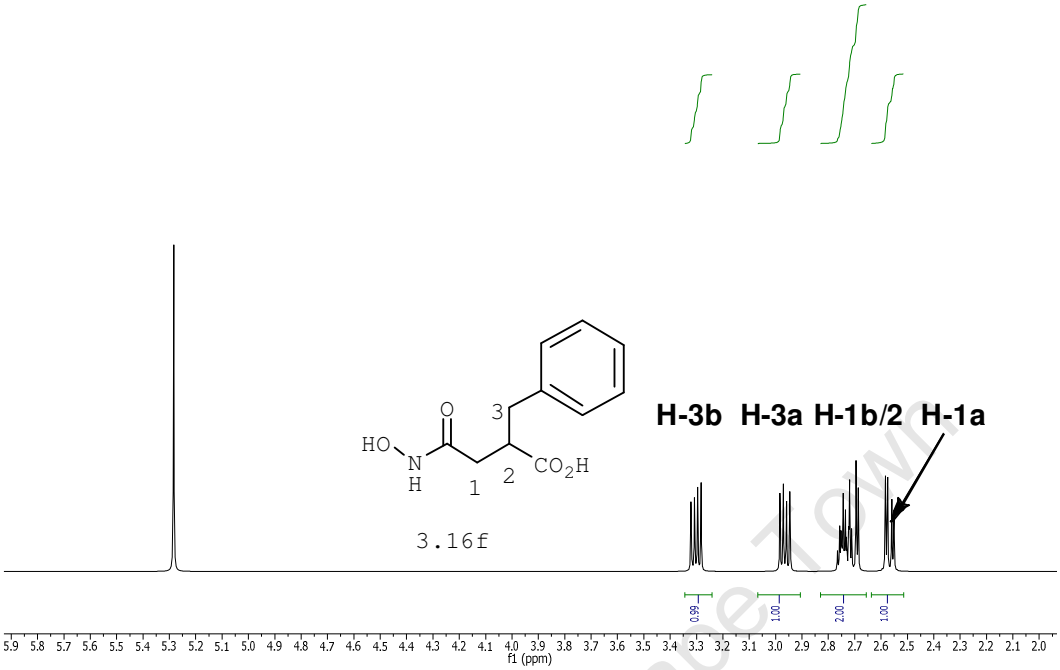
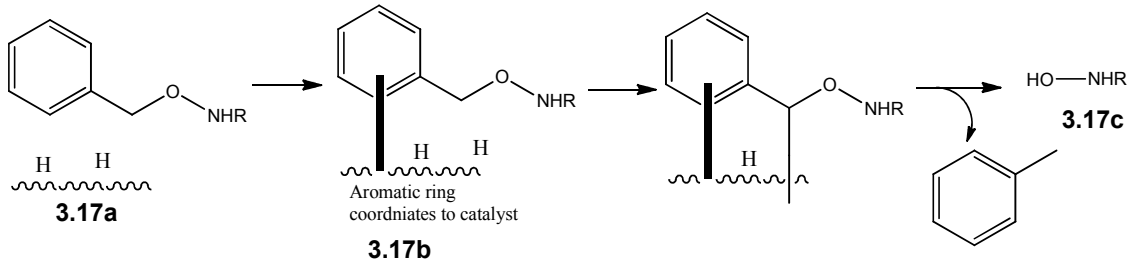


Figure 3.8:A section of the typical ^{1}HNMR of the free reactive group2-benzyl-4-(hydroxyamino)-4-oxobutanoic acid **3.16f** in $CDCl_3$.

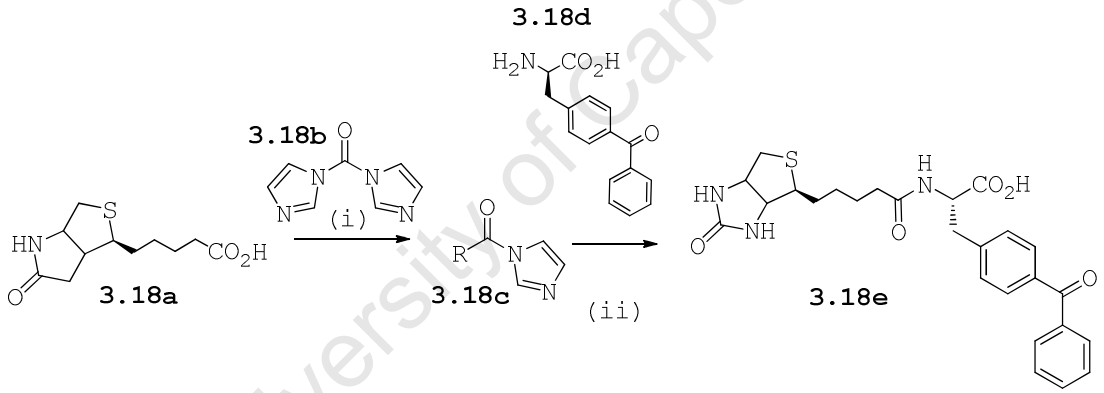
Mechanistically, the hydrogenolysis is believed to involve the cleavage of a C–O bond by addition of hydrogen in the presence of a metal catalyst. Plausibly, the hydrogen becomes adsorbed onto the metal catalyst surface. The hydroxybenzyl **3.17a** (scheme 3.17) then coordinates to the metal catalyst via the electron-rich aromatic ring. The C–O bond as depicted in **3.17b** now comes into close proximity with the palladium bound hydrogen atoms and is subsequently reduced. This leads to the free hydroxamate group of **3.17c** as well as a toluene molecule.



Scheme 3.17: Plausible mechanism of hydrogenolysis.[23]

3.2.5 Preparation of (2*S*)-3-(4-benzoylphenyl)-2-(5-((4*S*)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propanoic acid (Benzophenone-Biotin).

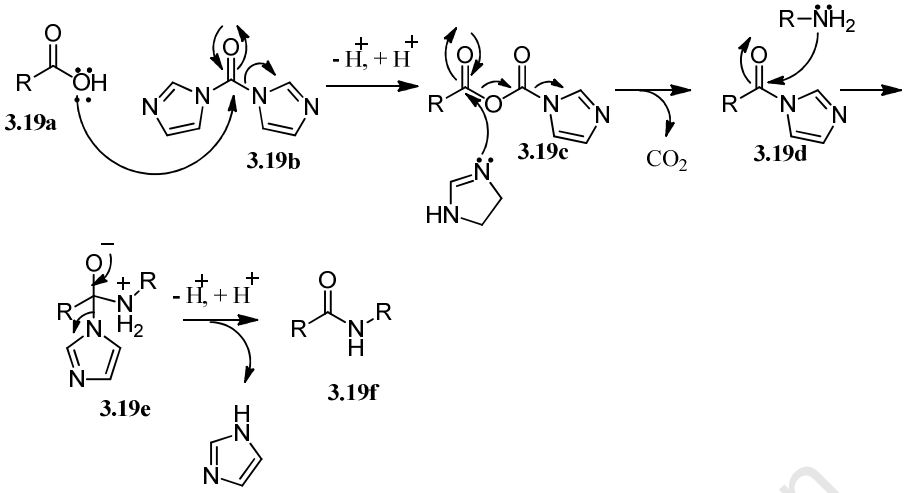
D-Biotin **3.18a** was activated with 1,1'-carbonyldiimidazole (CDI) **3.18b** in dry *N,N*-dimethylformamide (DMF) to give the imidazolide **3.18c**.(Scheme 15). Moisture and amines in DMF are deleterious as they lead to hydrolysis and aminolysis, respectively, of the reactive intermediate **3.18c**. Hence the use of dry, freshly distilled DMF is crucial to the success of the reaction. The activated biotin may be isolated and is stable under dry, inert conditions. However, without isolation, compound **3.18c** was coupled to 4-benzoyl phenylalanine **3.18d** in the same pot to give **3.18e** in 60% yield.



Scheme 3.18: Reagents and conditions. (i) DMF,DIEA, 70 °C, 30 min (ii) DMF, 70 °C-rt-50 °C, overnight, 60 %.

3.2.6 Mechanistic details of the cross-linking involving carbonyl diimidazole (CDI)

In a biomimetic fashion, the mechanism begins with the carboxylic acid **3.19a** lone pair attacking the carbonyl group of the CDI **3.19b**. A proton transfer results in the departure of one imidazole ring. The nucleophilic electron pair on this departed imidazole then returns and attacks the carboxylic carbonyl group of **3.19c**. Release of carbon dioxide and departure of an imidazole gives imidazolide **3.19d**.



Scheme 3.19: Mechanism of amide bond formation *via* the imidazolide

At the appropriate time, the amino group of the *p*-benzoyl phenylalanine then launches a nucleophilic attack on the carbonyl group of the imidazolide **3.19d** which culminates in the departure of the imidazole, a good leaving group, giving the product **3.19f.** Although theoretically, the carboxyl group of the *p*-benzoyl phenylalanine is also nucleophilic, the amino group is clearly a better nucleophile. Evidence of the successful synthesis of the intermediate (2*S*)-3-(4-benzoylphenyl)-2-(5-((4*S*)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propanoic acid **3.18e** is given by the ^{1}H NMR spectrum in figure 3.9.

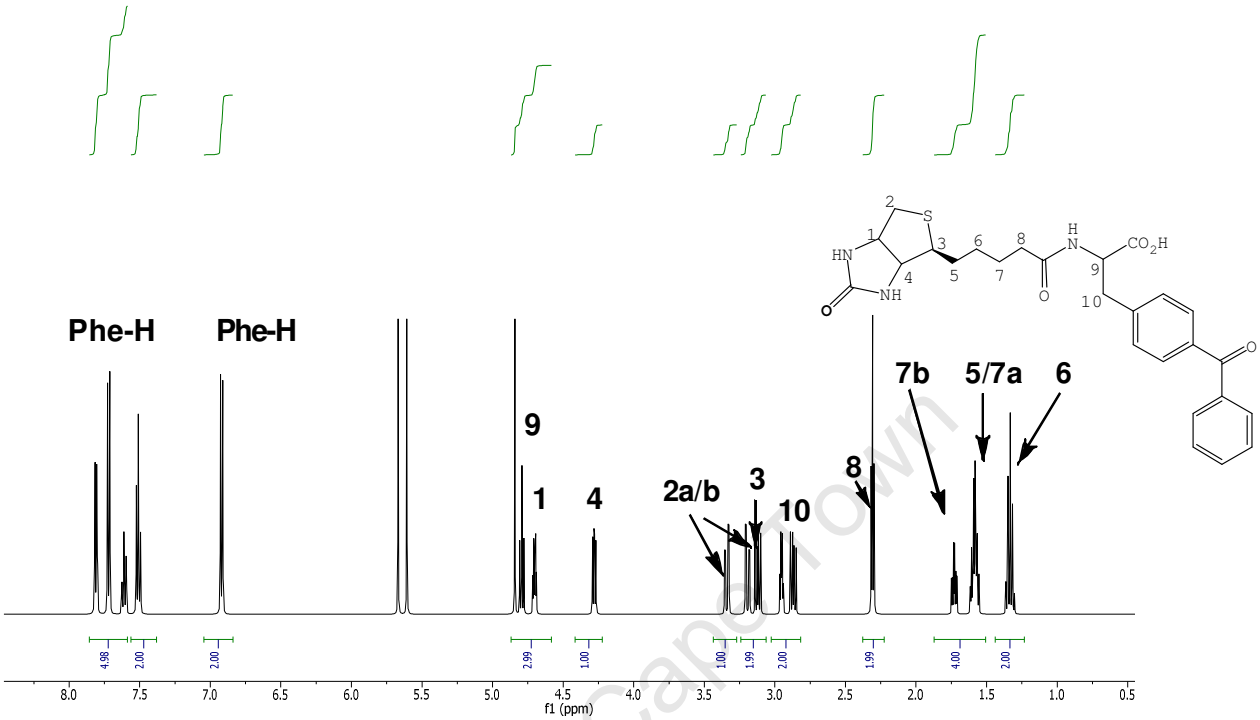


Figure 3.9: ^{1}H NMR of (2*S*)-3-(4-benzoylphenyl)-2-(5-((4*S*)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propanoic acid **3.18e** in $CDCl_3$.

Diagnostic features for this intermediate included a doublet of triplets at ~4.7 ppm and a doublets of doublets signal at ~4.3 ppm for the C-1 and C-4 protons respectively. Other key features were the doublet at ~3.1 ppm integrating for two protons assigned to the C-10 while a pair of doublets at 3.3 ppm and 3.2 ppm was assigned to each of the C-2 protons. The 6 protons on C-5,6 and 7 were assigned to the three multiplets appearing between 1.7 and 1.3 ppm.

3.2.7 Synthesis of 2,2'-(ethane-1,2-diylbis(oxy))diethanamine(PEG)$_3$ and 3,6,9,12,15,18,21-heptaoxatricosane-1,23-diamine (PEG)$_8$ Linkers

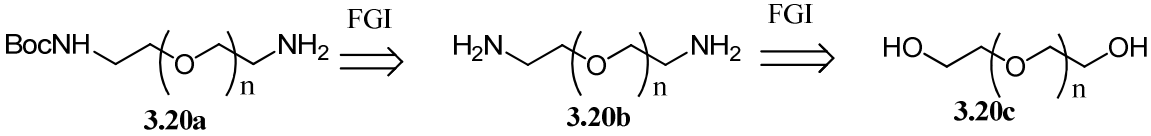
The improved pharmacological and biological properties that are associated with the covalent attachment of poly (ethylene glycol) (PEG) to therapeutically useful polypeptides is now an established fact. For this research and as described in Chapter 2, triethylene glycol and octaethylene glycol were chosen for probes to be deployed in the malaria and hypertension disease models respectively. The rationale for the choice of PEG spacers as well as the difference in chain length was discussed in more detail in chapter two of this thesis. The longer chain was desired in the hypertension model because for ACE, the active site is located at least 32 Å deep inside the protein. Therefore a probe with sufficient length spanning the active site and outside regions of the protein is ideal to ensure the photocrosslinker and biotin remain on the outside for photocrosslinking, visualization and affinity purification.

The polyethylene glycol linker is hydrophilic in nature which potentially minimizes non-specific interactions that could arise from the spacer group becoming buried in the hydrophobic pocket of proteins. Secondly, this spacer restricts hindrance between the peptide and reporter group, hence better avidin-biotin binding during affinity purification.

3.2.8 Retrosynthetic Strategy

Retrosynthetically, a functional group interconversion was planned to give way to the mono-boc protected homobifunctionalized PEG diamine adduct **3.20a** (Scheme 3.20). Commercially available polyethylene glycol **3.20c** was to be converted to the di-tosylate which would undergo S_N2

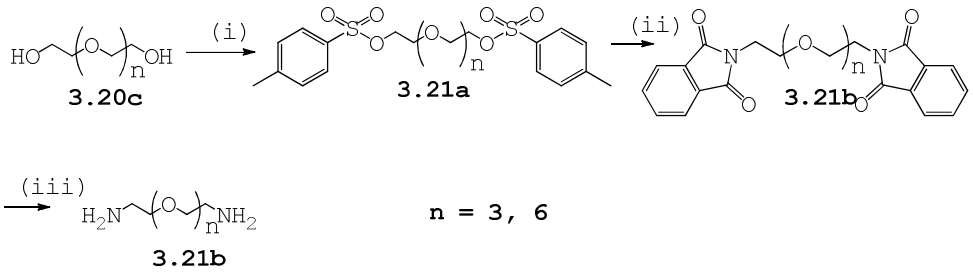
transformations because the tosylate is a good leaving group due to its ability to form a stable anion. This strategy would create an opportunity for substitution of the tosylate with an amine group which would be installed by an appropriate vehicle such as a pthalimide salt.



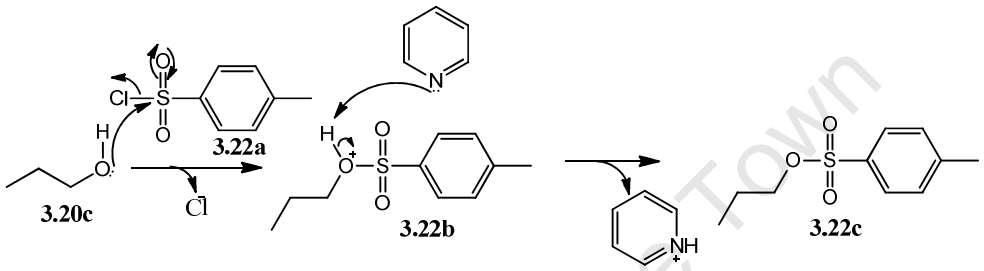
3.20a 3.20b 3.20c

Scheme 3.20: Retrosynthetic strategy for the PEG linker **3.20a**

Synthetically therefore, a modified Gabriel synthesis was used for the transformation of **3.20c** to the homobifunctionalized triethylene linker **3.20b.** This commenced with the activation of the terminal hydroxyl groups of PEG **3.20c** into the corresponding tosylate **3.21a** using tosyl chloride in the presence of pyridine. Tosylates are a good leaving group (the conjugate base of tosic acid has pKa = 2.8). This was then followed by an S_N2 reaction with potassium pthalamide in DMF forming the di-*N*-alkylphthalimide **3.21b.** The potassium phthalimide is a protected-NH_2 synthon which allows the preparation of primary amines by reaction with alkyls that have a good leaving group. After *N*-alkylation, the amine product may be unmasked by a cleavage reaction with a base or hydrazine. Hydrazinolysis was thus employed to liberate the diamine PEG linker **3.21b** in a typical Ing-Mnske procedure.[24] The product was isolated in 64% yield.



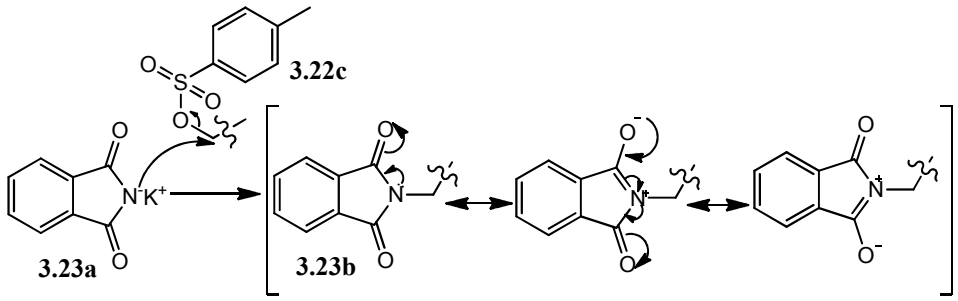
Scheme 3.21: Reagents and Conditions (i) TsCl, Pyridine, DCM 88% (ii) Potassium Pthalamide 68% (iii) Hydrazine monohydrate, 64%.



Scheme 3.22: Mechanistic details of tosylate ester formation

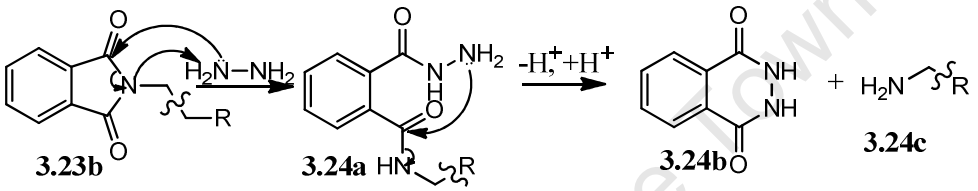
Mechanistically,the alcohol **3.20c** (Scheme 3.22) lone pair nucleophilically attacks the electrophilic centre of the tosyl chloride **3.22a** displacing the chloride giving way to the intermediate **3.22b.** The base, pyridine, is there to "mop up" the HCl by-product leading to the tosylate **3.22c.**

In the next step (Scheme 3.22) the tosylate **3.22c** reacts with the potassium phthalimide **3.23a** (Scheme 3.23 below) leading to the departure of the tosyl group and the desired *N*-alkylated resonance stabilized intermediate **3.23b.**



Scheme 3.23: Mechanistic details of nucleophilic *N*-phthalimide alkylation

To liberate the amine **3.24c** from the phthalyl residue **3.23b,** (scheme 3.24) the hydrazinolysis mechanism begins with a nucleophilic attack on the electrophilic carbonyl-carbon of **3.23b** by the hydrazine. This is followed by a proton transfer and ring opening resulting in intermediate **3.24a.** The second amine of the hydrazine also attacks on the other carbonyl carbon of **3.24a** and a proton-transfer culminates in the production of the phthalylhydrazide precipitate **3.24b** along with the primary amine **3.24c.**



Scheme 3.24: Mechanistic details of hydrazinolysis

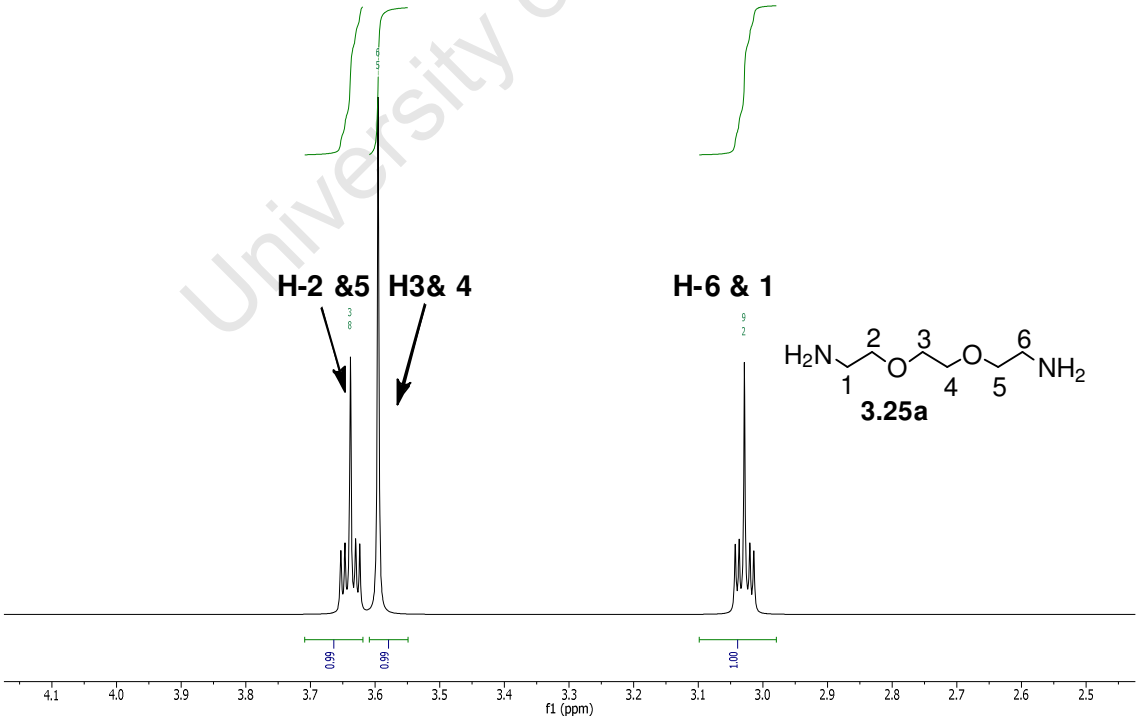
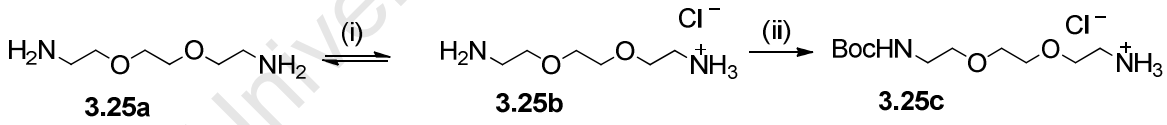


Figure 3.10: ^{1}H NMR of **3.25a** diamino product (n = 3) in $CDCl_3$

The multiplet at ~3.04 ppm was assigned to C-6 and C-1 protons, the singlet at 3.60 ppm to C-3 and C-4 protons whereas a multiplet at 3.63 ppm was for the C-2 and C-5 protons (Fig 3.10).

3.2.9 Mono-Boc-protectionof 2,2'-(ethane-1,2-diylbis(oxy))diethanamine(PEG)$_3$and 3,6,9,12,15,18,21,24-octaoxahexacosane-1,26-diamine (PEG)$_8$ Linker using a two-step fractional extraction.

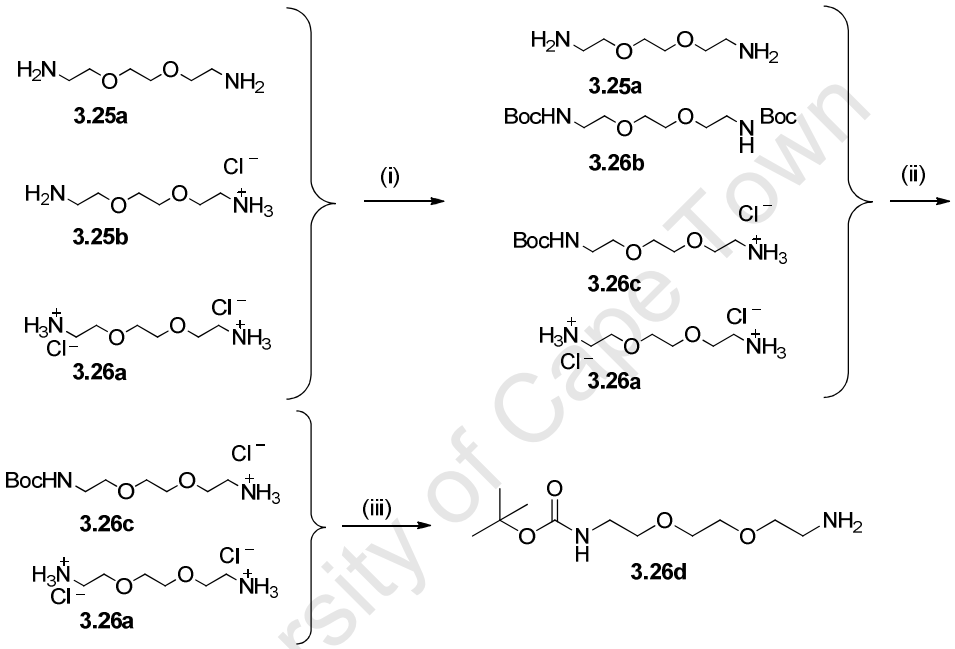
Mono-functionalization of symmetrical or unsymmetrical diamines is an essential step for the synthesis of biologically important pharmacophores and materials.[25] Mono-functionalized diamines hold functional molecules at the one end while the other end is free to be attached to other entities.[26] Among the various amine protecting groups, Boc- is the most popular, and its protected form is normally in a good physical state to be handled for the next reaction.[27] Subsequently, a reported facile route for mono-BOC protection of symmetrical diamines was employed to access the mono-protected diamine PEG linker **3.25c** from **3.25a**.[28](Scheme 3.25)



Scheme 3.25: Reagents and Conditions (i) HCl (g), MeOH aq., rt. 30 min (ii)$(BOC)_2O$

Using this route, the two amine groups are differentiated as an acid salt and a free base, which is ready for further functionalization. One molar equivalent of HCl gas was bubbled through a 50% aqueous ethanol solution of diamine **3.25a** bearing two basic amino groups. The solution was stirred for 30 min to reach full equilibrium. Mechanistically, three different species exist including starting diamine **3.25a** with

two free amines, **3.25b** as one free amine and one HCl salt as the dominant species, and **3.26a** with two amine salts in aqueous methanol (Scheme 3.26). Addition of 1 mol of $(BOC)_2O$ into this solution led to the reaction of the free amines in the species **3.25a** and **3.25b** giving way to possibly four different species **3.25a, 3.26a, 3.22b** and **3.22c** in the reaction as per the TLC. (Scheme 3.26)



Scheme 3.26: Reagents and Conditions (i) 1 mol $(BOC)_2O$, rt. 1h (ii) Et_2O/H_2O extraction (iii) Amberlyst A21 ion-exchange resin, 74%.

After the reaction was completed, an optimized two-step fractional extraction was employed. The MeOH was removed in *vacuo* and the remaining aqueous mixture was separated with diethyl ether removing the unreacted diamine as well as any di-Boc product **3.25a** and **3.26b** respectively. The freebase amines are then obtained by scavenging the crude reaction mixture with the basic Amberlyst A-21 ion-exchange resin giving crude mono-BOC product *tert*-butyl 2-(2-(2-aminoethoxy)ethoxy)ethylcarbamate **3.26d** in 74% yield. This was used in the next step without further purification because the TLC showed negligible amounts of the starting material **3.25a.**

A typical ^{1}HNMR of the mono-Boc species **3.26d** is given (fig. 3.11).

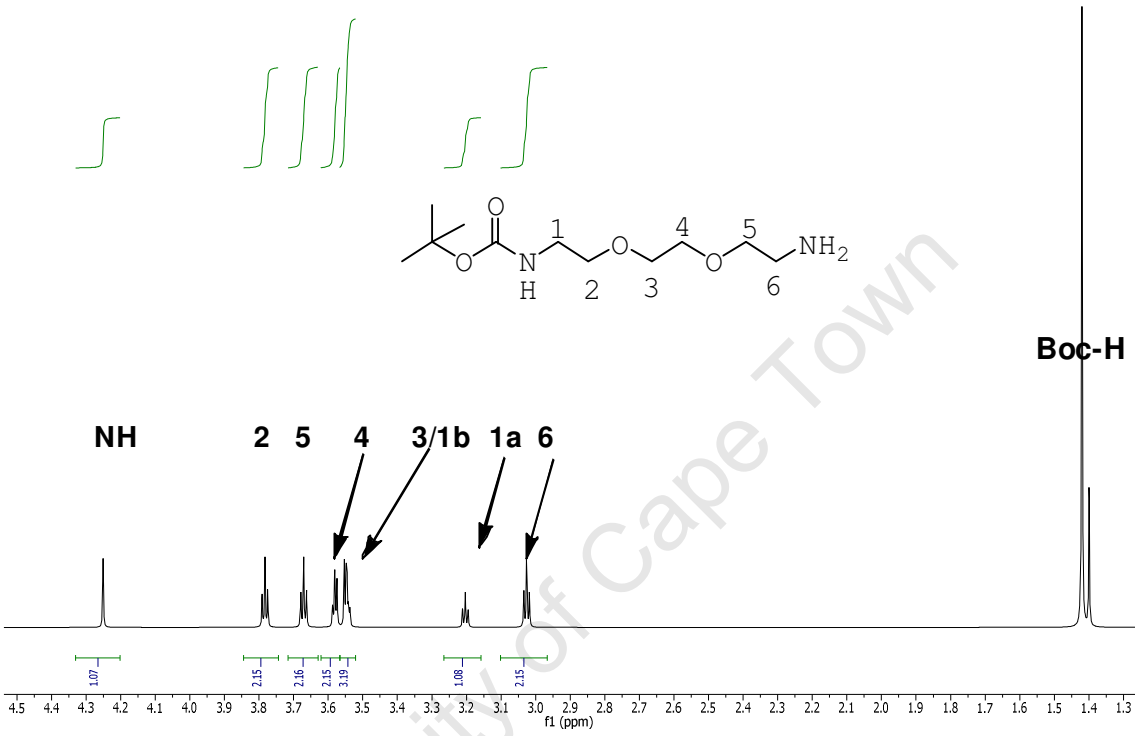


Figure 3.11: ^{1}H NMR of *t*-butyl 2-(2-(2-aminoethoxy) ethoxy) ethylcarbamate in $CDCl_3$ **3.26d**

3.2.10 Synthesis of PEG_3 hydroxamate probe

At this stage, all the sub-units or their respective appropriately protected derivatives proposed in the retrosynthetic strategy (Scheme 3.1) for the preparation of the hydroxamate probes were ready. The next step was the coupling of mono-Boc protected diamine **3.26d** to the **3.18e.** Typically, **3.26d** was crosslinked to 3-(4-benzoylphenyl)-2-(5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propanoic acid **3.18e** resulting in **3.27a.** The latter was then deprotected using TFA in DCM (1:1) which was followed by removal of the volatiles *in vacuo*. However,

although the deprotection appeared clean by TLC, the product yields after aqueous extraction to remove the residual TFA were poor. Aqueous workup was thus avoided by using a literature procedure[29] where the crude residue is dissolved in DCM and treated with amberlystA-21® ion-exchange resin giving freebase amine **3.27b**. The latter was used in the next step without further purification. A representative ^{1}HNMR of this intermediate is given in figure 3.12.

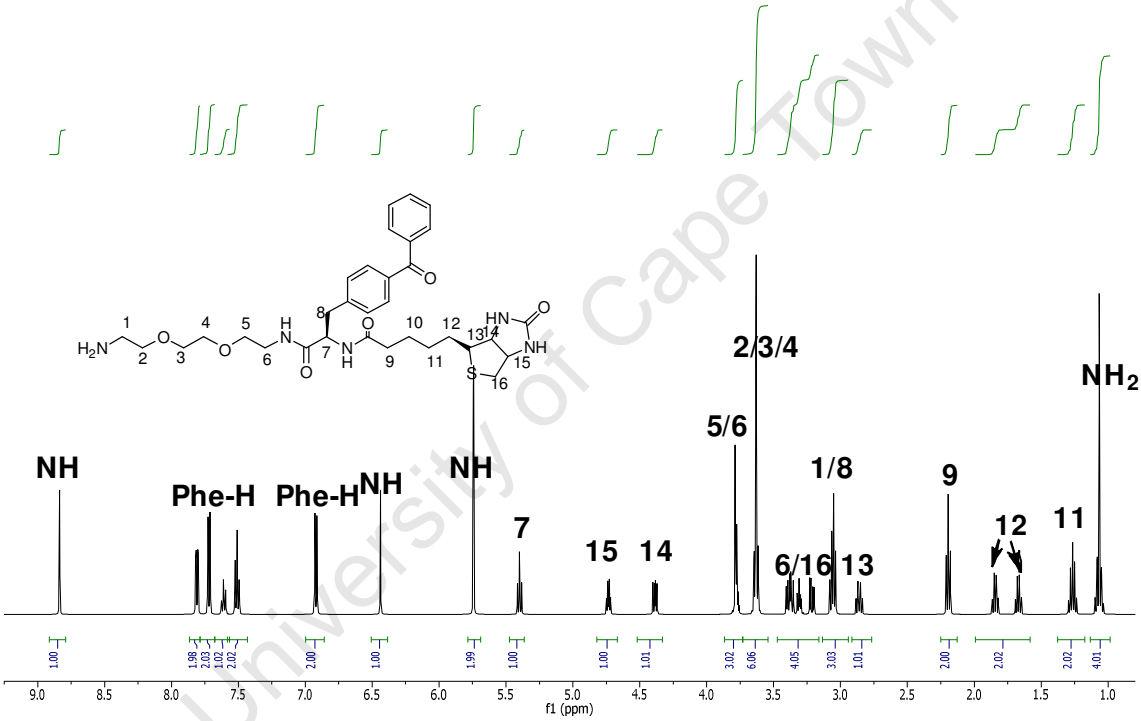
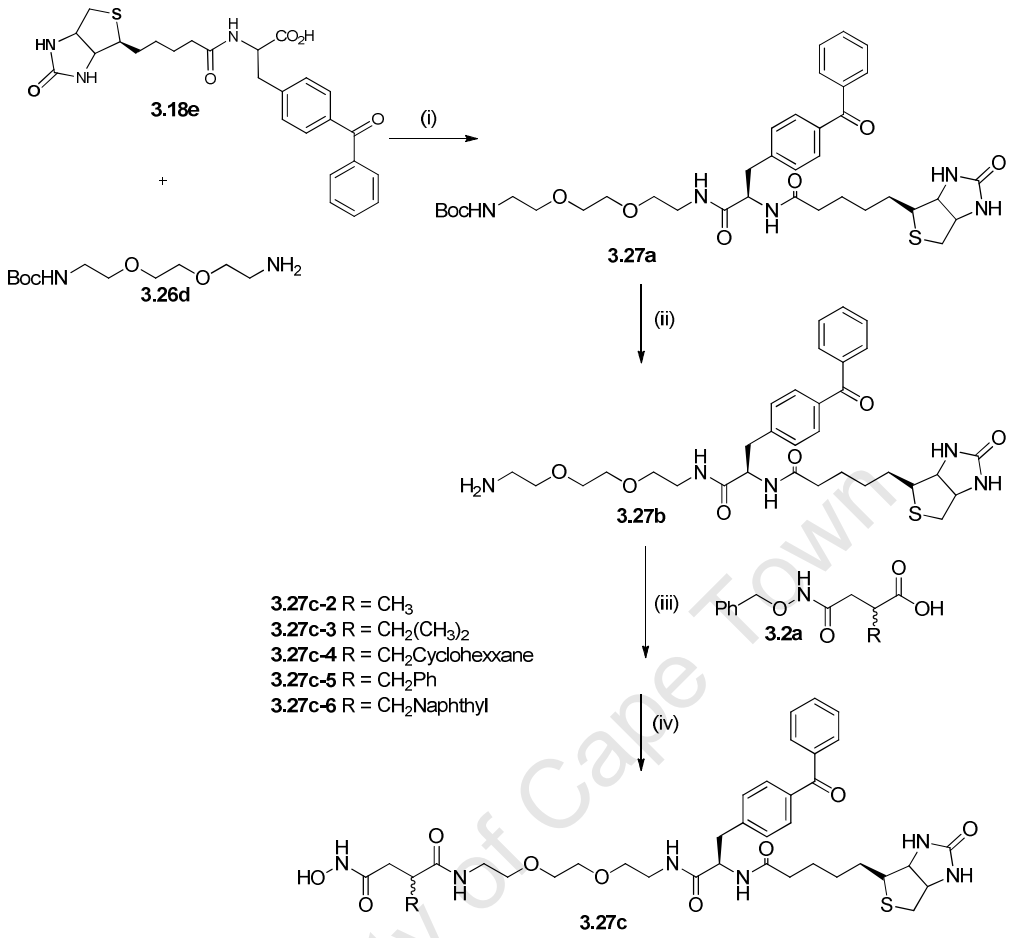


Figure 3.12: A section of ^{1}H NMR for **3.27b** in $CDCl_3$

The next step was a peptide coupling of **3.27b** to **3.2a** performed under standard conditions followed by unmasking of the hydroxamate "war head" in the final step by hydrogenolysis using H_2 in the presence of Pd-C in ethanol. This resulted in the probe series **3.27c-2, 3.27c-3, 3.27c-4, 3.27c-5** and **3.27c-6**(Scheme 3.27).



