Development of N-domain Selective Angiotensin-I Converting Enzyme (ACE) Inhibitors Using Computer Aided Drug Discovery (CADD)

A thesis submitted to the

University of Cape Town

in fulfilment of the requirements of the degree of

Doctor of Philosophy

By

Stephen Fienberg

Supervisors: Professor Kelly Chibale and Professor Edward Sturrock

University of Cape Town

Department of Chemistry

University of Cape Town

South Africa

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Acknowledgements

I wish to thank my supervisors Professor Kelly Chibale and Professor Edward Sturrock for making this PhD possible. Kelly Chibale has supervised me since my honours year in 2011 and inspired my interest in the field of medicinal chemistry. I must thank him for his belief in me and all his material support which has made this project possible. I would also like to thank Edward Sturrock for his encouragement and support in the supervision of this project. His gentle touch and insight played a crucial role in familiarising me with the fields of biochemistry and structural biology.

The first colleague I would like to thank is Dr. Grace Mugumbate. She mentored me for several years from my very first introduction to computer aided drug discovery until the beginning of this project. Her hard work and patience was invaluable to my introduction to this field.

Next I would like to thank Dr. Alexander Alex, a CADD expert who provided me with his crucial expertise while introducing me to the field.

I would next like to thank my computational colleagues Raban Masuka, Mohammed Albsheer, Razia Moorad and Elumalai Pavadai. They were always ready to lend a hand to and provide input when I was stuck.

I would like to thank the Novartis Institute for Biomedical Research (NIBR) for sponsoring my 4 month trip to Cambridge Ma. in 2013. There I was taught many of new CADD techniques and my understanding of the underlying principles would not be complete without the visit. Of the NIBR staff I would like to mention, Jeremy Beck, Rebecca Swett, Julian Levell, Daniel Mackay, Jose Duca and Camillo Vela for giving up their valuable time to share their expertise.

I would like to thank my colleagues in the synthetic lab who were also extremely patient while showing me the ropes while familiarising in the synthetic lab. These colleagues include Nicholas Njuguna, Malkeet Kumar, Tanya Paquet, Gurminder Kaur, Kawaljit Singh, Paul Njaria and Ferdinand Ndubi.

I would also like to thank the Sturrock group for their assistance with the inhibition assays. In particular I would like to thank Dr. Sylva Schwager for her amazing management of the lab and my colleagues Lizelle Lubbe, Vinasha Ramsamy, Kate Larmuth, Palesa Seele and Elaine Ferreira who were never shy to lend a hand.

I would especially like to thank Dr. Lauren Arendse for the close support she offered in both computational and biochemical aspects to this project as well as the special effort she made to proof-read this thesis.

Lastly, I would like to thank my friends and family who have supported me every step of the way. I would especially like to thank my friends John Woodland, Raissa Philibert and Timothy Povall who have stayed loyal friends through the duration of my decade long university career. The support they have provided has been invaluable.

I would also like to thank my family whose constant support and belief in me has guided me every step of the way and given me the freedom to pursue my doctoral studies

Abstract

Angiotensin-I (Ang-I) converting enzyme (ACE) is a zinc metalloprotease that plays a vital role in the Renin Angiotensin Aldosterone System (RAAS) and is a key antihypertensive drug target. In addition to Ang-I, ACE cleaves many other physiological substrates, thus extending its function beyond the regulation of blood pressure. Somatic ACE (sACE) consists of two structurally homologous yet distinct catalytic sites termed the N- and C-domains. The two catalytic domains of ACE have distinct substrate affinities and play different regulatory roles. The antifibrotic tetrapeptide Ac-SDKP is hydrolysed solely by the N-domain and thus is a potential target for interactions between the ligand and unique residues within the active site of the N- and C-domains, which need to be exploited to effect either N- or C-domain selectivity.

N-domain selective ACE inhibition has been demonstrated with peptides while crystallographic studies have shown that the N-domain to C-domain substitution of Arg381 with Glu403 within the S_2 subsite is integral to N-domain selective ACE inhibition. Three computer aided drug discovery (CADD) approaches were pursued to design N-domain selective drug-like ACE inhibitors (ACEi) with an acidic P_2 functional group that would confer N-domain selectivity via an interaction with Arg381 in the S_2 subsite.

Firstly, a fragment-based screening protocol was performed by running a set of chemical filters on 16000 drug fragment compounds (MW < 350), all of which contained a metal chelating group. 60 Ligands capable of binding to both the zinc metal and Arg381 in the S_2 subsite of the N-domain were tested for ACE inhibition against the two domains of ACE. Two of the fragments identified in this screen showed a modest ACE inhibition (IC_{50} +/- 200 µM), but no domain selectivity.

Secondly, a combinatorial library was created to explore the P_2 structure activity relationship (SAR) of a scaffold based on the core structure of the clinical ACEi, Enalaprilat. Over 400 variants were created to generate a combinatorial library. These compounds were docked against the two domains of ACE and a synthetic scheme was developed to synthesise compounds from this library. Using this scheme, one Enalaprilat analogue, **SF07** was synthesised as a mixture of diastereomers. **SF07** exhibited low micromolar N-domain inhibition with no C-domain inhibition observable below 100 µM.

For the third approach, 25 000 compounds containing biological data pertaining to ACE were extracted from the GVK BIO GOSTAR database. These compounds were filtered for drug-like properties and manually inspected for promising P_2 functionality. The N-domain selectivity of these compounds was then assessed via molecular docking against the two domains of ACE. This screen identified a series of diprolyl compounds with varied groups in the P_2 position. These compounds were subsequently

synthesised and tested *in vitro* for inhibition against both domains. The most N-domain selective compound from the series proved to be **SG6,** a diprolyl compound with an Asp group in the P_2 position. **SG6** displayed potent inhibition (K_i = 12 nM) and was 83-fold more selective towards the N-domain than the C-domain.

This study has demonstrated the N-domain selective inhibition of ACE by drug-like peptidomimetics. Two promising leads on drug-like N-domain selective ACE inhibitors, **SG6** and **SF07,** have been identified. These two compounds have the potential to pave the way for clinical N-domain selective ACEis and a novel treatment for cardiac and pulmonary fibrosis.

Abbreviations

ΔE_{MM}	Delta molecular mechanics energy
ΔG	Delta Gibbs Free Energy
ΔG_{bind}	Delta Energy of Ligand Binding
ΔG_{solv}	Delta Energy of Solvation
µM	micromolar
^1H	Proton
^1H NMR	Proton Nuclear Magnetic Resonance
Abz	4-(4-phenylazo)benzyloxybenzyl
ACE	Angiotensin Converting Enzyme
ACEi	Angiotensin Converting Enzyme inhibitor
AcOH	Acetyl Alcohol
AcSDKP	Acetyl-Serine-Aspartic Acid-Lysine-Proline
AMBER	Assisted Model Building with Energy Refinement
Ang-I	Angiotensin-I
Ang-II	Angiotensin-II
AT_1	Angiotensin type-1 receptor
BindingDB	Binding Database
BK	Bradykinin
Bn	Benzyl
Boc	*tert*-Butylcarbonyl
BPF	Bradykinin Potentiating Factor
bzt	blank zero time
C-18	Octodecyl carbon chain bonded silica
CADD	Computer Aided Drug Discovery
Cbz	Carboxybenzyl
CD_3OD	Deuterated Methanol
CFL	Chelator Fragment Database
CHARMM	Chemistry at Harvard Macromolecular Mechanics
ChEMBL	Chemical European MolecularBiology Laboratory
ChemScore	Chemical Score
ClogP	Calculated log partition coefficient
CPA	Carboxypeptidase A
CRO	Contract Research Organisation
d	doublet
D_2O	Deuterium Oxide
DCC	*N,N'*-dicyclohexylcarbodiimide
DCM	Dichloromethane
DMF	Dimethylformamide
Dnp	2,4-Dinitrophenol
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
Et_3N	Triethylamine
EtOAc	Ethyl acetate
Glide	Gride-based Ligand Docking with Energetics
Glide SP	Gride-based Ligand Docking with Energetics – Standard Precision
Glide XP	Gride-based Ligand Docking with Energetics – Extra Precision
Glp	pyroglutamate
GPCR	G-Protein Coupled Receptor
GUI	Graphical User Interface
Hbond	Hydrogen Bond contribution

HCl	Hydrogen Chloride
HHL	Hippuric acid Histidine Leucine
h-Phe	homophenylalanine
HOBt	*N*-hydroxy bezatriazole
HOSu	*N*-hydroxy succinimide
HPLC	High Performance Liquid Chromatography
HTS	High Throughput Screening
IC_{50}	half maximal inhibitory concentration
ICM	Internal Coordinate Modelling
IP	Intellectual Property
K_2CO_3	Potassium Carbonate
k_{cat}	Catalytic Rate Constant
KCN	Potassium Cyanide
Keto-ACE	5-S-5-benzamido-4-oxo-6-phenylhexanoyl-L-proline
KKS	Kalikrein Kinin System
K_m	Michealis-Menton constant
LB	Ligand Based
LC	Liquid Chromatography
LE	Ligand Efficiency
LH-RH	Leutenising Hormone Releasing Hormone
LiOH	Lithium Hydroxide
Lipo	Lipohillic contribution
Lis-W	Lisinopril-Tryptophan
LogP	Log Partition Coefficient
m	multiplet
MC	MonteCarlo
Mca	Mercaptan
MD	Molecular Dynamics
MeCN	Acetonitrile
MeOH	Methanol
MM	Molecular Mechanics
MM-GBSA	Molecular Mechanics – Generalised Born Surface Area
MMP	Matrix Metalloprotease
MM-PBSA	Molecular Mechanics – Poisson Boltzmann Surface Area
MOE-Dock	Molecular Operating Environment - Docking
$NaHCO_3$	Sodium Bicarbonate
NaOH	Sodium Hydroxide
NEP	Neutral Endopetidase
NH_4Cl	Ammonium Chloride
NIBR	Novartis Institute for Biomedical Research
nM	nanomolar
NMM	*N*-methylmorpholine
OPLS	Optimized Potentials for Liquid Simulations
PCR	Polymerase Chain Reaction
PDB	Protein Databank
Prep-HPLC	Preparative High Performance Liquid Chromatography
q	quartet
QSAR	Quantitative Structural Activity Relationship
RAAS	Renin Angiotensin Aldosterone System
RMSD	Root Mean Square Deviation
Rotb	Rotatable bonds

s	Singlet
sACE	somatic Angiotensin Converting Enzyme
SAR	Structure Activity Relationship
SB	Structure Based
sdf	structure data file
SHOP	Scaffold HOPping
SMILES	Simplified Molecular Input Line Entry System
S_N1	unimolecular nucleophilic substitution
S_N2	bimolecular nucleophilic substitution
t	triplet
T3P	Propylphosphonic anhydride
tACE	testis Angiotensin Converting Enzyme
THF	tetrahydrofuran
TLC	Thin Layer Chromatography
TMSCN	Trimethylsilyl cyanide
VS	Virtual Screening
VSAR	Virtual Structure Activity Relationship
VSGB	Variable Dielectric Surface Generalised Born
ZBG	Zinc Binding Group
ZFHL	Carboxybenzyl Phenyl Alanine Histidine Leucine
ZINC	ZINC is not commercial

Chapter 1 Contents

Chapter 1 – Literature Review

1.1 Introduction

The Angiotensin-I converting enzyme (ACE) is a zinc metalloprotease best known as an antihypertensive drug target.[1] ACE inhibitors (ACEi) have proven to be a popular treatment for chronic hypertension with at least 19 clinically approved drugs.

Since the elucidation of the structure of ACE, there has been renewed interest in the study of its inhibitors[2]. ACE was shown to contain two distinct homologous catalytic sites with a more diverse set of substrates than the principle vasoconstrictive Angiotensin-I substrate.[3] These two catalytic sites show different kinetic profiles towards a range of different substrates. While it has been shown that the inhibition of just one active site is sufficient for antihypertensive vasodilation,[4] the other active site is solely responsible for the hydrolysis of Ac-SDKP *in vivo*, a peptide responsible for collagen deposition and cardiac remodelling[5].

Previous studies suggest site-selective ACE inhibition could either reduce the side-effects of anti-hypertensive drugs or treat cardiac fibrosis, thus greatly expanding the original therapeutic scope of ACE inhibition.[5] Modern computer aided drug discovery (CADD) software and high resolution crystal structures lend themselves to a structure based approach for the optimisation of current ACEis into a site-specific inhibitor with either reduced anti-hypertensive effects or a novel anti-fibrotic drug.