Scheme 3.27: Reagents and Conditions (i) EDC-HCl, DIEA, HOBt, DCM, rt. Overnight. 47% (ii) TFA, DCM, rt. 3h, quant. (iii) EDC, DMAP, HOBt, DCM, rt. Overnight, (iv) Pd-C/H_2, 3 atm, rt.

As expected, it proved difficult to fully characterize the probes **3.27c–2,3, 4, 5** and **6** using ^{1}HNMR due to the number of protons and overlapping of peaks. A typical ^{1}H NMR is given in Fig. 3.14a. The respective M^+ peak was observed at 795.53amu (Fig. 3.14) and thus successful synthesis was concluded. After deprotection with H_2/Pd-C in methanol, the final product sample was immediately sealed in a dry sample vial to avoid chelation of the free hydroxamate to any metals.

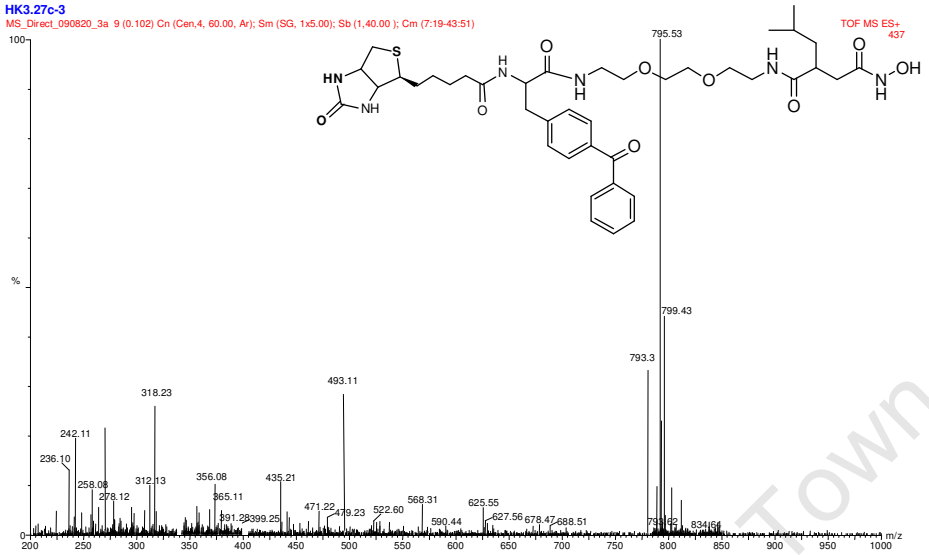


Figure 3.13: Electrospray ionization MS M+ peak for probe structure **3.27c–3.**

Figure 3.14: A section of ^{1}HNMR for **3.27c–3** in $CDCl_3$.

3.2.11 Synthesis of PEG_8 hydroxamate probe

According to Cravatt et al.,[30](Fig. 3.15) probes may be employed in selective inhibitor, substrate metabolic or enzyme discovery in order to achieve functional characterization. With the failure of our proposed hydroxamates (**3.16b** to **3.16g**) to inhibit ACE to satisfactory levels (discussed in detail in Chapter 5), we envisaged that replacement of the reactive group with known inhibitors Captopril and Enalaprilat would be a viable option with regards to studying their specificity. Therefore, their application in the hypertension disease model proteome would give an insight into whether they are specific to ACE or inhibit other related or unrelated enzymes as well which could explain some of the observed contra-indications.

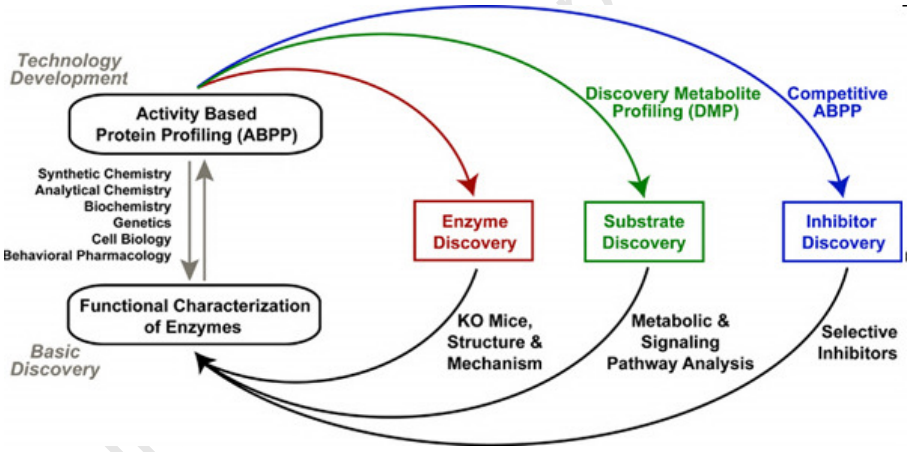
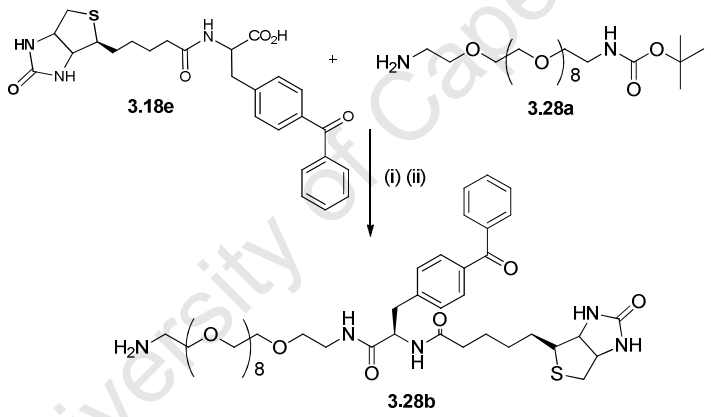


Figure 3.15 : Schematic summary of the importance of ABPP in drug discovery[30]

Subsequently, an Enalaprilat as well as Captopril based probe were synthesized.

3.2.12 Synthesis of N-((R)-1-amino-27-(4-benzoylphenyl)-25-oxo-3,6,9,12,15,18,21-heptaoxa-24-azaheptacosan-26-yl)-5-((4R)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide 3.24b

t-butyl(23-amino-3,6,9,12,15,18,21-heptaoxatricosyl)carbamate (*N*-Boc PEG_8) **3.28a** prepared in a similar fashion to **3.26d** (Scheme 3.26) was linked to the Biotin-Benzophenylalanine **3.18e** substructure under standard peptide coupling conditions using EDC-HCl and HOBt. Subsequent deprotection with 35% TFA in DCM, followed by treatment with Amberlyst A-21 ion exchange resin unmasked the free-base amine to give intermediate **3.28b.**

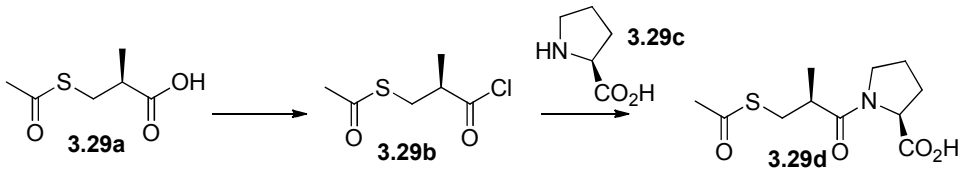


Scheme 3.28 : Reagents and Conditions (i) EDC-HCl, DIEA, HOBt, TEA, DCM, rt. Overnight. 47% (ii) TFA, DCM, rt. 3h, quant.

3.2.13 Synthesis of (S)-1-((S)-3-(acetylthio)-2-methylpropanoyl)pyrrolidine-2-carboxylic acid 3.29d

The strategy for the preparation of S)-1-((S)-3-(acetylthio)-2-methylpropanoyl)pyrrolidine-2-carboxylic acid (Acetylthio-Captopril) **3.29d** was to leave the prolyl carboxylic acid free for coupling to **3.28b.** Subsequently, the commercially available starting material(S)-3-(acetylthio)-2-methylpropanoic acid **3.29a** was activated to form an acyl

chloride **3.29b** using $SOCl_2$ in DCM. Nucleophilic acylation at the carbonyl of **3.29b** by the commercially available amino acid proline **3.29c** then gave the adduct **3.29d** (Scheme 3.29).

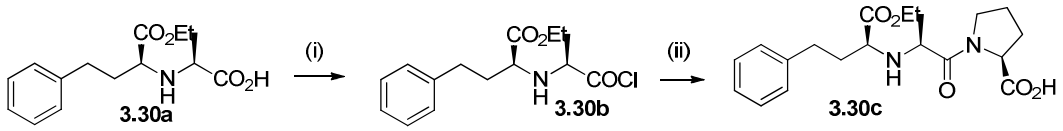


Scheme 3.29.:Reagents and Conditions: (i) $SOCl_2$, DCM, rt, on (ii) Proline, $NaHCO_3$, H_2O : Dioxane (1:1), 3h, 60%.

Mechanistically, the carboxylic acid activation is identical to that in methoxyester protection given in scheme 3.9 above. The only difference is that the attacking nucleophile is the prolyl amine.

3.2.14 Synthesis of (S)-1-((S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoyl)pyrrolidine-2-carboxylic acid (Enalapril) 3.30c

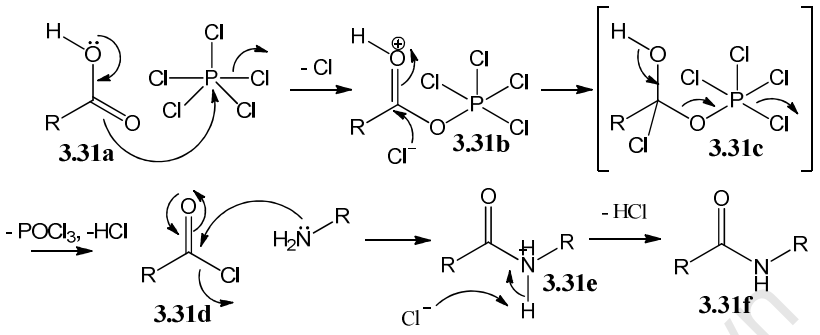
In a similar fashion, but utilizing PCl_5 commercially available **3.30a** was activated to the acid chloride **3.30b**. (*S*)-1-((S)-2-(((*S*)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoyl)pyrrolidine-2-carboxylic acid (Enalapril) **3.30c** was then synthesized by acylation of **3.30b** with proline. (Scheme 3.30) The pure product was isolated on silica gel.



Scheme 3.30: Reagents and conditions: (i) PCl_5, DCM, rt, on (ii) Proline, $NaHCO_3$, H_2O :Dioxane (1:1), 3h, 54%.

The mechanism of chloride activation by PCl_5 proceeds *via* a substitutive chlorination replacing the –OH of the carboxylate with the chloride. As depicted in scheme 3.31, an oxygen lone pair from acid **3.31a** attacks the phosphorous of PCl_5 leading to

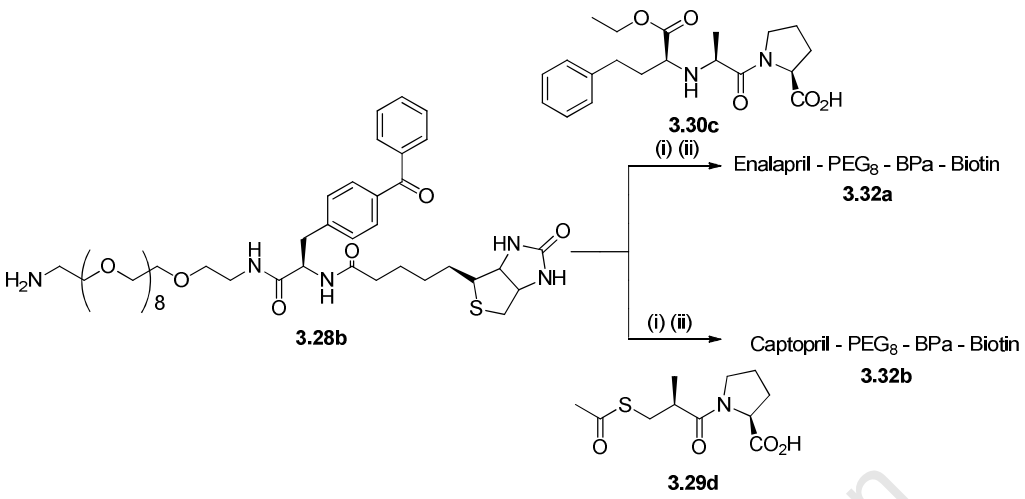
the dissociation of one Cl^-. The latter then attacks the electrophilic carbonyl carbon of intermediate **3.31b** leading to **3.31c.**



Scheme 3.31: Mechanism of PCl_5 activation and subsequent acylation

The next step from **3.31c** to **3.31d** is driven primarily by the formation of $POCl_3$ and also leads to the loss of HCl. The very reactive acid chloride intermediate **3.31d** is sufficiently electrophilic and is attacked by the amino-NH_2 lone pair to ultimately form the desired acylated product **3.31f** with the loss of HCl.

The ^{1}HNMR data of enalapril and acylthio-captopril were in agreement with the literature.[31,32] These were then coupled to intermediate **3.28b** respectively, and thereafter deprotected by saponification using NaOH to give target probes **3.32a** and **3.32b.** (Scheme 3.32)



Scheme 3.32: Synthesis of probes 3.23a and 3.23b. Reagents and conditions: EDC-HCl, DIEA, HOBt, DCM, rt. 72 hrs. 27% (ii) 2M NaOH aq. 3h, 88%.

Peak assignment for the ^{1}HNMR of probe **3.32a** and **3.32b** also proved to be a challenge due to overlapping and multiple peaks. Confirmation of the successful synthesis was observed by MS. Accordingly, molecular ion peaks for the compounds were observed at 1044.2 and 1174.1 amu (Fig.3.15a & b) respectively.

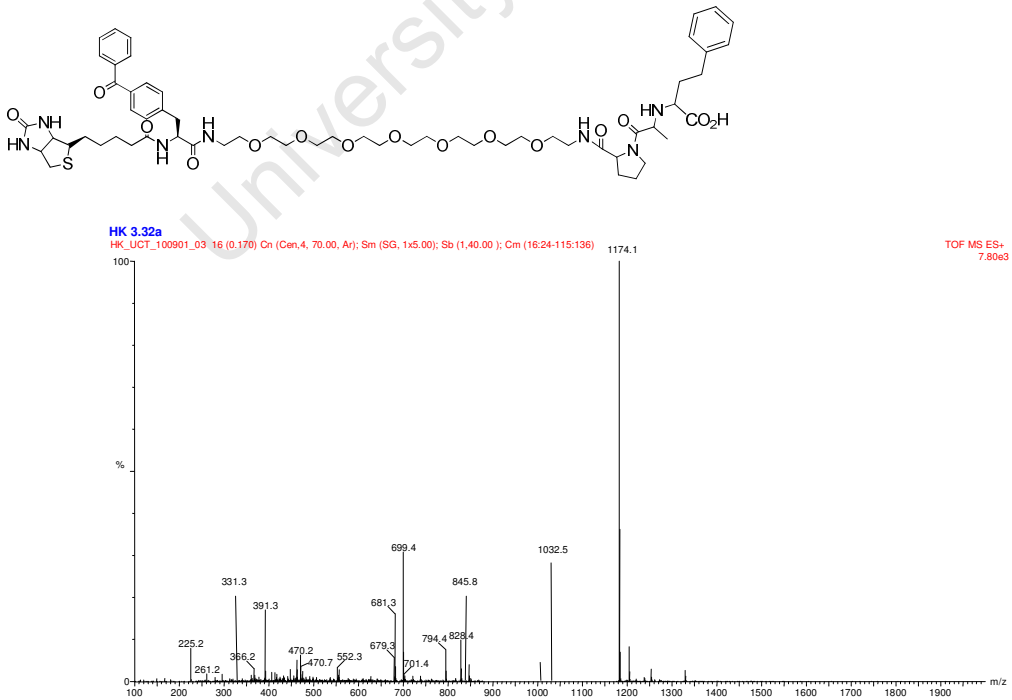


Figure 3.16a: Electrospray ionization MS M^{+} of enalaprilat probe **3.32a.**

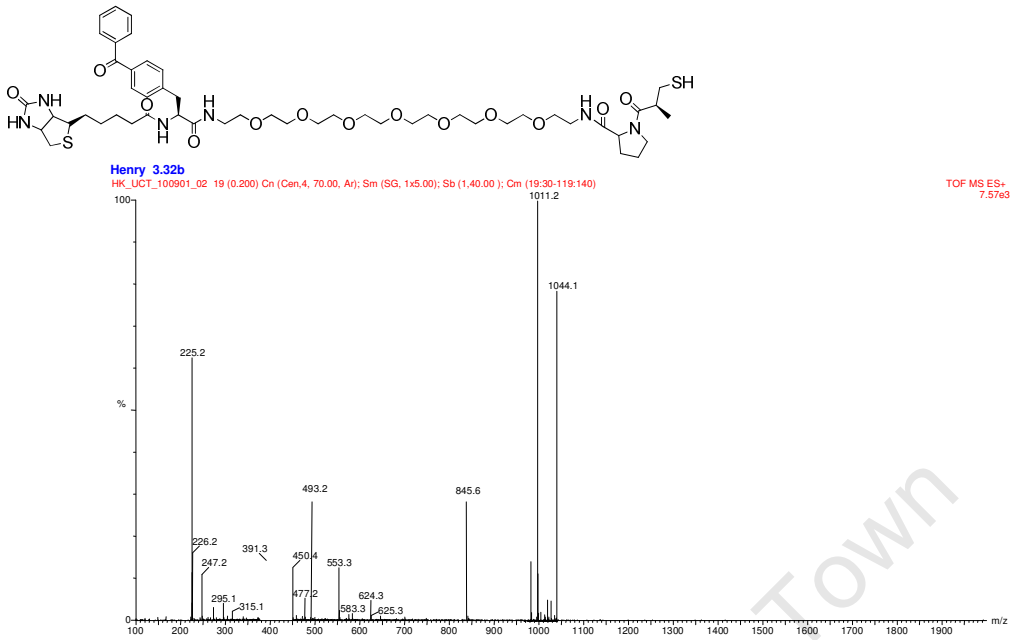


Figure 3.17b: Electrospray ionization MS M^+ of enalaprilat probe **3.32b.**

3.2.15 Summary

Following on from Chapter 2, this chapter has given a detailed retrosynthetic strategy followed by a synthetic discussion of the synthesis of the targeted probes structures. As stated in Chapters 1 and 2, the hydroxamate probes were synthesized for future application in malaria disease model whilst the sulfhydril (Captopril) and carboxylate (Enalapril) based ones were assayed in the hypertension disease model the results of which are described in Chapter 5. The latter will also present a discussion of the biological results of all the other compounds synthesized in this research.

References

[1] Bihovsky, R., Levinson, B. L., Loewi, R.C., Erhardt, P.W., Polokoff, M. A. *Journal of Medicinal Chemistry* **1995**; 38: 2119–2129.

[2] McNab, H. Meldrum's acid. *Chemical Society Reviews* **1978**;7: 345–358.

[3] Daqing, S.;Yucheng, W.;Zaisheng, L.;Guiyuan, D., *Synthetic Communications* **2000**; 30:713–726.

[4] Margaretha, P.; Polansky, O.E. *Tetrahedron Letters* **1969**; 4983– 4986.

[5] Mane, R.B.; Desai, D.G., *Chemical Industry***1982**, 809.

[6] Dhuru, S.P.; Mohe, N.U.; Salunkhe, M.M., *Synthetic Communications***2001**; 31:3653–3657.

[7] Chan, C.; Huang, X.,*Synthesis***1982**; 452–454.

[8] Wright, A.D.; Haslego, M.L.; Smith, F.Y., *Tetrahedron Letters***1979**; 2325–2326.

[9] Huang, X.; Xie, L., *Synthetic Communications***1986** ; 16:1701–1707.

[10] Huang, X.; Chan, C.; Wu, Q. *TetrahedronLetters***1982** ; 23:75–76.

[11] Toth, G.; Kover, K.E., *SyntheticCommunications***1995** ; 25 :3067–3074.

[12] Nutaitis, C.F.; Schultz, R.A.; Obera, J.; Smith, F.X.,

*Journal of OrganicChemistry***1980** ; 45 :4606-4608.

13 Hrubowchak, D.M.; Smith, F.X., *TetrahedronLetters***1983** ; 24:4951-4954.

14 Desai U.V.; Pore, D.M.; Mane, R. B.; Solabannavar, S. B.; Wadgaonkar, P. P., *Synthetic Communications*,**2004**, 34, 25-32.

15 Djerassi, C.; Engle, R.R. *Journal of the American Chemical Society***1953**;*75:*3838-3840.

16 Martin, V.S.;Palazón, J.M.; Rodríguez, C.M. *Encyclopedia of Reagents for Organic Synthesis* (Ed.: L. A. Paquette), Wiley,New York, **1995**, 6, 4415-4422.

17 Rossiter, B. E.; Katsuki, T.; Sharpless, K. B., *Journal of the American Chemical Society* **1981**; 103:464.

18 U.-S. Bäumer, U. S.; Schäfer, H. J.; *ElectrochimicaActa* **2003**;48: 489-495.

19 Cope, *A.* C., *Journal of the American Chemical Society,* **1941**, 63, 3455.

20 Lehnert, W., *Tetrahedron Letters,***1970**: 4723.

21 Gullotti, M.; Santagostini, L.; Pagliarin, R.; Granata, A.; Casella, L., *Journal of Molecular Catalysis A: Chemical* **2005**; 235: 271-284.

22 Cheng, Y., Deng, L. *Journal of the American Chemical Society***2001**;123: 11302-11303

23 Clayden, J., Greeves, N., Warren, S., Wothers, P. *Organic Chemistry (1st ed.)***2000**, Oxford University Press, USA

[24] Osby, J.O., Martin, M. G. Ganem, B., *Tetrahedron Letters* **1984**, 25,2093.

[25] Bolognesi, M. L.; Minarini, A.; Budriesi, R.; Cacciaguerra, S.; Chiarini, A.; Spampinato, S.; Tumiatti, V.; Melchiorre, C. *Journal of MedicinalChemistry.* **1998**, 41, 4150-4160.

[26] Kohn, M.; Wacker, R.; Peters, C.; Schroder, H.; Soulere, L.; Breinbauer, R.; Niemeyer, C. M.; Waldmann, H. *Angewantde ChemieInternational Edition***2003**; 42: 5830-5834.

[27] Greene, T. W.; Wuts, P. G. M. *Protective Groups in OrganicSynthesis, 3rd ed.*; John Wiley & Sons: New York, **1999**; 494-653.

[28] Lee, D. W.; Ha, H-J.; Lee, W. K. *Synthetic Communications* **2007**;37: 737-742.

[29] Srinivasan, N.;Yurek-George, A.;Ganesan, A. *Molecular Diversity* **2005**; 9: 291-293.

[30] http://www.scripps.edu/chemphys/cravatt/research.html

[31] Casy, A.F.; Dewar, G.H., *Journal of Pharmaceutical and Biomedical Analysis* **1994**; 12:855-861.

[32] Zoppi, A.; Linares, M.; Longhi, M., *Journal of Pharmaceutical and Biomedical Analysis***2005**; 37: 627-630.

CHAPTER 4

Potential Novel antihypertension agents with an Attenuated Zinc Binding Group

4.0 Introduction

The introduction in chapter one of this thesis described in brief the importance of the enzyme ACE in hypertension pathophysiology. This chapter will elaborate the full details as well as describe a novel approach towards potential new antihypertensive agents.

Playing a key role in the Renin Angiotensin system (RAS) cascade in the control of hypertension, ACE has been a therapeutic drug target. Treatment of hypertension *via* inhibition of this enzyme has been achieved quite successfully since the 1950's. However, this treatment has been associated with a panel of unpleasant side effects. These inhibitors were designed before the discovery of ACE as a two-domain enzyme, N- and C-domain and that each of these possesses an independent active site with different activation requirements. Current inhibitors inhibit both domains completely halting other important homeostatic functions of the enzyme. Consequently, an accumulation of substrates occurs leading to an array of side effects. Research has now established that only the C-domain of ACE is responsible for hypertension control. Inhibition of only the C-domain allowing the N-domain to continue its activity could possibly improve the side effect profiles and

provides a platform for research towards the next generation of antihypertensive ACE inhibitor drugs.

In the past few years, research towards C-domain specific inhibitors has focused on changes to the peptidomimetic backbone of existing as well as potential inhibitors whilst maintaining a strong Zinc-binding group (ZBG) with very little success. This part of the thesis, however, seeks to explore, using a small series of compounds, the effect of attenuating the characteristic ACE inhibitor ZBG. The hypothesis is that this effect will allow the more subtle and more specific interactions of the peptidomimetic backbone of potential inhibitors with the enzyme active site and may lead to the next generation of ACE inhibitors.

4.1 Hypertension

Hypertension, or elevated arterial blood pressure, is a condition that affects approximately 26% of the world's adult population. This figure, with its major socio-economic implications, is expected to increase to 29% by the year 2025, with at least two thirds of these patients having been shown to originate from economically developing countries.[1] This precipitates the urgent need to effectively treat such a condition with particular attention on the needs for treatment in sub-Saharan Africa. The treatment of hypertension also results in control of downstream cardiovascular conditions such as myocardial infarction and stroke.

4.2 The Renin Angiotensin System

The Renin-angiotensin system (Figure 4.1) is a complex, highly regulated pathway that is integral in the regulation of blood volume, electrolyte balance, and arterial blood pressure. This system consists of several enzymes that are responsible for the conversion of substrates into active peptides that play an important role in regulating blood pressure.

Renin, secreted from the kidney, is responsible for the cleavage of angiotensinogen to produce angiotensinI. Angiotensin I (Ang I), a decapeptide, is subsequently converted into angiotensin II (Ang II), an active octapeptide, by Angiotensin-converting Enzyme (ACE). [2] Binding of Ang II to appropriate receptors causes systemic vasoconstriction[3]that increases total peripheral resistance through a variety of mechanisms. Furthermore, it results in increased aldosterone secretion and water retention in the kidney.[2] Yet another substrate of ACE is the potent vasodilator bradykinin (BK), an intermediate in the Kallikrein-Kinin System. ACE cleaves the C-teminal dipeptide unit Phe-Arg inactivating BK resulting in decreased vasodilation. The plasma kallikrein-kinin system counterbalances the RAS.[4]

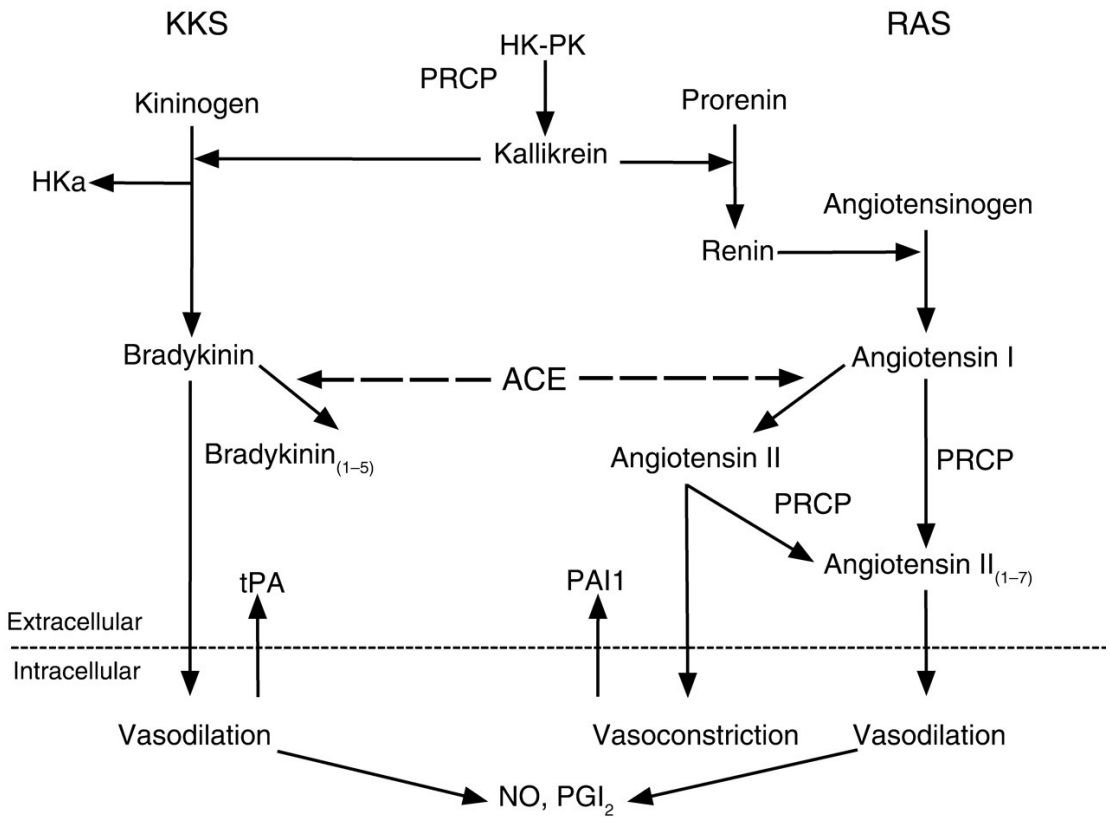


Figure 4.1: ***The interaction between the plasma KKS and RAS. Angiotensin-converting enzyme (ACE) converts inactive angiotensin I to the vasoconstrictor angiotensin II. At the same time ACE degrades active vasodilator bradykinin into bradykinin*** $_{(1-7)}$.[4]

4.3 Angiotensin-converting Enzyme (ACE) and its role in the RAS and KKS

ACE (EC 3.4.15.1) is a zinc dipeptidyl carboxypeptidase and, as described above, is responsible for the crucial conversion of Ang I to Ang II in the RAS. More specifically, this enzyme is responsible for the cleavage of a dipeptide C-terminal unit (HL) from the decapeptide substrate (DRVYIHPFHL) into the octapeptide form. From this stage, the cleaved product is subjected to further conversions of the RAS including ACE2 and aminopeptidases.[5] The various products of catalysis are capable of binding to their respective receptors and regulate cardiac function.

ACE is involved in another critical conversion, although not part of the RAS, whereby bradykinin (BK, RPPGFSPFR), a peptide part of the Kallikrein-Kinin System (KKS) with vasodilator characteristics, into its inactive form, bradykinin$_{1-7}$ (BK$_{1-7}$). (Figure 4.1) This serves as a complementary control for the increase in arterial blood pressure, whereby the peptides responsible for vasoconstriction (Ang II) are introduced and the peptides with vasodilation properties are degraded or inactivated.

Apart from Ang I and BK, ACE is also involved with other substrates such as peptide *N*-acetyl-seryl-aspartyl-lysyl-proline (AcSDKP), a peptide known for its haemoregulatory roles. Even in this case, ACE functions as a dipeptidyl carboxypeptidase, releasing the KP fragment in the plasma[6] . This together with others indicates that ACE plays a role in several other physiological systems apart from the RAS.

4.4 ACE characteristics

ACE has two isoforms, depending on tissue localisation. Somatic ACE (sACE) is ubiquitously expressed but predominantly on the plasma membrane of endothelial and epithelial cells.5 Molecular cloning revealed interesting characteristics of this enzyme. It has been shown that the cDNA consists of 1,306 residues, including the signal peptide, with a transmembrane domain near the C-terminal of sACE. The sACE gene consists of 26 exons which are translated 1 through to 26 excluding exon 13. Further, sACE was shown to have two domains, namely N- and C-

domains, that have high sequence similarities and both containing active sites.[7]

tACE is an isoform that is expressed exclusively in male germinal cells and is comparable to the C-domain of sACE.[8] The shortened protein results from the presence of a tissue specific promoter present in intron 12 of the sACE gene giving rise to a 2 435 nucleotide (732 amino acids) fragment. With the exception of a unique 36 amino acid residue region on the N-terminal, tACE is identical to the C-domain region of sACE.[9]

sACE consists of two catalytic domains, as described by the characteristic HEMGH zinc binding site. tACE is encoded by the same gene but transcribed with a tissue specific promoter present in intron 12. Hence tACE can be viewed as truncated relative to sACE. Both catalytic sites have a conserved Zn^{2+} but differs both in substrate specificities and in chloride activation profiles.[10,11] It has also been shown that the C-domain is more dependent on chloride ion for activation and subsequent substrate conversion, compared to the N-domain and both domains show similar kinetics for bradykinin breakdown.[10]

4.5 Three dimensional structure of ACE

4.5.1 Testis ACE

In 2003, the three dimensional crystal structure of tACE mutant (tACEΔ36NJ) co-crystallized with Lisinopril was

published.[12] The basic structural observations included the presence of a large abundance of α-helices and a 30Å central groove that divides the molecule approximately in half into designated sub-domains I and II (Figure 4.2). N-terminal helices α1, α2 and α3 form a cap-like structure that appears to partially cover the active site. This observation is interesting since it suggests that, in order for ACE to accommodate either substrate or inhibitor, a conformational change is required.5 The catalytic zinc ion is located in the active site, coordinated by residues His383 and His387 as predicted by the characteristic zinc binding motif.[12]

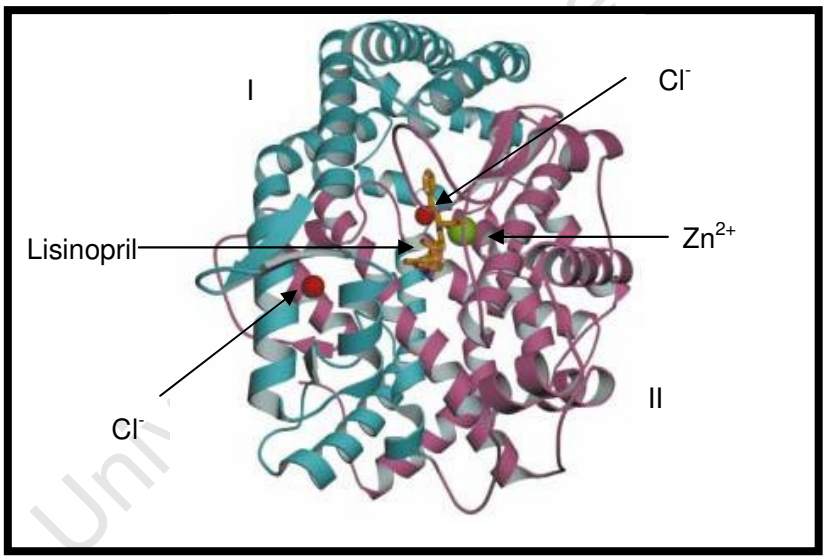


Figure 4.2: A structural view of testis ACE (tACE). The sub-domains (I and II) are given in green and purple respectively. Chloride ions are indicated and shown in red. Lisinopril and the catalytic zinc ion are shown in yellow and green respectively. The structure assumes an ellipsoid shape overall. [12]

In 2007, the three dimensional crystal structure of tACE mutant (tACEΔ36NJ) co-crytallized with RXPA380 (the only known C-domain specific inhibitor) was also solved. This

phosphinic peptide is a long molecule whose residues were revealed occupying S2, S1, S1' and S2'.[13] (Figure 4.3)

Figure 4.3: Cartoon of tACE (pink) showing RXPA380 (yellow)bound in the active site. Zinc and chloride ions are shown in grean and yellow respectively.

RXPA380 is reported to bind in the active site in an elongated conformation with the phosphinyl oxygens coordinating with the zinc. (Figure 4.4)

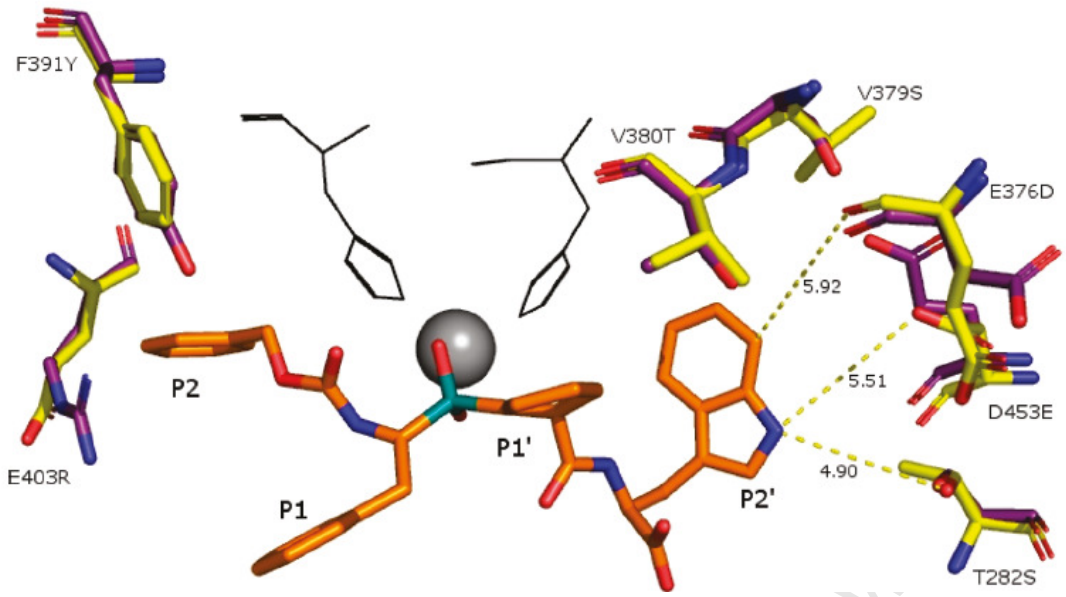


Figure 4.4: Stick representation of RXPA380 (orange) within the active site of tACE.[14]

The crystal structure revealed that despite the differences in zinc coordinating moieties in the different ACE inhibitors (carboxyl and sulfhydril) the inhibitor coordinating atoms are similarly positioned. The phosphinyl oxygens overlap almost precisely the same as those of carboxylate groups of enalaprilat and lisinopril with a distance of between 2–2.5Å away from the zinc atom. Furthermore, RXPA380 follows a conformation similar to captopril on the C-terminal side of the ZBG and four of the eight hydrogen bonding interactions in the active site being the same as both enalaprilat and captopril. Significant among these are the C-terminal carboxylate and the adjacent carbonyl oxygen moiety interaction with Lys511 and Tyr 520 and with His353 and His513 respectively. Chemical structures of RXPA380, enalaprilat and captopril with residue positions P_2, P_1, P_1' and P_2' in

the S_2, S_1, S_1' and S_2' subsites respectively are given in figure 4.5.

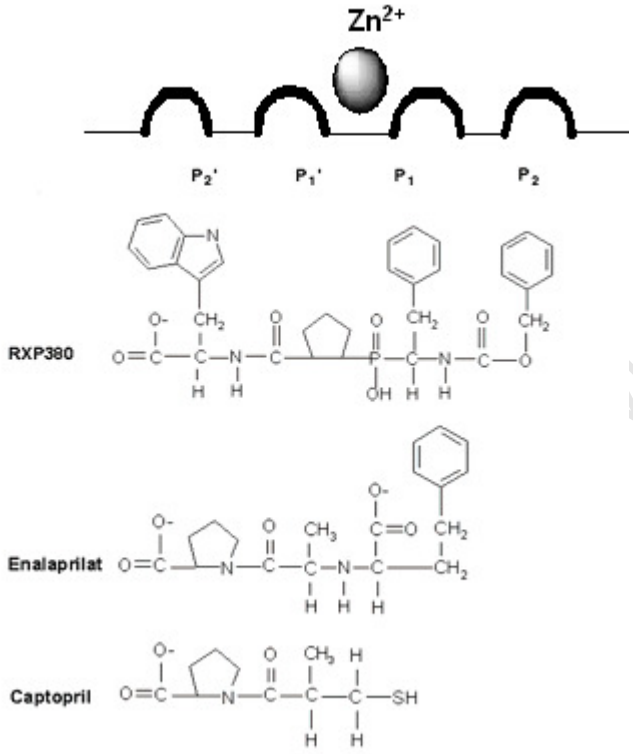


Figure 4.5: Chemical structures of RXPA380, enalaprilat and captopril with residue positions in the P_2, P_1, P_1' and P_2'

4.5.2 N-domain

The crystal structure of the N-domain of sACE has recently been solved. Many similarities can be observed when comparing the overall structures of the two domains, with the N-domain also being abundant in helices and having an equivalent topology to that of the C-domain.[15] The molecule consists of 18 α-helices and six short β-strands. The catalytic Zn^{2+} ion was found to be in the active site with the characteristic zinc binding motif found in the C-domain. Glycosylation sites and amount of glycosylation of

the protein also differ. Nevertheless, the overall conformational considerations between the two domains remain. A structural comparison of the two domains is given in figure 4.6.

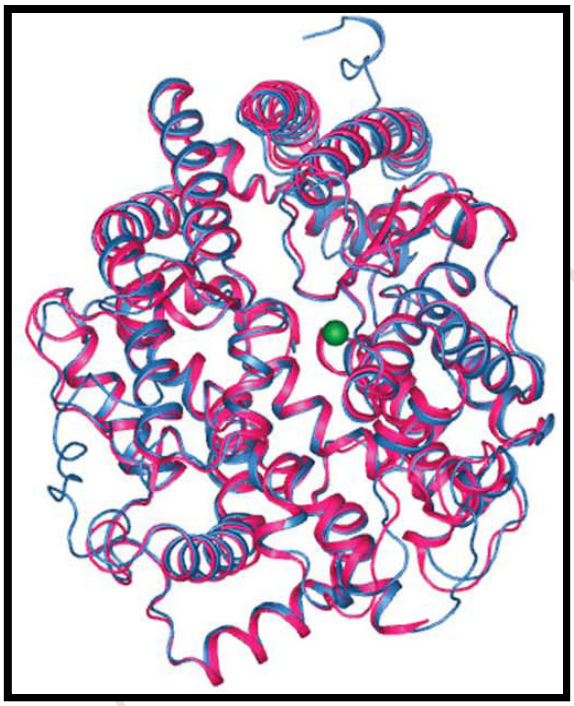


Figure 4.6: A structural comparison of the C- and N-domains. The domains, in overall structure, are very similar. The C-domain sequence is shown in pink while the N-domain is shown in blue. The catalytic zinc ion is shown in green.[15]

The tACE residues that form hydrogen bonds with RXPA380 are all conserved in the N-domain. However, the aromatic interaction with Phe391 is lost because it is replaced by Tyr369 suggesting a possibly different orientation of the inhibitor in this domain. The C-domain selectivity in RXPA380 is thus attributed to the lack of this interaction as well as reduced hydrophobicity in the S_2' subsite.[15]

4.5.3 Comparison of N- and C-domain active sites and inhibitor binding

The active site of the two ACE domains can be considered to be made up of four subsites (S_1, S_2, S_1', S_2') each containing different amino acid residues.[12] Comparisons can be drawn and made use of from the above crystal structure determinations, both in the binding of lisinopril in the active site and the interactions of this inhibitor with active site residues. Differences exist between the residues of the active sites of the N- and C-domains, with a larger number of differences existing particularly in the S_1' and S_2' subsites.[15] A summary of differences between the residues of the S_2' subsites in each domain are given in table 4.1.

Table 4.1: Summary of the differences between the S_2' residues of the active site

	S_2 subsite							
C-domain (tACE)	E403	F391	V518					
N-domain	R381	Y369	T496					
	S_1subsite							
C-domain (tACE)	V518	S516	E413					
N-domain	T496	N494	S119					
	S_1' subsite							
C-domain (tACE)	E162	N277	S284	E372	N374	E376	D377	V380
N-domain	D140	D255	E262	R350	T352	D354	N355	T358
	S_2' subsite							
C-domain (tACE)	T282	S284	E376	V379	V380	D453		
N-domain	S260	D262	D354	S357	T358	E431		

4.6 Historical Perspective of ACE Inhibitors and Inhibition

The story of the design and synthesis of the first orally active and potent inhibitors of ACE has been described as one of the first examples of rational drug design.[16] Involvement of a metal in the catalytic mechanism of ACE had been suspected from its discovery.[17] Confirmation of this fact in the 1970's [18] led to the mechanistic similarity assumption of ACE to Carboxypeptidase A (CPA).[19] The next breakthrough came with the discovery that bradykinin-potentiating peptides isolated from the venom of the Brazilian pit viper *Bothropsjararaca*[20] could specifically inhibit ACE.[21] Subsequent structural-activity studies showed optimal inhibitory capability of a C-terminal sequence of Phe-Ala-Pro[22] and led to the proposal that venom peptides were competitive substrate analogues. The next step was then to find orally active non-peptide analogues of the venom peptides.

A new by-product analogue concept based on work by Byers and Wolfenden[23] for CPA inhibitors led Cushman and Ondetti into recognizing the important role of active-site zinc coordination in binding affinity of inhibitors. Subsequently, their search for superior zinc-binding capabilities led to the replacement of the carboxyl group, typical in CPA inhibitors based on benzylsuccinic acid, with a sulfhydryl group, yielding Captopril[24] as the first orally active and clinically approved therapeutic agent for the treatment of hypertension and its related disorders.

Side effect related short-comings of Captopril, connected to the sulfhydryl moiety, led to reversion to carboxyl compounds and research into alternative functionalities. This led to the discovery of Enalaprilat, Lisinopril and all current generation ACE inhibitors as variations on the original theme, with most of the differences residing in the functionalities that bind the active site zinc and the subsite pockets.[25]

4.7 Classification of known ACE Inhibitors

Some current generation ACE inhibitors (Fig. 4.7) include sulfhydrils, ketones, carboxylates and phosphinates.

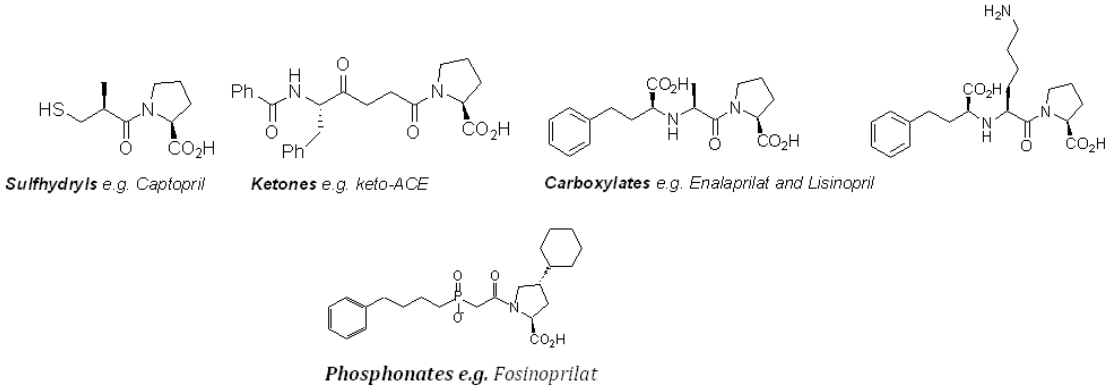


Figure 4.7: Current-generation ACE inhibitors

These compounds are tripeptide substrate analogs in which the C-terminal and penultimate amino acids are retained like in Captopril but the third amino acid is isosterically replaced by a substituted *N*-carboxymethyl group. The search for ACE inhibitors which lacked the sulfhydryl group also led to the investigation of phosphorous containing compounds.[26]

Generally, the mechanism of action of ACE inhibitors involves attenuating the effects of the Renin-angiotensin system (RAS) by inhibiting the conversion of angiotensin I to angiotensin II. They are selective in that they do not directly interfere with any other components of the RAS and have thus been a huge success clinically.[27]

4.8 Selectivity Profiles of Current generation ACE inhibitors

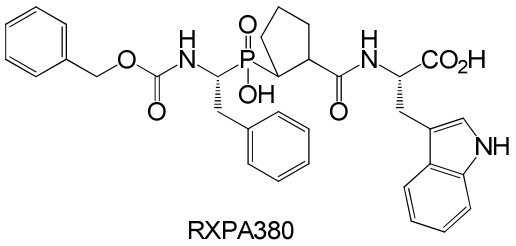
Table 4.2 below shows the selectivity profiles, with respect to the two ACE-domains of some inhibitors currently used in blood pressure treatment as well as research.

Table 4.2: Domain selectivity of some ACE inhibitors.[28]

Inhibitor	Residues	N-domain (nM)	C-domain (nM)
Enalaprilat	Phe-Ala-Pro	26.0	6.3
Captopril	Ala-Pro	8.9	14.0
Lisinopril	Phe-Lys-Pro	44	2.4
RXPA380	Phe-Phe-Pro-Trp	10,000.00	3.0
RXP407	Asp-Phe-Ala-Ala	2.0	2500

It is clear that the ACE inhibitors Captopril, Lisinopril and Enalaprilat show very modest degrees of domain-selectivity. The most C-selective inhibitor currently is the tetrapeptide RXPA380.(Figure 4.8) It is 3,000 times

more selective for the C- rather than the N-domain.[29] In designing this inhibitor, Divéet *al* exploited the known differences in the active site composition of the C- and N-domains. However, the challenges with regard to size (high molecular weight) and synthesis render the tetrapeptide not a good drug candidate. RXP407 on the other hand is 1000 times more N-domain selective.



RXPA380

Figure 4.8: Structure of RXPA380 the most C-domain selective inhibitor known

4.9 Adverse effects and the need for "second generation" ACE inhibitors

Current generation ACE-inhibitors have been contraindicated in many patients.[30] Clearly, the large number of contraindications underscores the need for newer and safer antihypertension drugs with better side effect profiles.

Current generation ACE inhibitors inhibit both the *N*- and *C*-domains. Therefore, taking the side effect profiles and the recent elucidation of the 3-D structure of ACE into consideration, highly specific single C-domain specific inhibitors of ACE offer an attractive alternative as antihypertension agents. The *N*- and *C*-domain active sites of ACE hydrolyze Ang I and BK at comparable rates *in vitro*, but *in vivo* it seems that the C-domain is primarily

responsible for regulating blood pressure.[31] This might indicate that a C-selective inhibitor would have an inhibitory profile comparable to current mixed inhibitors.

Whereas Ang I is hydrolyzed predominantly by the C-domain in vivo, BK is hydrolyzed by both domains[32,33] and therefore we envisage selective inhibition of the C-domain site will allow BK degradation to continue, catalyzed by the N-domain. This could be sufficient to prevent the excessive BK accumulation that has been observed during attacks of angioedema.[34] Thus a highly selective C-domain inhibitor has the potential for effective blood pressure control with reduced vasodilator-related side effects.

4.10 Study justification

The aim of this study therefore, was to pursue our proposed hypothesis with regards to achieving domain-selective ACE-inhibition. This hypothesis is based on the notion than an attenuated zinc-binding group (ZBG) may achieve C-domain selectivity.

Countless attempts have been made to synthesize novel C-domain specific ACE inhibitors based on changes in the inhibitor backbones surrounding a ZBG with little or no success in achieving the selectivity. This part of the research therefore focused on the synthesis of a small number of potential novel inhibitors with an attenuated ZBG. These were utilized in investigating the effect on selectivity and were based on structural features derived from the most C-domain selective inhibitor known, RXPA 380.

Current knowledge suggests that the remarkable selectivity of RXPA380 for this domain is due to the bulky tryptophan residue situated in its P_2'position.

Therefore, apart from attenuating the ZBG, the research investigated a simple structure activity relationship study of this tryptophan moiety effect by incorporating the residue into the P_2'position of two of the most commonly prescribed ACE-inhibitors, namely, Captopril and Enalaprilat. We envisaged that this study could give insight into whether the tryptophan moiety is the only determinant of the domain selectivity in RXPA380.

It is further envisaged that derivatization of current generation ACE inhibitors that show little or no selectivity for the two ACE domains as templates would be a good point to start in inhibitor selectivity studies.

4.11 ACE Zinc Binding Group

As observed in the above discussion, a prominent feature in current generation ACE inhibitors is the zinc binding group (ZBG).The main challenge presented by these drugs is the lack of selectivity between the two domains of ACE namely, N- and C- domain, the latter which is the only domain responsible for the control of hypertension and thus the therapeutic target.[35] Current generation ACEIs inhibit both domains, blocking the other homeostatic functions carried out by the *N*-domain, resulting in adverse side effects.[36]

ACE is a zinc metalloprotease enzyme which is closely related to matrix metalloproteases (MMPs).[37] ZBGs have been found to modulate selective inhibition of MMPs. Several novel MMP inhibitors have been reported and studies show that their selectivity is dependent on the nature of the ZBG. This is in contrast to most current matrix metalloprotease inhibitors (MMPI's), which obtain isoform selectivity solely from the peptidomimetic backbone portion of the compound. Different ZBGs have been appended to a common backbone and the inhibition efficiency of each inhibitor determined. These results show that the selectivity profile of each inhibitor is different as a result of the various ZBGs.[38] This study further demonstrated that the choice of ZBG can have a significant effect in a relevant patho-physiological endpoint.

Other studies have also demonstrated that strong ZBGs such as hydroxamates lack selectivity.[39] Their strong zinc (II)-binding capacity suppresses the overall effect of the weaker but more specific non-covalent interactions of the inhibitor backbone with the substrate specificity regions of the enzyme.[40]

MMPIs bearing the weak ZBG phosphinic acid have been found to efficiently discriminate between the different enzymes, especially through the increasedinteraction with both the primed and unprimed subsites of the protease active sites.[41] Furthermore, the placement of the ZBG is in the middle of a pseudopeptidic scaffold and not at its *N*- or *C*-terminus, as is the cases with hydroxamate and carboxylate inhibitors.

Therefore, specific, non-covalent interactions between a phosphinic peptide and a protease with the S2, S1, S1' and S2' can be significantly increased which can result in enhanced selectivity if these interactions are properly tuned. [42] Over the last several years, numerous MMPI's havebeen reportedthat actually do not possess a ZBG and hence do not bind the catalytic Zn^{2+} ion.[43, 44, 45, 46]

Current generation ACE inhibitors have ZBGs such as sulhydryl, carboxylates, hydroxamates [47] , ketones and phosphinic acids.[48] A plethora of other ZBGs exist for other members of the metalloprotease family.[46] Quantum chemistry calculations performed on simple models of interaction between the zinc ion and the various zinc-binding groups mentioned above predict that the hydroxamate function is the strongest ZBG group in this series, while the phosphoryl is the weakest.[49] These predictions fit with the extremely high potency displayed by many hydroxamates developed towards MMPs, as compared with the potency displayed by some phosphinic peptide inhibitors.[39]

This remarkable selectivity is attributed to the bulky backbone of the molecule and specifically on the tryptophan residue in the P_2'position. The molecule is more easily accommodated by the deeper and more hydrophobic C-domain S_1' and S_2' subsites than those in the *N*-domain.[15] Furthermore, in the position α- to the ZBG, there is a bulky five-membered cyclopentane structure which further conceals the ZBG and restricts free rotation about the Cα–Cβ bond. Domain selectivity assays of the analogues synthesized *enroute* to RXPA380 (Figure 4.9) revealed that there was significant C-

domain selectivity amongst the analogues with a group in the α'-position to the ZBG compared to those without. [50] These potency studies therefore strongly suggest that the selectivity of RXPA380 for the C-domain may not only be due to the tryptophan residue in the P_2', but the prolyl moiety in the P_1' as well. Studies have been performed before in this research group where bulkier hydrophobic residues inserted in the P2'did not effect domain selectivity in either the C- or N-domain.[51,52]

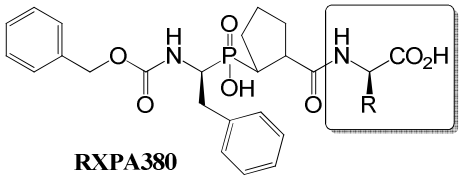
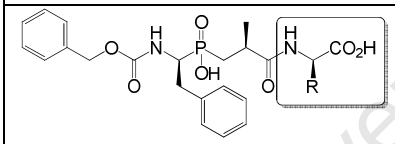
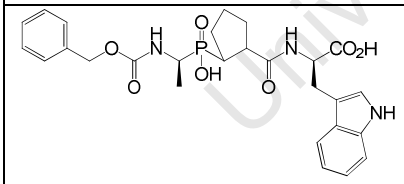
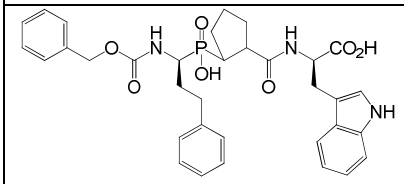
Compound	Amino Acid	Selectivity Ki for the C-domain
 RXPA380	PheProTrp	3300
	PheProAla	22
	PheProPro	15
	PheProArg	22
	PheAlaTrp	90
	PheAlaAla	1
	AlaProTrp	130
	HomoPheProTrp	138

Figure 4.9: Domain selectivity potency studies of RXPA380 and its analogues

Our hypothesis is therefore that a weaker and/or concealed zinc-binding group could present an opportunity for coming up with a C-domain selective ACE inhibitor resulting from enhanced contribution from the usually suppressed weaker but more specific non-covalent interactions.

It is also clear that the ZBG could have a significant effect on the selectivity of ACE inhibitors. From the plethora of evidence gathered in the study of the closely related MMPI's, a burning question to ask at this point is to what extent the ZBG affects domain-selectivity in ACE. A further and future question to consider for the next generation of domain-selective ACE inhibitors iswhether they should or should not have a ZBG. Consequently, a follow-up aspect of our hypothesis is to investigate whether the tryptophan residue in the RXPA380 structure is the only determinant of C-domain selectivity.

4.12 Summary of Rationale

The rationale for this strategy revolves around minimizing or eliminating the interaction with the catalytic Zn^{2+} ionto achieve domain selectivity, as the metal site is the most conserved feature between the N- and C-domain of ACE. Phophinates were thus chosen because of the advantageous placement of the ZBG in the middle of a pseudopeptidic scaffold and not at its N- or C-terminus, as in the cases of hydroxamate, carboxylate and sulfhydryl inhibitors. Less synthetically challenging phophinates were thus proposed and synthesized, including a varied R-group in the position

α- to the phosphinic acid, to investigate the effect on domain-selectivity. (Figure 4.10) It should be noted however, that compared to the ordinary phosphinic acids, the molecules have the NH group two bonds closer to the phosphinic group. This is a major backbone modification with possible implications on the overall alignment of the molecule in the active site bringing the proposed P_2' residue closer to the P_1' region.

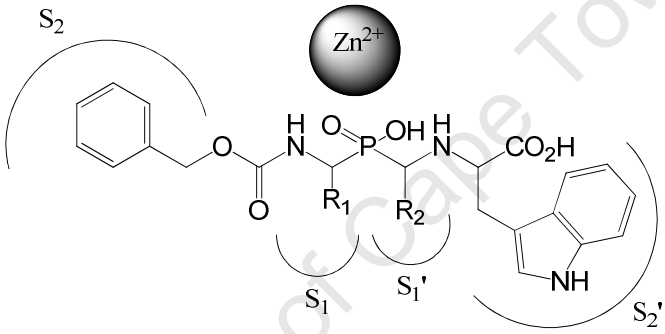


Figure 4.10: General structure for the design of target compounds depicted in the active site

Figures 4.10 gives the general structure for the design of target phosphinic acid based compounds depicted in the ACE-active site. The R-groups were selected from a different sized substituents ranging from methyl to benzyl groups with careful consideration of Lipinski's rule-of-five analysis. This analysis/guidelines predict that poor absorption or permeation of a orally administered compound are more likely if the compound meets the following criteria, molecular mass greater than 500 amu, high lipophilicity (expressed as cLogP greater than 5), more than 5 hydrogen bond donors, more than 10 hydrogen bond. These rules are based on the observation that most drugs are relatively small and lipophilic molecules.[53]Although the

molecular weights of the compounds ranges from 499 to 623 amuthey have 5 hydrogen bond donors and/or acceptors and generally, the compounds have cLogP's of less than 5. Moreover, the compound, RXPA380 on which the design was based itself has a molecular weight of 617 amu.

Given that an attenuated ZBG may influence domain selectivity and specificity, a further hypothesis was that the tryptophan residue of RXPA380 may not be the only structural determinant of C-domain selectivity. We therefore envisaged that replacement of the prolyl moiety of known ACE inhibitors with the bulkier tryptophan residue would provide insight into this domain selective-behaviour.A good precedence for this is the enhanced ACE *C*-domain selectivity observed with Keto-ACE analogues (Figure 4.11) whose structure also involved a ZBG in the middle of a pseudo-peptidomimetic scaffold.[54]

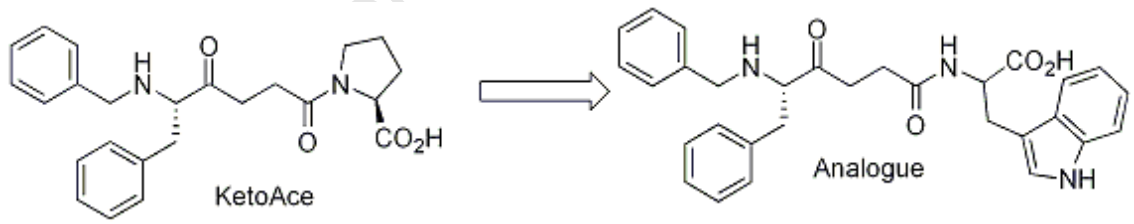


Figure 4. 11: KetoAce and its tryptophan based analogue

4.13 Objectives

The objectives for this part of the research were:

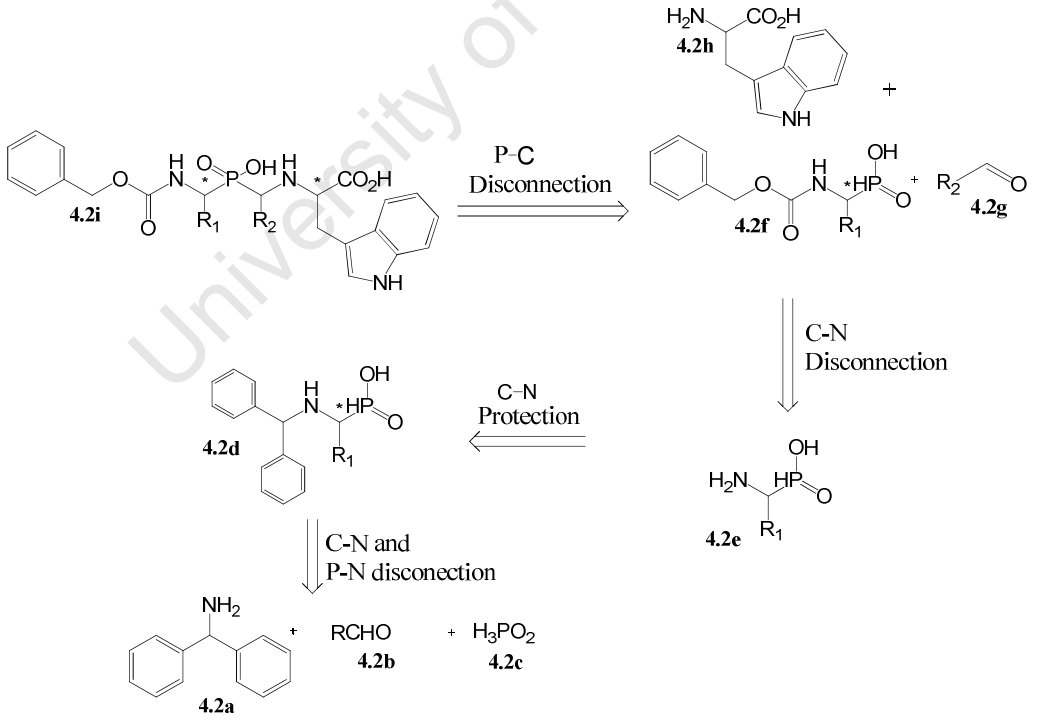
- To synthesize novel phosphinic acid-based ACE inhibitors with a substituent in the α-position relative to the phosphinate ZBG as potential novel antihypertension agents

- To synthesise new potential ACE-inhibitors incorporating a tryptophan residue in the P2'of the captopril and enalaprilat scaffold
- To evaluate the domain selectivity of these synthesised compounds *in vitro* using appropriate ACE inhibitory assays.

4.14 Synthetic Discussion

4.14.1 Retro-synthetic Strategy

Retrosynthetically, target molecule **4.2i** is synthesized in one pot from the amino acid tryptophan **4.2h** and an appropriate aldehyde **4.2g.** The latter is the protected derivative of α-aminophosphinic acid **4.2e** which is prepared from the deprotection of **4.2d.**

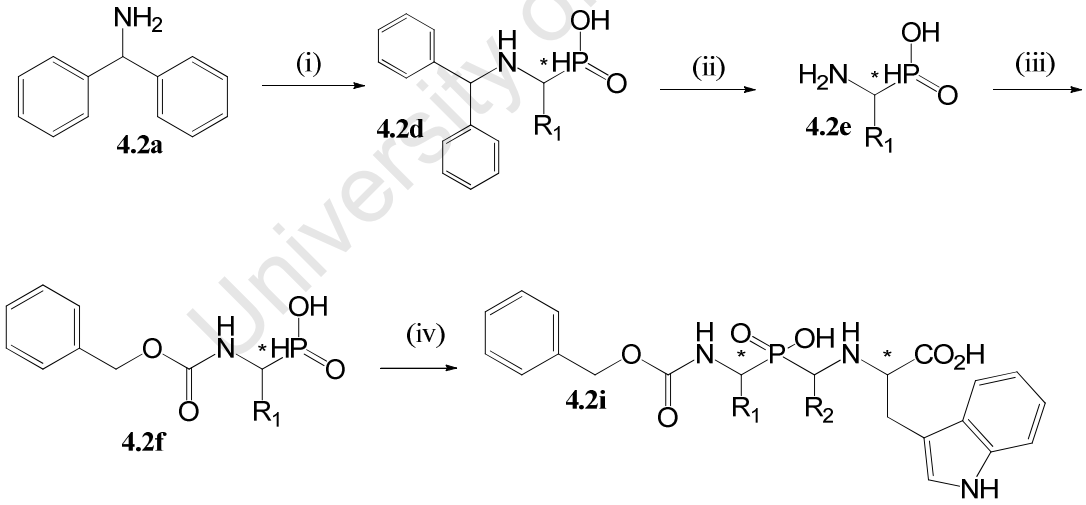


Scheme 4. 1: Retro-synthetic strategy for the preparation of the target compounds

A double disconnection across *C–N* and *P–N* bonds leads back to commercially available diphenylmethylamine **4.2a,** aldehyde **4.2b** and hypophosphorous acid **4.2c.**

4.14.2 Synthesis of (1-amino-2-phenylethyl)phosphinic acid 4.2e

Preparation of phosphinic acid **4.2d** was thus achieved using a well known literature procedure [55] by refluxing equimolar amounts of commercially available diphenylmethylamine hydrochloride **4.2a,** 50% aqueous hypophosphorous acid and an aldehyde in ethanol for 3h. The resulting diphenylmethylaminophosphonous acid **4.2d** precipitate was filtered off, dried and immediately dissolved in 48% hydrobromic acid and heated for 1 hour at 100 °C resulting in the cleavage of the diphenylmethyl group unmasking the amine group of α-amino-phosphinic acid **4.2e.** (Scheme 4.3)



Scheme 4.2:Reagents and conditions. (i) R_1CHO, H_3PO_2, EtOH, H_2O, reflux, 3h, 88% (ii) 48% HBr, H_2O, 100°C, 1h, quant. (iii) CBz–Cl, NaOH, $NaHCO_3$, H_2O, 4°C (iv)(a)R_2CHO, TrpOMe, 1,4-Dioxane, 5h (b) 4M NaOH, MeOH.

Precipitation was induced by dissolution in ethanol and the drop-wise addition of propylene oxide until complete. Purity was confirmed using melting point.

4.14.3 Preparation of 2-((1-((1-(((benzyloxy)carbonyl)amino)-2-phenylethyl)(hydroxy)phosphoryl)ethyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i.

The amine was then reacted with benzyl chloroformate in basic medium to yield **4.2f**. A reported one-step, three component reaction involving an appropriate aldehyde,(1-(((benzyloxy)carbonyl)amino)-2-phenylethyl)phosphinic acid **4.2f** (or its analogues) as well as methoxy-protected tryptophan gave way to 2-((1-((1-(((benzyloxy)carbonyl)amino)-2-phenylethyl)(hydroxy)phosphoryl)ethyl)amino)-3-(1H-indol-3-yl)propanoic acid **4.2i-4** (or its analogues).[56]

Table 4. 3: Percentage yields of 4.2i analogues

Compound	R_1	R_2	% Yield
4.2i – 1	CH_3	CH_3	ND* (Crude 81%)
4.2i – 2	$C(CH_3)_3$	CH_3	75
4.2i – 3	CH_2Ph	CH_3	ND* (Crude 43%)
4.2i – 4	$CH_2CH_2CH_3$	CH_3	66
4.2i – 5	CH_3	$CH_2CH_2CH_3$	68
4.2i – 6	$C(CH_3)_3$	$CH_2CH_2CH_3$	70
4.2i – 7	CH_2Ph	$CH_2CH_2CH_3$	72
4.2i – 8	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	68
4.2i – 9	CH_3	$C(CH_3)_3$	79
4.2i – 10	$C(CH_3)_3$	$C(CH_3)_3$	79
4.2i – 11	CH_2Ph	$(CH_3)_3$	88
4.2i – 12	$CH_2CH_2CH_3$	$(CH_3)_3$	82
4.2i – 13	CH_3	CH_2Ph	80
4.2i – 14	$C(CH_3)_3$	CH_2Ph	60
4.2i – 15	CH_2Ph	CH_2Ph	54
4.2i – 16	$CH_2CH_2CH_3$	CH_2Ph	50

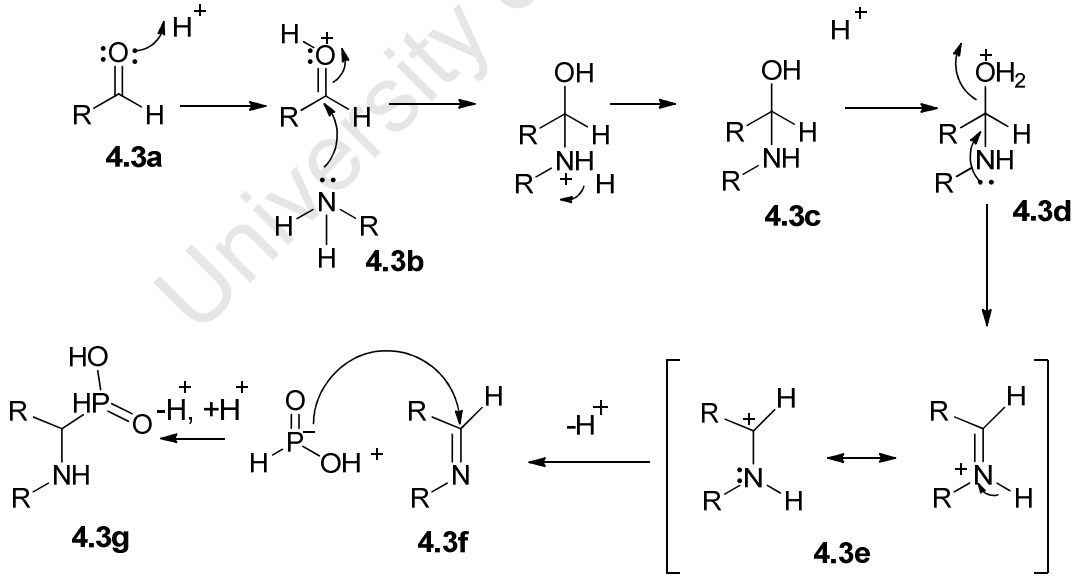
ND* – Yield not determined. M+ ion observed but compound extremely impure (HPLC). Very low yield and not enough material to isolate.

4.14.4 Mechanistic details for the preparation of 1-diphenylmethyl aminophophonous acid 4.4g.

There are two postulated mechanisms for this reaction. In the first one, it is said to proceed *via* the nucleophilic addition of hypophosphorous acid to the azomethine bond of 1,1-diphenyl-N-(2-phenylethylidene)methanamine Schiff base of phenyl acetaldehyde leading to the formation of 1-diphenylmethyl aminophophonous acid. The second possible mechanism includes attack of hypophosphorous acid on an aldehyde leading to α- hydroxyphosphonous acid, which

undergoes the nucleophilic substitution and it converts into the desired α-aminophosphonous acid. It is not known which mechanism is operative in this case but the more plausible one is the first. (Scheme 4.3)

The acidic media provides the carbonyl oxygen of **4.3a** with the catalytic proton resulting in the attack of the nucleophilic amine **4.3b** on the electrophilic carbonyl carbon followed by the loss of an H^+ from the now positively charged nitrogen. The hydroxyl oxygen of **4.3c** is protonated again forming an oxonium ion **4.3d** which then results in the departure of an H_2O, the driving force being formation of a resonance stabilized intermediate **4.3e.** Departure of a proton restores the nitrogen lone pair and forms the Schiff base **4.3f.** Nucleophilic addition of hypophosphorous acid to the azomethine position follows resulting in **4.3g.**



Scheme 4. 3: Mechanism for the preparation of diphenylmethyl aminophophonous acid 4.4g.

A typical H NMR for 4.2i-1 is given in figure 4.12.

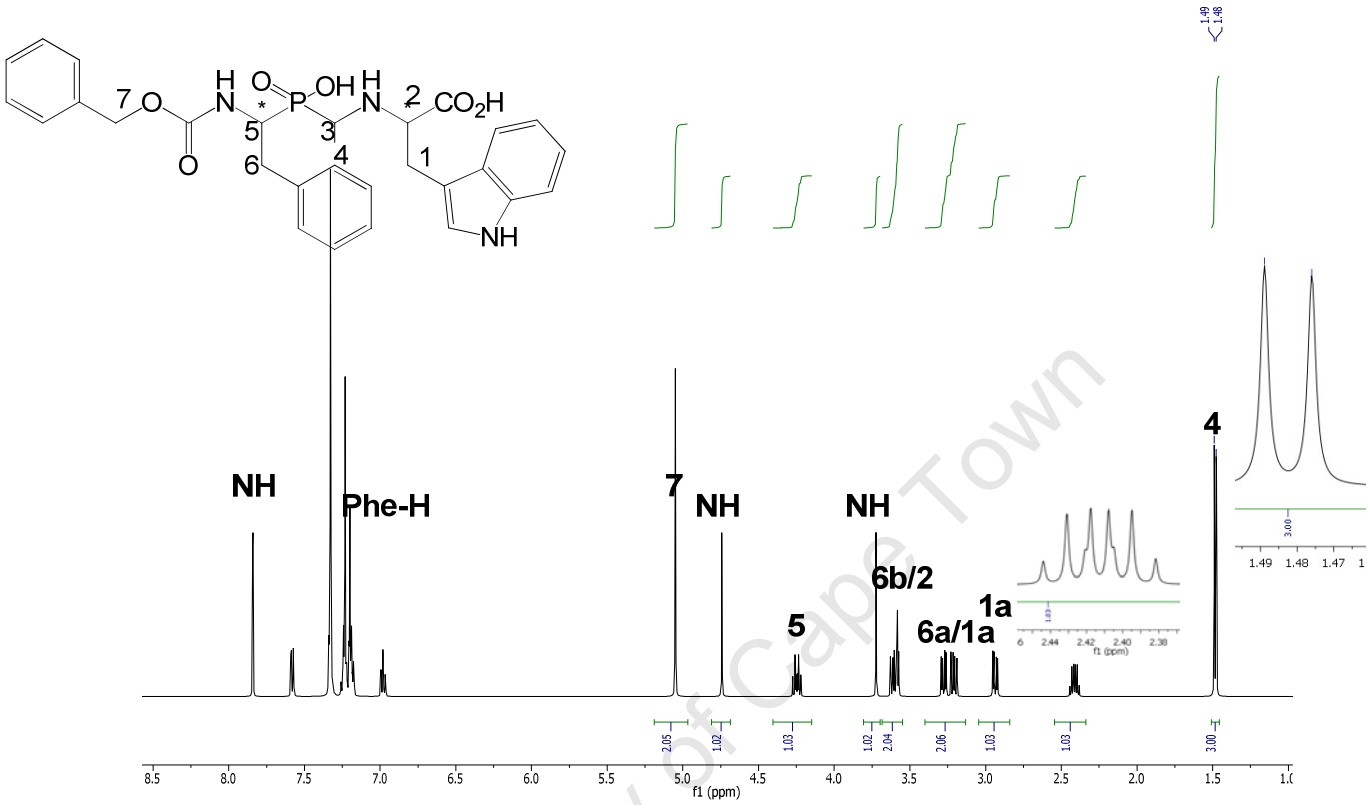
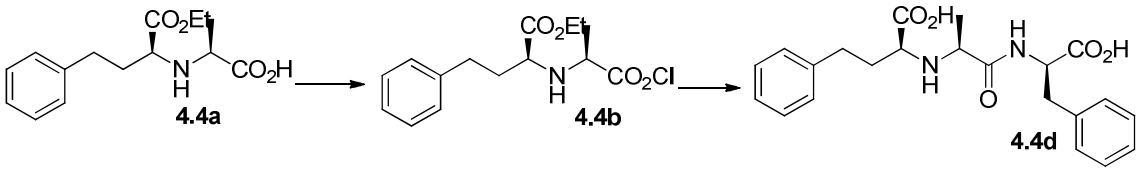


Figure 4.12: A section of the typical ^{1}H NMR for compound 4.2i-1 in $CDCl_3$

4.14.5Preparation of Enalaprilat and its analogues

The carboxylic acid (S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoic acid **4.4a** was activated by conversion to (S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoic hypochlorous anhydride **4.4b** using phosphorous pentachloride. Acyl chloride **4.4b** was then elaborated to compounds **4.4c,** (S)-2-(((S)-1-(((R)-1-carboxy-2-phenylethyl)amino)-1-oxopropan-2-yl)amino)-4-phenylbutanoic acid **4.4d** and (S)-2-(((S)-1-(((R)-1-carboxy-2-(1H-indol-3-yl)ethyl)amino)-1-oxopropan-2-yl)amino)-4-phenylbutanoic acid **4.4e** *via* a reaction with proline,

phenylalanine and tryptophan respectively each followed by saponification using 4M NaOH (Scheme 4.4).