Given its potential as a novel anti-fibrotic drug target, the full potential for site-selective ACE inhibition is yet to be explored. The extensive collection of ACEis and high resolution crystal structures provided a unique opportunity to derive site-specific ACEis from existing ACEis with no preference for either catalytic site. Modern CADD software is ideally suited to the task with the ability to accurately model and guide small chemical changes to these ACEis. The application of CADD software to the ACE target provides a unique opportunity to apply a structure based-approach to optimise the established set of ACEis into domain selective inhibitors.

1.2 Historical Aspects of ACE

1.2.1 The Discovery of ACE

ACE was first discovered by Skeggs et al[6] in 1956. The discovery began with the identification of two peptides responsible for elevated blood pressure.[7] These peptides were named hypertensin 1 and hypertensin 2 after the discovery of their vasoconstrictive properties[8]. Of the two peptides, only hypertensin 2 was shown to be capable of effecting a vasoconstrictive response on isolated aortic tissue strips. The peptides were then sequenced with hypertensin 2 observed to contain the same

sequence as hypertensin 1 minus the two C-terminal residues. Following this observation, Skeggs managed to isolate the enzyme responsible for this action from horse serum. This enzyme was shown to belong to the Zinc metalloprotease class of enzymes.[1] Today these peptides are known as angiotensin (Ang) I and II (Figure 1.1) while the enzyme is called angiotensin converting enzyme (ACE). This work followed on from studies performed by Goldblatt et al[9] a few years prior where the kidneys were linked to hypertension via the a renal enzyme named renin and an unknown substrate. The substrate of renin is now known to be angiotensinogen, a large peptide precursor to Angiotensin-I.

Ang I: Asp-Arg-Val-Tyr-Ile-His-Pro–Phe|His-Leu

Ang II: Asp-Arg-Val-Tyr-Ile-His-Pro–Phe

Figure 1.1: *The peptide sequences of Ang-I and Ang-II. The site of cleavage is indicated with a |.*

The discovery of Ang I and Ang II completed the picture of the blood pressure control mechanism first elucidated by Goldblatt. This pathway is known as the Renin Angiotensin Aldosterone System (RAAS). Lowered sodium levels and loss of blood volume triggers the release of angiotensinogen from the liver. The angiotensinogen is then cleaved in the kidneys by renin into Ang-I before returning to the bloodstream. While circulating, Ang-I is hydrolysed by ACE into Ang-II. From here Ang-II interacts with the G-protein coupled receptor (GPCR) angiotensin type 1 receptor (AT_1) to effect a cellular response leading to vasoconstriction, salt reabsorption and water retention. Figure 1.2 provides a simplistic overview of this pathway. Many decades of research have added several new axes and enzyme substrates to this pathway. Figure 1.2 summarises the best known axis of the RAAS.[10]

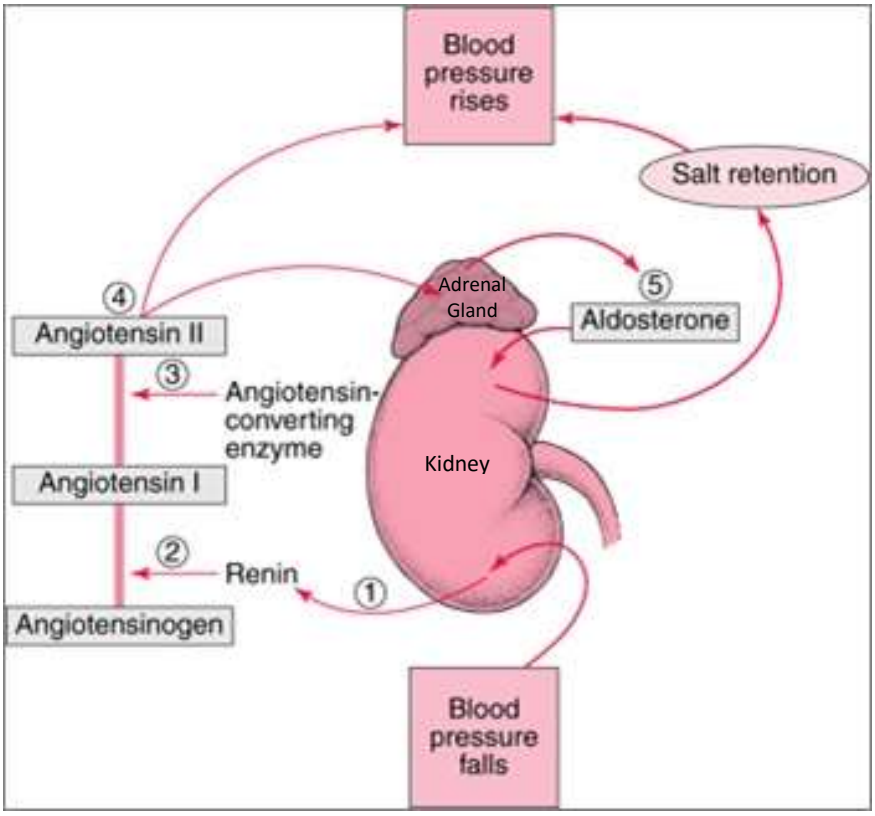


***Figure 1.2**: A Schematic outline of the major axis of the RAAS*

1.2.2 Bradykinin Potentiating Peptides

Bradykinin (BK) Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg, a well-known vasodilating nonapeptide is a crucial component of the kallikrein kinin system (KKS). BK was first discovered by Silva et al[11] in 1949 when studying the anticoagulating properties of the snake venom from the Brazilian pit viper *Bothrops jararaca. B. jararaca* venom was observed to be capable of lowering blood pressure by stimulating the release of a hypotensive nonapeptide which would later become known as BK. In reality, *B. jararaca* venom prevented the degradation of this native peptide as shown subsequently by Ferreira et al.[12] The BK preserving component of the venom was termed the Bradykinin potentiating factor (BPF). In 1968 Bakhle et al[13] demonstrated BPF to be capable of inhibiting the conversion of Ang-I to Ang-II. This was not just the first evidence of ACE inhibition, but also the first evidence of a link between ACE, BK and the KKS. Today the complete scope of the KKS is still not fully understood, but blood pressure regulation appears to be one of its primary functions. The KKS regulates blood pressure via the release of BK from kininogen allowing it to interact with the B_2 receptor, a GPCR which effects a vasodilatory cellular response in smooth endothelial muscle tissue.

To investigate the ACE inhibition mechanism of the BK potentiating factor, Ondetti et al separated the *B. jararaca* venom into its base components[14]. Using chromatographic techniques, six distinct peptides were isolated (

), all showing some degree of ACE inhibition. Following on, Collier et al[15] synthesised the similar artificial SQ 20881 nonapeptide, with a potent ACE IC_{50} of 0.56 μM, an order of magnitude stronger than the BPPs. Despite reasonable *in vivo* inhibition of ACE when administered parenterally, this large bulky nonapeptide has poor oral bioavailability due to its digestion by tryptic enzymes and a lack of permeability. A successful orally administered ACEi would therefore need to be small molecule resistant to tryptic enzyme degradation while mimicking some structural features of SQ20881.

***Table 1.1**: A list of the different peptides separated by Ondetti et al[14] and their associated ACE IC_{50} values.*

Sequence	IC_{50} (μM)
Glp-Trp-Pro-Arg-Pro-Thr-Pro-Gln-Ile-Pro-Pro	2
Glp-Trp-Pro-Arg-Pro-Gln-Ile-Pro-Pro	3
Glp-Asn-Trp-Pro-Arg-Pro-Gln-Ile-Pro-Pro	3
Glp-Asn-Trp-Pro-His-Pro-Gln-Ile-Pro-Pro	9
Glp-Ser-Trp-Pro-Gly-Pro-Asn-Ile-Pro-Pro	39
Glp-Gly-Gly-Trp-Pro-Arg-Pro-Gly-Pro-Glu-Ile-Pro-Pro	13
Glp-Trp-Pro-Arg-Pro-Gln-Ile-Pro-Pro-Gln-Ile-Pro-Pro (SQ20881)	0.56

To consolidate these findings, Dorer et al[16] investigated the ACE-catalysed hydrolysis of BK. It was found that there are two successive dipeptidase cleavages converting BK into $BK_{(1-7)}$ and then into $BK_{(1-5)}$. Unexpectedly, the optimal chloride concentration for ACE-mediated BK hydrolysis was shown to be a tenth of the optimal concentration for Ang-I hydrolysis. This anomaly would later be attributed to the dual domain nature of ACE. The vastly different sequences of Ang-I and BK as well as the three different products of dipeptide hydrolysis (His-Leu, Phe-Arg and Ser-Pro) suggested a broader substrate specificity of ACE than previously thought. This raised some doubts regarding its suitability as a therapeutic target. Such fears were later allayed during *in vivo* evaluation.

1.2.3 First Generation ACE Inhibitors

Captopril was the first orally administered ACE inhibitor to receive clinical approval. It was developed by Cushman and Ondetti in the mid 1970s[17] using an inspired piece of insight to guide structure activity relationship (SAR) studies in the absence of crystal structures. They started by examining the work of Shechter and Berger[18] where a variety of proteases were studied crystallographically. They categorised 3 classes of proteases and found them all to contain a central binding cleft with a catalytic site. Residue specific binding pockets were discovered to reside on either side of the central catalytic site. All subsites in the N-terminal direction along the peptide sequence from the catalytic site were named S_1, S_2 ... S_n while all the subsites in the C-terminal direction along the cleft were named S_1', S_2' … S_n'. Peptide residues termed P_n bind to the subsite S_n (Figure 1.3). With this formalism recently established, Cushman and Ondetti began to focus their attention on carboxypeptidase A (CPA), the only crystallised metalloprotease at the time.

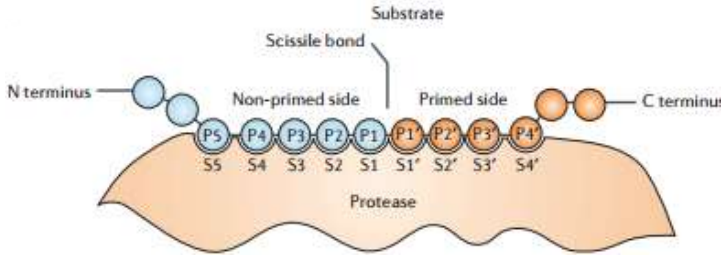


Figure 1.3 *The Schechter and Berger representation of a protease active site.* [18]

Byers and Wolfendonden[19] identified that the S_1' site of CPA had a preference for Phe residues and experimented with attaching a variety of carboxylic acid groups to Phe finding 2(*S*)-phenyl-succinic acid (Figure 1.4) to be a potent inhibitor. Cushman and Ondetti correctly assumed the carboxylic acid of 2(*S*)-phenyl-succinic acid to coordinate with the catalytic Zn atom thus blocking the enzyme binding site. The C-terminal carboxylic acid was also observed to be important in binding to CPA. Positively charged residues were therefore hypothesised to interact with this acid. The positioning of this positively charged residue helps define the end of the prime-side in the binding cleft, thus determining the number of C-terminal residues to be cleaved from the peptide substrate.