Compound	Amino Acid Residue	% Yield
4.4c	Pro	54
4.4d	Phe	60
4.4e	Trp	69

Scheme 4.4. General Reagents and Conditions: (i) PCl_5, DCM, rt. Overnight (ii) $NaHCO_3$, H_2O 1,4-Dioxane, Amino Acid(ii) NaOH, H_2O

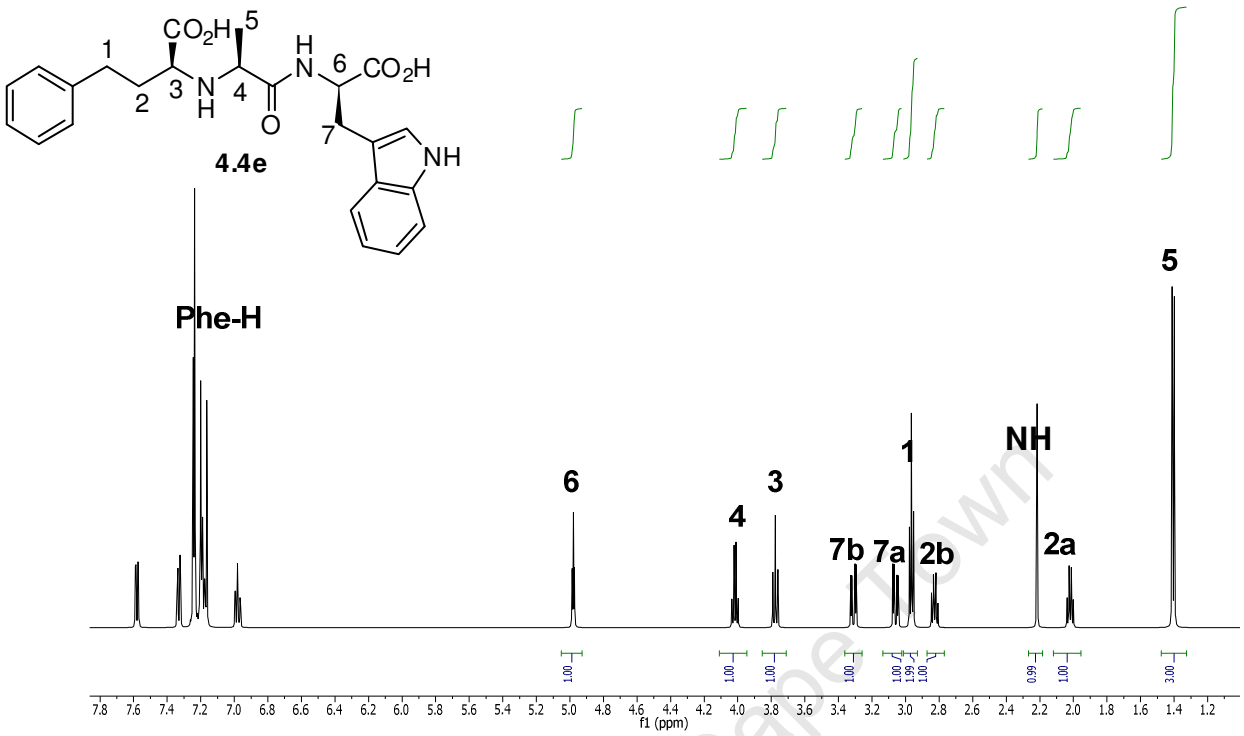
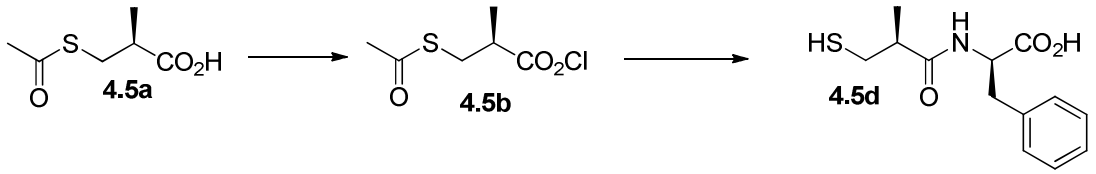


Figure 4.13: Typical HNMR for (S)-2-(((S)-1-(((R)-1-carboxy-2-(1H-indol-3-yl)ethyl)amino)-1-oxopropan-2-yl)amino)-4-phenylbutanoic acid (tryptophan analogue)**4.4e in $CDCl_3$.**

4.14.6 Preparation of (S)-1-((S)-3-mercapto-2-methylpropanoyl)pyrrolidine-2-carboxylicacid (Captopril)and its analogues

In a similar fashion to Enalaprilat and its analogues, the carboxylic acid of **4.5a** was activated by conversion to the acyl chloride **4.5b** using thionyl chloride. Acyl chloride **4.5b** was then elaborated to compounds **4.5c**, **4.5d** and **4.5e** *via* a reaction with the proline, phenylalanine and tryptophan respectively, followed by saponification using 4M NaOH. (Scheme 4.5). A typical ^{1}HNMR of this series of (R)-2-((S)-3-mercapto-2-methylpropanamido)-3-phenylpropanoic acid **4.5d** is given in Fig 4.14.



Compound	Amino Acid Residue	% Yield
4.5c	Pro	80
4.5d	Phe	64
4.5e	Trp	70

Scheme 4.5: General Reagents and Conditions: (i) $SOCl_2$, DCM, rt. Overnight (ii) $NaHCO_3$, H_2O 1,4-Dioxane, Amino Acid.HCl(ii) NaOH, H_2O.

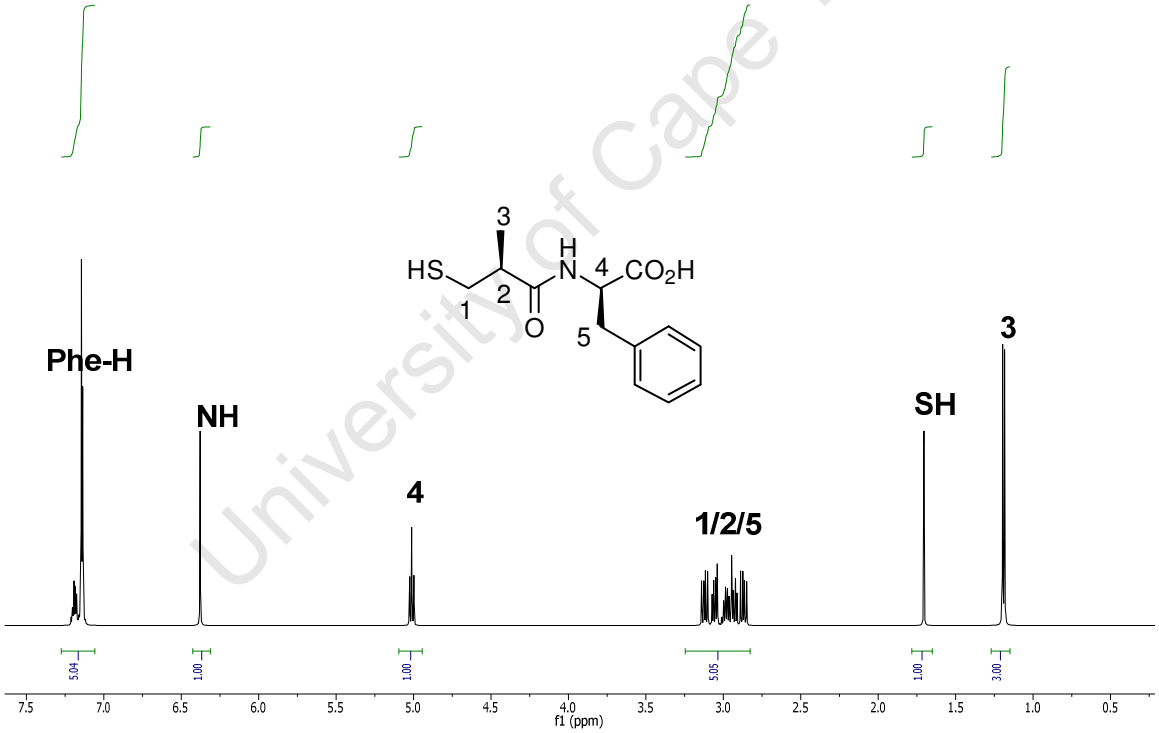
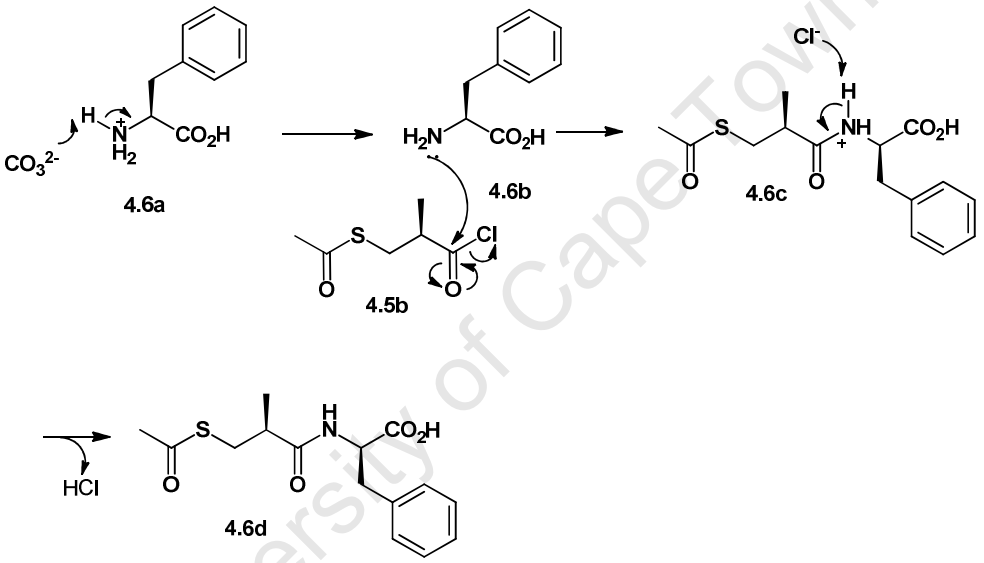


Figure 4.14: ^{1}HNMR of (R)-2-((S)-3-mercapto-2-methylpropanamido)-3-phenylpropanoic acid **4.5d** in Phenylalanine analogue) in $CDCl_3$

Mechanistically, intermediate **4.5b** is an active acid chloride that can easily undergo nucleophilic acyl

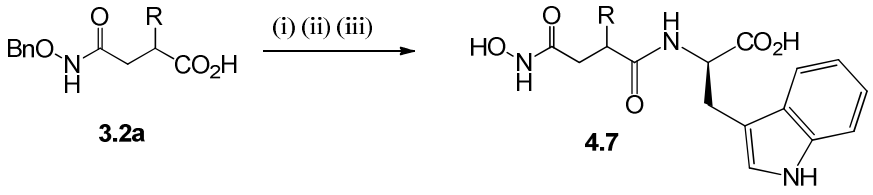
substitution (Scheme 4.8). An equivalent of base destroys the hydrochloride salt of phenylalanine **4.6a** and frees the nitrogen lone pair of electrons as depicted in **4.6b.** This lone pair then nucleophilically attacks the acyl carbonyl of **4.5b** resulting in the departure of a chloride anion. The amide bond results in adduct **4.6c** and the resulting hydrochloric acid formed in the reaction is taken care of by the excess sodium bicarbonate base leaving the product **4.6d.**



Scheme 4. 6: Mechanism of acyl substitution

4.14.7 Preparation of Tryptophan hydroxamates

The protected hydroxamates **3.2a–3** to **3.2a–6** were coupled to L–tryptophan methoxy–ester under standard coupling conditions followed sequentially by saponification and hydrogenolysis giving rise to a series **4.7a** to **4.7e.**



3.2a 4.7

Compound	R-group	% Yield
4.7a	CH_2Cyclohexyl	66
4.7b	$CH_2CH(CH_3)_2$	71
4.7c	CH_2Ph	70
4.7d	CH_2Napthyl	75
4.7e	CH_3	ND

ND: Sample difficult to purify

Scheme 4.7: Reagents and conditions (i) AA-Methyl ester, EDC-HCl, HOBt, DIEA (ii) NaOH, MeOH (iii) H_2/Pd-C, 3h.

A typical ^{1}HNMR for the compound **4.7c** is given in figure 4.15 below.

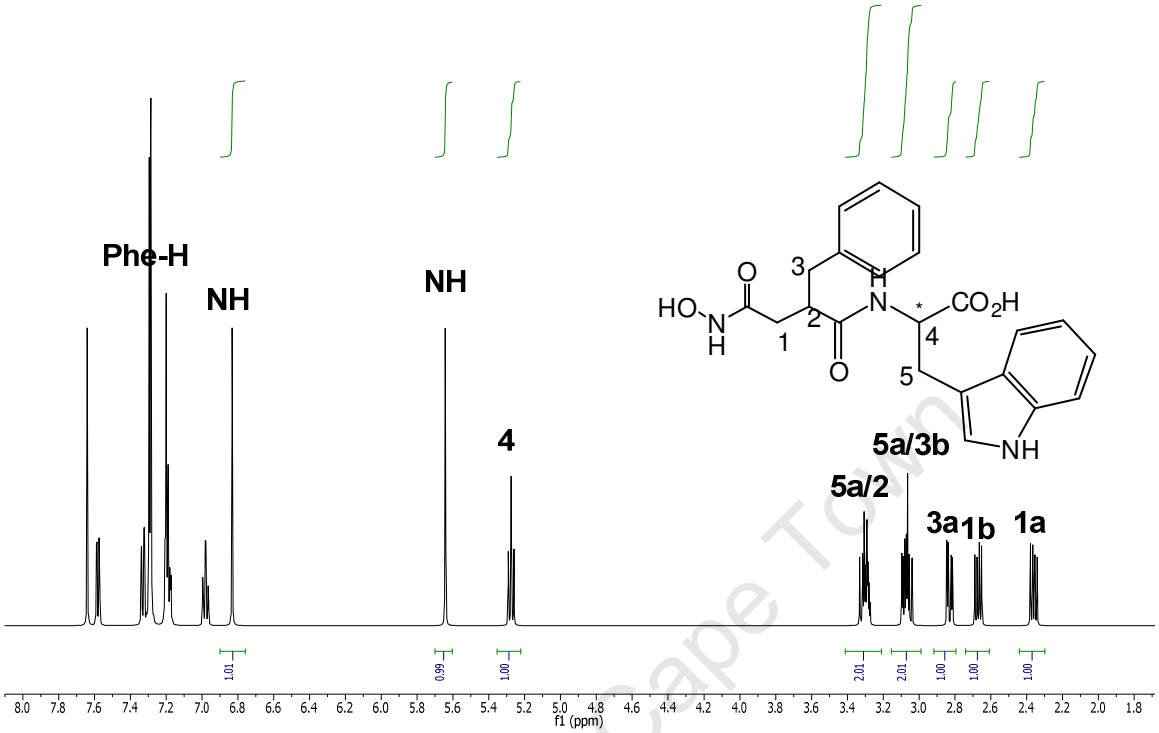


Figure 4.15:.A section of ^{1}HNMR for compound **4.7c** in $CDCl_3$.

4.15 Summary

Chapter 4 has discussed a novel arena of research in the drive towards the next generation domain-selective ACE inhibitors. Currently, there exist no ACE inhibitors that do not rely on the ZBG for their activity. Synthesis of potential novel ACE inhibitors with an attenuated, or indeed with an excluded ZBG could be the answer to domain selectivity.

The attenuated ZBG in this research is phosphinic acid which is a weak zinc binding group. Its zinc-binding ability is further inhibited by its being sterically

crowded with bulky groups in the α-position to the reactive group. Chapter 4 has presented how a series of inhibitors with an attenuated phosphinic acid ZBG have been synthesised. Chapter 5 will go on to discuss the biological results obtained from these compounds with respect to domain selectivity.

References

[1] Kearney, P. M. M.; Whelton, K.; Reynolds, *Lancet*, **2005**, *365*, 217 – 223.

[2] Guyton, A. C. (1971) Textbook of Medical Physiology, 4th Ed.W.B. Saunders Co.(Philadelphia), USA.

[3] Skeggs, L. T.; Jr., Lentz, K. E.; Kahn, J. R.; Shumway, N. P.; Woods, K. R., *Journal of Experimental Medicine* **1956**; 104:193–197.

[4] Schmaier, A.H., *Journal of Clinical Investigation* **2002**; 109: 1007–1009.

[5] Acharya, K. R.;Sturrock, E. D.; Riordan, J. F.; Ehlers, M. R.,*Nature Reviews Drug Discovery* **2003**; 2: 891–902.

[6] Rieger, K. J.;Saez–Servent, N.;Papet, M. P.; Wdzieczak–Bakala, J.;Morgat, J. L.; Thierry, J.; Voelter, W.; Lenfant, M. *Biochemical Journal* **1993**; 296 **(Pt 2)**, 373–378.

[7] Soubrier, F.;Alhenc–Gelas, F.; Hubert, C.;Allegrini, J.; John, M.;Tregear, G.;Corvol, P.*Proceeding of the National Academy of Science, U.S.A.***1988**;85: 9386–9390.

[8] Ehlers, M. R. W.; Fox, E. A.;Strydom, D. J.; Riordan, J.F. *Proceeding of the National Academy of Science, U.S.A.* **1989**;86:7741–7745.

[9] Howard, T. E.;Shai, S.–Y.; Langford, K. G.; Martin, B.

M.; and Bernstein, K. E. *Molecular and Cell Biology* **1990**; 10: 4294-4302.

[10] Jaspard, E.; Wei, L.;Alhenc-Gelas, F. *Journal of Biological Chemistry* **1993**;268:9496-9503.

[11] Wei, L., Alhenc-Gelas, F., Corvol, P., and Clauser, E. (1991) *Journal of Biological Chemistry* **1991**; 266:9002 9008.

[12] Natesh, R.;Schwager, S. L.;Sturrock, E. D.;Acharya, K. R. *Nature* **2003**;421:551-554.

[13] Corradi, H.R.; Chitapi, I.; Sewell, T.B.; Georgiadis, D.; Dive, V.; Sturrock, E.D.; Acharya, K.R.,*Biochemistry* **2007**;46:5473-5478.

[14] Kroger W. L.; Douglas R. G.; O'niell, H. G.; Dive, V.; Sturrock E. D. *Biochemistry* **2009**;48: 8405-8412.

[15] Corradi, H. R., Schwager, S. L., Nchinda, A. T., Sturrock, E. D., and Acharya, K. R. *Journal of Molecular Biology* **2006**;357: 964-974.

[16] Turner, A. J. & Hooper, N. M. *Trends in Pharmacological Sciences* **2002**;23: 177-183.

[17] Skeggs, L. T., Kahn, J. R.; Shumway, N. P.*Journal Of Experimental Medicine* **1956**; 103: 295-299.

[18] Cushman, D. W.; Cheung, H. S. *Biochemical Pharmacology* **1971**;20:1637-1648.

[19] Ondetti, M. A.; Rubin, B.;Cushman, D. W. *Science* **1977**; 196:441-444.

[20] Ferreira, S. H.; Rocha Silva, E. ; *Experientia* **1965**; 21: 347-349.

[21] Ng, K. K. & Vane, J. R. *Nature* **1967**: 216; 762–766.

[22] Ondetti, M. A., Williams, N. J., Sabo, E. F., Pluscec, J., Weaver, E. R., Kocy, O. *Biochemistry* **1971**; 10:4033–4039.

[23] Byers, L. D.,Wolfenden, R. *Biochemistry* **1973**;12: 2070–2078.

[24] Ondetti, M. A., Rubin, B.; Cushman, D. W. *Science* **1977**;196:441–444.

[25] Menard, J.; Patchet, A. A. *Advances in Protein Chemistry* **2001**; 56:13–75.

[26] Krapcho, J.; Turk, C.; Cushman, D. W.; Powell, J. R.; DeForrest, J. M.;Spitzmiller, E. R.;Karanewsky, D. S.; Duggan, M.;Rovnyak, G.; Schwartz, J. *Journal of Medicinal Chemistry* **1988**; 31: 1148–1160.

[27] Rang H.P.; Dale, M.M,; Ritter, J.M. *Pharmacology*. Churchill Livingstone, London. 4th ed. **1999** 290–296.

[28] Acharya, K. R.;Sturrock, E. D.; Riordan, J. F.; Ehlers, M. R. *Nature Reviews Drug Discovery* **2003**;2: 891–902.

[29] Georgiadis, D.; Beau, F.;Czarny, B.; Cotton, J.; Yiotakis, A.; Dive, V. *Circulation Research* **2003**; 93: 148–154.

[30] Okumura, H.;Nishimura, E.;Kariya, S.;Ohtani, M.; Uchino, K.; Fukatsu, T.; Odanaka, J.; Takahashi, T.; Watanabe, K.; Itoh T., *Yakugaku Zasshi* **2001**;121(3):253-7.

[31] Junot, C.; Gonzales, M.F.;Ezan, E.; Cotton, J.; Vazeux, G.; Michaud, A.;Azizi, M.;Vassiliou, S.;Yiotakis, A.;Corvol, P.; Dive, V.*Journal of Pharmacology and Experimental Therapeutics* **2001**; 38: 95-99.

[32] Esther, C. R.; Marino, E. M.; Howard, T. E.; Machaud, A.;Corvol, P.;Capecchi, M. R.; Bernstein, K. E. *Journal of Clinical Investigations* **1997**; 99: 2375-2385.

[33] Junot, C., Gonzales, M. F., Ezan, E., Cotton, J., Vazeux,G., Michaud, A., Azizi, M., Vassiliou, S., Yiotakis, A., Corvol, P., and Dive, V.*Journal of Pharmacology and Experimental Therapeutics* **2001**;297: 606-611.

[34] Messerli, F. H.; Nussberger, J.*Lancet* **2000**;356:608-609.

[35] Fuchs, S; Xiao, H. D.; Hubert, C.; Michaud, A.; Campbell, D.J.; Adams,J.W.; Capecchi, M.; Corvol, P.; Bernstein, K. *Hypertension* **2007**;51: 267- 274.

[36] Acharya, K.R.; Sturrock, E.D.; Riordan, J.F.; Ehlers M.R. *Nature Reviews Drug Discovery* **2003**; 2:891-902.

[37] Kara, I.; Ozkok, E.; Aydin, M.; Orhan, N. Cetinkaya,

Y.; Gencer, M.; Kilic, G.; Tireli, H. *Cephalalgia* **2007**;27: 235 – 243.

[38] Agrawal, A.; Romero-Perez, D.; Jacobsen, JA.; Villarreal, FJ.; Cohen, SM. *ChemMedChem* **2008**; 3: 812–820.

[39] Cuniasse, P.; Devel, L.; Makaritis, A.; Beau, F.; Georgiadis, D.; Matziari, M.; Yiotakis, A.; Dive, V. *Biochimie* **2005**; 87: 393.

[40] Rao, B. G. *Current Pharmaceutical Design* **2005**; 11 :295.

[41] Georgiadis, D.; Yiotakis, A. *Bioorganic & Medicinal Chemistry***2008**; 16:19, 8781– 8794.

[42] Dive, V.; Georgiadis, D.; Matziari, M.; Makaritis, A.; Beau, F.; Cuniasse, P.;Yiotakis, A. *Cellular and Molecular Life Sciences***2004**, 61, 2010.

[43] Johnson, AR.; Pavlovsky, AG.; Ortwine, DF.; Prior, F.; Man, C–F.; Bornemeier,DA.; Banotai, CA.; Mueller, WT.; McConnell, P.; Yan, C.; Baragi, V.; Lesch, C.;Roark, WH.; Wilson, M.; Datta, K.; Guzman, R.; Han, H–K.; Dyer, RD *Journal of Biological Chemistry***2007**:282:27781–27791.

[44] Li, JJ.; Nahra, J.; Johnson, AR.; Bunker, A.; O'Brien, P.' Yue, W–S.; Ortwine,DF.; Man, C–F.; Baragi, V.; Kilgore, K.; Dyer, RD.; Han, H–K. *Journal of Medicinal Chemistry***2008**; 51: 835–841.

[45] Morales, R. ; Perrier, S. ; Florent, J–M. ; Beltra, J.; Dufour, S.; De Mendez, I.; Manceau, P.; Tertre, A.; Moreau, F.; Compere, D.; Dublanchet, A–C. ; O'Gara,M. *Journal of Molecular Biology* **2004** 341; 1063–1076.

[46] Jacobsen, JA.; Jourden JLM.; Miller, TM.; Cohen, SM.

*Biochimica et Biophysica Acta 2***009**,doi: 10.1016/j.bbamcr.2009.08.00.

47 Turbanti, L. Cerbai, G.; Di Bugno, C; Giorgi, R; Garzelli, G.; Criscuoli, M.;Renzetti, A.R.; Subissi, A.; Bramanti, G.; DePriest, S.A. *Journal of Medicinal Chemistry* **1993**; 36: 699-707.

48 Krapcho, J.; Turk, C.; Cushman, D. W.; Powell, J. R.; DeForrest, J. M.;Spitzmiller, E. R.; Karanewsky, D. S.; Duggan, M.; Rovnyak, G.; Schwartz, J.*Journal of Medicinal Chemistry* **1988** 31, 1148-1160.

49 Cheng, F.; Zhang, R.; Luo, X.; Shen, J.; Li, X.; Gu, J.; Zhu, W.; Shen, J.; Sagi,I.; Ji, R.; Chen, K.; Jiang, H. *Journal of Physical Chemistry* **2002 106**, 4552-4559.

50 Georgiadis, D.; Beau, F.; Czarny, B.; Cotton, J.; Yiotakis, A.; Dive, V. *Circulation Research* **2003** 93,148-154.

51 Tanya Paquet, MSc Thesis. 2007.

52 Henry Kambafwile, MSc Thesis, 2008.

53 Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., *Advanced Drug Delivery Review* **2001** 46: 3-26.

54 Nchinda A. N.;Chibale, K.;Redelinghuys, P.; Sturrock, E.D. *Bioorganic & Medicinal Chemistry Letters* **2006**, 16(17), 4612-4615.

55 Baylis, E. K; Campbell, C.D.; Dingwall, J.G. *Journal of theChemical Society, Perkin Transactions 1: Organic and Bio-Organic Chemistry1972-1999*,**1984**; 12: 2845-53.

56 Wu, J; Sun, W; Xia, H-G; Sun, X. *Organic& Biomolecular Chemistry* **2006**; 4:1663-1666.

Chapter 5

Biological Assays

5.1 Introduction

This thesis has thus far described the synthesis of various hydroxamate reactive groups and their respective probes. It has also given a synthesis of a series of new inhibitors with an attenuated phosphinic acid ZBG. Further, it has described the synthesis of analogues of current generation ACE inhibitors Captopril and Enalaprilat. As clearly outlined in Chapter 4, the discovery of current generation ACE inhibitor was not all rational but involved elements of serendipity. Another famous example of the latter was the discovery of penicillin in 1928 by Alexander Fleming[1] and therefore, it remains an integral part of drug discovery. Consequently, a small number of hydroxamate inhibitors were synthesized as described in Chapter 4 from the condensation of the hydroxamate reactive groups meant for the probes to tryptophan. The biological results for these compounds are also reported.

As discussed in earlier Chapters, it should be noted that the biological assay results given herein include all except those for the hydroxamate based probes for the malaria disease model. Further, with the exception of the Captopril and Elalaprilat analogues, all the compounds were screened as stereoisomeric mixtures. Pure stereoisomers could not be obtained because there was no appropriate equipment to carry out the separations. However, this was not an isolated case as stereoisomers have been screened in assays in the literature.

The importance of ACE in human homeostasis as well as the relevance of ACE inhibitors has been discussed briefly in Chapter 1 and given in greater detail in Chapter 4. The biological results of all the assays carried out on the C- and N-domain constructs of this enzyme with all the synthesized compounds will be given in this Chapter in four sections beginning with screening of the suitability of the initially proposed hydroxamate reactive groups as well as those coupled to tryptophan. This will be followed by the assays carried out to establish whether the synthesized ABPs for the hypertension model would bind and inhibit ACE. Thirdly, an examination of the inhibitory and domain selectivity profiles of the new compounds in the attenuated phosphinic acid series will be carried out. Fourthly, the inhibitory and domain selectivity profiles of Enalaprilat and Captopril analogues modified by the replacement of the P_2' prolyl moiety with a tryptophan will examined.

5.2 Preparation of the Inhibitor Compounds

The inhibitors (approx. 1mg) were dissolved in an appropriate volume of DMSO to give a 20mM inhibitor stock solution. The solution was then further diluted with DMSO to yield a 50uM stock solution. An appropriate volume of Phosphate buffer pH 8.3 containing 0.1M K_2HPO_4/KH_2PO_4, 0.3M NaCl and 10 μM $ZnSO_4$ was then added resulting in a working solution of 250nM of inhibitor.

5.3 Enzymes and substrate

For the evaluation of ACE selectivity and activity, the inhibitors were tested with *t*ACEΔ36-g13. [2] This is a variant of testis ACE lacking the 36 N-terminal residues unique to

this isoform as well as the transmembrane and cytosolic regions. Thus, it is identical to the C-domain of somatic ACE) and purified *N*-domain (a minimally glycosylated construct of the N-domain of somatic ACE generated by site-directed mutagenesis)[3] protein constructs. Both these proteins used in this study were expressed and purified by Ms S.L.U. Schwager of the Zinc Metalloprotease Research Group. A 1.05mg/mL stock solution of *N*-domain enzyme was diluted in an appropriate volume of buffer to give a 10 nM enzyme working solution while the C-domain working solution was diluted in buffer to a working solution of 25nM concentration.

5.4 Z-FHL Substrate Preparation

ACE activity was assayed using the substrate benzyloxycarbonyl-phenylalanyl-histidyl-leucine (Z-FHL) (Bachem, Budendorf, Switzerland) as described previously. [4] Z-FHL is a synthetic peptide, a substrate cleaved efficiently by both N- and C-domains. A 20mM stock solution of *Z*-FHL was made up by dissolving 110 mg of *Z*-FHL in 1ml 0.28 M NaOH and 9 ml distilled water was added to make a final volume of 10ml. The working stock solution of 1mM was prepared by dilution with an appropriate volume of stock buffer solution (0.5M K_2HPO_4; 0.5M KH_2PO_4, pH 8.3; 1.5 M NaCl, 10 μM $ZnSO_4$) and distilled water to which 2ml Z-FHL stock solution was added.

5.5 ACE Assay

ACE activity was determined using a spectrofluorimetric assay. It employs the tripeptide substrate *Z*-FHL. The assay is based on the principle that ACE cleaves penultimate

peptide bond to yield an HL moiety that forms a fluorescent adduct upon addition of o-pthaldialdehyde.[4] The amount adduct produced is directly proportional to the amount of dipeptide cleaved by ACE and compared to activity of the enzyme with a known inhibitor as well as enzyme only with no inhibitor as controls.

The screening assay at 250nM is a semi-highthroughput screen which gives a read-out of the percentage reduction of enzyme activity in the presence of the potential inhibitor compared to control experiments which are without any inhibitor as a negative control (100% control) as well as a positive control using a known inhibitor such as Captopril or Lisinopril for comparative purposes. Individual ACE domains namely C- and N-domain are incubated with a particular inhibitor 40uL of inhibitor working solution was pre-incubated with 40uL of individual ACE domains namely C- and N-domain of enzyme and incubated at room temperature for 15 minutes.

Each experiment was done in triplicate by adding 20uL of the pre-incubated solution to each well. 20uL of the 1mM Z-FHL was then added to each well. The plate was then incubated at 37°C in a water bath for 15 minutes. Thereafter, each reaction was stopped with 185uL stop solution (170uL 0.28M NaOH and 15uL o-pthaldialdehyde 20mg/ml in methanol) and derivatised for 10 minutes at room temperature. Derivatisation was stopped with 25uL of 3N HCl. Fluorescent intensities were then measured using a Cary Eclipse Varian Fluorescence Spectrophotometer with excitation setting at 360 nm and emission at 485 nm (slit width 10x10).

5.6 Data Analysis

Graphs of percentage remaining activity of each individual ACE domain in the presence of potential inhibitor were plotted and analyzed using GraphPad Prism 4.01 (GraphPad Software, La Jolla, CA) and/or Microsoft Excel®. The enzyme without inhibitor was used as a control and deemed to be 100% enzyme activity. Experiments were repeated and each assay was performed in triplicate. The value resulting from the percentage activity remaining for the N-domain as numerator divided by that of the C-domain as denominator describes the selectivity for the C-domain for each inhibitor. The larger this figure is, the more C-domain selective the inhibitor appears to be.

5.6.1 Introduction to Assay Results

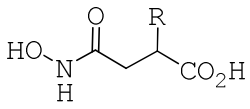
These assays served to determine the inhibitory profiles and/or C-domain selectivity of the synthesized compounds at a fixed 250nM compound concentration with each domain of ACE. Bar charts depicting percentage activity of control (with no inhibitor) compared with of the 250 nM of each compounds incubated with the enzyme were plotted. Here, a reduction in enzyme activity to 90% for instance is described as 10% inhibition. This provided comparison of the ACE inhibition potencies of the compounds as well as selectivity for either the C- or N-domain. Compounds exhibiting good inhibitory potency at this concentration were to then be carried through for IC_{50} determination with individual ACE-domains.

5.6.2 *In Vitro* Assay results for hydroxamate reactive groups 3.15–1 to 3.15–6.

The title compounds have been evaluated for their biological activity competitively inhibiting ACE at 250 nM concentration with reference to incubation with no compound. The percentage remaining activity of the enzyme was calculated and is presented in Table 5.1. The results indicated that almost all these potential hydroxamate "warheads" had low to moderate inhibition of ACE at this screening concentration. The compound showing the best inhibition of enzyme activity was found to be **3.15–6** with 30% inhibition of the N-domain activity (equal to 70% remaining enzyme activity) and 15% of the C-domain compared to Captopril which had 97% control in the N- and 98.5% in the C-domain. Compound **3.15–6** had a napthyl subtituent in the R-position. The rest of the compounds namely **3.15–1** to **3.15–5** (with Cyclohexyl, Isopropyl, methyl and phenyl R-substituent) all generally had less than 15% inhibition in both domains with no significant preference for either domain. This inhibition was not as robust as desired in a reactive group to be employed in a probe. It was concluded that although hydroxamate compounds with related structural features have been reported to be active against P*f*A-M1 (and were used in Malaria probes in this thesis) they would not be effective in the hypertension targeting ACE.

Table 5.1: Effect of compounds 3.15-1 to 3.15-6 (250nM) on ACE percentage activity of control compared to no compound at all.

Compound	**% Activity remaining**		**Selectivity**
	N	**C**	
Enzyme only	100	100	1.0
3.15-1	85	95	0.89
3.15-2	95	95	1.0
3.15-3	95	95	1.0
3.15-4	90	90	1.0
3.15-5	90	100	0.9
3.15-6	70	85	0.82
Captopril	3	1.45	0.5



3.15

3.15-1 R = H
3.15-2 R = Me
3.15-3 R = Isopropyl
3.15-4 R = CH2Cyclohehyl
3.15-5 R = Benzyl
3.15-6 R = CH2Napthyl

Figure 5.1: General structure of compounds **3.15-1 to 3.15-6**

5.6.3 *In Vitro* Assay results for hydroxamate series 4.7a to 4.7d.

The title compounds have been evaluated for their biological activity competitively inhibiting ACE at 250nM as well as 5uM concentration with reference to incubation of enzyme with no compound or with known ACE inhibitor Lisinopril as negative and positive controls respectively. The percentage remaining activity of the enzyme at this concentration for these hydroxamic acid derivatives were calculated and are presented in Tables 5.2 and 5.2b. For all the compounds the results indicated very low (less than 25% inhibition) to moderate inhibition (more than 25 but less than 50%). Further, there was no significant

difference observed between results obtained in either the N- or C-domain.

Table 5.2: Z-FHL assay of hydroxamate series cross-linked to tryptophan at 250nM inhibitor concentration

Compound	% Activity Remaining		Selectivity
	N-Domain	C-domain	
Lisinopril	3	1.8	1.67
4.7a	95	88	1.08
4.7b	90	100	0.9
4.7c	95	80	1.19
4.7d	95	60	1.58
Enzyme	100	100	1

Table 5.2b: Z-FHL assay of hydroxamate series crosslinked to tryptophan at 5uM inhibitor concentration

Compound	% Activity Remaining		Selectivity
	N-Domain	C-domain	
Lisinopril	ND	1.8	ND
4.7a	ND	82	ND
4.7b	ND	100	ND
4.7c	ND	91	ND
4.7d	ND	14	ND
Enzyme	ND	100	ND

ND: Not determined.