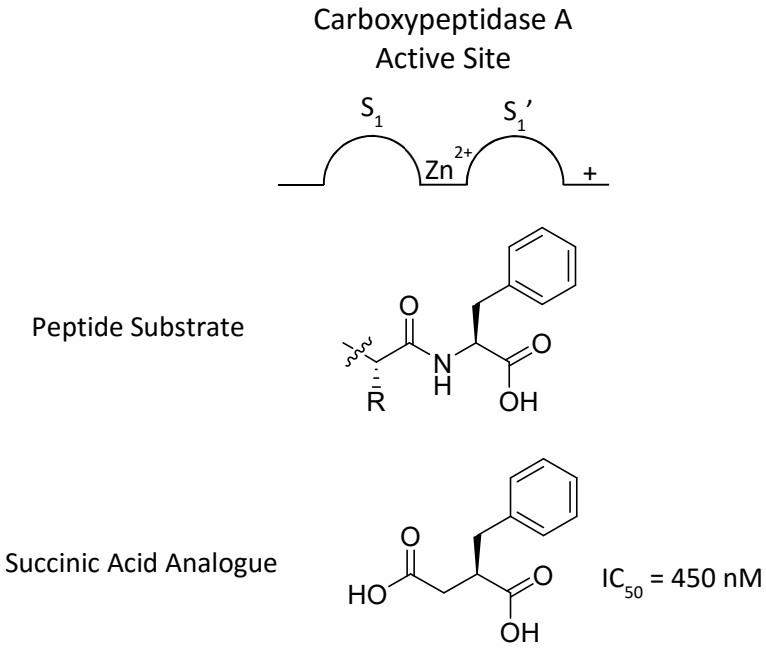


Figure 1.4: *Schematic of the binding site of CPA. The S1 and S1' subsites bind to both its native peptide substrate and the inhibitor 2(S)-phenyl-succinic acid*

Since ACE is a dipeptidase rather than a carboxypeptidase, two subsites were assumed to exist between the Zn atom and the positively charged carboxylic acid binding site. To fit this model, Cushman and Ondetti began coupling succinic acid to a variety of amino acids. The chelating carboxylic acid was then substituted with a thiol, a much stronger Zn chelator. This substitution increased ACE inhibition by over 1000-fold in many cases, thus confirming the succinyl carboxylic acid was indeed coordinating with the Zn. The strongest succinyl amino acid inhibitor was succinyl-Pro while a thiol

substitution of the succinyl acid yielded a 2000-fold increase in potency. This observation was consistent with the *B. jararaca* venom peptide sequences as they contain a C-terminal Pro. The next step in the SAR exploration was P_1' methyl substitution. Methyl groups were added to both succinyl and thiol Pro with a variety of different stereochemical configurations in both the 2 and 3 positions. Variations in chain length were also tested. The best inhibitor proved to be thiol-2-methyl-Pro. This small molecule mimics an Ala-Pro fragment binding in the S_1' and S_2' subsites and has a strong metal coordinating thiol. This compound became known as Captopril (Figure 1.5) and was a pioneering example of the peptidomimetic class of small molecule drugs.[17]

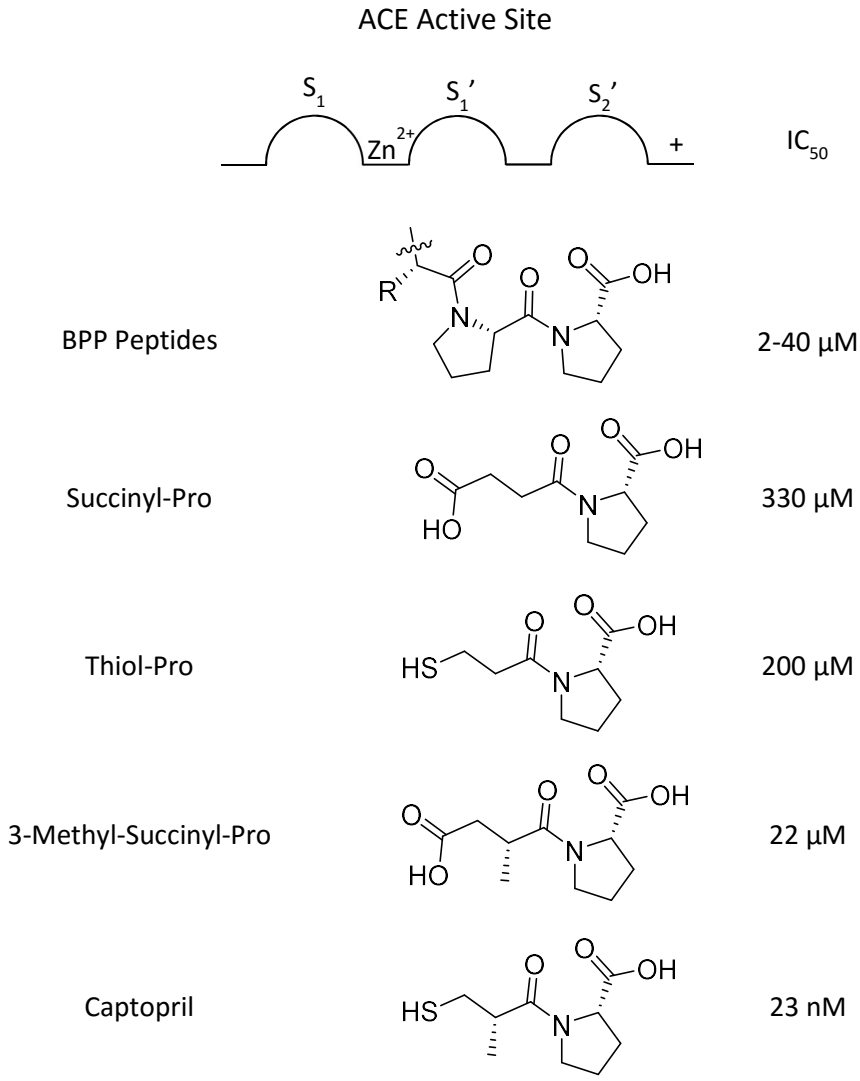


Figure 1.5*: The SAR evolution of Captopril from the C-terminal Pro-Pro to Captopril.*

Captopril was clinically approved as an effective antihypertensive agent in 1981. It was, however, plagued by side-effects such as skin-rash, angioedema and loss of taste. This prompted Patchett et al[20] to improve on Captopril and develop a new line of ACE inhibitors. It was suggested that most of the Captopril side-effects were caused by the thiol group. A weaker Zn coordinating carboxylic acid group

was consequently revisited. With a weaker Zn coordination, more interactions within the ACE binding pocket would need to be exploited. The two interactions targeted were a H-bond between the amide of the scissile peptide bond and an additional peptidomimetic side-chain moiety to interact with the S_1 subsite. After replacing the thiol moiety with a carboxylic acid, a secondary amino group was added to a position β to the carboxylic acid, mimicking the H-bond donor characteristic of the native amide. As expected, this compound displayed reduced ACE inhibition with an IC_{50} of 2.4 μM compared to the IC_{50} of 23 nM of Captopril.

In an attempt to improve inhibition, SAR in the P_1 position was explored. Examining the sequence of Ang-I and the BPPs reveals the S_1 subsite to be highly specific towards hydrophobic side-chain moieties such as Phe and Ile. Hydrophobic side-chain mimics were attached to the P_1 position with the best moiety proving to be a phenyl group attached to a two carbon alkyl chain. Since a new stereocentre was created during this substitution, the two different diastereomers originating from this substitution were separated and tested individually. The *S* stereoisomer at this centre proved to be more potent (IC_{50} = 1.2 nM) than the diastereomeric mixture (IC_{50} = 3.8 nM). In this molecule with three chiral centres, the *SSS* stereochemical configuration mimics the stereochemistry of a natural Phe-Ala-Pro peptide emphasising the importance of stereochemistry in peptidomimetics. Since previous studies had shown the S_1' subsite of ACE to bind to a wide variety of side-chains, Patchett et al also explored the P_1' SAR by switching methyl side-chains of Ala with a variety of both natural and synthetic side-chain moieties. The introduction of a Lys side-chain into this molecule gave rise to a potent inhibitor with an IC_{50} of 1.2 nM. Both these new ACEis achieved improved potencies over Captopril.

The Phe-Ala-Pro like compound suffered from poor oral bioavailability due to its net charge in solution of -1 which compromised cell permeability. This problem was rectified by esterifying the one carboxylic acid and delivering it in the prodrug form. These two compounds went on to pass clinical trials and are now known as Lisinopril and Enalaprilat. Figure 1.6 illustrates the SAR evolution of these compounds.

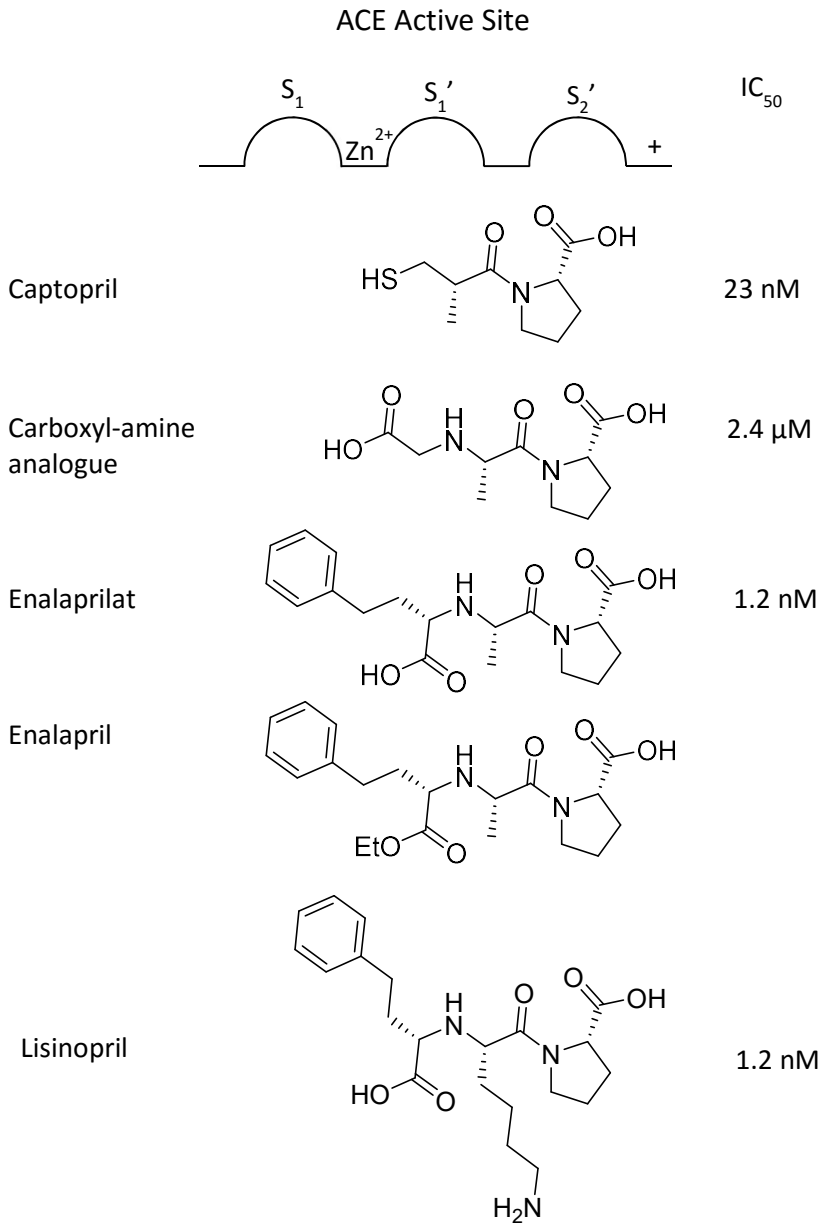


Figure 1.6: *The SAR evolution of Enalaprilat and Lisinopril. The thiol of Captopril was substituted for a carboxylic acid and a secondary amine was added to create an H-bond donor analogue. An h-Phe group was then introduced onto the carbon adjacent to the new carboxylic acid to create Enalaprilat which was esterified to make the Enalapril prodrug. Lisinopril came about by substituting the Ala mimicking methyl group with a Lys mimicking amino-butyl group.*

Enalaprilat and Lisinopril proved to be hugely successful drugs and laid out a useful guideline for designing small molecule metallo-protease inhibitors. The Zn coordinating carboxylic acid mimics the tetrahedral oxyanionic transition state of the hydrolysis reaction involved in the cleavage of a peptide bond (Figure 1.7). The C-terminal end of the drug then requires a carboxylic acid to interact with an electropositive pocket, which is important for ensuring peptides bind in the correct orientation. A good ACEi also needs good side-chain mimicking groups, natural or unnatural, while a secondary amine helps to improve the pharmacokinetic properties of the compound by raising its net charge to counter the negatively charged carboxylic acids while providing a H-bond donor. Unlike Lisinopril,

Enalapril is delivered in a prodrug form with one esterified acid to raise its net charge to 0 improving cell permeability.