The potential hydroxamate "war heads" **3.15-1** to **3.15-6** presented poor biological results. This could possibly be due to their inability to interact effectively in the S_1-, S_1'- and S_2'-subsite pockets of the inhibitors. The R-group present may only be interacting with the S_1' but the lack of any interaction especially with the all important S_2'-

subsite could have rendered them inactive. Prompted by these results, a tryptophan residue was cross-linked to these structure resulting in preparation of **4.7a, 4.7b, 4.7c** and **4.7d.** We envisaged that the added S_2' interaction would result in better biological results which would then make them suitable for the preparation of future hydroxamate based probes to apply in the hypertension (ACE) model. The results obtained with these compounds however also presented a low to moderate inhibitory profile. Compound **4.7d** showed the best promise with a 60 percent reduction in ACE activity at 250nM. As presented above, at a concentration of 5uM, this compound had a percentage reduction of activity of 86%. This compound **4.7d** has a bulky Napthyl R-group in the P_1 position where its analogues **4.7a, 4.7b** and **4.7c** had a cyclohexyl, isoleucine, and a phenylalanine moiety in the same position. A recent study by Akif *et al*[5] suggests that very bulky groups situated in the P_1' position are unexpectedly oriented towards the deeper S_2' pocket which may rationalize the accommodation of such a large side chain. In this case, **4.7d** has a considerably bulkier R-group which could force its way into the S_2' explaining why this compound exhibited slightly better control over the C-domain in the assay. However, the difference here is that Akif *et al* had a bicyclic structure with isoxazole heteroatoms nitrogen and oxygen. These were found to have interactions with an Asp^{415} and His^{383} respectively.

5.6.4 *In Vitro* Assay results for ABP's 3.32a and 3.32b

The title compounds have been evaluated for their biological activity competitively inhibiting C-domain at 5 uM concentration with reference to incubation of enzyme

with no compound. The percentage remaining activity of the enzyme at this concentration for the Enalaprilat probe **3.32a** and Captopril probe **3.32b** was calculated and is presented in Table 5.3. In this case, inhibitor Captopril, the enzyme on its own and a buffer only were also used as controls to observe change in activity.

The results at 5 uM indicated that the Enalaprilat probe **3.32a** had very low inhibition of ACE (10%) whereas the Captopril probe **3.32b** had moderate (60%)inhibition leaving an apparent 40% ACE activity after 90 minutes incubation.

Table 5.3: Z-FHL assay of enalaprilat probe **3.32a** and Captopril probe **3.32b** at 5 uM

Compound	**% Activity remaining** **C-domain**
Enzyme Only	100
Enalaprilat probe **3.32a**	90
Captopril Probe **3.32b**	40
Captopril	5

From these results, the title compounds were further evaluated for their biological activity competitively inhibiting *t*-ACE at 250nM concentration with reference to appropriate negative and positive controls. The reduction in concentration was done mainly to investigate whether a similar inhibitory pattern would be observed at this lower concentration. The percentage remaining activity of the enzyme at this concentration for the Enalaprilat probe **3.32a** and Captopril probe **3.32b** was calculated and is presented in Table 5.4. Like previously, inhibitor Captopril and the enzyme on its own with buffer only were also used as controls to observe change in activity.

The results indicated that the Enalaprilat probe **3.32a** had completely no control of *t*-ACE whereas the Captopril probe **3.32b** had very low control leaving >90% ACE activity after 90 minutes incubation.

Table 5.4: Z-FHL assay of Enalaprilat probe **3.32a** and Captopril probe **3.32b** at 250nM

Compound	**% Activity remaining C-domain**
Enzyme Only	100
Enalaprilat probe **3.32a**	100
Captopril Probe **3.32b**	90
Captopril	5

As highlighted in **Chapter 1** of this thesis, principally, the function of an ABP relies on the inhibitor moiety. However, the linker chain, the benzophenone photo-crosslinker and biotin reporter tag inevitably modify the structure and physicochemical features of the inhibitors Enalaprilat and Captopril, leading to obvious changes in size, bulk and hydrophobicity of the molecules. Furthermore, the linker was attached to the C-terminus of the inhibitor where a free carboxylate has been described to have an important interaction with enzyme active-site Lys511. The latter and the former may have reduced the inhibitor ability to access and bind the enzyme active-site, hampering its application for *in vitro* binding resulting in low or no inhibition observed in the screening assay.

To overcome this hurdle, it is proposed that a click-chemistry probe be designed and would be more appropriate. This should be an inhibitor-linker combo ABP which is untagged but with a small chemical adapter for the

attachment of a reporter tag after labeling. This "tag-free" ABP would maximally resemble the prototype inhibitor and minimally affects target binding *in vitro*. Alkyne and azide are two examples of such adapters. Following covalent labeling of protein targets, the probe may be ligated to the reporter tag *in vitro*, through appropriate chemical techniques.[6],[7] Chemical adapters such as those mentioned above are inert in assay aqueous media while their coupling is quick, efficient and specific under mild conditions in the presence of water. Another advantage of this two-step labeling strategy is that it would allow one probe to accept diverse reporter tags, and simplifies probe design and synthesis.

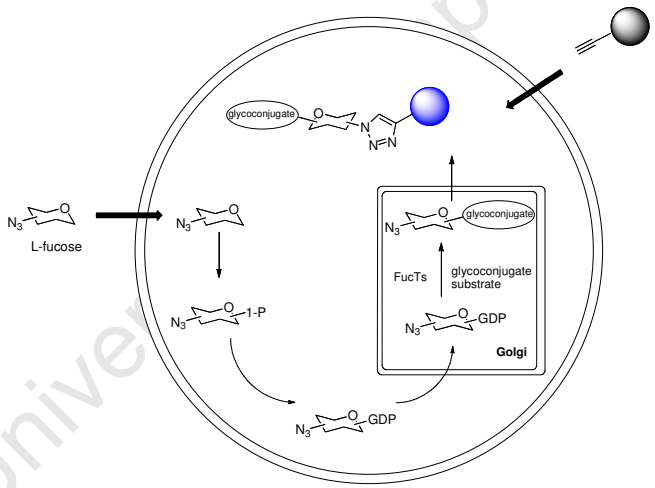


Figure 5.2: An illustration of the application of a click chemistry probe.[7]

5.6.5 *In Vitro* Assay results for Enalaprilat and Captopril analogues 4.4d, 4.4e, 4.5d and 4.5e.

The title compounds have been evaluated for their biological activity competitively inhibiting the N- or C-domains of ACE at 250nM concentration with reference to a

negative and positive controls, enzyme with no compound and Captopril. The percentage remaining activity of the enzyme at this concentration for the Captopril and Enalaprilat analogues **4.4d**, **4.4e**, **4.5d** and **4.5e** were calculated and are presented in Table 5.5. All four compounds exhibited approximately 20% activity reduction in the N-domain at this concentration. In the C-domain however, the figures observed were 0, 30, 20, 0% percentage control of activity for **4.4d**, **4.4e**, **4.5d** and **4.5e** respectively. In terms of significance of inhibition, these results were very low compared to the control, Captopril. Furthermore, as highlighted above, there is no significant preference for either the N- or C-domain in all the compounds at 250nM concentration.

Table 5.5: Percentage activity remaining of ACE compared to Captopril as a standard

Compound	**% Activity remaining**		**Selectivity**
	N-domain	**C-domain**	
Enzyme	100	100	1.0
4.4d	80	100	0.8
4.4e	80	70	1.14
4.5d	80	80	1.0
4.5e	80	100	0.8
Captopril	<1	<1	1

One hypothesis of this research was that the tryptophan residue of RXPA380 may not be the only structural determinant of C-domain selectivity. Consequently, established ACE inhibitors Enalaprilat and Captopril were modified by replacement of the prolyl moiety with the bulkier tryptophan residues as is found in RXPA380 as well as replaced with a phenylalanine residue. Any change in

selectivity between the N- and C-domain would provide insight into this domain selective-behaviour and effect of these residues. A previous study in our group (Zinc Metalloprotease Group) on Lisinopril involved a similar strategy as well as molecular docking studies.[8] It was found that the S_2' subsite of the enzyme readily accommodated the tryptophan moiety of the Lisinopril derivative a finding which confirmed earlier studies highlighting the importance of a deep S_2' pocket as a determinant of C-domain-selectivity. [9] It was expected therefore that with the tryptophan moietiy, additional hydrophobic and hydrogen-bonding interactions with S2' residues Thr282, Val379, Val380, and Asp453 of the C-domain active site (Fig. 2), and may increase the C-domain-selectivity of these compounds as compared with Enalaprilat and Captopril.

However, the results indicated very low to moderate inhibition and no significant difference between the N- or C-domain selectivity. These results could be attributed to the inevitable changes in the 3-dimensional structure and physicochemical properties of these compounds compared to Enalaprilat and Captopril. The latter could have caused different molecular orientation in the active site and sub-site residues resulting in little or no inhibition.

5.6.6 *In Vitro* Assay results for Phosphinic acids 4.2i-1 to 4.2i-16.

The rationale behind the design of these phosphinic acids was based on minimizing or eliminating the interaction with the catalytic Zn^{2+} ion to achieve domain selectivity. The metal site is the most conserved feature between the N- and

C-domains of ACE. Phosphinates were chosen because of the advantageous placement of the ZBG in the middle of a pseudopeptidic scaffold and not at its N- or C-terminus, as in the cases of hydroxamate, carboxylate and sulfhydryl inhibitors. Then R-groups were strategically placed in the positions adjacent to the ZBG to partially mask and consequently attenuating the ZBG. Structural diversity was introduced by placing various groups in the P_1 and P_1' moeities of the potential inhibitors. The title compounds were evaluated for their biological activity competitively inhibiting ACE at a fixed concentration of 250nM with control reference to incubation of enzyme with no compound as negative control and Lisinopril being the positive control. The percentage remaining activity of the enzyme at this concentration for these phosphinic acid derivatives were calculated and are presented in Table 5.6.

These inhibition studies for the compounds showed that **4.2i-7** had the best efficiency to inhibit the C-domain at 53%, and 21% in the N-domain. Compounds **4.2i-4**, **4.2i-10**, **4.2i-11**, **4.2i-14**, **4.2i-15** and **4.2i-16** exhibited 33, 34, 32, 39, 36 and 36% inhibitory efficiency in the C-domain compared to 0, 4, 11, 22, 18 and 27% respectively in the N-domain. Compounds **4.2i-5** and **4.2i-8** showed generally little or no activity exhibiting between 0 and 4 % inhibition in the C- and 0 and 8% in the N-domain. Others were compounds **4.2i-6**, **4.2i-9**, **4.2i-12** which exhibited minimal inhibition (27, 19 and 25%) in the C-domain and completely no inhibition in the N-domain.

It is worth noting that although all the compounds did not have excellent inhibitory potency, it was observed

generally that the compound with the best efficiency to inhibit the C-domain of ACE, namely, **4.2i-7** and those with efficiency closely related to it such as **4.2i-11** and **4.2i-15** all have the phenylalanine moiety in the P_1-position. From the X-ray crystalstructure of RXPA380 in the C-domain, this bulky group interacts with the S_1 subsite of the C-domain and possibly hydrophobic interactions occurring with the Val518 residue which is replaced by a Tyr496 in the N-domain. Generally, the S_1 in the C-domain construct is more hydrophobic than that in the N-domain. This phenylalanine feature of the molecule is therefore expected to be better accommodated by the C-domain. However, at best, only a 1.68 preference was observed for the C- over the N-domain in **4.2i-7**. This could also explain the reason why compounds such as **4.2i-4**, **4.2i-10**, **4.2i-14** and **4.2i-16** also exhibited a slight preference of the C- over the N-domain. The latter compounds had either a propyl chain or a tert-butyl moiety which was possibly stretching more to the part of the S_1 subsite where the VAL518 is located. **4.2i-5** and **4.2i-9** which both have only a methyl group in the same position showed the least measure of control as well as difference in domain preference.

Both in terms of selectivity as well as inhibition, it seemed the best combination of R-groups was the phenyl alanine in the P_1 position and a propyl moiety in the P_1' position. Generally however, bulky hydrophobic groups performed better than the smaller methyl groups in the P_1-position as was seen with compounds **4.2i-2**, **4.2i-4**, **4.2i-7**, **4.2i-11**, **4.2i-10**, **4.2i-14**, **4.2i-16**. Compounds with the smaller methyl groups in the P_1 position such as **4.2i-5** and **4.2i-9** had no inhibition in the N- domain and approximately

4% control in the C-domain. In the P1' position however, no particular pattern was observed and both bulkier and smaller groups were reasonably accommodated with activity control in the C-domain of up to 40% at 250nM concentration being observed in compounds with both a small and/or larger moiety. Going by this observation, it may be suggested that the S1'subsite is able to tolerate a wider range of substituents and it would be worthwhile in future to conduct an SAR study with groups of diverse electronic configuration such as heteroatoms as well as electron donating and withdrawing group.

Table 5.6: Z-FHL assay of phosphinic acid series at 250nM inhibitor concentration

Compound	R1	R2	% Activity Remaining		Selectivity
			N-	C-domain	
Lisinopril			**3.2**	**1.8**	**~1.8**
4.2i-1	CH_3	CH_3	**ND**	**ND**	**ND**
4.2i-2	$C(CH_3)_3$	CH_3	**100**	**60**	**1.67**
4.2i-3	CH_2Ph	CH_3	**ND**	**ND**	**ND**
4.2i-4	$CH_2CH_2CH_3$	CH_3	**100**	**67**	**1.49**
4.2i-5	CH_3	$CH_2CH_2CH_3$	**100**	**96**	**~1**
4.2i-6	$C(CH_3)_3$	$CH_2CH_2CH_3$	**100**	**73**	**1.37**
4.2i-7	CH_2Ph	$CH_2CH_2CH_3$	**79**	**47**	**1.68**
4.2i-8	$CH_2CH_2CH_3$	$CH_2CH_2CH_3$	**91**	**100**	**0.91**
4.2i-9	CH_3	$C(CH_3)_3$	**100**	**81**	**1.2**
4.2i-10	$C(CH_3)_3$	$C(CH_3)_3$	**96**	**66**	**1.45**
4.2i-11	CH_2Ph	$(CH_3)_3$	**89**	**68**	**1.31**
4.2i-12	$CH_2CH_2CH_3$	$(CH_3)_3$	**100**	**75**	**1.36**
4.2i-13	CH_3	CH_2Ph	**ND**	**ND**	**ND**
4.2i-14	$C(CH_3)_3$	CH_2Ph	**78**	**61**	**1.28**
4.2i-15	CH_2Ph	CH_2Ph	**82**	**64**	**1.28**
4.2i-16	$CH_2CH_2CH_3$	CH_2Ph	**73**	**64**	**1.14**
Enzyme			**100**	**100**	**1**

ND: Not determined. Insufficient purity

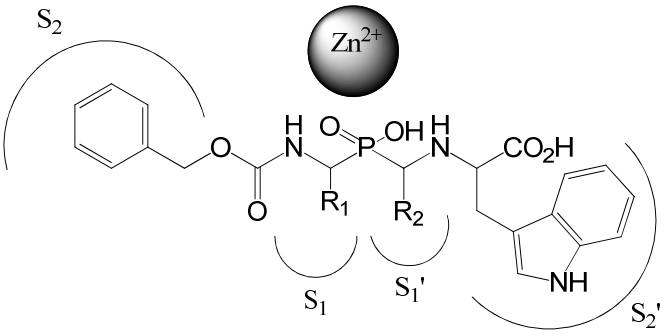


Figure 5.3: General structure for the design of target compounds depicted in the active site with varied R-groups as per the Table 5.6.

From the foregoing, it may further be suggested that ZBG is an essential feature for effective binding in the ACE active site. This being the case, it is proposed that in future, exploration and structure activity relationship studies be pursued with cyclic bioisosteres such as 1,2,4- and 1,3,4-Oxadiazoles and Oxathiazoles in attempts to retain the zinc binding capability while reducing the dependence on it for binding. Another angle would be the inclusion of a non sterically hindered (or attenuated) analogue without groups in the positions adjacent to the ZBG with an identical carbon chain backbone that is a good inhibitor as a positive control.

Generally for existing ACE inhibitors, it would also be important to carry out an SAR investigation on the influence of heteroaromatic groups as well as a small variety of electron donating and withdrawing substituents in the R-groups as these would possibly pick up more interactions in the amino acid residues of the active site. Consideration of such changes are highly recommended in drug discovery. [10] Another proposal would be to carry out

high through-put biological screens of compound libraries with the separate domains of ACE. The latter has the potential of revealing possible domain selective compounds and/or resulting in completely new scaffolds on which to base future hypertension ACE-inhibitor drug discovery. Furthermore, carrying out a fragment based modelling may lead to identification of potential novel fragments.

5.7 Future Biological Work

Validation of assay results above was attempted at different concentrations. 5uM concentration was also attempted in order to try and observe enhanced inhibition of the enzyme by the inhibitors. The phosphinic acid derivatives were problematic in that they precipitated out of buffer solution at concentrations higher than 250nM. Care was taken to ensure not more that 1% DMSO was in the assay media and more of the buffer was used, but this generally resulted in the compounds precipitating out of solution. This work was not pursued further due to time constraints. However, as a way forward in the design of these and new inhibitors, appropriate solubility studies be carried out. The solubility problems faced in this research may be approached from several angles. Firstly, appropriate modification of the dilution protocol can be made, or secondly detailed assessment of compound solubility can be conducted, thirdly, the ACE assay can be optimized for less soluble compounds such as these. Finally during the inhibitor design, moieties with hydrogen bond donors and acceptors or other groups that can enhance solubility should be incorporated in the scaffold.

Furthermore, an SAR study is recommended for the hydroxamates cross-linked to a tryptophan residue above and when compounds are identified that would reduce the percentage remaining activity to the region of 10% or lower, IC_{50} values would then be determined. The IC_{50} is a term used to describe the inhibitor concentration required to cause a 50% decrease in enzyme activity. The lower the IC_{50} value of a compound, the greater the inhibition. A sigmoidal-dose response curve analysis is used to determine the IC_{50} values.

This chapter has given biological assay results as well as offered a biological discussion with regards to all compounds assayed. Although very good inhibition amy not have been observed, in-roads have been established on the design and synthesis of new probes for use in the hypertension disease model. Apart from this, new compounds have been synthesized and assayed and have revealed the effects of carrying out certain modifications on existing as well as new inhibitors. Furthermore, synthesized aminophosphinates have been employed successfully in biological evaluation of the C- and N-domain constructs of Angiotensin-converting enzyme.

References

1 *Encyclopedia Britannica* Vol 9. Fleming, Sir Alexander. Chicago, Ill; London, UK; Toronto, Canada; Geneva, Switzerland; Sydney, Australia; Tokyo, Japan; Manila, the Philippines: William Benton; 1969;437.

2 Gordon K.; Redelinghuys, P.; Schwager, S.L.U.; Ehlers, M.R.;Papageorgiou, A.C.; Natesh, R.; Acharya, K.R.; Sturrock, E.D. *Biochem J.* **2003** 15; 371 (Pt 2) 437–442.

3 Anthony, C.S.; Corradi, H.R.; Schwager, S.L.U.; Redelinghuys, P.; Acharya, K.R.; Sturrock, E.D. *J. Biol. Chemistry* **2010** 285(46), 35685–93.

4 Friedland, J.; Silverstein, E. *American Journal of Clinical Pathology* **1976** 66,416–424.

5 Akif, M., Schwager, S.L., Anthony, C.S., Czarny, B., Beau. F., Dive, V., Sturrock, E.D., Acharya, K.R. *Biochem. J.* **2011** 436 (1) :53–9.

6 Speers, A. E.; Cravatt, B. F., *Chemistry & Biology*, **2004** 11, 535–546.

7 Wong, C–H., *Proceedngs of the National Academy of Sciences, USA* 2006, 103, 12371–12376.

8 Nchinda, A. T.; Kelly Chibale, K. ; Redelinghuys, P.; Sturrock, E. D., Bioorganic & Medicinal Chemistry Letters, **2006**, 16, 4616–4619

9 Georgiadis, D.; Cuniasse, P.; Cotton, J.; Yiotakis, A.; Dive, V. *Biochemistry* **2004**, 43, 8048.

10 Celassie, C.D. *Burger's Medicinal Chemistry and Drug Discovery, 2003,* Sixth Edition, Volume 1: Drug Discovery Edited by Donald J Abraham ISBN 0–471–27090–3 © 2003 John Wiley & Sons, Inc.

Chapter 6

Conclusion

6.0 Conclusion

The simplicity of the chemistry employed during the synthesis, coupled with the commercial availability of reagents in all synthetic areas makes the synthesis of the probes, phosphinic acids, hydroxamates, Enalaprilat and Captopril derivatives an attractive route in the development of domain-selective inhibitors.

ABPs have been introduced in the arena of hypertension research through a captopril-linked probe among others that can be potentially employed to label whole proteomes in the elevated hypertension disease condition compared to the normal condition. These newly synthesized probes did not show significant inhibitory results probably due to altered physicochemical properties and bulk. Proposals for future work to cross these hurdles have been forwarded in the biological discussion given in Chapter 5 in the form of "click chemistry" probes. Synthetic routes for the synthesis of new probes for future use in malaria research have also been successfully developed.

New phosphinic acid-based ACE inhibitors with a substituent in the α-position relative to the phosphinate ZBG as potential novel antihypertension agents have been synthesized via three-component pseudo-Kabachnik-Fields reaction involving an aldehyde, amine and phosphoryl compound. It should be noted however that the purification of the resultant phosphinate derivatives was challenging

because some of the reactions never went to completion and resulted in multiple products forming. The desired compounds had to be carefully isolated using prep TLC or reversed phase HPLC using appropriate solvent systems.

New potential ACE-inhibitors incorporating a tryptophan residue in the P2'of the captopril and enalaprilat scaffold have also been prepared. Although the results ranged from low inhibition to no inhibition at all, biological assays have been carried out to evaluate the domain selectivity of all synthesised compounds *in vitro* using appropriate ACE inhibitory assays.

The sulfhydryl and carboxylate ABP's, hydroxamates as well as the phosphinic acids were evaluated with enzymatically active purified C- and N-domain constructs. Although the carboxylic acid probe **3.32a** showed no inhibition, the sulfhydryl probe **3.32b** exhibited an apparent IC_{50} of 25 µM.

For the hydroxamate inhibitors, compound **4.7d** was found to inhibit ACE by 86% at 5 µM. Amongst the phosphinic acid compounds with the attenuated zinc-binding group, **4.2i-7** inhibited the C-domain construct by 53% whereas **4.2i-2**, **4.2i-14**, and **4.2i-16** inhibited by 40, 39 and 36% respectively at 250 nM The different groups adjacent to the ZBG's had varying degrees of effect on the activity of the compounds with the most active **4.2i-7** having a pseudo-phenylalanine moiety in the α- and a propyl group in the α'-position relative to the phosphinic acid. On the other hand, compared to the 99% inhibition by Captopril in both domains, Enalaprilat analogue **4.4e** inhibits by 30 and 20% in the C- and N domains where Captopril analogue **4.5d**

inhibit both domains by 20% respectively at 250 nM. None of the compounds showed significant selectivity for the C- or N-domains.

Finally, in retrospect, a lesson learnt from this research is that instead of carrying out biological assays at the end of the synthesis, an iterative cycle of synthesis and design, guided by biological results would have been more instructive.

Chapter 7

General Experiment

All commercially available chemicals were purchased from either Sigma-Aldrich or Merck, South Africa, unless otherwise stated. Reactions were monitored by thin layer chromatography using Merck F_{254}aluminum-backed pre-coated silica gel plates and were visualized by ultraviolet light or iodine vapor. Silica gel chromatography was performed using Merck kieselgel 60: 70-230 mesh for gravity columns. Melting points were determined on a Reichert-Jung Thermovar hotstage microscope and are uncorrected. Microanalyses were determined using a FisonsEA 1108 CHNO-S instrument. Low Resolution Mass Spectra were recorded on an API 4000 triple quadropole (Applied Biosystems), fitted with a Turbo Ion spray source, samples were introduced via Harvard Syringe Pump at 10ul/min. High resolution MS data are not given as the facility was not available.

Purity was routinely measured by HPLC on a Thermo Separation Products (Spectra Physics) reversed phase HPLC instrument with a P200 Gradient Pump and a UV2000 UV/VIS detector fitted with a Technokroma C_{18} semi-preparative Tracer EXcel 120 ODSB 5mm 20 x 1.0 column.Mobile phases: 0.1% TFA in millipore purified water (A) and 0.1% TFA and 75% HPLC grade acetonitrile in Millipore purified water (B). A gradient was formed from 80% of A in 5 minutes and held at 80% for 30 min, then 100 B for 3 minutes and back to 80% A for two minutes; flow rate was 0.3 mL/min and injection volume was 50 µL; retention times (HPLC, tR) are

given in minutes. Compound HPLC purity was determined by monitoring at 215 and 280 nm.

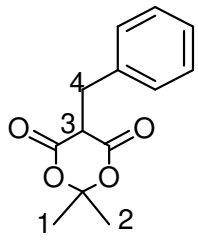
^{1}H NMR spectra were recorded on a Varian Mercury spectrometer at 300 MHz, a Varian Unity Spectrometer at 400 MHz and a Bruker Systems Avance III 400MHz instrument. All spectra were recorded in deuterochloroform, deuterodimethylsulfoxide, deuteromethanol or D_2O downfield from tetramethylsilane (TMS) as an internal standard. All chemical shifts are recorded in ppm. H-bonding between compounds containing NH_2 or OH groups and the deuterated methanol or deuterated water resulted in a reduction of hydrogen signals observed in ^{1}H NMR spectra.**^{1}HNMR coupling constants are measured in Hertz and are rounded off resulting in exactly the same coupling constants for coupling partners.**

^{13}C NMR spectra were measured on a Varian Mercury spectrometer at 75 MHz or a Varian Unity spectrometer at 100 MHz or a Bruker Systems Avance III 400MHz instrument. **The format used for recording ^{13}C NMR data is that accepted by most international journals (including American Chemical Society journals). In this format chemical shift values are simply listed without specific assignment to carbon atoms.**

Dry solvents were prepared by appropriate techniques. Tetrahydrofuran was dried over lithium aluminium hydride and subsequently distilled over sodium wire prior to using benzophenone as indicator. Methylene chloride was dried over phosphorous pentoxide and distilled prior to use. Unless otherwise stated, all other solvents used were AR grade solvents.

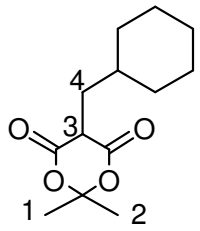
General procedure for the synthesis of compounds 3.3b-3, 3.3b-4, 3.3b-5 and 3.3b-6

To a stirred solution of 2,2-Dimethyl-1,3-dioxane-4,6-dione (2.16g, 15.0 mmol) in 40 ml EtOH was added K_3PO_4 (0.6 g, 2.82 mmol) followed by a solution of the appropriate aldehyde (15.0 mmol) in 5 ml EtOH. Mechanical stirring was continued until completion of the reaction (TLC) and to it $NaBH_4$ (1.2g, 31.50 mmol) was added. On completion of the reaction (TLC) the solvent was removed in vacuo. H_2O (50 ml) was added and the mixture acidified to pH_4 with 1*N*HCl. This mixture was then separated with EtOAc. The organic layer was washed several times with H_2O (4 x 50 ml) and dried over $MgSO_4$. After evaporation of the solvent in vacuo and recrystallization from EtOAc/Hexanes 5-substituted-2,2-dimethyl-1,3-dioxane-4,6-dione's **3.3b-3, 3.3b-4, 3.3b-5** and **3.3b-6** were obtained.



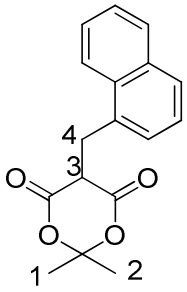
(±)5-Benzyl-2,2-dimethyl-1,3-dioxane-4,6-dione 3.3b-5

Yield:91%; Colourless crystals from EtOAc-Hexane; m.p. 80-81 °C (Lit.[1] m.p. 82-84); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.7 (s, 6H, H-1/2), 3.4 (d, *J* = 12 Hz, 2H, H-4), 4.0 (t, *J* = 7 Hz, 1H, H-3), 7.25 (m, 5H, PhH).



(±)5-(Cyclohexylmethyl)-2,2-dimethyl-1,3-dioxane-4,6-dione 3.3b-4

Yield: 79 %; colourless crystals from EtOAc-Hexane; m.p. 110-112 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H0.95 (m, 2H, Cyclo-CH), 1.3 (m, 4H, Cyclo-CH), 1.7 (s, 6H, H-1/2), 1.8 (m, 5H, Cyclo-CH), 2.0 (dd, *J* = 11, 4 Hz, 2H, H-4), 3.6 (t, *J* = 7 Hz, 1H, H-3); ^{13}C NMR (100 MHz, $CDCl_3$) δ_C 26.0, 26.3, 27.1, 27.3, 28.6, 32.9, 33.3(2C), 33.8, 44.1, 104.5, 168.1, 168.6; m.p. 110-112 ˚C.



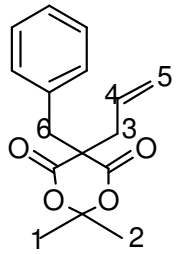
(±)2,2-Dimethyl-5-(naphthalen-1-ylmethyl)-1,3-dioxane-4,6-dione 3.3b-6

Yield: 90 %; Colourless crystals from EtOAc-Hexane; m.p. 91 - 94 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H1.7 (s, 6H, H-1/2), 3.15 (d, *J* = 9 Hz, 2H, H-4), 4.1 (t, *J* = 7 Hz, 1H, H-3), 7.3 (m, 7H, Naphthyl-*H*); MS m/z 284.31.

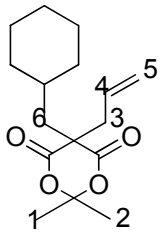
General procedure for the preparation of compounds 3.5a-3, 3.5a-4, 3.5a-5 and 3.5a-6

5-substituted-2,2-dimethyl-1,3-dioxane-4,6-dione (11.00

mmol) was dissolved in DMF (15 ml). K_2CO_3 (2.073, 15.00 mmol) was then added and stirred for 10 min at r.t. Allyl bromide (1.3 ml, 1.815 g, 15.00 mmol) was added to the mixture and stirred overnight. H_2O was added and the mixture separated with EtOAc. The organic layers were repeatedly washed with water (3 x 50 mL) and dried over $MgSO_4$. The solvent was removed under reduced pressure and the crude product was purified on silica gel, eluting with 20% EtOAc–Hexane giving 5-bis-substituted-2,2-dimethyl-1,3-dioxane-4,6-dione derivatives **3.5a-3, 3.5a-4, 3.5a-5 and 3.5a-6.**



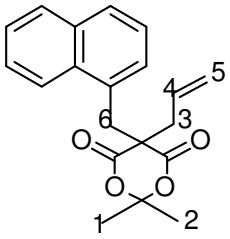
5-Allyl-5-benzyl-2,2-dimethyl-1,3-dioxane-4,6-dione 3.5a-5.
Yield: 75 %; Pale yellow oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H1.6 (s, 6H, H-1/2), 2.6 (d, *J* = 7.6 Hz, 1H, H-3), 3.2 (s, 2H, H-6), 4.9 (dd, *J* = 12, 5 Hz, 2H, H-5), 5.5 (m, 1H, H-4), 7.2 (m, 5H); Anal. Calcd. for $C_{16}H_{18}O_4$: C, 70.06; H, 6.61. Found: C, 70.26; H, 6.64.; MS calculated for $C_{16}H_{18}O_4$ m/z 274.1 [M+1], found 274.2.



5-Allyl-5-(cyclohexylmethyl)-2,2-dimethyl-1,3-dioxane-4,6-

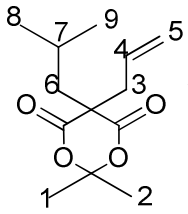
dione 3.5a-4.

Yield: 80 %; Clear colourless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H0.8 (m, 2H, Cyclo-CH), 1.1 (m, 1H, Cyclo-CH), 1.3 (m, 2H, Cyclo-C*H*), 1.6 (s, 6H, H-1/2), 1.7 (m, 6H, Cyclo-C*H*),1.75 (d, *J* = 9 Hz, 2H, H-6), 2.7 (d, *J* = 8 Hz, 2H, H-3), 5.0 (m, 2H, H-5), 6.2 (m, 1H, H-4).



5-Allyl-2,2-dimethyl-5-(naphthalen-1-ylmethyl)-1,3-dioxane-4,6-dione 3.5a-6

Yield:90 %; A light brown oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H1.7 (s, 6H, H-1/2)2.7 (d, *J* = 8 Hz, 1H, H-3), 3.7 (s, 2H, H-6), 5.1 (dd, *J* = 12,5Hz, 2H, H-5), 6.2 (t, *J* = 7 Hz, 1H, H-4), 7.3 (m, 7H); Anal. Calcd. for $C_{20}H_{20}O_4$: C, 74.06; H, 6.21. Found: C, 74.26; H, 6.64.; MS calculated for $C_{20}H_{20}O_4$m/z 324.2 [M+], found 323.1.

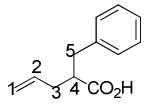


5-Allyl-5-isobutyl-2,2-dimethyl-1,3-dioxane-4,6-dione 3.5a-3.

Yield: 88 %; Light yellow solid m.p. 91-94 ˚C; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.91 (s, 6H, H-8/9), 1.4 (m, 1H, H-7), 1.7 (s, 6H, H-1/2) 2.0 (d, *J* = 9 Hz, 2H, H-6), 2.6 (d, *J* = 9 Hz, 2H, H-3), 4.9 (m, 2H, H-5), 5.9 (m, 1H, H-4).

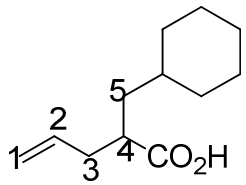
General Procedure for the preparation of compounds 3.7a-3, 3.7a-4, 3.7a-5, 3.7a-6

A suspension of 5-Allyl-5-substituted-2,2-dimethyl-1,3-dioxane-4,6-dione **(**2.10 mmol) in NaOH (2M, 10 mL) and 1,4-dioxane (5 mL) was stirred at 100 °C for 3 h. The resulting clear solution was concentrated, acidified to pH 1 by adding 6 M HCl, and extracted with EtOAc (25 mL). The extract was dried over $MgSO_4$, filtered, and concentrated. The resulting residue was dissolved in DMSO (10 mL), and the solution was stirred at 130 °C for 3 h. The reaction mixture was taken up in EtOAc (80 mL), washed with water (40 mL x 2), dried over $MgSO_4$, filtered, and concentrated. The residual material was purified by chromatography on silica gel eluting with 20 % EtOAc in hexanes to give the 2-substituted-4-pentenoic acids **3.7a-3**, **3.7a-4**, **3.7a-5**, **3.7a-6**.



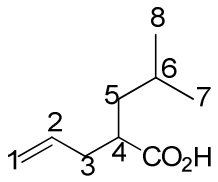
(±)-2-Benzyl-4-pentenoic acid 3.7a-5.

Yield: 76 %; Acolourless waxy oil;^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.2 (m, 1H, H-3a), 2.6 (m, 1H, H-3b),2.8(m, 1H, H-4), 2.9 (dd, *J* = 12, 7 Hz, 1H, H-5a), 3.1 (dd, *J* = 12, 7 Hz, 1H, H-5b), 5.0 (dd, *J* = 12, 5 Hz, H-1) 5.9 (m, 1H, H-2), 7.3 (m, 5H, Ph-H);Anal. Calcd for $C_{12}H_{14}O_2$: C, 75.76; H, 7.42; found: C, 75.50; H, 7.50. MS calculated for $C_{12}H_{14}O_2$ m/z 190.3 [M+H], found 189.1.



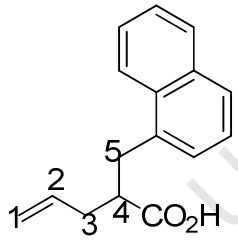
(±)-2-(Cyclohexylmethyl)-4-pentenoic acid3.7a-4.

Yield: 67 %; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.9 (m, 2H, Cyclo-C*H*), 1.1(m, 1H, Cyclo-C*H*), 1.3(m, 4H, Cyclo-C*H*),1.6(t, *J* = 7 Hz, 1H, H-5), 1.7(m, 3H, H-Cyclo-CH), 1.8 (m, 3H, H-5/Cyclo-CH), 2.1(m, 1H, H-4), 2.3(m, 1H, H-3a), 2.4(m, 1H, H-3b),5.1(dd, *J* = 12, 8 Hz, 2H, H-1), 5.8(m, 1H, H-2).



(±)-2-Isobutyl-4-pentenoic acid 3.7a-3.

Yield: 70 %; Clear colourless oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.9(d, *J* = 12Hz, 6H, H – 7/8), 1.5(m, 1H, H-6), 1.6(t, *J* = 7Hz, 1H, H-5a), 1.7(t, *J* = 7Hz, 1H, H-5b), 2.3(m, 2H, H-3a/4), 2.4 (m, 1H, H-3b) 5.1(m, 2H, H-1), 5.9(m, 1H, H-2).

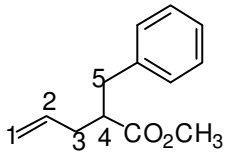


(±)-2-(Naphthalen-1-ylmethyl)-4-pentenoic acid 3.7a-6.

Yield: 88 %; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.3(m, 1H, H-3a), 2.7(m, 1H, H-3b), 3.1(m, 2H, H-4/5a), 3.4 (dd, J = 12, 6 Hz, 1H, H-5b) 5.1 (m, 2H, H-1), 6.0(m, 1H, H-2), 7.3 (m, 7H, Ph-H). MS calculated for $C_{16}H_{16}O_2$m/z 240.6 [M+1], found 239.2.

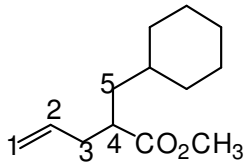
General Procedure for the preparation of compounds 3.7b-3, 3.7b-4. 3.7b-5, 3.7b-6

The carboxylic acid (1.58 mmol) was dissolved in MeOH (8 mL) and the mixture cooled to 0 ^{O}C in an ice bath. $SOCl_2$ (0.28 g, 0.173 ml, 2.37 mmol) was added dropwise over 10 min. The reaction was stirred at 0 ^{O}C for one hour and allowed to warm to room temperature. It was allowed to stir further overnight. Thereafter, the solvent was removed in vacuo leaving the products **3.7b-3, 3.7b-4. 3.7b-5, 3.7b-6** which were used in the next step without further purification.



(±) Methyl-2-benzyl-4-pentenoate 3.7b-5.