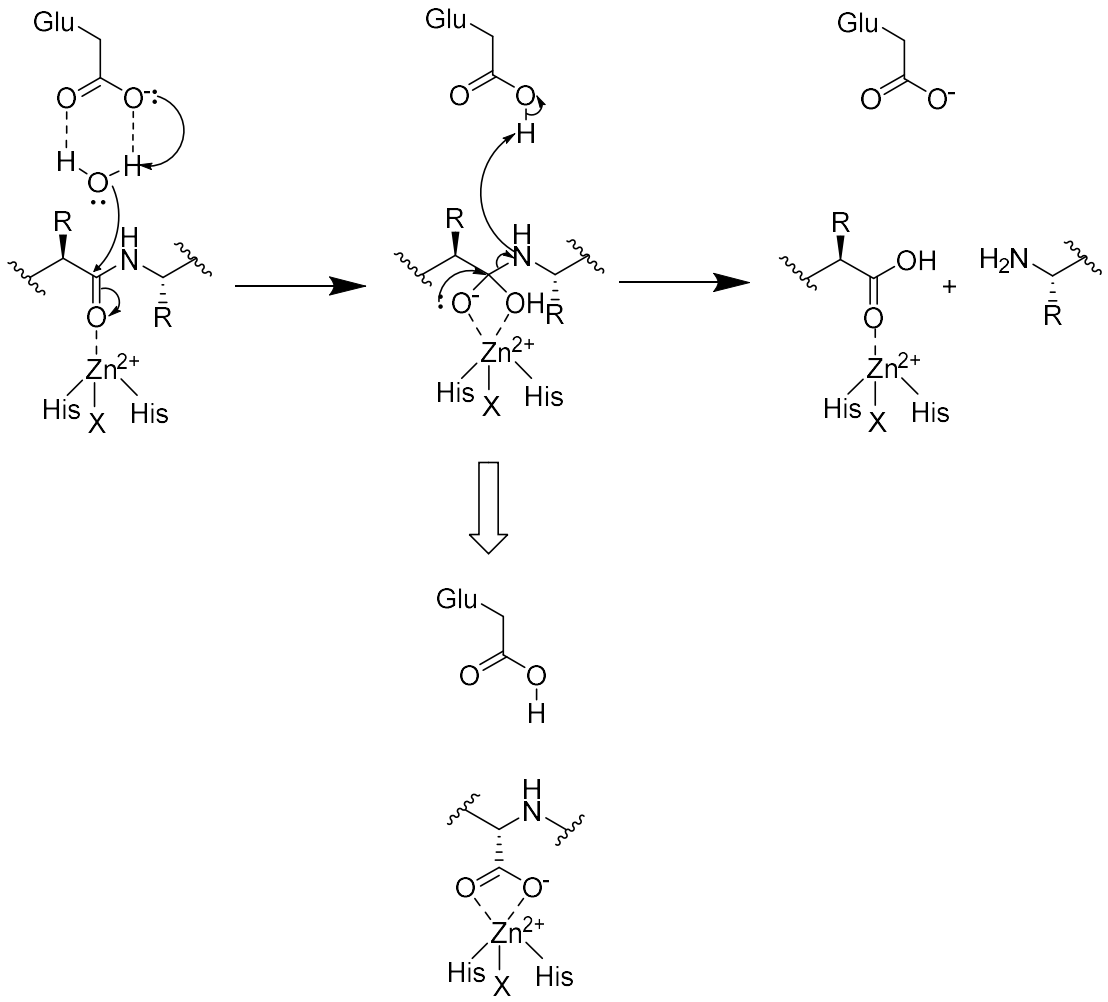


Figure 1.7: *Mechanism of metalloprotease catalysed peptide hydrolysis. A carboxylic ZBG mimics the tetrahedral intermediate of peptide hydrolysis.*

Since Lisonopril and Enalapril entered the market, at least 17 additional ACEis have been approved (Table 1.2). All of them follow the same design pricinciples with a core Zinc binding group (ZBG), a terminal P2' carboxylic acid, 2-3 side-chain mimicking groups and/or a secondary amine H-bond donor. Only Captopril, Zofenopril, Alacepril and Rentiapril utilise a thiol ZBG. Attempts to minimise thiol toxicity were made with Zofenopril and Alacepril by capping this group to make a thioester prodrug.

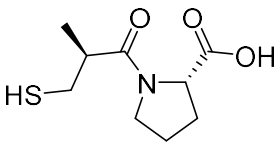
Almost all of the remaining ACEis have a carboxylic acid ZBG, a P_1' Ala/methyl and a P_1 Phenyl group. For reasons described earlier, substituting the thiol ZBG with a carboxylic acid neccesitates the introduction of a H-bond donor amine in a position adjacent to the acid. In a similar manner to

Enalaprilat, the introduction of both a second carboxylic acid and an amine to an ACEi in most cases gave the molecule a net charge of -1, making the compound too polar for optimal bioavailability. These drugs have therefore, had their bioavailabilty optimised with an ethyl ester capping a carboxylic acid for the oral formulation of the drug. Most of the carboxylic acid class of ACEis differ in the P_2' position. The original ACEis had a Pro in this position while the later examples often have much larger ring systems. These ring systems include a wide variety of hydrophobic cycloalkyl, aromatic, heteroaromatic and heterocyclic moieties.

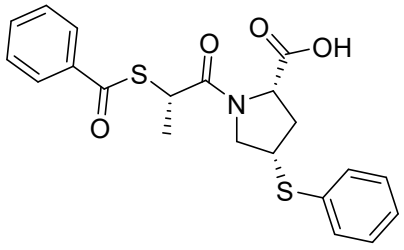
Fosinopril and Ceranopril are notable exceptions as they contain a phosphinic acid ZBG. Phosphinic acids are not popular in oral drugs as their high polarity often compromise permeability and bioavailibility. Fosinopril is therefore delivered with a large and bulky lipophillic phosphoester to improve its permeability. Ceranopril lacks this prodrug phosphoesterification as its P_1' Lys moiety is positively charged in solution countering the acidity of the phosphinic acid. These two drugs are analagous to the esterified Enalapril and the non-esterified Lisinopril. Since the market is saturated with potent ACEis, further research on this target might have seemed pointless. That was until the structure of the target was investigated.

***Table 1.2**:* The structures of all 19 clinically approved ACEis categorised according to their respective ZBGs.

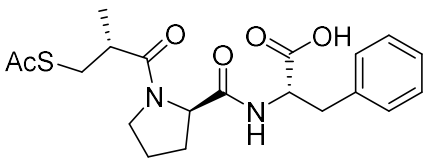
Thiolate ZBG



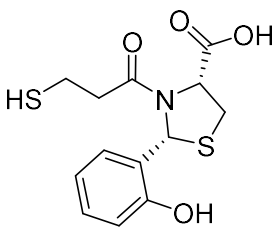
Captopril[17]



Zofenopril[21]

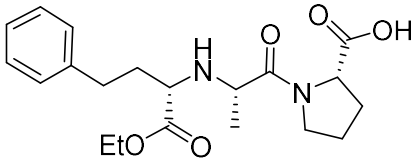


Alacepril[22]

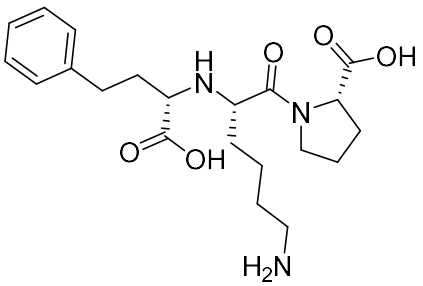


Rentiapril[23]

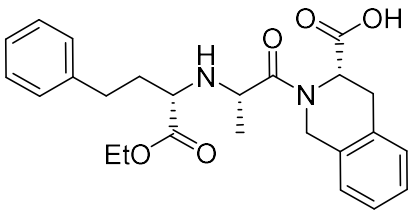
Carboxylic Acid ZBG



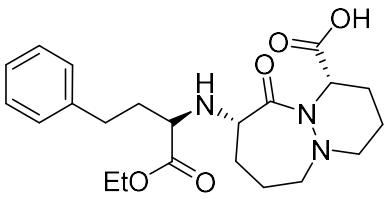
Enalaprilat[20]



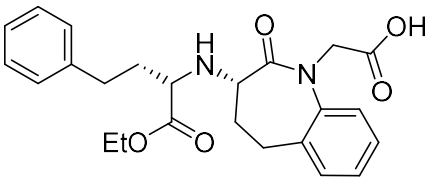
Lisinopril[20]



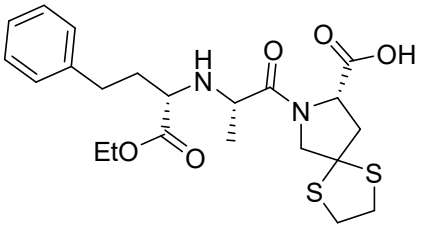
Quinapril[24]



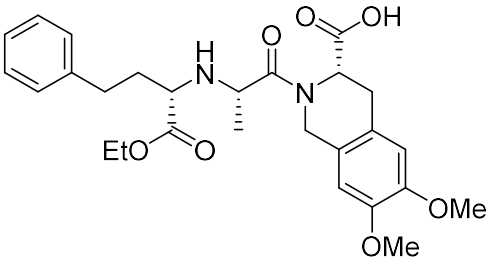
Cilazapril[25]



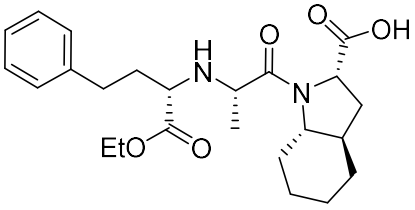
Benazepril[26]



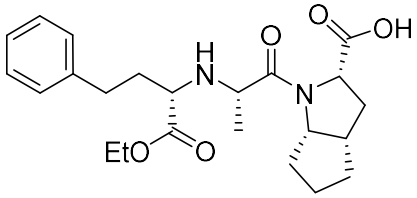
Spirapril[27]



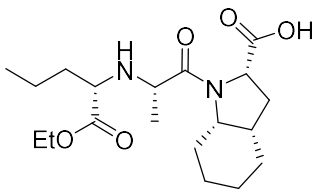
Moexipril[28]



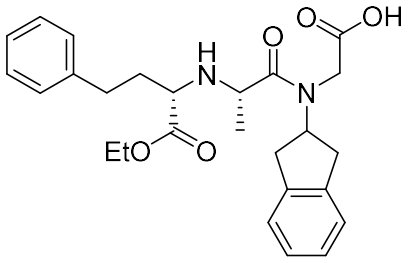
Trandolapril[29]



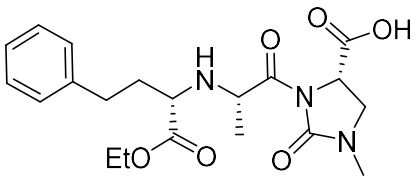
Ramipril[30]



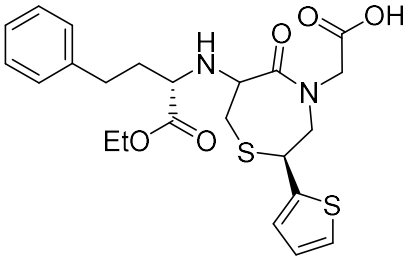
Perindopril[31]



Delapril[32]

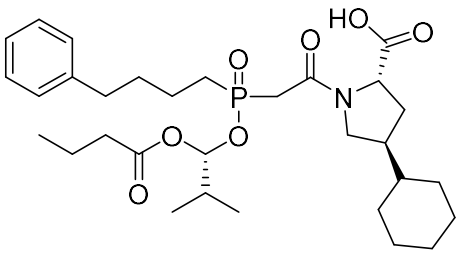


Imidapril[33]

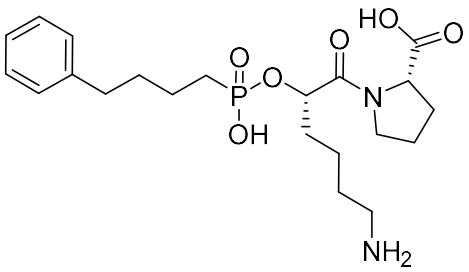


Temocapril[34]

Phosphinate ZBG



Fosinopril[35]



Ceranopril[36]

1.3 The Two Catalytic Domains of ACE

1.3.1 The Discovery of the Two Domains of ACE

The first detailed enzymatic characterisation of ACE was performed by Das and Soffer in 1975.[37] This study established ACE as a glycosylated protein with a MW of 129 kDa. Furthermore, kinetic studies revealed ACE activity to be chloride-dependent. A few years later El-Dorry et al[38] isolated what appeared to be a very similar enzyme from rabbit testis tissue using Lisinopril-Sepharose affinity chromatography. This enzyme was shown to be significantly smaller than ACE at 94 kDa, but showed a remarkably similar kinetic profile with respect to the cleavage of Hip-His-Leu and Ang-I substrates. In addition, both enzymes could be successfully immunoprecipitated using the same antibodies. mRNA comparisons of the two enzymes proved that this testicular enzyme was not a post translational

product of ACE but rather that the two proteins originate from different promoters within the same gene. The role of this ACE-like enzyme was not understood at the time since Ang-II had no known targets in the testicular tissue while its expression appeared to be hormonally controlled.

The 1980s saw massive advances in the field of genetics and molecular biology with the advent of polymerase chain reaction (PCR) and DNA sequencing technology. This new technology helped to fast-track the structural studies on ACE. Soubrier et al[2] sequenced somatic ACE for the first time. From this sequence, a C-terminal lipophilic region for the membrane anchor was identified. The sequence of ACE was then overlaid with that of thermolysin and neutral endopeptidase (NEP), two Zn metallopeptidases whose structures were known. Surprisingly, ACE was shown to possess not one but two catalytic sites. In a protein of 1277 residues, two homologous regions of 357 residues (termed N- and C-domains) were identified with an overall sequence homology of 67.7%.

Sequencing also solved the mystery of the similar ACE-like enzyme found in testicular tissue. This enzyme was shown to be an exact copy of the C-domain region from the full enzyme. ACE was therefore shown to be expressed in two different isoforms depending on its location within the body. The testicular isoform of this enzyme is designated testis ACE (tACE) while the isoform expressed throughout the rest of the body is designated somatic ACE (sACE) (Figure 1.8). This structural breakthrough regarding the domain structure of the protein paved the way for studies investigating the two active sites of the enzyme.