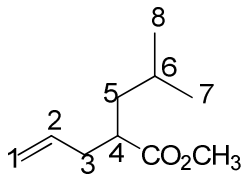
Yield: 76 %; A colourless waxy oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.2 (m, 1H, H-3a), 2.6 (m, 1H, H-3b),2.8(m, 1H, H-4), 2.9 (dd, *J* = 12, 7 Hz, 1H, H-5a), 3.1 (dd, *J* = 12, 7 Hz, 1H, H-5b),3.7(s, 3H, OC*H*$_3$), 5.0 (dd, *J* = 12, 5 Hz, H-1) 5.9 (m, 1H, H-2), 7.3 (m, 5H, Ph-H).



(±) Methyl-2-(cyclohexylmethyl)-4-pentenoate 3.7b-4.

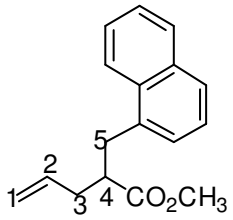
Yield: quant.; Colourless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H1.0 (m, 2H, Cyclo-C*H*), 1.3(m, 4H, Cyclo-C*H*), 1.5(t, *J* = 8 Hz, 1H, H-5a) 1.7 (m, 4H, H-5b/Cyclo-C*H*), 1.8(m, 2H, Cyclo-CH), 2.1(m, 1H, H-4), 2.2(m, 1H, H-3a), 2.4(m, 1H, H-3b),

3.6(s, 3H, *OCH*$_3$), 5.1(d, *J* = 9 Hz, 2H, H-1), 5.8(m, 1H, H-2).



(±) Methyl-2-isobutyl-4-pentenoate 3.7b-3.

Yield: 70 %; Clear colourless oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.9(d, *J* = 12Hz, 6H, H - 7/8), 1.5(m, 1H, H-6), 1.6(t, *J* = 7Hz, 1H, H-5a), 1.7(t, *J* = 7Hz, 1H, H-5b), 2.3(m, 2H, H-3a/4), 2.4 (m, 1H, H-3b), 3.6(s, 3H, OCH_3), 5.1(m, 2H, H-1), 5.9(m, 1H, H-2).



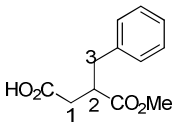
(±) Methyl-2-(naphthalen-1-ylmethyl)-4-pentenoate 3.7b-6.

Yield: Quant; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.3(m, 1H, H-3a), 2.7(m, 1H, H-3b), 3.1(m, 2H, H-4/5a), 3.4 (dd, *J* = 12, 6 Hz, 1H, H-5b), 3.7(s, 3H, OCH_3), 5.0 (m, 2H, H-1), 6.0(m, 1H, H-2), 7.3 (m, 7H, Ph-H). MS calculated for $C_{17}H_{18}O_2$ m/z 254.3 [M+1], found 253.2.

General Procedure for the preparation of compounds 3.7c-3, 3.7c-4, 3.7c-5, 3.7c-6

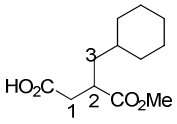
To a solution of alkene (1.42 mmol) in acetonitrile (10 mL), CCl_4 (5 mL) and H_2O (5 mL) were added $NaIO_4$ (1.82 g,

8.52 mmol) and a catalytic amount of ruthenium oxide (0.3 eq., 0.06 g, 0.43 mmol). The mixture was stirred at room temperature for 20 h. The reaction mixture was filtered, and the filtrate was concentrated. The resulting mixture was extracted with EtOAc (40 mL x 2), and the combined extracts were washed with brine (40 mL), dried over $MgSO_4$, filtered, and concentrated. The residual material was purified by silica gel chromatography eluting with 30 % EtOAc in Hexane to give the carboxylic acid products **3.7c-3, 3.7c-4, 3.7c-5, 3.7c-6.**



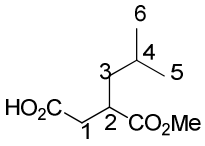
(±)3-Benzyl-4-methoxy-4-oxobutanoic acid 3.7c-5

Yield: 65 %; Clear colourless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.7 (dd, *J* = 13, 5 Hz, 1H, H-1a), 2.9 (dd, *J* = 12, 5 Hz, 1H, H-1a), 3.1 (dd, *J* = 12, 8 Hz, 1H, H-3a), 3.2(m, 1H, H-2), 3.5 (m, 1H, H-3b), 3.7 (s, 3H, OCH_3), 7.1 (m, 5H, Ph); m/z = 222.1 [M+1]. Anal.Calcd.for $C_{12}H_{11}O_4$: C, 64.85; H, 6.35. Found: C, 64.61; H, 6.20%.



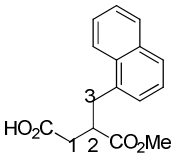
(±)3-(Cyclohexylmethyl)-4-methoxy-4-oxobutanoic acid 3.7c-4

Yield: 55 %; Clear colourless; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H1.0 (m, 2H, Cyclo-C*H*), 1.1(m, 1H, Cyclo-C*H*), 1.3 (m, 4H, H-Cyclo-C*H*), 1.6(m, 1H, H-3a), 1.7(m, 3H, Cyclo-C*H*), 1.8(m, 2H, Cyclohexyl-C*H*), 2.0 (t, *J* = 8 Hz, 1H, H-3b), 2.5(m, 1H, H-2), 2.7 (dd, *J* = 12, 5 Hz, 1H, H-1a), 3.0 (m, 1H, H-1a)3.7 (s, 3H, O*CH*$_3$).



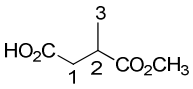
(±)3-(Methoxycarbonyl)-5-methylhexanoic acid 3.7c-3

Yield: 61 %; Clear colourless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 0.9(d, *J* = 12Hz, 6H, H – 5/6), 1.6 (m, 1H, H-3a/4), 2.1 (t, *J* = 7 Hz, 1H, H-3b) 2.5 (m, 1H, H-2), 2.7(dd, *J* = 12, 5Hz, 1H, H-1a), 3.0 (dd, *J* = 12, 5 Hz, 1H, H-1b),3.7(s, 3H, *OCH*$_3$).



(±)4-Methoxy-3-(naphthalen-1-ylmethyl)-4-oxobutanoicacid 3.7c-6.

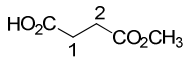
Yield: 60 %; Brown waxy oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.7 (m, 1H, H-1a), 3.0(m, 1H, H-1b),3.3(m, 1H, H-3a), 3.5(m, 2H, H-3b/2), 3.7(s, 3H, OC*H*$_3$), 7.1-7.8(m,7H, Naphthyl-*H*).



(±)4-Methoxy-3-methyl-4-oxobutanoic acid 3.13b-1

A stirred suspension of 3-methyldihydrofuran-2,5-dione (1.0 g, 8.76 mmol) and anhydrous methanol (0.43 mL, 0.34 g, 10.5 mmol) was refluxed for 1 h. After the mixture had become homogeneous, heating was continued for an additional half an hour. After cooling, the excess methanol was removed under vacuum, and the desired compound isolated on silica gel (EtOAc/Hexane 1:9) yielding a white solid of 4-methoxy-3-methyl-4-oxobutanoic acid **3.13b-1** (0.448, 3.07 mmol, 35 %). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.3 (d, *J* = 12 Hz, 3H, H-3)

2.6 (m, 1H, H-1a), 2.7 (dd, *J* = 13, 7Hz, 1H, H-1b), 3.1 (m, 1H, H-2), 3.7 (s, 3H).



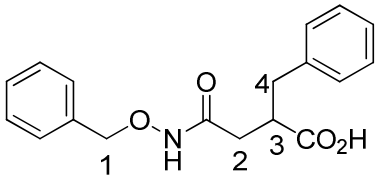
(±)4-Methoxy-3-methyl-4-oxobutanoic acid 3.13b-2

A stirred suspension of dihydrofuran-2,5-dione (1.0 g, 9.99 mmol) and anhydrous methanol (0.49 mL, 0.38 g, 11.9 mmol) was refluxed for 1 h. After the mixture had become homogeneous, heating was continued for an additional half an hour. After cooling, the excess methanol was removed under vacuum, yielding a white solid of 4-methoxy-4-oxobutanoic acid **3.13b-2** (1.19 g, 8.99 mmol, 90 %).^{1}H NMR (400 MHz, $CDCl_3$): δ_H2.7 (t, *J* = 7 Hz, 2H, H-1), 2.8 (t, *J* = 7 Hz, 2H, H-2),3.6(s, 3H, OCH_3).

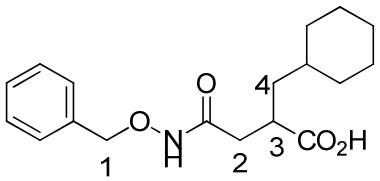
General Procedure for the preparation of compounds 3.2a-1, 3.2a-2, 3.2a-3, 3.2a-4, 3.2a-5 and 3.2a-6.

The carboxylic acid(0.45 mmol) was dissolved in DMF (10 mL) and the mixture cooled to 0 ^{O}C. To the solution were added EDC (0.086 g, 0.45 mmol), O-benzylhydroxylamine hydrochloride (0.072 g, 0.45 mmol), diisopropylethylamine (0.060 g, 0.45 mmol), and 1-HOBt (0.07 g, 0.45 mmol). The reaction mixture was stirred overnight.Water (100 mL) was added to the mixture and then it was separated with EtOAc (50 mL x 3). Brine (20mL) was added to break up the emulsion that formed. The organic layer solution was washed with HCl (0.1 *M*, 40 mL x 1), $NaHCO_3$(aq) (5 %, 50 mL x1) and brine (40 mL x 1), dried over $MgSO_4$, filtered, and concentrated. The residue was dissolved in MeOH (10 mL) and

to this was added NaOH (aq) (4 M, 10 mL) and stirred for 3 h. Most of the MeOH was removed in vacuo and the remaining acqueous solution was separated with diethyl ether (20 mL x 1) The organic layer was discarded and the mixture acidified to pH 2 with 2 *N* HCl. The mixture was then separated with EtOAc (50 mL x 3) and the combined organics washed with H_2O (50 mL x 1), brine (50 mL x 1), dried over $MgSO_4$ and the solvent removed in vacuo. The title compounds were recrystallized from MeOH giving products **3.2a-1, 3.2a-2, 3.2a-3, 3.2a-4, 3.2a-5 and 3.2a-6.**

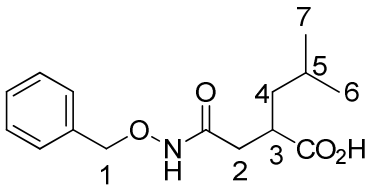


(±)2-Benzyl-4-((benzyloxy)amino)-4-oxobutanoic acid 3.2a-5. Yield: 75%;Colourless crystals from MeOH; m.p. 122 – 124 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 2.5 (dd, *J* = 12, 3 Hz, 1H, H-2a) 3.2 (m, 2H, H-2b/3), 3.3 (dd, *J* = 12, 5 Hz, 1H, H-4a) 3.6 (dd, *J* = 12, 6 Hz, 1H, H-4b), 4.8 (s, 2H, H-1), 7.4 (m, 10H);^{13}C NMR (100 MHz, $CDCl_3$): δ 35.7, 36.6, 44.9, 76.4, 126.1, 127.1 (3C),128.1 (2C), 128.9 (2C), 129, 138, 141.2, 169.9, 178.4; Calcd for $C_{18}H_{19}NO_4$: C, 63.15; H, 5.35, N, 4.40 Found: C, 62.03; H, 5.27, N, 4.36%. MS calculated for $C_{18}H_{19}NO_4$ m/z = 313.13 [M+H], found 313.1.



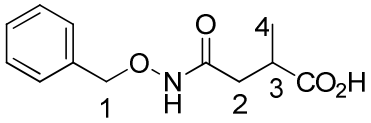
(±)4-((Benzyloxy)amino)-2-(cyclohexylmethyl)-4-oxobutanoic acid 3.2a-4.

Yield: 72%; Colourless crystals from MeOH; m.p. 91 – 93 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9 (m, 2H, Cyclo-C*H*), 1.3(m, 5H, Cyclo-C*H*), 1.7(m, 6H, Cyclo-C*H*, H-4), 2.3(d, *J* = 11 Hz, 1H, H-2a), 2.7(d, *J* = 13, 1H, H-2b), 2.8(m, 1H, H-3), 4.9(s, 2H, H-1), 7.4(m, 5H, Ph*H*). Calcd for $C_{18}H_{25}NO_4$: C, 67.69; H, 7.89; N, 4.39. Found: C, 67.11; H, 7.75; N, 4.28%. MS calculated for $C_{18}H_{25}NO_4$ m/z = 319.13 [M+H], found 318.1.



(±)2-(2-((Benzyloxy)amino)-2-oxoethyl)-4-methylpentanoic acid 3.2a-3.

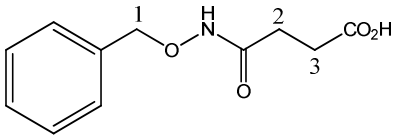
Yield: 70 % yield; White crystalline material from MeOH; m.p. 101 – 104 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9 (d, *J* = 9 Hz, 6H, H-6/7) 1.5 (m, 1H, H-5), 1.6 (dd, *J* = 12, 9 Hz, 1H, H-4a), 1.9 (dd, *J* = 12, 8 Hz, 1H, H-4b), 2.1(m, 1H, H-2a), 2.7 (m, 2H, H-2b/3) 4.8 (s, 2H, H-1), 7.4 (m, 5H, Ph*H*). Calcd for $C_{15}H_{21}NO_4$: C, 60.11; H, 6.99; N, 4.55. Found: C, 59.04; H, 7.15; N, 4.28%. MS calculated for $C_{15}H_{21}NO_4$ m/z = 280.15 [M+H], found 279.5.



(±)4-((Benzyloxy)amino)-2-methyl-4-oxobutanoic acid 3.2a-2

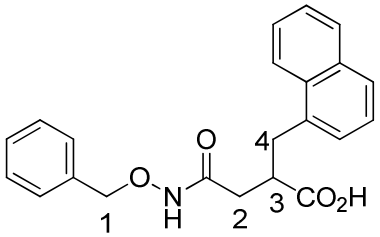
Yield: 60%; Colourless crystals from MeOH; m.p. 123 – 125 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.2 (d, *J* = 9 Hz, 3H, H-4) 2.6 (m, 1H, H-3), 2.8 (d, *J* = 8 Hz, 2H, H-2), 4.8 (s, 2H,

H-1), 7.4 (m, 5H). Calcd for $C_{12}H_{15}NO_4$: C, 60.75; H, 6.37; N, 5.90. Found: C, 60.61; H, 6.42; N, 5.75%. MS calculated for $C_{12}H_{15}NO_4$ m/z = 237.01 [M+H], found 237.1.



(±)4-((Benzyloxy)amino)-4-oxobutanoic acid 3.2a-1

Yield: 90 %; White crystals from MeOH; m.p. 113 – 117 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 2.6 (t, *J* = 7 Hz, 2H, H-3), 2.7 (t, *J* = 7 Hz, 2H, H-2), 4.8 (s, 2H, H-1), 7.4 (m, 5H). MS calculated for $C_{11}H_{13}NO_4$ m/z = 223.1 [M+H], found 223.2.



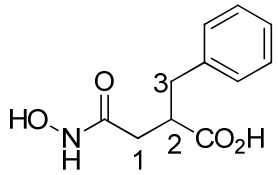
(±)4-((Benzyloxy)amino)-2-(naphthalen-1-ylmethyl)-4-oxobutanoic acid 3.2a-6.

Yield: 84 %; Brown waxy oil; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 2.8 (dd, *J* = 12, 4 Hz, 1H, H-2a) 3.1 (m, 1H, H-2b), 3.2 (m, 2H, H-4a/3), 3.5 (m, 1H, H-4b), 4.8 (m, 2H, H-1), 7.4 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 34.7, 37.6, 45.9, 77.4, 125.1(5C), 127.1 (3C),127.3 (2C), 128.9 (2C), 130.1, 137.0, 142.2, 169.9, 178.4; MS calculated for $C_{22}H_{21}NO_4$ m/z = 362.4 [M+], found 362.2.

General Procedure for the preparation of compounds 3.16b.3.16c. 3.16d, 3.16e, 3.16f and 3.16g.

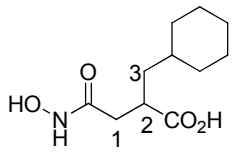
Hydroxylamine O-benzylated compound (0.179 mmol)was

dissolved in MeOH (10 mL). Pd-C (0.06 g, 0.564 mmol) was added to the mixture and the reaction carried out under hydrogen atmosphere (3 atm) for 2 h. The reaction mixture was filtered through celite and solvent removed in vacuo to leave the product free hydroxamate **3.16b. 3.16c. 3.16d, 3.16e, 3.16f** and **3.16g** whichwere stored in a sealed plastic vial.



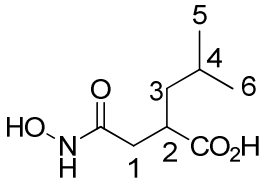
(±)2-Benzyl-4-(hydroxyamino)-4-oxobutanoic acid 3.16f

Yield: 41 %;^{1}H NMR (400 MHz, $CDCl_3$): δ_H 2.4 (dd, *J* = 13,6 Hz, 1H, H-1a) 2.6 (dd, *J* = 13, 6 Hz, 1H, H-1b), 2.8 (dd, *J* = 12, 7 Hz, 1H, H-3a), 3.1 (m, 1H, H-2), 3.2 (dd, *J* = 12, 7 Hz, 1H, H-3b), 7.3 (m, 5H);^{13}C NMR (100 MHz, $CDCl_3$): δ 34.7, 37.6, 44.7, 127.1, 128.1 (2C),128.9, 129, 137.1, 170.9, 178.7; Calcd for $C_{11}H_{13}NO_4$ m/z = 223.4 [M+], found 222.1.



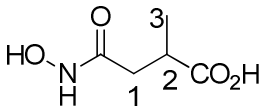
(±)2-(Cyclohexylmethyl)-3-(hydroxyamino)-3-oxopropanoic acid 3.16e

Yield: 60 %; White crystalline powder; m.p. 89 – 92 ^{o}C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9 (m, 2H, Cyclo-C*H*), 1.2 (m, 3H, Cyclo-C*H*), 1.3 (m, 2H, Cyclo-C*H*), 1.7 (m, 6H, Cyclo-C*H*, H-3), 2.3 (dd, *J* = 12, 7 Hz 1H, H-1a), 2.5 (dd, *J* = 12, 7 Hz, 1H, H-1b), 2.7 (m, 1H, H-2).^{13}C NMR (100 MHz, $CDCl_3$): δ 25.8 (2C), 26.4, 33.1, 33.7 (2C) 35.1 , 36.4, 47.2, 170.1, 178.1.



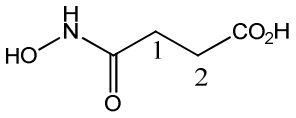
(±)2-(2-(Hydroxyamino)-2-oxoethyl)-4-methylpentanoic acid 3.16d

Yield:58 %; White crystalline solid; m.p. 91 – 94 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H0.9 (d, *J* = 8 Hz, 3H, H-5) 0.95 (d, *J* = 8 Hz, 3H, H-6), 1.2 (m, 1H, H-4), 1.7 (t, *J* = 7 Hz, H-3) 2.2 (d, *J* = 9 Hz, 2H, H-1) 2.7 (m, 1H, H-2).



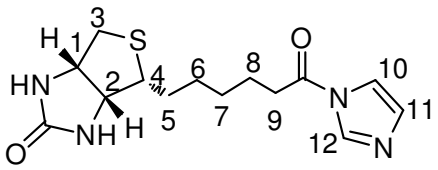
(±)3-(Hydroxyamino)-2-methyl-3-oxobutanoic acid 3.16c

Yield: 60 %; White crystalline powder; m.p. 76 – 78 %; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.3 (d, *J* = 11 Hz, 3H, H-3) 2.2 (d, *J*= 9 Hz, 1H, H-1a), 2.8 (m, 2H, H-2 and H-1b).



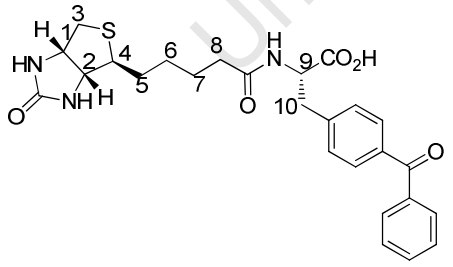
3-(Hydroxyamino)-3-oxobutanoic acid 3.16b

Yield: 55 %; White crystalline powder; m.p. 93 – 95%;^{1}H NMR (400 MHz, $CDCl_3$):δ_H 2.3 (t, *J* = 7 Hz, 2H, H-2), 2.7 (t, *J* = 7 Hz, 2H, H-1).



(4S)-4-(5-(1H-Imidazol-1-yl)-5-oxopentyl)tetrahydro-1H-thieno[3,4-d]imidazol-2(3H)-one 3.18c

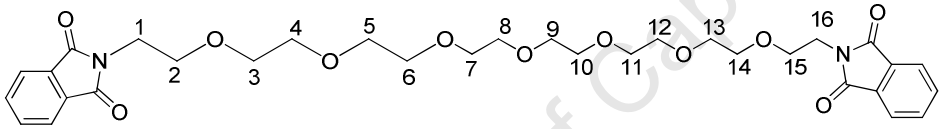
5-((4S)-2-oxohexahydro-1H-thieno[3,4-b]pyrrol-4-yl)pentanoic acid (0.2 g, 0.82 mmol) **3.18a** was dissolved in anhydrous DMF (15 mL) at 70 °C in a round bottom flask with magnetic stirrer. The solution was cooled to room temperature, and 1,1'-carbonyldiimidazole **3.18b** (0.2 g, 1.23 mmol) was added with stirring at room temperature until a white precipitate formed (30 min). The crude precipitate was filtered and washed with dry chloroform (20 mL). The filter cake was re-suspended in chloroform (10 mL) and refluxed for 30 min. The suspension was cooled to room temperature and filtered to obtain pure (4S)-4-(5-(1H-imidazol-1-yl)-5-oxopentyl)tetrahydro-1H-thieno[3,4-d]imidazol-2(3H)-one **3.18c**(0.206 g, 0.699 mmol, 85 % yield) as a white solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ_H 1.5 (m, 4H, H-6/7),1.8 (m, 4H, H-5/8), 2.1 (t, 2H, *J*=7 Hz, H-9), 2.6 (dd, 1H, *J*= 12, 5 Hz, H-2a), 2.8 (dd, 1H, *J*= 12, 5 Hz, H-2b), 3.4 (m, 1H, H-4), 4.19 (m, 1H ,H-1), 4.28 (m, 1H, H-2), 6.4 (s, 2H, N*H*), 7.1 (m, 1H, H-10), 7.5 (m, 1H, H-11), 7.9 (m, 1H, H-12). ^{13}C NMR (DMSO-d6) δ_C 23.5, 27.7, 27.9, 34.0, 39.7, 55.2, 59.1, 60.9, 116.3, 130.1, 136.8, 162.5, 171.4.In agreement with the literature.[2]



(4S)-4-(6-(1H-Imidazol-1-yl)-6-oxohexyl)tetrahydro-1H-thieno[3,4-b]pyrrol-2(3H)-one 3.18e.

3.18c (4S)-4-(5-(1H-imidazol-1-yl)-5-oxopentyl)tetrahydro-1H-thieno[3,4-d]imidazol-2(3H)-one (0.19 g, 0.65 mmol) and **3.18d**(R)-2-amino-3-(4-benzoylphenyl)propanoic acid (0.174

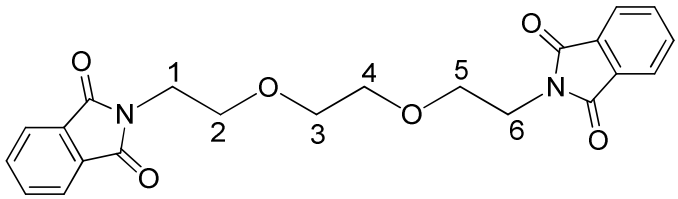
g, 0.645 mmol)) were dissolved in anhydrous THF (10 ml) and stirred overnight. The solvent was removed in vacuo leaving crude **3.18e**(2S)-3-(4-benzoylphenyl)-2-(5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propanoic acid (0.192 g, 0.387 mmol, 60 % yield) which was used in the next step without further purification.^{1}H NMR(400 MHz, DMSO-d_6): δ_H1.30 (m, 2H, H-6), 1.6 (m, 3H, H-5a/7), 1.7 (m, 1H, H-5b), 2.3 (t, *J* = 7 Hz, 2H, H-8), 2.8 (dd, *J* = 12, 5 Hz, 1H, H-10a), 2.9 (m, 1H, H-4), 3.2 (m, 1H, H-10b), 3.3 (dd, *J* = 13, 7 Hz, H-3), 4.30 (d, *J* = 8,1H, H-2), 4.7 (t, *J* = 6 Hz, 1H, H-1), 4.8 (t, *J* = 6 Hz, 1H, H-9), 5.7 (br s, 2H, N*H*), 7.3-7.9 (m, 9H); MS Calcd for $C_{26}H_{29}N_3O_5$Sm/z = 495.1 [M+H], found 495.1.



2,2'-(3,6,9,12,15,18,21-Heptaoxatricosane-1,23-diyl)bis(isoindoline-1,3-dione) 3.21a-1

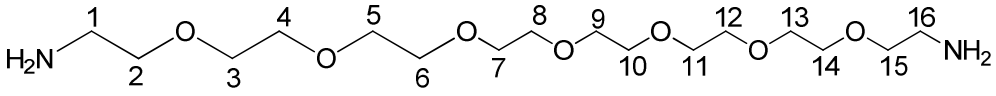
3,6,9,12,15,18,21-heptaoxatricosane-1,23-diol (1 g, 2.70 mmol) was dissolved in DCM (20 ml). Tosyl-Cl (1.08 g, 5.67 mmol)and pyridine (0.437 ml, 5.40 mmol) were added to the mixture and the reaction was stirred for 24 h. The volatiles were removed in vacuo and the residue taken up in $CHCl_3$ (15 mL). Potassium phthalamide (1.05 g, 5.67 mmol)was added slowly over 15 minutes and the reaction stirred for 12 h. $CHCl_3$ (50 ml)was added and the mixture was washed with $NaHCO_3$ (30 mL), brine (30 mL) and water (50 mL). The organic phase was dried over $MgSO_4$, solvent removed under reduced pressure leaving the product 2,2'-(3,6,9,12,15,18,21-heptaoxatricosane-1,23-diyl)bis(isoindoline-1,3-dione) 3.21a-1(1.018 g, 1.620 mmol, 60 % yield)as a white solid.^{1}H

NMR (400 MHz, $CDCl_3$): δ_H3.25 (t, *J* = 6 Hz, 4H, H-1/16), 3.4 (t, *J* = 6 Hz, 4H, H-2/15), 3.6 (m, 24H, OCH_2CH_2O),7.9 (m, 8H, Phthalyl-H).MS Calcd for $C_{32}H_{40}N_2O_{11}$ m/z = 628.1 [M+H], found 628.1.



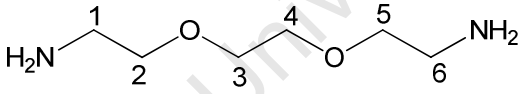
2,2'-((Ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(isoindoline-1,3-dione) 3.21a-2

2,2'-(ethane-1,2-diylbis(oxy))diethanol (8.0 g, 53.3 mmol) was dissolved in DCM (20 ml). Tosyl-Cl (21.33 g, 112 mmol)and pyridine (8.62 ml, 107.00 mmol) were added to the mixture and the reaction was stirred for 24 h. The volatiles were removed in vacuo and the residue taken up in $CHCl_3$ (15 mL). Potassium phthalamide (20.72 g, 112.00 mmol)was added slowly over 15 minutes and the reaction stirred for 12 h. $CHCl_3$ (50 mL) was added and the mixture was washed with $NaHCO_3$ (50 mL), brine (50 mL) and water (50 mL). The organic phase was dried over $MgSO_4$, solvent removed under reduced pressure leaving the product 2,2'-((ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(isoindoline-1,3-dione)**3.21a-2**(10.88 g, 26.6 mmol, 50 % yield).^{1}H NMR (400 MHz, $CDCl_3$): δ_H3.5 (m, 4H, H-3/4),3.9 (s, 8H, H-1/2/5/6), 7.8 (m, 8H, Phthalyl-H).MS Calcd for $C_{22}H_{20}N_2O_6$ m/z = 408.1 [M+H], found 408.1.



3,6,9,12,15,18,21-Heptaoxatricosane-1,23-diamine 3.25c-1

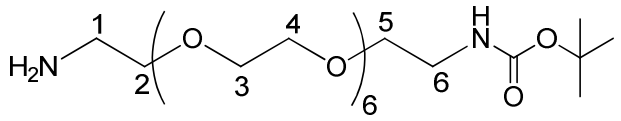
2,2'-(3,6,9,12,15,18,21-heptaoxatricosane-1,23-diyl)bis(isoindoline-1,3-dione)**3.21a-1** (1.35 g, 2.15 mmol) was dissolved in excess hydrazine hydrate (0.107 g, 2.15 mmol) (6 ml) and stirred overnight. A white precipitate that formed was filtered off and the solvent removed. $CHCl_3$ (50 mL) and $NaHCO_3$ (10mL, 10mM) was slowly added to the mixture. The aqueous layer was separated from the organic layer and retained. The aqueous layer was extracted with chloroform (2 x 20 mL) and once with Et_2O (20ml). The aqueous layer was brought to ~ pH 7 by the drop-wise addition of HCl (1 M) and dried on the lyophilizer. The remaining residue was taken up in MeOH (20 mL) and filtered. The MeOH was removed under reduced pressure leaving the final product 3,6,9,12,15,18,21-heptaoxatricosane-1,23-diamine **3.21c-1**(0.51 g, 1.37 mmol, 64 % yield) as a clear oily residue. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H3.2 (t, *J* = 6 Hz, 4H, H-1/16), 3.5 (m, 24H, OCH_2CH_2O), 3.8 (t, *J* = 7 Hz, 4H, H-2/15);^{13}C NMR (75.5 MHz, $CDCl_3$) δ_C 41.8 (2C), 73.1 (2C), 71.5 (12C), 74.4 (2C).



2,2'-(Ethane-1,2-diylbis(oxy))diethanamine 3.21c-2

2,2'-((ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl))bis(isoindoline-1,3-dione) **3.21a-2** was dissolved in hydrazine hydrate (20 mL) and the reaction was stirred for 24 h. The precipitate that formed was filtered off and the solvent reduced *in vacuo* and the residue taken up in $CHCl_3$ (50 mL) and $NaHCO_3$ (10mM) was slowly added to the mixture. The aqueous layer was separated from the organic layer and retained. The aqueous layer was extracted with chloroform

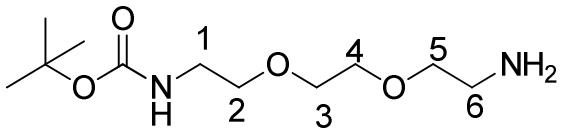
(2 x 20 mL) and once with Et_2O (20ml). The aqueous layer was brought to ~ pH 7 by the drop-wise addition of HCl (1 M) and dried on the lyophilizer. The remaining residue was taken up in MeOH (20 mL) and filtered. The MeOH was removed under reduced pressure leaving the final product 2,2'-(ethane-1,2-diylbis(oxy))diethanamine (4.5 g, 119.00 mmol, 70 % yield) as a clear oily residue.^{1}H NMR (400 MHz, $CDCl_3$): δ_H3.0 (t, *J* = 7 Hz, 4H, H-1/6), 3.6 (s, 4H, H-3/4),3.7 (m, 4H, H-2/5), 5.2 (br s, 4H);^{13}C NMR (100 MHz, $CDCl_3$): δ_c 43.1 (2C), 70.1 (2C), 75.1 (2C).[3]



Tert-butyl(23-amino-3,6,9,12,15,18,21-heptaoxatricosyl)carbamate 3.28a

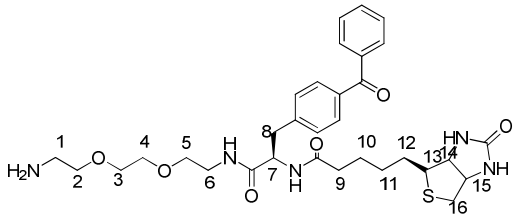
HCl(g) (52 mg, 1.4 mmol, H_2SO_4/NaCl) was bubbled through cooled MeOH (10 ml, 0 °C) under closed vessel conditions. The mixture was stirred for 15 min at 0 °C and was carefully added to 3,6,9,12,15,18,21-heptaoxatricosane-1,23-diamine **3.21c-1**(0.50 g, 1.37 mmol). The mixture was stirred for 15 min at room temperature before adding 10 mL of H_2O and stirring for another 0.5 h. To the solution BOC_2O (0.32 ml, 1.36 mmol) in 10 mL of MeOH was added at room temperature over 10 min, and the resultant solution was stirred for 1 h. The mixture was concentrated *in vacuo*. Unreacted diamine was removed by diethyl ether (20 mL x 2). The residue was treated with 2 N NaOH (20 mL) solution. The product in the organic layer was extracted with CH_2Cl_2 (50 mL x 3). The combined extracts were washed with brine, dried over anhydrous MgSO4, and concentrated in vacuo to yield tert-butyl (23-amino-3,6,9,12,15,18,21-heptaoxatricosyl)carbamate **3.28a** (0.382 g, 0.814 mmol, 60 %

yield) the mono-BOC product as a colorless oil. ^{1}H NMR, (400 MHz, D_2O): δ_H 1.4 (s, 9H, Boc-H), 2.8 (m, 2H, H-1)3.1 (m, 2H, H-6), 3.5 (t, *J* = 7 Hz, 4H, H-2/5), 3.7 (s, 24H), 5.3 (bs, 2H, NH_2).



Tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate 3.26d

HCl (1.03 g, 28.01 mmol) was bubbled through cooled MeOH (150 ml, 0 °C). The mixture was stirred for 15 min at 0 °C and was carefully added to 2,2'-(ethane-1,2-diylbis(oxy))diethanamine **3.21c-2**(4.20 g, 28.30 mmol). The mixture was stirred for 15 min at room temperature before adding 25 mL of H_2O and stirring for another 0.5 h. To the solution BOC_2O (6.58 ml, 28.30 mmol) in 100 mL of MeOH was added at room temperature over 10 min, and the resultant solution was stirred for 1 h. The mixture was concentrated in vacuo. Unreacted diamine was removed by diethyl ether (100 mL x 2). The residue was treated with 2 N NaOH (200 mL) solution. The product in the organic layer was extracted with CH_2Cl_2 (150 mL x 3). The combined extracts were washed with brine, dried over anhydrous MgSO4, and concentrated in vacuo to yield tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate **3.26d** (5.21 g, 20.97 mmol, 74 % yield) the mono-BOC product as a colorless oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.40 (s, 9H, t-butyl),2.9(m, 2H, H-1),3.2(m, 2H, H-6), 3.5 (t, *J* = 7 Hz, 4H, H-2/5), 3.6 (s, 4H, H-3/4), 5.2 (br s, 2H, N*H*$_2$), 8.0 (br s, 1H, N*H*). In agreement with the literature.[3]

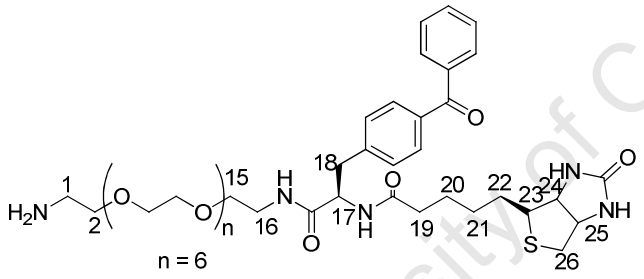


N-((R)-1-((2-(2-(2-Aminoethoxy)ethoxy)ethyl)amino)-3-(4-benzoylphenyl)-1-oxopropan-2-yl)-5-((4R)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide 3.21b-2

(2S)-3-(4-benzoylphenyl)-2-(5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propanoic acid **3.18e** (0.100 g, 0.201 mmol) was dissolved in DMF (5 mL) and the mixture cooled to 0 ^{O}C. To the solution were added EDC (0.039 g, 0.201 mmol), tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate **3.25c-2** (0.05 g, 0.20 mmol), DIEA (0.035 ml, 0.201 mmol), and HOBt (0.031 g, 0.20 mmol). The reaction mixture was stirred overnight.Water (50 mL) was added to the mixture and then it was separated with EtOAc (50 mL x 3). Brine (20mL) was added to break up the emulsion that formed. The organic layer solution was washed with HCl (0.1 *M*, 40 mL x 1), $NaHCO_3$ (aq) (5 %, 50 mL x1) and brine (40 mL x 1), dried over $MgSO_4$, filtered andconcentrated. The material, a light yellow oilwasdissolved in DCM (5 mL) and the mixture cooled to 0 ^{O}C. TFA (3 ml, 38.9 mmol) was added to the solution and the mixture stirred overnight. The volatiles were removed in vacuo and the material re-dissolved in DCM (10 ml). Amberlyst A21 resin (1 g) were put in the mixture and it was stirred for 3 min. The resins were filtered off and the DCM removed in vacuo and the product was isolated on silica gel eluting with 40% EtOAc in hexanes leaving the product*N*-((S)-1-((2-(2-(2-aminoethoxy)ethoxy)ethyl)amino)-3-(4-

benzoylphenyl)-1-oxopropan-2-yl)-5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide **3.21b-2**(0.061 g, 0.097 mmol, 80 % yield)as a clear brown oil. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H1.0 (m, 2H, H-10), 1.3 (m, 2H, H-11), 1.8 (m, 2H, H-12), 1.6 (m, 2H, H-10), 2.2 (t, *J* = 7 Hz, 2H, H-9), 2.9 (m, 1H,H-13), 3.1 (m, 3H, H-1/8a), 3.3 (m, 4H, H-6a/8b/16), 3.6 (m, 6H, H-2/3/4), 3.5 (m, 3H, H-5/6b), 4.4 (dd, *J* = 12, 5 Hz, 1H, H-14), 4.7 (m, 1H, H-15), 5.4 (t, *J* = 6 Hz, 1H, H-7), 7.0 – 7.8 (m, 9H, Ph-H). MS Calcd for $C_{32}H_{43}N_5O_6S$m/z = 625.4 [M+H], found 625.1.

Compound **3.28b** below was synthesized using the same procedure as above.