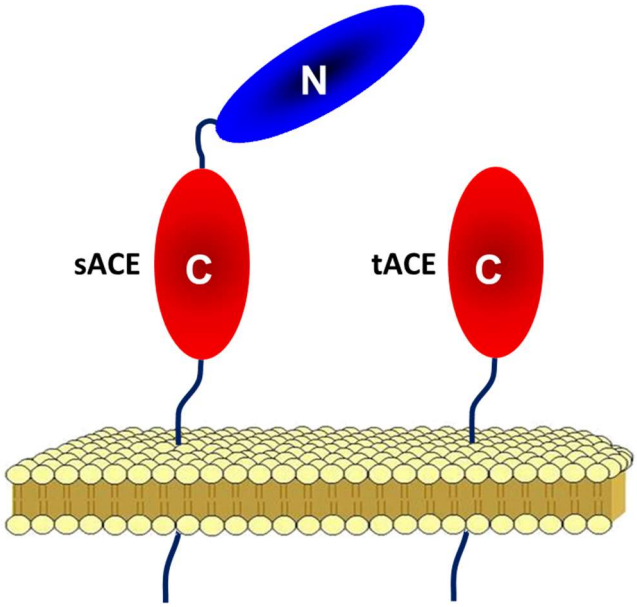


***Figure 1.8**: A schematic representation of two different isoforms of ACE; sACE and tACE. sACE is the ubiquitously expressed isoform with both the N- and C-domains present along with a transmembrane and linker region. tACE is a form of ACE expressed only in testicular tissue with just the C-domain and the transmembrane anchor present.*

1.3.2 Kinetic Properties of the Two Domains of ACE

The first line of research into the different catalytic domains was directed towards comparing the differential rates of substrate hydrolysis between the two domains. Wei et al[3a] expressed truncated forms of ACE containing either just the C-domain or the N-domain active site mutants with zinc-binding sites disrupted. Under assay conditions, both the N- and C-domains showed almost identical K_m values towards Ang-I. However, N-domain knockout mutants showed a 3-fold increase in the k_{cat} values as opposed to the C-domain knock out mutants. The k_{cat}s of the individual domains add up to the k_{cat} of sACE suggesting no synergistic effect between the two domains is present. The rates of hydrolysis were also compared between the two domains using the Hip-His-Leu (HHL) substrate. It was found that the C-domain was responsible for 90% of the HHL hydrolysis of sACE. The implication here is that Ang-I and HHL hydrolysis is primarily controlled by the C-domain hence all inhibitors were developed against C-domain inhibition without much information regarding their N-domain binding.

The chloride dependence of the different domains was also investigated[3b]. sACE was known to be highly chloride dependent with an optimal chloride concentration of 800 mM. The optimum chloride concentration for the individual domains was determined to be 10 mM and 800 mM for the N- and C-domains respectively with only the C-domain activity proving to be chloride dependent. *In vivo* chloride concentration is related to water volume, a kidney regulated function thus suggesting C-domain chloride dependence to be another link between renal function and the RAAS. Conversely, the N-domain appears to operate independently of *in vivo* chloride concentrations. These findings further emphasise that the majority of the known ACE characteristics at that time were a result of the C-domain function while the precise role of the N-domain remained unclear.

As the differences between the two domains were elucidated, a more comprehensive study of the domains and their substrates was performed. Jaspard et al[39] expanded on the previous two studies to investigate the activity of ACE using three additional peptides, BK, substance P, (Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH_2) and luteinising hormone releasing hormone (LH-RH), Glp-His-Trp-Ser–Tyr-Gly-Leu-Arg-Pro-Gly-NH_2. Both domains showed similar K_m and k_{cat} values for the hydrolysis of BK. In wild type sACE, the total BK k_{cat} proved to be roughly equivalent to the sum of the individual k_{cat}s implying that both domains hydrolyse BK with similar efficiencies.

In the case of substance P, sACE was able to cleave both the C-terminal dipeptide and tripeptide, a rare observation for an enzyme, which had been assumed to be an exclusive dipeptidase. In this instance, sACE cleaved the terminal tripeptide of Substance P at three times the rate of its terminal

dipeptide. This is presumably because Phe-Gly-Leu is a better fit for the S_1, S_1' and S_2' subsites. The N-domain catalysed this peptide at 1.5 times the rate of the C-domain. In the case of LH-RH, ACE was able to cleave both the C- and N-terminal tripeptides. Here the reaction was catalysed by the N-domain at 10 times the rate of the C-domain. While this study showed that some substrates are selective towards the N-domain, sACE showed very high K_m values of 265 µM and 6 mM towards LH-RH and substance P respectively. Such high dissociation constants are unlikely to translate into any observable activity towards these substrates *in vivo.* Another point to note is the C-terminal amidation of these two peptides. It was originally postulated that a C-terminal acid was essential for a peptide to bind to sACE yet substance P and LH-RH are clear exceptions. Despite this exception it is possible that the C-terminal amidation of these peptides is responsible for their high dissociation constants.

1.3.3 AcSDKP – An N-Domain Selective Substrate

The new discovery of the two homologous active sites of ACE posed almost as many new questions as it answered. While it had been shown that the C-domain is responsible for the majority of Ang-I hydrolysis and more than half of the BK degradation, no distinct role for the N-domain had been identified until the discovery of AcSDKP. AcSDKP is a small tetrapeptide first isolated by Lenfant et al[40] from foetal calf marrow. It was identified as a signalling molecule responsible for the inhibition of pluripotent haematopoietic stem cell proliferation after irradiation. It acts by arresting the cell's entry into the S-phase of the cell cycle. Further investigation by Robinson et al[41] showed the role played by AcSDKP in the cell cycle arrest to be indirect, suggesting it to inhibit the release of a S-phase stimulator. This new pathway was implicated in the prevention of tumour growth and cell death in irradiated cell lines hence it showed the potential as a novel cancer therapeutic.

AcSDKP was subsequently shown to be rapidly hydrolysed by ACE,[42] providing further evidence that its role extends beyond blood pressure regulation. Examining the hydrolysis by the individual domains, it was found that the N and C-domains have similar binding affinities towards AcSDKP with K_ms of 31 and 39 µM respectively.[43] The N-domain, however, catalyses AcSDKP at about 40 times the rate of the C-domain (N-domain k_{cat} = 16s^{-1}; C-domain k_{cat} = 0.40 s^{-1}) suggesting the *in vivo* ACE hydrolysis of AcSDKP occurs exclusively at the N-domain. This was the first pathway identified where the N-domain of ACE plays a primary role. Later studies showed that ACE inhibition was a viable *in vivo* treatment for maintaining AcSDKP levels and arresting the cell cycle of irradiated haemopoietic stem cells.[44]

At the same time a link between ACE and cardiac remodelling/fibrosis was being established. Sun et al[45] showed ACE to play a crucial role in the deposition of collagen in cardiac tissue, the lead cause of cardiac fibrosis. Brooks et al[46] then showed that captopril was capable of treating fibrosis induced

myocardial infarction. AcSDKP was subsequently shown to mitigate the fibrotic effects of increased collagen deposition in cardiac tissue.[47] Furthermore, Peng et al[48] linked the antifibrotic effect of ACEis to raised AcSDKP concentrations *in vivo.* Thus, ACE was suggested to be a viable antifibrotic therapeutic target in addition to an antihypertensive one.

1.4 Selective Inhibition of the Two Domains

1.4.1 ACEis and the Two Domains of ACE

To complement the recent kinetic studies on the individual domains of ACE, Wei et al[3b] completed an inhibitor study on these individual domains. The K_i*s* of four ACEis were reassessed against the individual domains using Hip-His-Leu as a substrate (Table 1.3).

Table 1.3: The K_i values of 4 inhibitors against the N- and C-domains of ACE.[3b]

	K_i (nM)			
ACE Form	**Captopril**	**Enalaprilat**	**Lisinopril**	**Trandolaprilat**
Wild type	1.3	0.65	0.39	0.045
N-domain	0.89	2.6	4.4	0.31
C-domain	1.4	0.63	0.24	0.029

Since the C-domain is predominantly responsible for the cleavage of Hip-His-Leu, it is not surprising that each ACEi displayed similar K_is towards both sACE and its C-domain. In contrast, the N-domain K_is were different to that of sACE in each case. Of the four ACEis, Lisinopril and Trandolaprilat were the most C-domain selective, showing 18 and 10-fold selectivities towards the C-domain over the N-domain respectively. Surprisingly, Captopril proved to be marginally N-domain selective. While the inhibition profiles of the two domains appear distinct, all four inhibitors are potent against both domains suggesting a high degree of structural similarity despite their functional differences.

1.4.2 Selective Inhibition of the Two Domains

With the success of ACEis as a treatment for hypertension, there have been several side-effect drawbacks. The most notable of these side-effects are persistent cough and angioedema which, despite being uncommon, is potentially life threatening.[49] When the cause of angioedema and cough was identified as excess BK accumulation,[50] the idea of developing domain specific ACEis became an attractive one. The rationale being that a C-domain selective ACEi could sufficiently reduce Ang-I hydrolysis to treat hypertension but would leave the N-domain free to degrade BK, thus mitigating the cause of angioedema and cough. Conversely N-domain selective ACE inhibition has therapeutic

potential in the treatment of cardiac and pulmonary fibrosis/remodelling without affecting blood pressure.

While the first generation of ACEis were shown to potently inhibit both domains, the different kinetic profiles of the two domains suggested domain selective inhibition was possible. Michaud et al[51] examined the differences in inhibition by Captopril, Lisinopril and Fosinoprilat against the two domains using the two native substrates, Ang-I and AcSDKP. Captopril was the only one of the three inhibitors to display a slight selectivity towards the inhibition of AcSDKP hydrolysis as opposed to Ang-I hydrolysis. Since AcSDKP hydrolysis is entirely controlled by the N-domain, the N-domain selectivity of Captopril was assumed to be responsible. Deddish et al[52] went on to examine the differential activity of ACE on many different peptides and found Ang_{1-7} to be of particular interest. The N-domain hydrolysed Ang_{1-7} much faster than the C-domain with a k_{cat} of 27 min^{-1} as opposed to 0.36 min^{-1}. The second observation was that Ang_{1-7} is a 10 to 15-fold stronger inhibitor of C-domain Ang-I hydrolysis than the N-domain. These measurements were later contradicted by Thomas et al[53] measuring these k_{cat}s for the N- and C-domain at 3.0 and 1.4 s^{-1} respectively. Despite the disagreement over these kinetic parameters, the sequence of Ang_{1-7} played an integral role in the design of the first C-domain selective ACEi.

Ang_{1-7} has the sequence Asp-Arg-Val-Tyr-Ile-His-Pro. Since all ACEis at the time had mimicked the terminal two or three residues on the C-terminal of a peptide, the search for a small molecule inhibitor mimicking this sequence was conducted. An ACEi published in 1981 by Weare et al[54] called Keto-ACE (Figure 1.9) stood out with its peptide mimicking pseudo Phe-Phe-Gly-Pro sequence sharing some chemical similarity with the sequence of the four C-terminal residues of Ang_{1-7}. When tested *in vitro,* keto-ACE showed a remarkable 50-fold C-domain selectivity when tested against the BK substrate while a 30-fold selectivity was observed when tested against the Ang-I substrate. Thus the first small molecule C-domain selective inhibitor was confirmed.

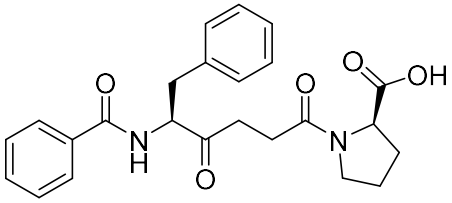


Figure 1.9: *Chemical structure of keto-ACE*

This techinique of studying the inhibition and kinetics of native peptides to customise a peptidomimetic inhibitor also proved useful for designing an N-domain selective inhibitor. In the mid 1990s, Jiracek et al[55] discovered several metalloprotease inhibitors based on a central phosphinate

ZBG with two amino acid residues on either side. N-terminal acetylation and C-terminal amidation was also tested on these peptides. A high throughput solid-phase peptide synthesis protocol was used to synthesise various permutations of this phosphinic acid tetrapeptide template. Using the sequence of AcSDKP as a guideline, Dive et al found Ac-Asp-Phe-(PO_2CH_2)-Ala-Ala-NH_2 to be a potent N-domain selective ACE inhibitor with a 2000-fold selectivity towards the N-domain.[56] This compound was named **RXP407** (Figure 1.10).

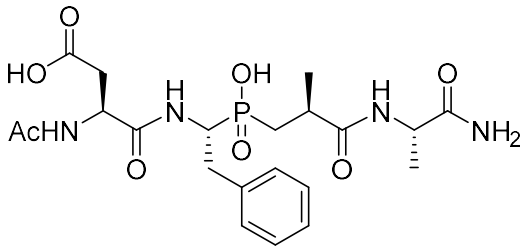


Figure 1.10: *Chemical structure of* ***RXP407***

Table 1.4*:* Inhibition of the two ACE catalytic domains by RXP407 and its analogues

Name	Sequence	ACE-N K_i (nM)	ACE-C K_i (nM)
RXP407	Ac-Asp-PheΨ(PO_2CH_2)-Ala-Ala-NH_2	12	25000
Compound II	Ac-Asp-PheΨ (PO_2CH_2)-Ala-Ala-OH	2	7
Compound III	NH_2-Asp-PheΨ (PO_2CH_2)-Ala-Ala- NH_2	5	800
Compound IV	Ac-Ala-PheΨ (PO_2CH_2)-Ala-Ala- NH_2	15	200

Table 1.4 details the inhibition of **RXP407** and its derivative compounds against the two domains of ACE. Substituting the C-terminal amide for a carboxylic acid in **Compound II** abrogates the N-domain selectivity of **RXP407**. A 3600-fold decrease in N-domain K_i caused by substituting the C-terminal carboxylic acid with an amide is extremely surprising as both domains are dipeptidases with assumed positively charged regions in the S_2' prime subsite. A non-uniform change in binding affinity between the two domains over this carboxylic acid to amide substitution was inconsistent with the existing model for the two domains of ACE. **Compound III** is the N-terminal deacetylated form of **Compound I**. This deacetylation produces a 30-fold reduction in N-domain selectivity. Lastly in **Compound IV**, the P_2 Asp is replaced with an Ala. This substitution effects a 100-fold reduction in N-domain selectivity. The SAR conclusions drawn from this study showed that all four of these phosphinic peptides are potent ACE N-domain inhibitors. The selectivity towards the N-domain over the C-domain hinged on a carboxy terminal amide, a terminal amine acetylation and a P_2 carboxylic acid, all of which appear to be crucial to the disruption of C-domain binding.

Georgiadis et al[57] followed a similar approach to develop a C-domain selective phosphinic acid tripeptide. The most potent ACE C-domain selective inhibitor derived from this study was Cbz-Phe-(PO_2CH_2)-Pro-Trp which was later named **RXPA380** (Figure 1.11)

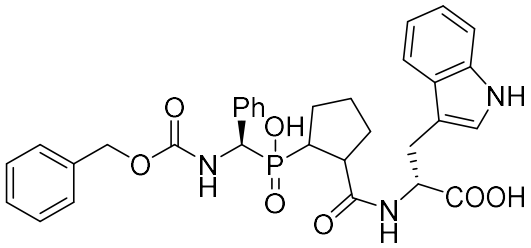


Figure 1.11: *Chemical structure of* ***RXPA380***

In a Mca-Ala-Ser-Asp-Lys-DPa-OH competitive inhibition assay, the K_i of **RXPA380** towards the ACE C-domain was 3 nM while no inhibition of the ACE N-domain was observed for inhibitor concentrations up to 10 µM. The C-domain selectivity of **RXPA380** was suggested to stem from either the entropic constraints brought about by the P_1' pseudo-Pro or the large hydrophobic indole group of the Trp in the S_2' pocket.

Table 1.5: An SAR series probing for C-domain selectivity at the P_1' and P_2 positions.

Compound	Structure	C-domain K_i (nM)	N-domain K_i (nM)	C Selectivity (fold)
RXPA380	Cbz-Phe-(PO_2)-Pro-Trp-OH	3	> 10000	> 3300
7	Cbz-Phe-(PO_2)-Ala-Trp-OH	0.5	45	90
8	Cbz-Phe-(PO_2)-Pro-Ala-OH	20	450	22.5
9	Cbz-Phe-(PO_2)-Ala-Ala-OH	0.8	0.8	1
10	Cbz-Phe-(PO_2)-Pro-Pro-OH	4	60	15
11	Cbz-Phe-(PO_2)-Pro-Arg-OH	9	200	22.2
12	Cbz-Ala-(PO_2)-Pro-Trp-OH	60	8000	133
13	Cbz-(h)Phe-(PO_2)-Pro-Trp-OH	65	9000	138

Georgiadis et al[58] examined the chemical basis for the C-domain selectivity of **RXPA380** and synthesised a series of **RXPA380** analogues to probe the contribution made by each amino acid side-

chain towards this domain-selectivity. Compounds **7-13** from Table 1.5 each have one or two side-chain substitutions. From this SAR study, it emerged that both the P_1' pseudo-Pro and the P_2' Trp are essential for C-domain selectivity. Substituting either moiety for an Ala came at a significant cost but not a complete loss in C-domain selectivity. This would seem to suggest that both the P_1' Pro and P_2' Trp are deleterious to N-domain binding with the P_1' Pro more so than the P_2' Trp. Considering all the BPPs contain a P_1' Pro, this moiety alone cannot be responsible for a deleterious N-domain interaction. The P_2' Trp would seem to be a much better candidate for the deleterious N-domain interaction. Intriguingly, the Pro-Trp combination appear to be crucial for C-domain selectivity. This suggests that the Pro locks the Trp in a conformation forcing an unfavourable interaction with the S_2' pocket of the N-domain.

The development of **RXP407** and **RXPA380** was a major breakthrough in designing domain-specific ACEis. Both compounds display poor drug-like properties due to their large molecular weights, their highly polar phosphinic acid groups and the presence of several peptide bonds. These molecules, however, do serve as useful templates for further SAR studies and the development of small molecule drug-like domain-selective ACE inhibitors.

1.4.3 *In vivo* Domain Specific Inhibition

In vivo domain specific ACE inhibition has been investigated to determine the pharmacological and therapeutic efficacy of inhibiting catalytic domains of ACE individually. Since **RXPA380** and **RXP407** have poor oral bioavailability, Georgiadis et al[57] tested their effects intravenously by injecting each inhibitor into mice and then infusing them with radiolabelled Ang-I, BK or AcSDKP substrate. The ratio of radiolabelled hydrolysed substrate was then measured at regular intervals. **RXPA380** and **RXP407** behaved as expected when inhibiting the hydrolysis of these three radiolabelled substrates. The **RXPA380** inhibition curve of Ang-I hydrolysis showed two inflection points. It was capable of inhibiting 75% of the Ang-I hydrolysis at low concentrations and was only able to inhibit it completely at very high concentrations. Almost the reverse was observed with the dual inflection inhibition profile of **RXP407**. **RXP407** was shown to be capable of inhibiting about 30% of Ang-I hydrolysis at low concentrations with full inhibition achieved again at very high concentrations. In the case of BK, only partial inhibition of BK hydrolysis was achieved with each inhibitor when administered individually yet full inhibition was achieved in the presence of both inhibitors when administered in combination. In the case of AcSDKP, **RXPA380** had no effect on its degradation while **RXP407** was capable of inhibiting AcSDKP hydrolysis completely.

The above findings were entirely consistent with previous observations relating to the role of the two domains of ACE and the three substrates. However, since exogenous substrates were used, the mouse model did not show the effect of domain selective inhibition on the endogenous levels of these three substrates and the associated physiological responses. To study the effect of domain selective ACE inhibition on endogenous Ang-I, BK and AcSDKP, Fuchs et al developed transgenic mice where either the N- or C-domain was inactivated.[5]

In their first study, Fuchs et al bred ACE mutant mice with their N-domain disabled. The serum from these mice displayed about 75% of the Ang-I cleavage ability of wild type mice without the ability to hydrolyse Ac-SDKP. However, *in vivo,* the N-domain inactivated mice showed identical Ang-II/Ang-I ratios as wild type mice. Even more surprising was the response of the mice to an Ang-I infusion. The N-domain knockout mice again showed an identical response to an Ang-I infusion as the wild type mice. N-domain knockout mice on the other hand displayed negligible AcSDKP hydrolysis in comparison to the wild type. These findings suggest that N-domain inactivated mice can compensate for the lower Ang-I hydrolysis further upstream in the RAAS via increased hydrolysis of angiotensinogen. It also infers that N-domain selective ACE inhibition will have no effect on blood pressure.

In the corollary study, Fuchs et al[4] bred ACE mutant mice with inactivated C-domains. As expected, serum from these mutant mice could only hydrolyse the substrate HHL at a small fraction of the wild type while its AcSDKP hydrolysis was identical to wild type serum. *In vivo,* the mutant mice displayed a 4.2-fold reduction in the Ang-II/Ang-I ratio compared with wild type while $BK_{(1\text{-}7)}$/BK ratios were similar to wild type mice. Furthermore, when injected with Ang-I, the hypertensive response was half the magnitude in the C-domain mutant mice than the wild type while a similar blood pressure reduction was observed in response to a BK infusion in both the wild type and C-domain knockout mice.

These experiments indicate that selective inhibition of the C-domain is more than sufficient to achieve the desired anti-hypertensive response from inhibiting both catalytic sites of ACE. Furthermore, selective C-domain inhibition leaves the N-domain free to hydrolyse some BK which can negate some of the side-effects like angioedema and persistent cough caused by dual domain inhibition. This study, therefore, successfully demonstrated the benefits of individually inhibiting both domains of ACE while showing them to be promising drug targets in their own rights.

1.5 ACE Structural Determination

1.5.1 Crystal Structures of ACE

During the second half of the 20th century, protein crystallography developed into an important tool for drug discovery and proved invaluable for studying small molecule-protein interactions. The first crystal structure of ACE was solved by Natesh et al[59] in 2003, nearly 30 years after the discovery of Captopril, making the achievement of Cushman and Ondetti even more remarkable. Many unsuccessful attempts were made over the decades to crystallise ACE until the advent of recombinant expression and DNA sequencing revealed the reason for these failures to some extent. Sequencing showed ACE to contain a lipophilic membrane anchor and a highly flexible linker region connecting the two domains. Both the membrane anchor and the linker region are highly disordered in solution. In addition to these disordered regions, ACE is highly glycosylated posing a hindrance to crystal packing. These physical features meant ACE could not be crystallised without a few manipulations.

Natesh et al[59] crystallised a truncated variant of the tACE isoform lacking the N-terminal 36 residues, the TM and the cytoplasmic region. Since tACE consists of a single globular domain with an exact replica of the ACE C-domain, there was no disordered linker region or N-domain to contend with. This truncated construct still appeared to retain all of the original enzymatic activity of tACE and was co-crystallised with Lisinopril (PDB code: 1O86, resolution: 2.0 Å) and in the apo form (PDB code: 1O8A, resolution: 2.0 Å).