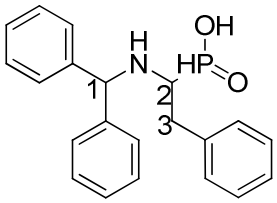


N-((R)-1-((2-(2-(2-Aminoethoxy)ethoxy)ethyl)amino)-3-(4-benzoylphenyl)-1-oxopropan-2-yl)-5-((4R)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide 3.28b

Yield: 47 %; Yellow oil; ^{1}H NMR (400 MHz, $CDCl_3$):δ_H 1.1 (m, 2H, H-21), 1.6 (m, 2H, H-22), 1.7 (m, 2H, H-20), 2.1 (t, *J* = 7Hz, H-19), 2.6 (d, *J* = 9 Hz, 1H, H-26a), 2.8 (d, *J* = 9 Hz, 1H, H-26b), 3.0 (m, 2H, H-1/16), 3.2 (t, *J* = 8 Hz,1H, H-23), 3.3 (d, *J* = 6 Hz, 2H, H-18), 3.5 (m, 4H, H-2 and H-15), 3.6 (s, 24H, OCH_2CH_2O), 4.2 (s, 1H, H-24), 4.4 (s, 1H, H-25), 4.7 (d, *J* = 6 Hz, 1H, H-17), 6.9 (m, 2H, PhH), 7.2 (m, 5H, PhH), 7.7 (m, 2H, PhH). MS m/z 845.1 [M+1] Observed 844.1.

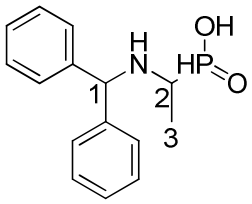
General Procedure for the preparation of compounds 4.2d-1, 4.2d-2, 4.2d-3 and 4.2d-4.

50 % aqueous Hypophosphorous acid (1.801 g, 28.00 mmol, 2.5 ml) is added to a suspension of diphenylmethanamine (5 g, 27.30 mmol) in EtOH (100 ml) and heated to 90 ^{O}C. At that temperature, the aldehyde (27.30 mmol) in EtOH (20 ml) is added slowly over 10 min and heating is continued for 3 h followed by stirring at ambient temperature overnight. The product is isolated as a precipitate which is filtered off, washed with cold EtOH and diethyl ether and used in the next step without further purification.



(±) (1-(Benzhydrylamino)-2-phenylethyl)phosphinic acid 4.23d-3

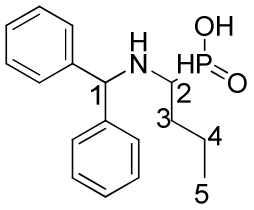
Yield: 8.44 g, 24.01 mmol, 88 %; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 2.5 (d, *J* = 11 Hz, 2H, H-3), 3.1 (t, *J* = 8 Hz, 1H, H-2), 5.2 (s, 1H, H-1), 7.3 (m, 15H, PhH).MS Calcd for $C_{21}H_{22}NO_2P$m/z = 351.4 [M+H], found 351.2.



(±) (1-(Benzhydrylamino)ethyl)phosphinic acid 4.2d-1

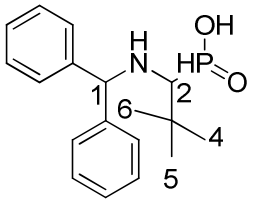
Yield: 2.70 g, 9.82 mmol, 60 % yield; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 1.2(d, *J* = 9 Hz, 3H, H-3), 2.1(bs, 1H, NH), 2.8(m, 1H, H-2), 5.1(s, 1H, H-1), 7.2(m, 10H, PhH); $C_{15}H_{18}NO_2P$ MS

m/z [M+1] calculated 275.28. Observed [M+1] 275.1.



(±) (1-(Benzhydrylamino)butyl)phosphinic acid 4.2d-4

Yield:2.68 g, 8.84 mmol, 54 % yield; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 0.9(t, *J* = 7 Hz, 3H, H-5) 1.3(m, 2H, H-4), 1.6 (q, *J* = 6 Hz, 2H, H-3), 2.1(bs, 1H, NH), 2.5(m, 1H, H-2), 5.1(s, 1H, H-1), 7.3(m, 10H, PhH); $C_{17}H_{22}NO_2P$ MS m/z [M+1] calculated 303.33. Observed [M+1] 303.1.



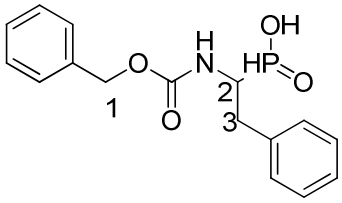
(±) (1-(Benzhydrylamino)-2,2-dimethylpropyl)phosphinic acid 4.2d-2

Yield: 2.8 g, 8.95 mmol, 60%; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 0.9(s, 9H, H-3/4/5), 2.0(bs, 1H, NH), 2.7(s, 1H, H-2), 5.1(s, 1H, H-1), 7.3(m, 10H, PhH); $C_{18}H_{24}NO_2P$ MS m/z [M+1] calculated 317.36. Observed [M+1] 317.1.

General Procedure for the preparation of compounds 4.2f-1, 4.2f-2, 4.2f-3 and 4.2f-4.

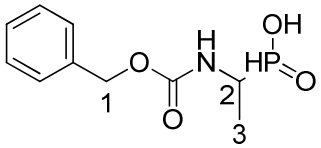
The phosphinic acid (14.23 mmol) was dissolved in 48% HBr (10 ml, 184 mmol) and heated to 120 $^{\circ}$C for 2 h. The mixture was then concentrated under reduced pressure and the residue diluted with water and extracted with diethyl ether

(3 x 50 ml). The pH of the mixture was then adjusted to pH 10 with NaOH (4 M) and brought to 0 ºC on an ice-bath. Benzyl chloroformate (2.438 ml, 17.08 mmol) was added slowly over 5 minutes and the mixture stirred at 0 ºC for one hour and at ambient temperature overnight and extracted with diethyl ether (2 x 50 ml). The aqueous phase is then acidified to pH 2 by the addition of 6 M HCl. The mixture was then extracted with EtOAc (3 x 50 ml), washed with H_2O, dried over $MgSO_4$ and the solvent removed in vacuo leaving the benzyloxyphosphinic acid's **4.2f-1**, **4.2f-2**, **4.2f-3** and **4.2f-4**.



(±)(1-(((Benzyloxy)carbonyl)amino)-2-phenylethyl)phosphinic acid 4.2f-3

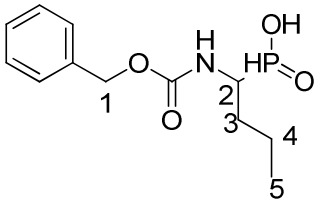
Yield: 2.91 g, 64 %; White chrystalline solid; m.p. 131-133 °C;^{1}H NMR (400 MHz, $DMSO_6$): δ_H 2.9 (d, *J* = 8 Hz, 2H, H-3), 4.1 (t, *J* = 8 Hz, 1H, H-2), 5.2 (s, 2H, H-1), 7.3 (m, 10H, PhH).$C_{16}H_{18}NO_4P$ MS m/z [M+1] calculated 319.36. Observed [M+1] 319.1.



(±) (1-(((Benzyloxy)carbonyl)amino)ethyl)phosphinic acid 4.2f-1

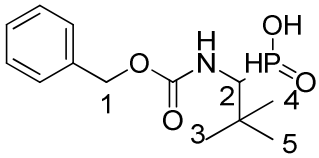
Yield: 2.83 g, 64 %; White chrystalline powder; m.p. 101 - 104 °C; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 1.4 (d, *J* = 9 Hz, 3H, H-

3), 3.6 (m, 1H, H-2), 5.2 (s, 2H, H-1), 7.3 (m, 5H, PhH).$C_{10}H_{14}NO_4P$ MS m/z [M+1] calculated 243.19. Observed [M+1] 243.1.



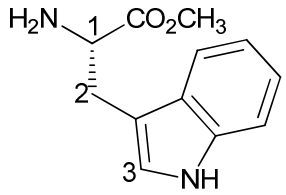
(±) (1-(((Benzyloxy)carbonyl)amino)butyl)phosphinic acid 4.2f-4

Yield: 0.626 g, 70 %; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 0.9(t, *J* = 7 Hz, 3H, H-5), 1.3 (m,2H, H-4), 1.7(m, 2H, H-3), 3.6 (t, *J*=7 Hz, 1H, H-2), 5.2 (s, 2H, H-1), 7.3 (m, 5H, PhH), 8.1(m, 1H, NH).$C_{12}H_{18}NO_4P$ MS m/z [M+1] calculated 271.19. Observed [M+1] 271.1.



(±)(1-(((Benzyloxy)carbonyl)amino)-2,2-dimethylpropyl)phosphinic acid 4.2f-2

Yield: 0.626 g, 70 %; White crystalline powder; m.p. 99 – 103 °C; ^{1}H NMR (400 MHz, $DMSO_6$): δ_H 0.9(s, 9H, H-3/4/5), 3.4 (d, *J* = 12 Hz, 1H, H-2), 5.2 (s, 2H, H-1), 7.3 (m, 5H, PhH), 8.1(s, 1H, NH).$C_{13}H_{20}NO_4P$ MS m/z [M+1] calculated 285.34. Observed [M+1] 285.5.



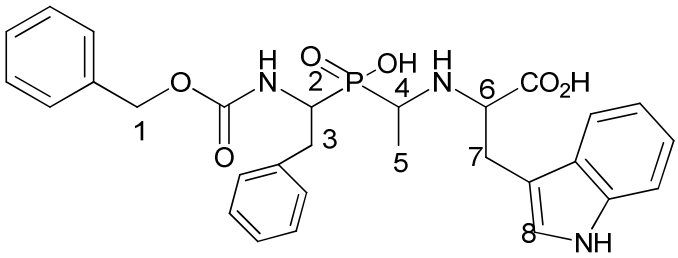
(S)-methyl 2-amino-3-(1H-indol-3-yl)propanoate (Tryptophan methyl ester)

Tryptophan (5.0 g, 24.50 mmol) was suspended in MeOH (20 mL) and the mixture was cooled to 0 °C on an ice-bath. $SOCl_2$ (2.7 ml, 36.7 mmol) was added drop-wise over 20 min. Upon completion of addition and when all the tryptophan had dissolved into the mixture, the reaction was allowed to warm to room temperature and was stirred overnight. The volatiles were removed under reduced pressure leaving the product. Tryptophan methyl ester (5.3 g, 24.0 mmol, 99%) which was used in the next step without further purification.^{1}H NMR (400 MHz, $CDCl_3$): δ_H3.2 (d, *J* = 11 Hz, 2H, H-2), 3.6 (s, 3H, OC*H*$_3$), 4.2 (t, *J* = 7 Hz, 1H, H-1), 7.1(s, 1H, H-3), 7.3 (m, 5H, PhH).$C_{12}H_{14}N_2O_2$ MS m/z [M+1] calculated 217.16. Observed [M+1] 217.1.

General Procedure for the preparation of compounds 4.2i-1 up to 4.2i-16.

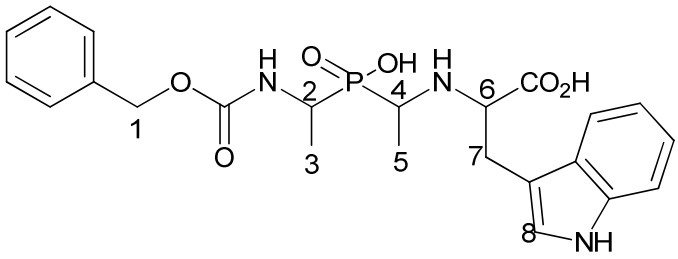
Tryptophan methyl ester (0.14 g, 0.63 mmol), an aldehyde (0.626 mmol) and the phosphinic acid (0.63 mmol)were dissolved in MeOH (5 ml) and 2 drops of HCl were added. The reaction mixture was refluxed for 16 h. The solvent was removed under reduced pressure and then dissolved in MeOH (5 ml). NaOH (3 ml, 4 M) was added and the mixture stirred overnight. The MeOH was removed under reduced pressure and the mixture separated with diethyl ether (2 x 20 ml). The aqueous phase was acidified to pH 2 (6 N HCl) and then

extracted with EtOAc (2 x 30 ml), dried over $MgSO_4$and reduced under pressure. The crude product was recrystallized from MeOH giving the products **4.2i-1** up to **4.2i-16.**



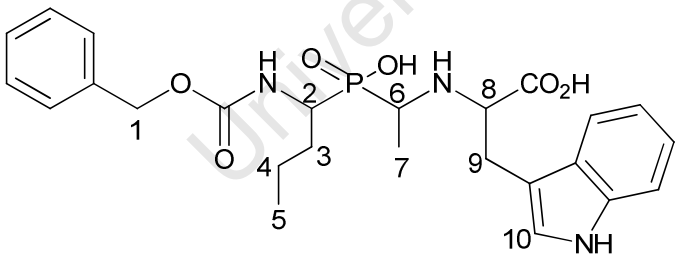
(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)-2-phenylethyl)(hydroxy)phosphoryl)ethyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-3

Yield: 0.147 g, 43 %; Brown crystalline powder from MeOH; m.p. 113 – 115 ^{o}C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.3 (d, *J* = 9 Hz, 3H, H-5), 3.0 (m,2H, H-4/7a), 3.2 (dd, *J* = 11, 6 Hz, 1H,H-7b), 3.3 (dd, *J* = 12, 5 Hz, 1H, H-3a), 3.6 (m, 1H, H-3b), 4.4 (t, *J* = 7 Hz, 1H, H-6), 4.5 (m, 1H, H-2), 5.2 (s, 2H, H-1)7.1 (s, 1H, H-8), 7.3 (m, 13H, PhH); ^{13}C NMR, ($CDCl_3$)19.9, 28.2, 40.3, 55.8, 59.3, 65.9, 66.8, 110.7, 112.1, 118.8, 119.8, 121.7, 123.0, 125.9, 127.1 (2C), 127.4, 127.6, 127.7 (2C), 128.6(2C), 128.9(2C), 136.1, 136.5, 137.7, 155.9, 171.5; $C_{29}H_{32}N_3O_6P$ MS m/z [M+1] calculated 549.1. Observed [M+] 548.1; HPLC Purity: 85%; t_R' = 25.1 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)ethyl)(hydroxy)phosphoryl)ethyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-1

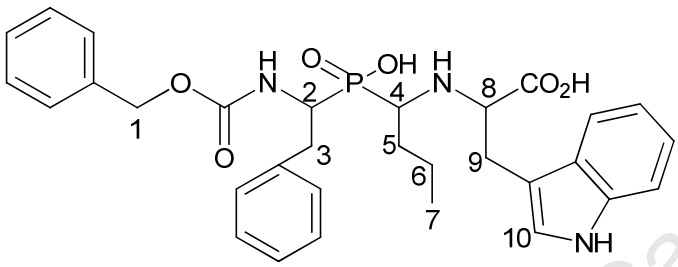
Yield:0.141 g,81 %; Tan coloured crystalline material from MeOH; m.p. 91 – 95 ^{o}C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.4(m, 6H, H-3/5), 2.8 (m, 1H, H-4), 2.9(bs, 1H, NH), 3.12 (dd, *J* = 12, 5 Hz, 1H, H-7a), 3.3 (dd, *J* = 12. 5 Hz, 1H, H-7b), 4.2 (m, 1H, H-2),4.5 (t, *J* = 7 Hz, 1H, H-4), 5.2(s, 2H, H-1), 7.1 (s, 1H, H-8), 7.3 (m, 10H, PhH), 8.1(s, 1H, NH); ^{13}C NMR, (100 MHz, $CDCl_3$);18.7, 19.3, 28.2, 54.7, 55.5, 68.9, 69.8, 109.7, 114.1, 120.8, 121.8, 122.1, 123.0, 127.1(4C), 128.9(2C), 136.1(2C), 155.9, 171.5; $C_{23}H_{28}N_3O_6P$ MS m/z [M+1] calculated 471.4. Observed [M+] 470.2. HPLC Purity: 75%; t_R' = 23.1 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)butyl)(hydroxy)phosphoryl)ethyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-4

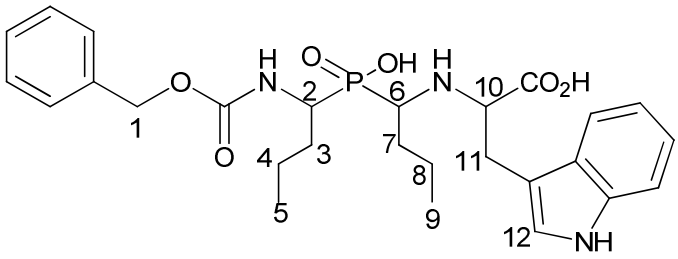
Yield:0.101 g, 66 %; Brown crystalline powder from MeOH; m.p. 112 – 115 ^{o}C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9(t, *J* = 7 Hz, 3H, H-5), 1.2(d, *J* = 12 Hz, 2H, H-4), 1.5 (d, *J* = 12 Hz, 3H, H-7), 2.0(m, 1H, H-3a), 2.1(m, 1H, 3b), 2.6 (m, 1H,

H-6),3.0 (dd, *J* = 12, 5 Hz, 1H, H-9a), 3.1(m, 2H, H-8/9b), 4.0(m, 1H, H-2), 5.1 (s, 2H, H-1), 7.1 (s, 1H, H-10), 7.3 (m, 9H, PhH), 8.0(s, 1H, NH); ^{13}C NMR, ($CDCl_3$);14.9, 18.9, 21.3, 26.6, 27.2(2C), 54.9, 65.9, 66.8, 107.7, 109.4, 118.8, 119.8, 121.7, 125.0, 128.1(2C), 128.4(2C), 128.9 (2C), 136.1(2C), 155.9,175.5; $C_{25}H_{32}N_3O_6P$ MS m/z [M+1] calculated 500.5. Observed [M+] 499.2.HPLC Purity: >95%; t_R' = 24.3 min.



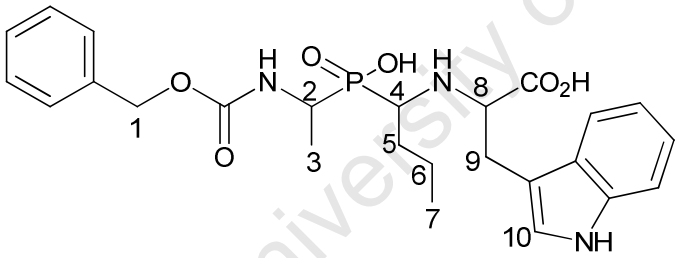
(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)-2-phenylethyl)(hydroxy)phosphoryl)butyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-7

Yield: 0.130 g,72 %; White crystalline material from MeOH; m.p. 134 – 137 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9(t, *J* = 8 Hz, 3H, H-7), 1.5 (m, 2H, H-6), 1.9 (m, 2H, H-5), 2.5(m, 1H, H-4), 2.7(dd, *J* = 12, 5 Hz, 1H, H-9a), 3.1 (dd, J = 12, 5 Hz, 1H, H-9b), 3.3 (dd, *J* = 12, 5 Hz, 1H, H-3a), 3.6 (m, 2H, H-3a/8), 4.3 (m, 1H, H-2), 5.2(s, 2H, H-1), 7.3 (m, 15H, PhH), 8.0(s, 1H, NH); ^{13}C NMR, (100 MHz, $CDCl_3$);13.8, 16.1, 18.2, 27.2(2C), 39.3, 57.6(2C), 66.8(2C), 109.7(2C), 119.8(2C), 121.7, 123.0, 125.9, 127.1(6C),128.6(4C),136.1(2C), 137.7, 171.5.; $C_{31}H_{36}N_3O_6P$ MS m/z [M+1] calculated 576.5. Observed [M+1] 575.1. HPLC Purity: >95%; t_R' = 29.1 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)butyl)(hydroxy)phosphoryl)butyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-8

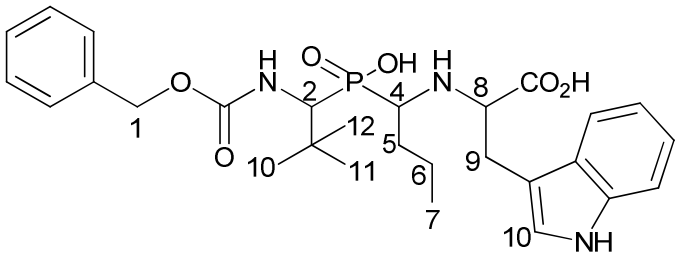
Yield: 0.066 g, 68%; White crystalline material from MeOH; m.p. 91 – 94 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9 (m, 6H, H-5/9), 1.2 (m, 2H, H-4), 1.6(m, 2H, H-8), 1.8 (m, 2H, H-3a/7b), 2.2 (m, 1H, H-3b), 2.4 (bs, 1H, NH), 2.8(d, *J* = 6 Hz, 1H, H-6), 2.9 (dd, *J* = 12, 6 Hz, 1H, H-11a), 3.4 (dd, *J* = 12, 6 Hz, 1H, H-11b), 3.9 (d, *J* = 9 Hz, 1H, H-2), 4.4 (t, *J* = 7 Hz, 1H, H-10), 5.2(s, 2H, H-1), 7.3 (m, 10H, PhH), 7.9(s, 1H, NH);$C_{27}H_{36}N_3O_6P$ MS m/z [M+1] calculated 527.6. Observed [M+] 527.1.HPLC Purity: >90%; t_R' = 25.9 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)ethyl)(hydroxy)phosphoryl)butyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-5

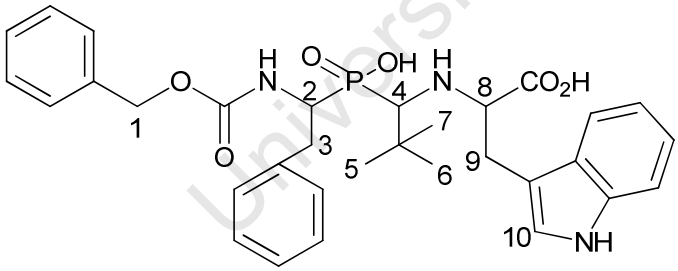
Yield: 0.070 g, 68 %; White crystalline powder from MeOH; m.p. 101 – 103 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9(t, *J* = 8 Hz, 3H, H-7), 1.4 (d, *J* = 9 Hz, 2H, H-3), 1.6 (m, 2H, H-6), 1.9 (m, 2H, H-5), 2.0(bs, 1H, NH), 2.7 (t, *J* = 8 Hz, 1H, H-4), 2.9 (d, *J* = 12 Hz, 2H, H-9),3.9(m, 1H, H-2), 4.3 (t, 1H, *J* = 7 Hz, H-8), 5.2(s, 2H, H-1), 7.3 (m, 10H, PhH), 8.2(s, 1H, NH);$C_{25}H_{32}N_3O_6P$ MS m/z [M+1] calculated 501.5.

Observed [M+] 500.2.HPLC Purity: >95%; t_R' = 24.9 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)-2,2-dimethylpropyl)(hydroxy)phosphoryl)butyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-6

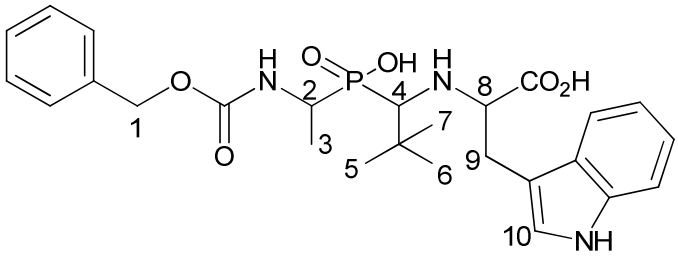
Yield: 0.066 g, 70 %; Tan crystalline powder from MeOH; m.p. 113 – 115 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9 (t, J = 7 Hz, 3H, H-7), 1.1(s, 9H, H-10/11/12), 1.4 (m, 4H, H-6), 1.8 (m, 2H, H-5), 2.0(bs, 1H, NH), 2.5(t, *J* = 8 Hz, 1H, H-4), 3.1 (d, *J* = 9 Hz, 2H, H-9), 4.2 (s, 1H, H-2), 4.5 (t, 1H, *J* = 7 Hz, H-8), 5.2(s, 2H, H-1), 7.3 (m, 10H, PhH), 8.2(s, 1H, NH);$C_{28}H_{38}N_3O_6P$ MS m/z [M+1] calculated 541.6. Observed [M+1] 541.1.HPLC Purity: >94%; t_R' = 28.2 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)-2-phenylethyl)(hydroxy)phosphoryl)-2,2-dimethylpropyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-11

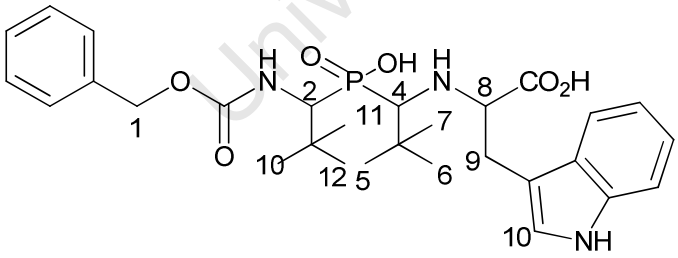
Yield: 0.076 g, 82 %; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9(s, 9H, H-5/6/7), 2.1(bs, 1H, NH), 2.6(m, 1H, H-4), 2.9 (dd, *J* = 11.5 6 Hz, 1H, H-9a), 3.1 (dd, *J* = 11.5, 6 Hz, 1H, H-9b), 3.2 (m, 1H, H-3a), 3.4 (dd, *J* = 11.5, 6 Hz, 1H, H-3b), 4.4 (m, 1H, H-8), 4.8 (m, 1H, H-2), 5.2(s, 2H, H-1), 7.3 (m,

15H, PhH), 8.1(s, 1H, NH);$C_{32}H_{38}N_3O_6P$ MS m/z [M+1] calculated 590.6. Observed [M+] 589.1.HPLC Purity: >90%; t_R' = 29.8 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)ethyl)(hydroxy)phosphoryl)-2,2-dimethylpropyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-9

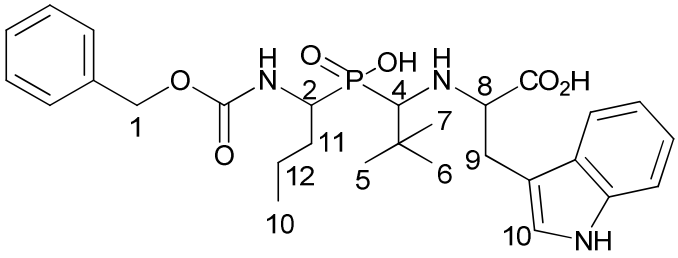
Yield: 0.083 g, 79 %; Brown crystalline powder from MeOH; m.p. 116 – 119 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.1(s, 9H, H-5/6/7), 1.6(d, *J* = 9 Hz, 2H, H-3), 2.0(bs, 1H, NH), 2.6(s, 1H, H-4), 3.0 (dd, *J* = 12, 5 Hz, 1H, H-9a),3.3 (m, 1H, H-9b), 3.9(m, 1H, H-2/8), 5.2(s, 2H, H-1), 7.3 (m, 10H, PhH), 8.0(s, 1H, NH);$C_{26}H_{34}N_3O_6P$ MS m/z [M+1] calculated 514.5. Observed [M+] 513.3.HPLC Purity: >95%; t_R' = 28.4 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)-2,2-dimethylpropyl)(hydroxy)phosphoryl)-2,2-dimethylpropyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-10

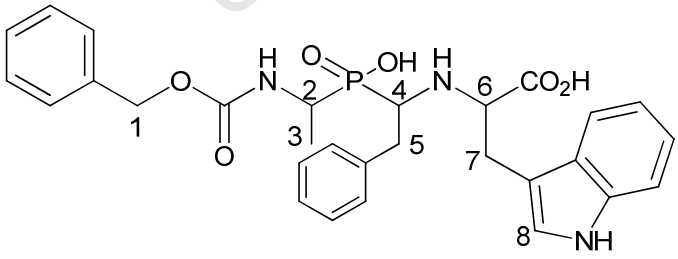
Yield: 0.077 g, 79 %; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.1 (m, 18H, H-5/6/7/10/11/12), 2.0(bs, 1H, NH), 2.6(s, 1H, H-4),

3.0 (d, *J* = 12 Hz, 2H, H-9), 4.1 (s, 1H, H-2), 4.4 (t, *J* = 7 Hz, 1H, H-8), 5.2(s, 2H, H-1), 7.3 (m, 10H, PhH), 8.0(s, 1H, NH);$C_{29}H_{40}N_3O_6P$ MS m/z [M+1] calculated 556.6. Observed [M+1] 555.1.HPLC Purity: >95%; t_R' = 29.1 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)butyl)(hydroxy)phosphoryl)-2,2-dimethylpropyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-12

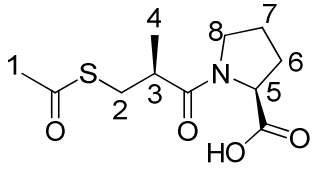
Yield: 0.088 g, 88 %; Brown crystalline solid from MeOH; m.p. 77 - 80 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9 (t, *J* = 6 Hz, 3H, H-10), 1.1 (s, 9H, H-5/6/7), 1.2 (m, 2H, H-12), 1.9 (m, 1H, H-11a), 2.2 (m, 1H, H-11b), 2.6 (s, 1H, H-4), 3.0 (dd, *J* = 13, 6 Hz, 1H, H-9a), 3.4 (m, 1H, H-9b), 4.0 (m, 2H, H-2/8), 5.2 (s, 2H, H-1), 7.3 (m, 10H, PhH), 8.0(s, 1H, NH);$C_{28}H_{38}N_3O_6P$ MS m/z [M+1] calculated 541.6. Observed [M+] 540.1.HPLC Purity: >95%; t_R' = 27.9 min.



(±)2-((1-((1-(((Benzyloxy)carbonyl)amino)ethyl)(hydroxy)phosphoryl)-2-phenylethyl)amino)-3-(1H-indol-3-yl)propanoic acid 4.2i-13

Yield: 0.09 g, 80 %; Light brown crystalline solid from

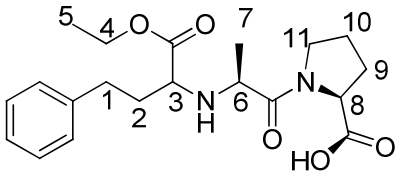
MeOH; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.1 (d, *J* = 7 Hz, 3H, H-3), 3.0(m, 1H, H-7a), 3.4 (m, 4H, H-4/5/7b), 4.0 (m, 1H, H-2), 4.4 (m, 1H, H-6), 5.2(s, 2H, H-1), 7.3 (m, 15H, PhH), 8.0(s, 1H, NH);$C_{29}H_{32}N_3O_6P$ MS m/z [M+1] calculated 548.6. Observed [M+] 547.1.HPLC Purity: 70%; t_R' = 26.1 min.



(S)-1-((S)-3-(Acetylthio)-2-methylpropanoyl)pyrrolidine-2-carboxylic acid 3.29d

(*S*)-3-(acetylthio)-2-methylpropanoic acid **3.29a** (0.100 g, 0.616 mmol) was dissolved in DCM (10 ml) and cooled to 0 ^oC on ice. $SOCl_2$ (0.054 ml, 0.740 mmol) was added dropwise over 3 min. Upon completion of $SOCl_2$ addition, the reaction mixture was heated at 70 ^oC for 4 h. The reaction mixture was then allowed to cool to ambient temperature and then the solvent was removed in vacuo. The residue was then dissolved in dioxane (10 ml) and then water (10 ml) was added. The mixture was cooled to 5 ^oC then $NaHCO_3$ (0.078 g, 0.925 mmol) was added. The mixture was stirred until all the solids dissolved completely and then proline (S)-pyrrolidine-2-carboxylic acid (0.071 g, 0.616 mmol) was added. The mixture was allowed to warm to ambient temperature and then stirred overnight. Most of the dioxane was removed under reduced pressure and water (20 ml) was added and the mixture was separated with EtOAc (50 ml x 2) and the organic layer discarded. The aqueous layer was then acidified to pH 2 with HCl (6 *N*). The mixture was separated with EtOAc and the organic layer washed with water (2 x 50

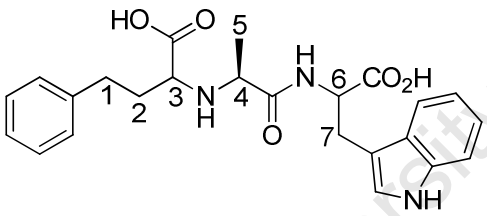
ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure leaving (S)-1-((S)-3-(acetylthio)-2-methylpropanoyl)pyrrolidine-2-carboxylic acid **3.29d**(0.128 g, 0.493 mmol, 80 % yield) as a clear colourless residue which was used in the next step without further purification.^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.1(d, *J* = 8 Hz, 3H, H-4), 1.9(m, 3H, H-6a/7), 2.1(m, 1H, H-6b), 2.4(s, 3H, Acetyl-H), 3.5(m, 3H, H-2a/3/8a), 3.6(m, 1H, H-8b), 3.8(m, 1H, H-2a), 4.7(t, *J* = 7 Hz, 1H, H-5).



(S)-1-((S)-2-(((S)-1-Ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoyl)pyrrolidine-2-carboxylic acid (Enalapril) 3.30c

(S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoic acid **3.30a**(0.10 g, 0.36 mmol) was dissolved in DCM (10 ml) and cooled to 0 ºC on ice and PCl_5(0.08 g, 0.36 mmol) was added slowly over 3 min. Upon completion of PCl_5 addition, the reaction mixture was stirred at ambient temperature for 5 h. The solvent was removed in vacuo. The residue was then dissolved in 1,4-dioxane (10 ml) and then water (10 ml) was added. The mixture was cooled to 5 ºC then $NaHCO_3$(0.045 g, 0.537 mmol) was added. The mixture was stirred until all the solids dissolved completely and then proline (S)-pyrrolidine-2-carboxylic acid (0.041 g, 0.358 mmol) was added. The mixture was allowed to warm to ambient temperature and then stirred overnight. Most of the 1,4-dioxane was removed under reduced pressure and water (20 ml) was added and the mixture was separated with EtOAc (50 ml x 2) and the

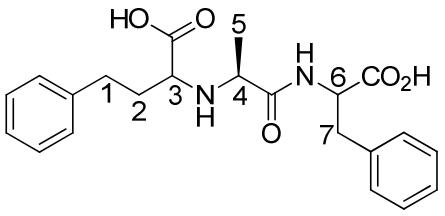
organic layer discarded. The aqueous layer was then acidified to pH 2 with HCl (6 *N*). The mixture was seperated with EtOAc and the organic layer washed with water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure leaving (S)-1-((S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoyl)pyrrolidine-2-carboxylic acid **3.30c**(0.081 g, 0.215 mmol, 60 % yield) as a clear colourless residue which was used in the next step without further purification.^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.2(t, *J* = 7 Hz, 3H, H-5), 1.4(d, *J* = 11 Hz, 3H, H-7), 1.9(m, 3H, H-2a/10), 2.1(m, 1H, H-9a), 2.3(m, 2H, H-2b/9b), 2.7(t, *J* = 7 Hz, 2H, H-1), 3.1(m, 1H, H-11a), 3.5(t, *J* = 7 Hz, 1H, H-3), 7.2(m, 5H, Phe-H). $C_{20}H_{28}N_2O_5$ MS m/z [M+1] calculated 376.4. Observed [M+] 375.2.



(2S)-2-(((2S)-1-((1-Carboxy-2-(1H-indol-3-yl)ethyl)amino)-1-oxopropan-2-yl)amino)-4-phenylbutanoic acid 4.4e

(S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoic acid **4.4a**(0.10 g, 0.36 mmol) was dissolved in DCM (10 ml) and cooled to 0 ºC on ice and PCl_5 (0.08 g, 0.36 mmol) was added slowly over 3 min. Upon completion of PCl_5 addition, the reaction mixture was stirred at ambient temperature for 5 h. The solvent was removed in vacuo. The residue was then dissolved in dioxane (10 ml) and then water (10 ml) was added. The mixture was cooled to 5 ºC then $NaHCO_3$ (0.045 g, 0.537 mmol) was added.

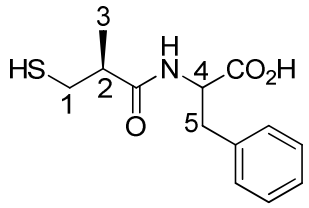
The mixture was stirred until all the solids dissolved completely and then phenylalanine methyl ester was added. The mixture was allowed to warm to ambient temperature and then stirred overnight. Most of the 1,4-dioxane was removed under reduced pressure and water (20 ml) was added and the mixture was separated with EtOAc (50 ml x 2) and the organic layer was washed with HCl (50 ml, 0.05 M), $NaHCO_3$ (aq) (50 ml, 5 %) water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure. The residue was then dissolved in MeOH (10 ml) and NaOH (15 ml, 4M) was added and the mixture stirred overnight. The reaction mixture was separated with EtOAc and the organic layer discarded. The aqueous layer was then acidified to pH 2 with HCl (6 *N*) and separated with EtOAc, washed with water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and the solvent removed under reduced pressure. The residue was purified on silica gel (EtOAc: Hexane: 1:5) and the product was collected as a white solid **4.4e**(0.108 g, 0.247 mmol, 69 % yield); ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.2 (d, *J* = 7 Hz, 3H, H-5), 1.9 (m, 2H, H-2), 2.7 (t, *J* = 7 Hz, 2H, H-1), 2.9 (t, *J* = 7 Hz, 1H, H-3), 3.0 (m, 2H, H-4/7a), 3.3(dd, *J* = 12, 5 Hz, 1H, H-7b), 5.1 (t, *J* = Hz, 1H, H-6), 7.1(m, 8H, Phe-H/Indolyl-H), 7.3(s, 1H, Indolyl-H), 7.6(s, 1H, Indolyl-H). $C_{24}H_{27}N_3O_5$ MS m/z [M+1] calculated 437.5. Observed [M+1] 437.1.



(2S)-2-(((2S)-1-((1-Carboxy-2-phenylethyl)amino)-1-oxopropan-2-yl)amino)-4-phenylbutanoic acid 4.4d

(S)-2-(((S)-1-ethoxy-1-oxo-4-phenylbutan-2-yl)amino)propanoic acid **4.4a**(0.100 g, 0.358 mmol) was dissolved in DCM (10 ml) and cooled to 0 $^{\circ}$C on ice and PCl_5 (0.075 g, 0.358 mmol) was added slowly over 3 min. Upon completion of PCl_5 addition, the reaction mixture was stirred at ambient temperature for 5 h. The solvent was removed in vacuo. The residue was then dissolved in 1,4-dioxane (10 ml) and then water (10 ml) was added. The mixture was cooled to 5 ^{O}C then $NaHCO_3$ (0.045 g, 0.537 mmol) was added. The mixture was stirred until all the solids dissolved completely and then phenylalanine methyl ester methyl 2-amino-3-phenylpropanoate (0.064 g, 0.358 mmol) was added. The mixture was allowed to warm to ambient temperature and then stirred overnight. Most of the 1,4-dioxane was removed under reduced pressure and water (20 ml) was added and the mixture was separated with EtOAc (50 ml x 2) and the organic layer was washed with HCl (50 ml, 0.05 M), $NaHCO_3$ (aq) (50 ml, 5 %) water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure. The residue was then dissolved in MeOH (10 ml) and NaOH (15 ml, 4M) was added and the mixture stirred overnight. The reaction mixture was separated with EtOAc and the organic layer discarded. The aqueous layer was then acidified to pH 2 with HCl (6 *N*) and separated with EtOAc, washed with water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over MgSO4 and the solvent removed under reduced pressure.The residue was purified on silica gel (EtOAc: Hexane: 1:5) and the product was collected as a white crystalline solid 4.4d (0.086 g, 0.215 mmol, 60 % yield); ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.4 (d, *J* = 9 Hz, 3H, H-5), 2.0 (m, 1H, H-2a), 2.2 (m, 1H, H-2b), 2.7(t, *J* = 7 Hz, 1H, H-3), 2.8 (t, *J* = 7 Hz, 2H,

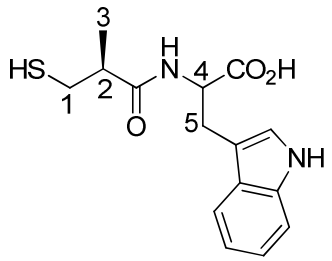
H-1), 2.9 (dd, *J* = 12, 5 Hz, 1H, H-7a), 3.1 (dd, *J* = 12, 5 Hz, 1H, H-7b), 3.6 (q, *J* = 6 Hz, 1H, H-4), 5.0 (t, *J* = 7 Hz, 1H, H-6), 7.3(m, 5H, Phe-H); $C_{22}H_{26}N_2O_5$ MS m/z [M+1] calculated 398.45. Observed [M+] 397.1.



2-((S)-3-Mercapto-2-methylpropanamido)-3-phenylpropanoic acid 4.5d.

(S)-3-(acetylthio)-2-methylpropanoic acid **4.5a**(0.100 g, 0.616 mmol) was dissolved in DCM (10 ml) and cooled to 0 ^{o}C on ice. $SOCl_2$ (0.054 ml, 0.740 mmol) was added dropwise over 3 min. Upon completion of $SOCl_2$ addition, the reaction mixture was heated to 70 ^{o}C for 4 h. The reaction mixture was then allowed to cool to ambient temperature and then the solvent was removed in vacuo. The residue was then dissolved in dioxane (10 ml) and then water (10 ml) was added. The mixture was cooled to 5 ^{o}C then $NaHCO_3$ (0.078 g, 0.925 mmol) was added. The mixture was stirred until all the solids dissolved completely and then phenylalanine 2-amino-3-phenylpropanoic acid (0.102 g, 0.616 mmol) was added. The mixture was allowed to warm to ambient temperature and then stirred overnight. Most of the dioxane was removed under reduced pressure and water (20 ml) was added and the mixture was separated with EtOAc (50 ml x 2) and the organic layer discarded. MeOH (10 ml) and NaOH (aq) (15 ml, 4 M) was added and the mixture stirred overnight. The mixture was separated with EtOAc (50 ml x 2) and the organic layer discarded. The aqueous layer was then

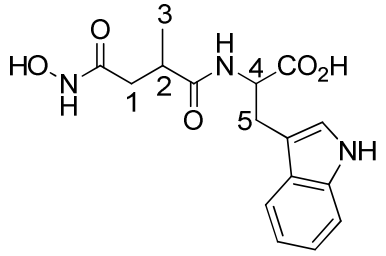
acidified to pH 2 with HCl (6 *N*). The mixture was separated with EtOAc and the organic layer washed with water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure leaving a clear yellow residue. Purification on silica gel with EtOAc and Hexane (1:4) gave the pure product 2-((S)-3-mercapto-2-methylpropanamido)-3-phenylpropanoic acid **4.5d**(0.105 g, 0.395 mmol, 64 % yield) as a clear yellow solid.^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.3 (d, *J* = 8 Hz, 3H, H-3), 1.4(s, 1H, SH), 2.5 (dd, *J* = 12, 5 Hz, 1H, H-1a), 2.9(m, 1H, H-5a), 3.0 (dd, *J* = 12, 5 Hz, 1H, H-1b), 3.0 (dd, *J* = 12, 5 Hz, 1H, H-1b), 3.1(m, 1H, H-4), 3.2 (dd, *J* = 12, 5 Hz, 1H, H-5b), 4.8 (t, *J* = 6 Hz, 1H, H-4), 7.2(m, 5H, Phe-H). $C_{13}H_{17}NO_3S$ MS m/z [M+1] calculated 267.34. Observed[M+] 266.3.



3-(1H-Indol-3-yl)-2-((S)-3-mercapto-2-methylpropanamido)propanoic acid 4.5e

(S)-3-(Acetylthio)-2-methylpropanoic acid **4.5b** (0.100 g, 0.616 mmol) was dissolved in DCM (10 ml) and cooled to 0 ^oC on ice. $SOCl_2$ (0.054 ml, 0.740 mmol) was added dropwise over 3 min. Upon completion of $SOCl_2$ addition, the reaction mixture was heated to 70 ^oC for 4 h. The reaction mixture was then allowed to cool to ambient temperature and then the solvent was removed in vacuo. The residue was then dissolved in dioxane (10 ml) and then water (10 ml) was added. The mixture was cooled to 5 ^oC then $NaHCO_3$ (0.078 g, 0.925 mmol) was added. The mixture was stirred until all

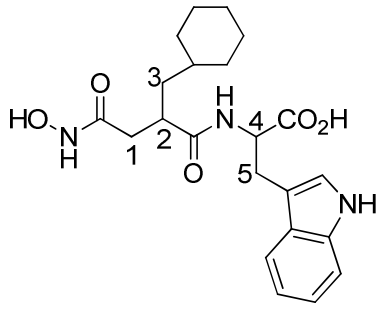
the solids dissolved completely and then tryptophan 2-amino-3-(1H-indol-3-yl)propanoic acid (0.126 g, 0.616 mmol) was added. The mixture was allowed to warm to ambient temperature and then stirred overnight. Most of the 1,4-dioxane was removed under reduced pressure and water (20 ml) was added and the mixture was separated with EtOAc (50 ml x 2) and the organic layer discarded. MeOH (10 ml) and NaOH (aq) (15 ml, 4 M) was added and the mixture stirred overnight. The mixture was separated with EtOAc (50 ml x 2) and the organic layer discarded. The aqueous layer was then acidified to pH 2 with HCl (6 *N*). The mixture was separated with EtOAc and the organic layer washed with water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure leaving a clear yellow residue. Purification on silica gel with EtOAc and Hexane (1:4) gave the pure product 3-(1H-indol-3-yl)-2-((S)-3-mercapto-2-methylpropanamido)propanoic acid **4.5e** (0.121 g, 0.395 mmol, 64 % yield) as a clear yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.3(d, *J* = 9 Hz, 3H, H-3), 2.9 (m, 1H, H-1a), 3.1(m, 2H, H-1b/2), 3.1(dd, *J* = 13, 7 Hz, 1H, H-5a), 3.3(dd, *J* = 13, 7 Hz, 1H, H-5b) 4.9 (t, *J* = 6 Hz, 1H, H-4), 7.1(m, 3H, Indolyl-H), 7.3(s, 1H, Indolyl-H), 7.6(s, 1H, Indolyl-H). $C_{15}H_{18}N_2O_3S$ MS m/z [M+1] calculated 306.34. Observed [M+] 307.2.



2-((R)-4-(Hydroxyamino)-2-methyl-4-oxobutanamido)-3-(1H-indol-3-yl)propanoic acid 4.7e

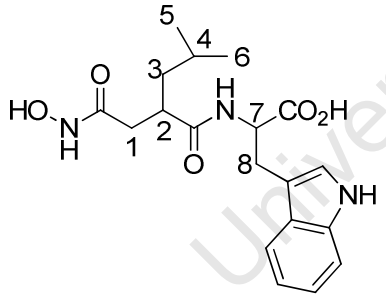
(R)-4-((benzyloxy)amino)-2-methyl-4-oxobutanoic acid **3.2a-1** (0.05 g, 0.211 mmol) was dissolved in DMF (10 ml) and cooled to 0 ^{O}C in an ice bath. EDC (0.040 g, 0.211 mmol), HOBT (0.032 g, 0.211 mmol), DIEA (0.037 ml, 0.211 mmol) were added. The reaction mixture was stirred overnight. Water (100 ml) was added and the mixture was separated with EtOAc (50 ml x 3) and the organic layer was washed with HCl (50 ml, 0.05 M), NaHCO3 (aq) (50 ml, 5 %) water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and then solvent removed under reduced pressure. The residue was dissolved in MeOH (10 ml) and NaOH (15 ml, 4M) was added and the mixture stirred overnight. The reaction mixture was separated with EtOAc and the organic layer discarded. The aqueous layer was then acidified to pH 2 with HCl (6 *N*) and separated with EtOAc, washed with water (2 x 50 ml) and brine (1 x 50 ml). It was then dried over $MgSO_4$ and the solvent removed under reduced pressure. The residue was purified on silica gel (EtOAc: Hexane: 1:5) and the product was collected as a white solid. The white solid was dissolved in MeOH (20 ml) and Pd-C (0.09 g, 0.846 mmol) was added. The mixture was shaken on a Parr hydrogenator for three hours. The mixture was filtered through celite and MeOH was removed under reduced pressure leaving the product 2-((R)-4-(hydroxyamino)-2-methyl-4-oxobutanamido)-3-(1H-indol-3-yl)propanoic acid (0.046 g, 0.139 mmol, 66 % yield) as white solid.^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.2(d, *J* = 7 Hz, 3H, H-3), 2.5(dd, *J* = 12, 5 Hz, 2H, H-1a), 2.9(m, 1H, H-2), 3.1 (m, 2H, H-1b/5b), 3.3 (dd, *J* = 12, 6 Hz, 1H, H-5b), 3.3 (t, *J* = 6 Hz, 1H, H-4), 7.1 (m, 3H, Indolyl-H), 7.3 (s, 1H, Indolyl-H), 7.6 (s, 1H, Indolyl-H). $C_{16}H_{19}N_2O_2$ MS m/z [M+1] calculated 333.33. Observed [M+] 332.1.

This is the general method that was employed in the preparation **4.7a**, **4.7b**, **4.7c** and **4.7d**.



(±)2-((R)-2-(Cyclohexylmethyl)-4-(hydroxyamino)-4-oxobutanamido)-3-(1H-indol-3-yl)propanoic acid 4.7a

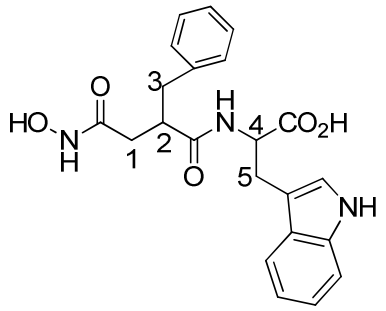
Yield: 0.046 g, 71 %; White crystalline material; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.4(m, 13H, H-3/Cyclohexyl-H), 2.6(d, *J*= 9 Hz, 2H, H-1), 2.9(m, 1H, H-2), 3.4(d, *J* = 6 Hz, 1H, H-5) 4.5(t, *J* = 6 Hz, 1H, H-4), 7.1(m, 3H, Indolyl-H), 7.3(s, 1H, Indolyl-H), 7.6(s, 1H, Indolyl-H). $C_{22}H_{29}N_3O_5$ MS m/z [M+1] calculated 415.48. Observed [M+] 414.1.



(±)2-((R)-2-(2-(Hydroxyamino)-2-oxoethyl)-4-methylpentanamido)-3-(1H-indol-3-yl)propanoic acid 4.7b

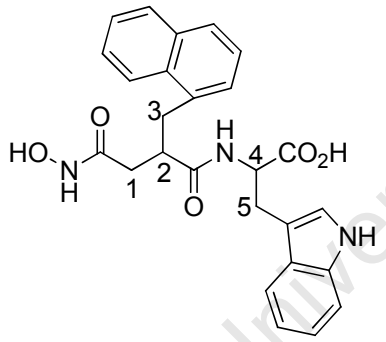
Yield: 0.022 g, 70 %;^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9(d, *J* = 9 Hz, 6H, H-5/6), 1.4 (m, 1H, H-3a), 1.8 (m, 1H, H-4), 1.9 (t, *J* = 7 Hz, 1H, H-3b), 2.2(m, 1H, H-1a), 2.5 (m, 1H, H-1b), 3.1 (m, 2H, H-2/8a), 3.3 (m, 1H, H-8b), 4.8 (t, *J* = 7 Hz, 1H, H-7), 7.1 (m, 3H, Indolyl-H), 7.3 (s, 1H, Indolyl-H), 7.6 (s, 1H, Indolyl-H). $C_{19}H_{25}N_3O_5$ MS m/z [M+1]

calculated 375.4. Observed [M+] 374.1.



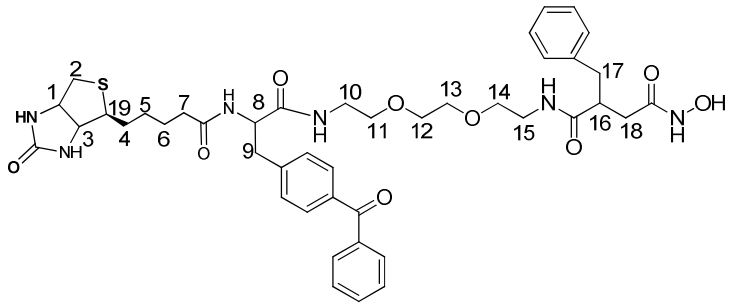
(±)2-(2-Benzyl-4-(hydroxyamino)-4-oxobutanamido)-3-(1H-indol-3-yl)propanoic acid 4.7c

^{1}H NMR (400 MHz, $CDCl_3$):δ_H 1.2 (d, *J* = 7 Hz, 3H, H-3), 2.5 (d, *J*= 8 Hz, 2H, H-1), 2.9 (m, 1H, H-2), 3.3 (d, *J* = 7 Hz, 1H, H-5) 4.5(t, *J* = 6 Hz, 1H, H-4), 7.1 (m, 3H, Indolyl-H), 7.3 (s, 1H, Indolyl-H), 7.6(s, 1H, Indolyl-H). $C_{22}H_{23}N_3O_5$ MS m/z [M+1] calculated 409.44. Observed [M+] 408.1.



(±)2-(4-(Hydroxyamino)-2-(naphthalen-1-ylmethyl)-4-oxobutanamido)-3-(1H-indol-3-yl)propanoic acid 4.7d

^{1}H NMR (400 MHz, $CDCl_3$): δ_H 1.3(d, *J* = 9 Hz, 3H, H-3), 2.3 (d, *J* = 8 Hz, 2H, H-1), 3.0 (m, 1H, H-2), 3.3 (d, *J* = 7 Hz, 1H, H-5) 4.6(t, *J* = 6 Hz, 1H, H-4), 7.1 (m, 3H, Indolyl-H), 7.3(m, 5H, Naphth-H/Indolyl-H), 7.6 (s, 1H, Indolyl-H). $C_{26}H_{25}N_3O_5$ MS m/z [M+1] calculated 459.49. Observed [M+] 458.1.



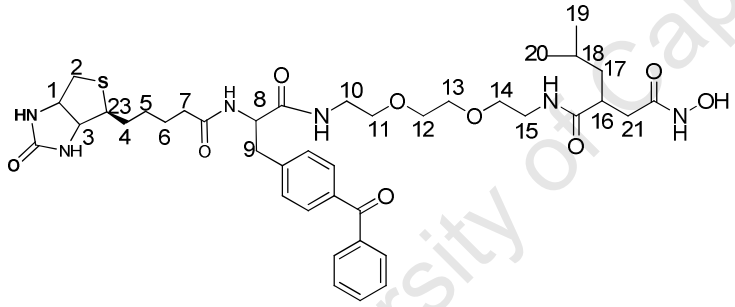
N1-((11S)-11-(4-Benzoylbenzyl)-10,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6-dioxa-9,12-diazaheptadecyl)-2-benzyl-N4-hydroxysuccinamide 3.27c-5

2-benzyl-4-((benzyloxy)amino)-4-oxobutanoic acid (7.51 mg, 0.024 mmol) was dissolved in DMF (2 mL) and the mixture cooled to 0 ^{O}C. To the solution were addedEDC (4.60 mg, 0.024 mmol),N-((S)-1-((2-(2-(2-aminoethoxy)ethoxy)ethyl)amino)-3-(4-benzoylphenyl)-1-oxopropan-2-yl)-5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (0.015 g, 0.024 mmol), DIEA (4.19 µl, 0.024 mmol), and HOBT (3.67 mg, 0.024 mmol). The reaction mixture was stirred overnight. Water (20 mL) was added to the mixture and then it was separated with EtOAc (10 mL x 3). Brine (5mL) was added to break up the emulsion that formed. The organic layer solution was washed with HCl (0.1 *M*, 5 mL x 1), $NaHCO_3$ (aq) (5 %, 5 mL x1) and brine (10 mL x 1), dried over $MgSO_4$, filtered, and concentrated. The residue was dissolved in MeOH (5 mL) and Pd-C (0.026 g, 0.248 mmol) added. The reaction mixture was stirred under H_2 overnight. The mixture was filtered through celite and the solvent removed under reduced pressure leaving the product N1-((11S)-11-(4-benzoylbenzyl)-10,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6-dioxa-9,12-diazaheptadecyl)-2-benzyl-N4-hydroxysuccinamide (0.014 g, 0.017 mmol, 80 % yield) as a yellow solid.

^{1}H NMR (400 MHz, $CDCl_3$): δ_H1.4 (m, 2H, H-5), 1.7 (m, 1H, H-4a), 1.8 (m, 2H, H-6), 2.3 (m, 1H, H-4b), 2.4 (m, 3H, H-

9/18a), 2.7 (m, 3H, H-15a/17a/18b), 2.99 – 3.5 (m, 8H, H-2a/9/10a/15a/16a/17b/19), 3.56 (m, 2H, H-13), 3.7 (m, 2H, H-11), 3.74 (m, 2H, H-12), 3.85 – 4.00 (m, 4H, H-2b/10b/14), 4.6 (dd, J = 7.7, 1.1 Hz, 1H, H-3), 4.9 (m, 1H, H-1), 5.2 (t, J = 6.0 Hz, 1H, H-8), 6.9 – 7.9 (m, Phe-H). ^{13}C NMR (75.5 MHz, $CDCl_3$): δ_c 24.4, 25.3, 28.3, 34.5, 36.1, 37.1(2C), 41.1(3), 46.3, 55.6, 58.7, 62.4, 63.7, 70.1(4C) 126.0, 127.4(2C), 128.1(6C), 130.2(4C), 132.4, 135.6, 138.0, 138.4, 140.3, 164.7, 171.7, 173.9, 174.6, 175.1, 194.3; $C_{43}H_{54}N_6O_9$ MS m/z [M+1] calculated 830.99. Observed [M+] 829.8.

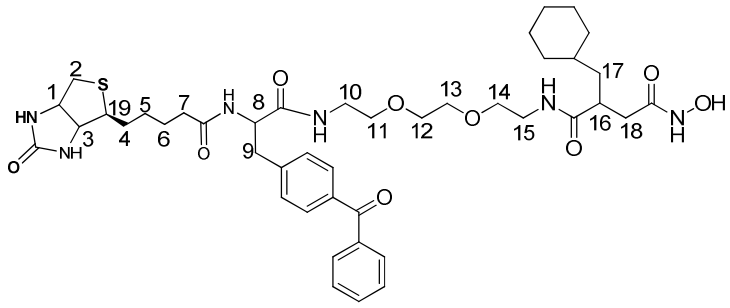
This was the general method employed for the synthesis of **3.27c-2**, **3.27c -3**, **3.27c -4** and **3.27c** below.



N1-(11-(4-Benzoylbenzyl)-10,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6-dioxa-9,12-diazaheptadecyl)-N4-hydroxy-2-isobutylsuccinamide 3.27c-3

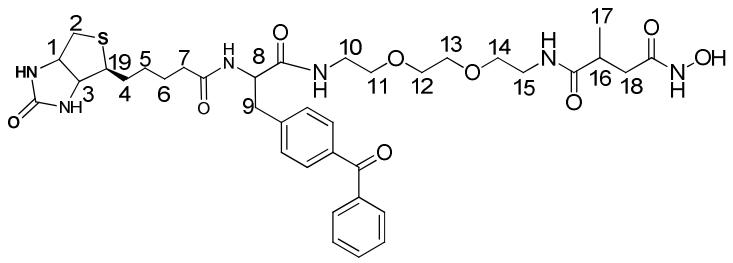
Yield: 0.014 g, 80 %; Yellow glassy material; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H 0.9(d, J = 9 Hz, 6H, H-19/20), 1.0 (m, 1H, H-17a) 1.4 (m, 3H, 5/18), 1.6 (m, 1H, H-17b),1.8 (m, 1H, H-4a), 1.9 (m, 2H, H-4b/6), 2.3 (m, 2H, H-7), 2.7 (m, 2H, H-21), 2.8 (m, 1H, H-16), 3.0 (m, 2H, H-9a/23), 3.2 (m, 1H, H-2a), 3.5(m, 10H, H-2b/9b/10/12/13/15), 3.9 (m, 4H, H-11/14), 4.3(m, 1H, H-3), 4.7(m, 1H, H-1), 5.2 (m, 1H, H-8), 7.0-8.0 (m, 9H, Phe-H);^{13}C NMR (75.5 MHz, $CDCl_3$): 22.8(2C),

24.2(2C), 25.2, 28.3, 36.1, 37.0, 37.4, 40.2, 41.1(4C), 55.6, 58.7, 62.4, 63.7, 70.1(4C), 127.4(2C), 128.4(2C), 130.2(4C), 132.4, 135.6, 138.4, 140.3, 164.7, 169.9, 171.7(2C), 175.0, 194.3; $C_{40}H_{56}N_6O_9S$ MS m/z [M+1] calculated 796.9. Observed [M+] 795.5.



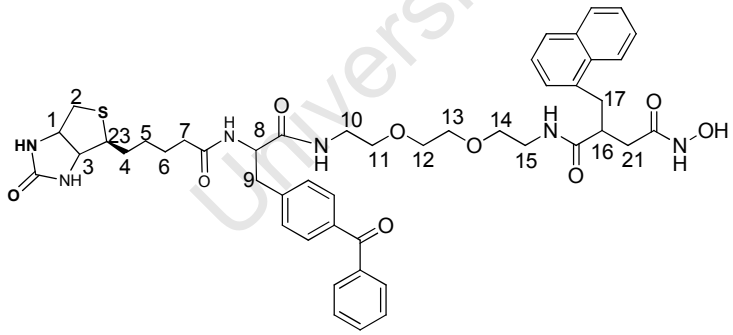
N1-((11S)-11-(4-Benzoylbenzyl)-10,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6-dioxa-9,12-diazaheptadecyl)-2-(cyclohexylmethyl)-N4-hydroxysuccinamide 3.27c-4

Yield:0.014 g, 80 %; Yellow glassy solid from MeOH, ^{1}H NMR (400 MHz, $CDCl_3$): 1.4-1.8 (m, 6H, H-4/5/Cyclohexyl-H), 2.0 (m, 2H, H-6), 2.4 (m, 4H, H-7/17), 2.8 (m, 2H, H-16/18a), 2.9 (m, 3H, H-9a/18b/19), 3.2 (m, 4H, H-10a/2a/11), 3.5 (m, 9H, H-2b/9b/10b/12/13/15), 3.9 (m, 2H, H-14), 4.3 (m, 1H, H-3), 4.7 (m, 1H, H-1), 5.0 (m, 1H, H-8), 7.0 - 7.7 (m, 9H, PheH); ^{13}C NMR (75.5 MHz, $CDCl_3$): 24.4, 25.2, 25.8(3C), 28.3, 32.3, 33.2(2C), 35.5, 36.1, 37.0, 37.4, 41.1, 41.3(3C), 55.6, 58.7, 62.4, 63.7, 70.1(4C), 127.4(2C), 128.4(2C), 130.2 (4C), 132.4, 135.6, 138.4, 140.3, 164.7, 169.9, 171.7, 173.9, 174.5, 194.3; $C_{43}H_{60}N_6O_9S$ MS m/z [M+1] calculated 837.1. Observed [M+] 836.5.



N1-((11S)-11-(4-Benzoylbenzyl)-10,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6-dioxa-9,12-diazaheptadecyl)-N4-hydroxy-2-methylsuccinamide3.27c-2

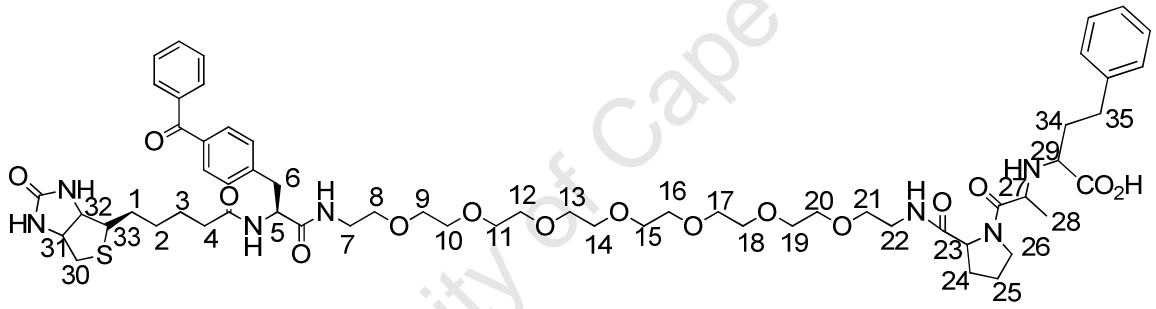
Yield: 0.024 g, 90%; Yellow oil; ^{1}H NMR (400 MHz, $CDCl_3$):1.4(m, 5H, H-5/17), 1.8 (m, 1H, H-4a), 2.0 (m, 2H, H-6), 2.1 (m, 1H, H-4), 2.7 (m, 3H, H-16/18), 3.1-3.7 (m, 17H, H-2/9/10/11/12/13/14/15/19), 4.4 (dd, *J* = 9, 5 Hz, 1H, H-3), 4.8 (m, 1H, H-1), 5.4 (m, 1H, H-8), 7.0 – 7.8 (m, 9H, Phe-H); ^{13}C NMR (75.5 MHz, $CDCl_3$): 16.5, 23.4, 25.2, 27.3, 35.1, 36.4, 37.9, 39.1, 41.1(2C), 42.6, 56.6, 59.7, 62.4, 63.7, 71.1(4C), 127.4(4C), 130.2(4C), 133.4, 136.6, 139.4, 141.3, 165.7, 169.9, 170.7, 172.9, 173.0, 191.3; $C_{37}H_{50}N_6O_9S$ MS m/z [M+1] calculated 754.34. Observed [M+1] 754.1.



N1-((11S)-11-(4-Benzoylbenzyl)-10,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6-dioxa-9,12-diazaheptadecyl)-N4-hydroxy-2-(naphthalen-1-ylmethyl)succinamide 3.27c-6

Yield:39 mg, 66; Yellow oil; ^{1}H NMR (400 MHz, $CDCl_3$): δ_H1.4 (m, 2H, H-5), 1.8 (m, 1H, H-4a), 1.9 (m, 3H, H-4b/6), 2.3

(t, *J* = 8 Hz, 2H, H-7), 2.8 (m, 2H, H-21a/17a), 2.9 (m, 3H, 9a/21b/23), 3.3 (m, 2H, 2a/15a), 3.3 – 3.7 (m, 13H, H-2a/9a/10/12/13/14/15a/16/17b), 3.8 (t, *J* = 7 Hz, 2H, H-11), 4.3 (dd, *J* = 9, 6 Hz, 1H, H-3), 4.7 (m, 1H, H-1), 5.21 (t, *J* = 7 Hz, 1H, H-8), 6.9-7.9 (m, Phe-H); ^{13}C NMR (75.5 MHz, CD_3OD): 24.4 , 25.2, 28.3, 34.5, 35.4, 36.1, 37.6, 41.1, 41.3, 41.6, 46.7, 55.6, 58.7, 62.4, 63.7, 70.1(2C), 70.2(2C), 120.6, 125.6, 125.8, 126.1, 126.5, 126.9, 127.4(2C), 128.4, 128.5, 128.6, 130.2(3C),, 131.3, 131.7, 132.4, 132.6, 135.3, 135.6, 138.4, 140.3, 164.7, 171.7, 173.9, 174.6, 175.2, 192.3; $C_{47}H_{56}N_6O_9S$ MS m/z [M+1] calculated 881.05 Observed [M+] 880.1.

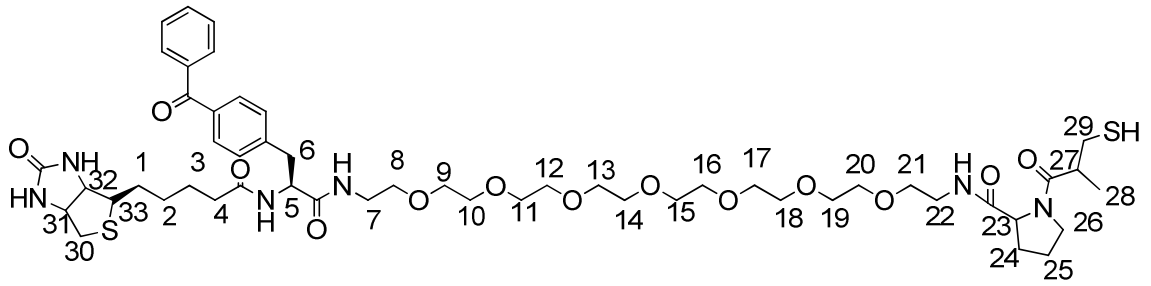


(2S,4S)-5-(2-(((26S)-26-(4-Benzoylbenzyl)-25,28-dioxo-32-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6,9,12,15,18,21-heptaoxa-24,27-diazadotriacontyl)carbamoyl)pyrrolidin-1-yl)-4-methyl-5-oxo-2-phenethylpentanoic acid 3.32a

1-((2S,4S)-4-(ethoxycarbonyl)-2-methyl-6-phenylhexanoyl)pyrrolidine-2-carboxylic acid (0.036 g, 0.095 mmol) was dissolved in DMF (2 mL) and the mixture cooled to 0 ^{O}C. To the solution were added EDC (0.018 g, 0.095 mmol), N-((S)-1-amino-27-(4-benzoylphenyl)-25-oxo-3,6,9,12,15,18,21-heptaoxa-24-azaheptacosan-26-yl)-5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (0.08 g, 0.095 mmol), DIEA (0.017 ml, 0.095 mmol), and HOBT (0.014 g, 0.095 mmol). The reaction mixture was stirred

overnight.Water (20 mL) was added to the mixture and then it was separated with EtOAc (20 mL x 3). Brine (20mL) was added to break up the emulsion that formed. The organic layer solution was washed with HCl (0.1 *M*, 40 mL x 1), $NaHCO_3$ (aq) (5 %, 50 mL x1) and brine (40 mL x 1), dried over $MgSO_4$, filtered, and concentrated. The residue was dissolved in MeOH (5 mL) and to this was added NaOH(aq) (4 M, 5 mL) and stirred for 3 h. Most of the MeOH was removed in vacuo and the remaining acqueous solution was separated with diethyl ether (20 mL x 1) The organic layer was discarded and the mixture acidified to pH 2 with 2 *N* HCl. The mixture was then separated with EtOAc (20 mL x 3) and the combined organics washed with H_2O (20 mL x 1), brine (20 mL x 1), dried over $MgSO_4$ and the solvent removed in vacuo. The title compound (2S,4S)-5-(2-(((26S)-26-(4-benzoylbenzyl)-25,28-dioxo-32-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6,9,12,15,18,21-heptaoxa-24,27-diazadotriacontyl)carbamoyl)pyrrolidin-1-yl)-4-methyl-5-oxo-2-phenethylpentanoic acid (0.067 g, 0.057 mmol, 60 % yield) was a yellow oil;. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.3 (d, *J* = 7 Hz, 3H), 1.5 (m, 2H,), 1.6 (m, 2H), 1.7 (m, 2H), 1.9 (m, 2H), 2.1 (m, 2H), 2.3 (m, 1H), 2.4 (m, 1H), 2.6 (t, *J* = 7, 2H), 2.7 (t, *J* = 7 Hz, 2H), 2.9 (t, *J* = 7.1 Hz, 1H), 3.16 (m, 2H), 3.2 (m, 2H), 3.7 (m, 2H), 3.4 (t, *J* = 3.7 Hz, 1H), 3.5 (m, 1H), 3.7 (m, 21H), 3.8 (m, 10H), 3.9 (m, 1H), 4.8 (m, 1H), 5.1 (t, *J* = 4.1 Hz, 1H), 5.2 (m, 1H), 5.4 (m,1H), 5.7 (d, *J* = 6 Hz, 2H), 6.3 (s, 1H), 6.9 (m, 2H), 7.3 (m, 5H), 7.5 (m, 2H), 7.6 (m, 1H), 7.7 (m, 2H), 7.9 (m, 4H), 8.8 (s, 1H). ^{13}C NMR (75.5 MHz, $CDCl_3$): 23.7, 24.1, 24.4, 25.2, 28.3, 29.5, 30.3, 32.1, 36.1, 37.4, 41.1, 41.3(2C), 49.5, 55.6, 57.2, 58.7, 62.4, 63.6(2C), 68.2, 70.1(14C) 126.0, 127.4(2C), 128.1(4),

128.8(2C), 130.2(4C), 132.4, 135.6, 138.4, 140.3, 142.0, 164.7, 171.0, 171.7, 173.3, 173.9, 174.7, 192.3; $C_{60}H_{85}N_7O_{15}S$ MS m/z [M+1] calculated 1174.3. Observed [M+] 1174.1.



N-((26S)-26-(4-Benzoylbenzyl)-25,28-dioxo-32-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6,9,12,15,18,21-heptaoxa-24,27-diazadotriacontyl)-1-((R)-3-mercapto-2-methylpropanoyl)pyrrolidine-2-carboxamide

3.32b

1-((R)-3-(acetylthio)-2-methylpropanoyl)pyrrolidine-2-carboxylic acid (0.025 g, 0.095 mmol) was dissolved in DMF (2 mL) and the mixture cooled to 0 ^{o}C. To the solution were added EDC (0.018 g, 0.095 mmol), N-((S)-1-amino-27-(4-benzoylphenyl)-25-oxo-3,6,9,12,15,18,21-heptaoxa-24-azaheptacosan-26-yl)-5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (0.08 g, 0.095 mmol), DIEA (0.017 ml, 0.095 mmol), and HOBt (0.014 g, 0.095 mmol). The reaction mixture was stirred overnight. Water (20 mL) was added to the mixture and then it was separated with EtOAc (20 mL x 3). Brine (20mL) was added to break up the emulsion that formed. The organic layer solution was washed with HCl (0.1 *M*, 40 mL x 1), $NaHCO_3$ (aq) (5 %, 50 mL x1) and brine (40 mL x 1), dried over $MgSO_4$, filtered, and concentrated. The residue was dissolved in MeOH (5 mL) and to this was added NaOH (aq) (4 M, 5 mL) and stirred for 3 h. Most of the MeOH was removed in vacuo and the remaining

acqueous solution was separated with diethyl ether (20 mL x 1) The organic layer was discarded and the mixture acidified to pH 2 with 2 *N* HCl. The mixture was then separated with EtOAc (20 mL x 3) and the combined organics washed with H_2O (20 mL x 1), brine (20 mL x 1), dried over $MgSO_4$ and the solvent removed in vacuo. The title compound N-((26S)-26-(4-benzoylbenzyl)-25,28-dioxo-32-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6,9,12,15,18,21-heptaoxa-24,27-diazadotriacontyl)-1-((R)-3-mercapto-2-methylpropanoyl)pyrrolidine-2-carboxamide (0.049 g, 0.047 mmol, 50 % yield) was a yellow oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 1.1 (m, 2H), 1.3 (m, 5H), 1.4 (s, 1H), 1.7 (m, 1H), 2.0 (m, 4H), 2.1 (m, 1H), 2.2 (t, *J* = 8 Hz, 2H), 2.9 (m, 2H), 3.1 (m, 2H), 3.2 (dd, *J* = 13, 5 Hz, 1H), 3.4 (m, 4H), 3.6 (m, 26H), 3.8 (m, 6H), 4.1 (m, 1H), 4.4 (m, 1H), 4.8 (m, 2H), 5.4 (t, *J* = 8 Hz, 1H), 5.6 (s, 1H), 5.7 (s, 1H), 5.71 (s, 1H), 6.5 (s, 1H), 6.9 (m, 2H), 7.5 (m, 2H), 7.6 (m, 1H), 7.7 (m, 2H), 7.8 (m, 2H), 8.9 (s, 1H). ^{13}C NMR (75.5 MHz, $CDCl_3$):16.4, 24.1, 24.4, 24.8, 25.2, 28.3, 29.5, 36.1, 37.4, 41.1(4C),49.8, 55.6, 58.7, 62.4, 63.1, 68.5, 70.1(5C), 127.4, 128.4(2C), 130.2(2), 130.3(2C), 132.4, 135.6, 138.4, 140.3, 164.7, 171.0, 171.7, 173.9, 178.8, 194.3;$C_{51}H_{76}N_6O_{13}S_2$ MS m/z [M+1] calculated 1044.31. Observed [M+] 1044.1.

References

[1] Desai, U. V.; Pore, D. M.; Mane, R. B.; Solabannavar, S. B.; Wadgaonkar, P.P.,*Synthetic Communications* **2004**, 34(1), 25-32.

[2]Nandkishor S. Chindarkar and Andreas H. Franz ARKIVOC **2008** (xv) 21-33

[3] Lee, DW.; Ha, H-J.; Lee, WK. *Synthetic Communications* **2007**;37: 737-742