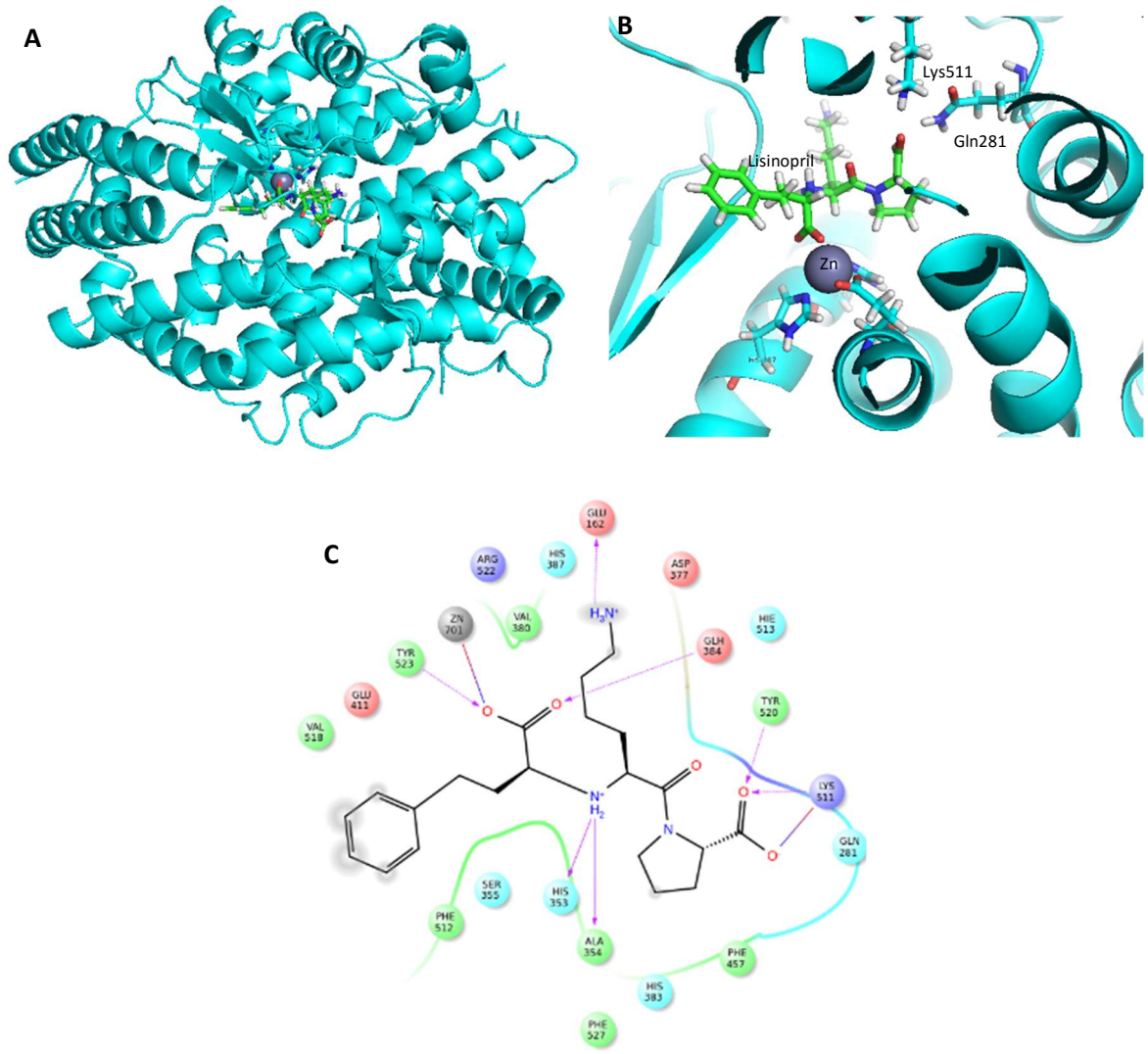


Figure 1.12: *The first crystal structure of tACE (PDB entry 1O86).* ***A****. A representation of the entire tACE crystal structure with the binding cleft visible. Inside the binding cleft the catalytic Zn metal is clearly coordinated to Lisinopril.* ***B****. A close-up of Lisinopril inside this C-domain binding site. Here it can be seen that Zn coordinates with 3 residue side-chains (His384, His387 and Glu411). The central carboxylic acid of Lisinopril also coordinates with the Zn while its terminal carboxylic acid fits into an electropositive hole created by the adjacent Lys511 and Gln281 residues. Its Lys and Phe side-chains appear to be firmly fixed within distinct subsites.* ***C****. A 2D ligand interaction map summarising all the residue side-chain interactions experienced by Lisinopril within this binding site.*

The crystal structure 1O86 confirmed and contextualised many of the features of the ACE C-domain (Figure 1.12). The chloride dependency was explained by the presence of two chloride binding sites within electropositive pockets. These chloride ions appear to be essential for correct protein folding. The structure was largely globular with a central binding cleft. The C-domain could be further split into two sub-domains on either side of the binding cleft. As expected, the Zn was centrally located within the binding cleft coordinated to two His and one Glu residues (His383, His387 and Glu411). The terminal carboxylic acid of Lisinopril was anchored in an electropositive hole in the S_2' pocket created by the positively charged Lys511 and the polar H-bond donating Gln281 residues. The S_1 and S_1'

pockets are deep and lined with several hydrophobic residues while the amine of Lisinopril acts as a donor in an H-bond with a backbone carbonyl.

Crystallising the N-domain proved to be more of a challenge, most likely due to the increased number of glycans on the surface of the N-domain. Nevertheless, a truncated N-domain construct was crystallised after some of the N-glycan networks were enzymatically trimmed.[60] Unfortunately, crystallisation was still compromised by a few highly disordered loops, which restricted the resolution of the structure to 3.0 Å. At this resolution, the two ACE domains were easily superimposed with a Cα RMSD of 0.45 Å between 82 active site residues. The Lisinopril poses were virtually identical in these two structures (Figure 1.13).

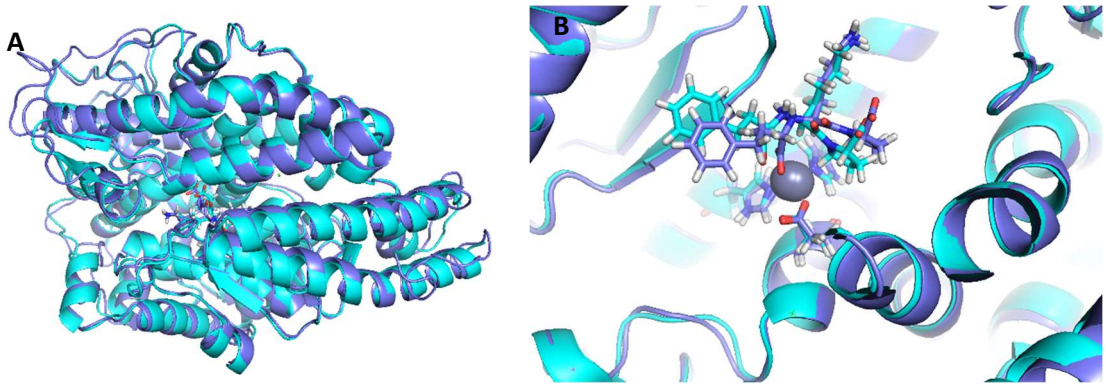


***Figure 1.13**: An overlay of the crystal structures of the two domains of ACE. The N-domain (cyan) and the C-domain (blue) are derived from the PDB entries 2C6N and 1O86 respectively. **A** shows the entire crystal structure overlays of both domains. All the α-helices are clearly present in both structures while there is some deviation in the disordered loop regions. **B** shows a close-up of the Lisinopril ligand from both crystals which clearly bind with identical poses in both domains.*

The *N*-glycosylation of the N-domain was examined in more detail by Anthony et al[61] with a series of glycosylation site mutants to establish the sites essential for maintaining enzyme activity. Figure 1.14 shows a map of the *N*-glycosylation sites of the two domains. It was found that the *N*-domain with only the 3, 8 and 9 *N*-glycosylation sites intact (Ndom389G) was the minimally glycosylated form of the enzyme necessary to preserve both its stability and enzymatic activity. This glycosylation mutant was crystallised in a stable complex with **RXP407** and its structure was solved to a resolution of 1.99 Å (PDB code: 3NXQ).

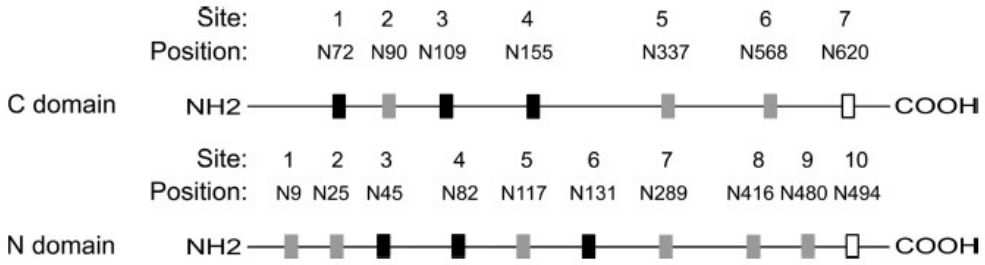


***Figure 1.14**: An N-glycosylation map of the two active site domains of ACE. The 7 sites in the C-domain and the 10 sites in the N-domain are indicated. The black boxes indicate conserved sites, the grey boxes indicate the glycosylation sites unique to that domain and the white boxes are unglycosylated sites.*

To date (Dec 2016), there are 35 entries in the PDB for *H. sapien* ACE using either the tACE (tACEΔ36NJ) construct or the G13 mutant for the C-domain and the Ndom389G construct for the N-domain. Of these 29 entries, 14 are from the N-domain and the remaining 21 are from the C-domain. (Table 1.6).

***Table 1.6**: A table cataloguing all the published ACE crystal structures. The fields provided are the construct used, resolution and the ligand cocrystal.*

PDB Code	Construct	Resolution (Å)	Ligand
	C-domain		
1O8A	tACEΔ36NJ	2.00	Apo
1086[59]	tACEΔ36NJ	2.00	Lisinopril
1UZE	tACEΔ36NJ	1.82	Enalaprilat
1UZF[62]	tACEΔ36NJ	2.00	Captopril
2OC2[60]	tACEΔ36NJ G13	2.25	**RXPA380**
2IUL	tACEΔ36NJ G13	2.01	Apo
2IUX[63]	tACEΔ36NJ G1234	2.80	Apo
2XY9[64]	tACEΔ36NJ G13	1.97	**FII**
2YDM[65]	tACEΔ36NJ G13	2.44	Se-Captopril
3BKK	tACEΔ36NJ G13	2.17	**kAF**
3BKL[66]	tACEΔ36NJ G13	2.18	**kAW**
3L3N[67]	tACEΔ36NJ G13	2.30	**lisW**

4APH	tACEΔ36NJ G13	1.99	Ang-II
4APJ[68]	tACEΔ36NJ G13	2.6	BPPb
4CA5[69]	tACEΔ36NJ G13	1.85	**FI**
4C2R	tACE G13 R522Q	2.30	Apo
4C2Q	tACE G13 R522K	2.40	Apo
4C2P	tACE G13 R522K	1.99	Captopril
4C2O	tACE G13 D465T	1.80	Apo
4C2N[70]	tACE G13 E403R	2.59	Apo
4BZR[71]	tACEΔ36NJ G13	1.84	**K-26**
	N-domain		
2C6F	Ndom wt	3.01	Apo
2C6N[60]	Ndom wt	3.00	Lisinopril
2XYD[64]	Ndom389	2.15	**FII**
3NXQ[61]	Ndom389	1.99	**RXP407**
4BXK[72]	Ndom389	2.20	**33RE**
4BZS[71]	Ndom389	2.10	**K-26**
4CA6[69]	Ndom389	1.91	**FI**
4UFB	Ndom389	1.80	Lys-Pro
4UFA[73]	Ndom389	1.80	Ac-SD
5AM8	Ndom389	1.90	Amyloid-β 4-10
5AM9	Ndom389	1.80	Amyloid-β 10-16
5AMA	Ndom389	1.80	Amyloid-β 1-16
5AMB	Ndom389	1.55	Amyloid-β 35-42

5AMC	Ndom389	1.65	Amyloid-β 4-10

1.5.2 Structural Determinants of C-Domain Selectivity

As already discussed, Georgiadis et al investigated the effect of each side-chain moiety on the C-domain selectivity of **RXPA380**[58]. In this study, some molecular modelling experiments were performed to determine the most important interactions to explain the C-domain selectivity of **RXPA380**. Molecular Mechanics (MM) simulations were performed on **RXPA380** in the binding pocket of the crystal structure 1O86 to predict bioactive conformations of **RXPA380**. These simulations suggested that the indole moiety of **RXPA380** sits deeper in the S_2' pocket of ACE than the Pro group of Lisinopril. Here the indole is brought into contact with the residues Val379 and Val380 of the 1O86 crystal structure. These two hydrophobic residues serve to create a hydrophobic region in the vicinity of the indole moiety of **RXPA380.** These interactions were confirmed in the crystal structure 2OC2.[74] Superimposing the N-domain and C-domain structures showed that Val379 and Val380 of the C-domain are replaced with a Ser357 and Thr356 respectively (Figure 1.15). These polar residues create a hydrophilic region in the vicinity of the hydrophobic indole moiety of **RXPA380**. This leads to an unfavourable interaction in the S_2' subsite of the N-domain of ACE. It therefore explains the lack of binding to the N-domain, hence the C-domain selectivity observed with **RXPA380**.

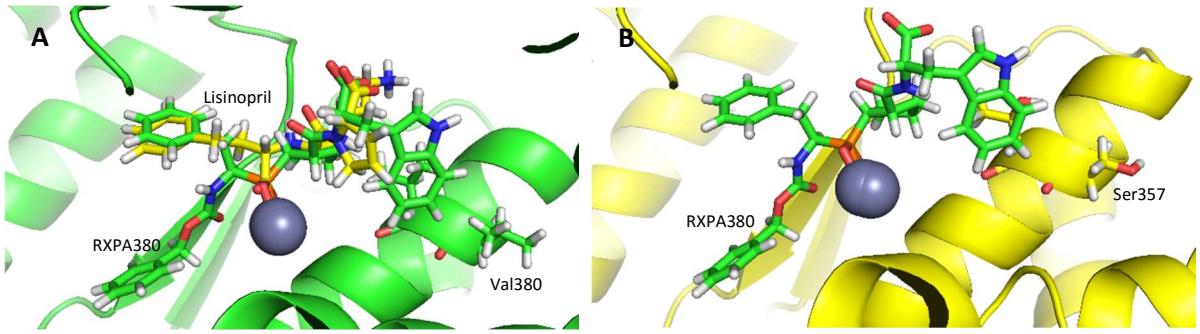
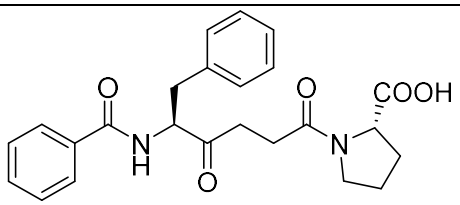
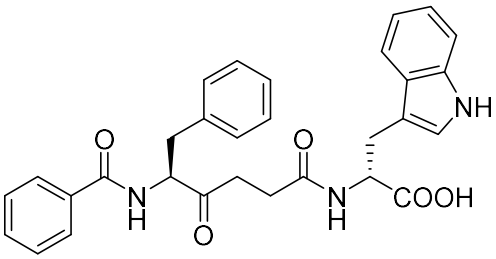
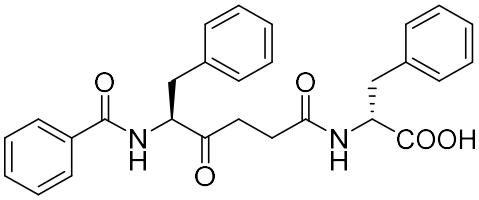
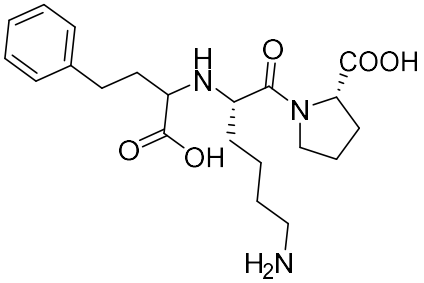
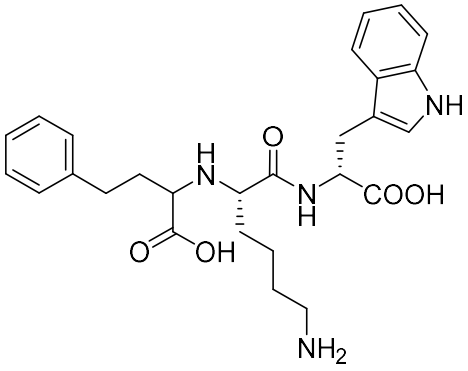


***Figure 1.15**: Structural determinants of the C-domain selectivity experienced by ACE. A. **RXPA380** (green) bound to the C-domain and overlaid with Lisinopril (yellow) from 1O86. The indole moiety of **RXPA380** sits much deeper in the S_2' subsite where it can interact with Val379 and Val380. It must also be noted that the Cbz moiety of **RXPA380** sits in the S_2 subsite which Lisinopril cannot access. B. **RXPA380** is placed within the active site of the N-domain (2C6N). In the N-domain, Val379 and Val 380 have been replaced with a Ser and Thr residue respectively. These polar side-chains interact unfavourably with the Indole moiety.*

The SAR of a bulky P_2' hydrophobic group was also explored in two other series; a series based on **keto-ACE** which has a P_2' Pro and a series based on Lisinopril (Table 1.7). In the case of the **Keto-ACE** series,[60] replacing the P_2' Pro with a Trp and Phe destroyed any remaining N-domain inhibition. This trend was also observed with the Lisinopril derivative. Substituting the C-terminal Pro moiety of Lisinopril for a Trp to create Lis-W greatly reduced its N-domain inhibition (400-fold) but maintained its C-domain inhibition. **LisW**[67] seemed like a promising drug lead owing to its similarity to Lisinopril;[75] however, it displayed poor oral bioavailability.[76] Despite the failure of **LisW**, Lisinopril P_2' analogues are still the subject of ongoing research into C-domain selective ACE inhibition.

***Table 1.7**: A table of two chemical series exploring P_2' SAR in ACEis. The two series shown are derivatives of either Keto-ACE (kAP) or Lisinopril*

Compound	Structure	C-domain K_i (nM)	N-domain K_i (nM)	C Selectivity (fold)
Keto-ACE		50	1500	30
kAW		679	854200	1258
kAF		830	>500000	>600
Lisinopril		1.2	4.8	4
LisW		6.6	1700	258

1.5.3 Structural Determinants of N-Domain Selectivity

N-domain selectivity has already been discussed within the context of the discovery of **RXP407**. Based on the sequence of five phosphinic acid tetrapeptides, it was concluded that N-domain selectivity was caused by both the presence of a C-terminal amide and an N-terminal Asp. The successful cocrystallisation of **RXP407** in the N-domain by Anthony et al[61] challenged the findings of Dive et al with respect to the importance of the C-terminal amide. Modelling **RXP407** within the active site of the original N-domain crystal[60] (PDB code: 2C6N) did not provide any direct evidence to support the theory that a C-terminal amidation was in part responsible for the selective N-domain binding. Both domains have clearly defined positively charged cavities at the edge of the S_2' pocket facilitating the binding of the anionic C-terminal carboxylic acid. The positively charged hole is lined with the side-chains of a Lys and Gln, which occupy identical positions in both domains. The notion of a C-terminal amidation being responsible for a massive deleterious C-domain interaction (three orders of magnitude loss in binding affinity) would therefore seem improbable. It is rare to find such a large change in binding energy during SAR studies, which can be solely attributed to such a small change.

The interaction between the acidic P_2 group and the Arg381 residue in the $S_{2,}$ pocket is more likely to be responsible for conferring N-domain selectivity.[77] In the crystal structure of the N-domain in complex with Lisinopril (2C6N), Arg381 is pointing away from the active site. Kroger et al[77] correctly predicted Arg381 to reposition itself to form a salt bridge with the aspartate of **RXP407** as was observed in the PDB structure 3NXQ.[61] The Arg to Glu substitution in the C-domain would result in a clash between two negatively charged groups. The proposed relationship between C-terminal amidation and a deleterious C-domain interaction is presumably related to the mechanism of **RXP407** entering the active site cleft rather than its final resting position.

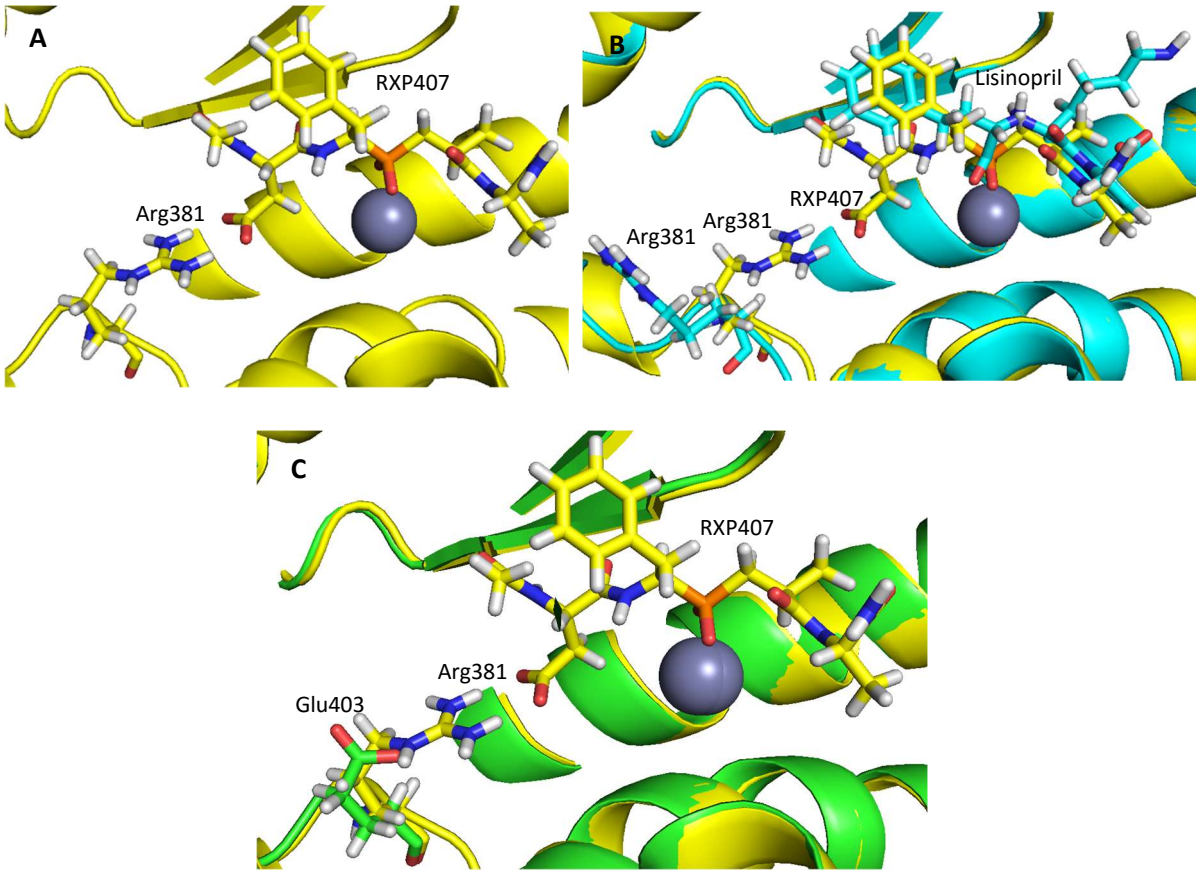


***Figure 1.16**:The crystal structure of **RXP407** in complex with the N-domain. **A**. The **RXP407** ligand bound to the ACE N-domain from the crystal 3NXQ. Here a salt bridge is clearly visible between **RXP407**'s P2 carboxylic acid and Arg381. **B**. When overlaid with the N-domain structure from 2C6N (cyan), a huge difference in the position of Arg381 is observed. In the absence of a P_2 carboxylic acid, Arg381 is capable of swinging out and reorientating itself away from the binding cleft. **C**. **RXP407** in the 3NXQ crystal structure (yellow) overlaid with the C-domain of 1O86 (green). Here it is shown that Arg381 is substituted for Glu403 in the C-domain. An acid-acid clash in the S_2 subsite can explain the deleterious effect this moiety has on the binding of **RXP407** to the C-domain.*

The contribution of an acidic P_2 moiety towards disrupting the ligand binding with the C-domain and conferring C-domain selectivity has been reaffirmed with a recent scaffold hopping exercise by Douglas et al.[72] Using the SHOP[78] methodology, the substitution of the P_2 Asp moiety of **RXP407** with a tetrazole was predicted to preserve N-domain selective ACE inhibition. This substituted **RXP407** analogue was synthesised and named 33RE (Figure 1.16) and the K_i of this compound towards the N- and C-domains of ACE were measured at 11.2 nM and 10.4 µM respectively. These K_i values translate to an N-domain selectivity factor of 927, comparable to that of **RXP407**. While **33RE** reaffirms the contribution towards N-domain selectivity made by a P_2 acidic moiety, it is still a phosphinic tetratpeptide like ACE. To date no drug-like N-domain selective ACEis have been reported but **RXP407** and **33RE** provide a promising template for the design of such an inhibitor.

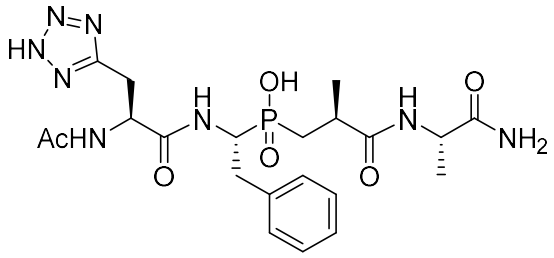


***Figure 1.17**: 33RE*

1.6 Summary and Research Objectives

Despite the wide ranging therapeutic success of ACE inhibitors, new evidence has come to light regarding ACE and its role within the RAAS. As previously mentioned, there are clear advantages to targeting both catalytic domains of ACE individually with domain specific inhibitors. There is now an extensive collection of crystal structures available for both domains. Modern CADD software and high performance computers make it possible to precisely target the structural differences between these two domains as opposed to the more traditional methods of designing peptidomimetics based on peptide substrates.

The research objective of this project was therefore to use a variety of CADD tools and crystal data to guide the design of novel drug-like N-domain selective ACE inhibitors. Several techniques were employed to develop such N-domain selective inhibitors. These include:

- Fragment screening
- Combinatorial libraries
- Database mining
- Protein-ligand docking
- Molecular Dynamics

Once the most promising hits were decided on, they were purchased or synthesised and then subjected to competitive inhibition assays *in vitro* against both domains to test for ACE domain selectivity. Subsequent chapters will describe this approach and the results thereof.

Chapter 2 – Fragment-Based Screening

2.1 Introduction

2.1.1 Docking Background

The primary objective of this study was to discover N-domain selective ACE inhibitors using computer aided drug discovery (CADD) methods. CADD or *in silico* drug discovery is the use of computer software to predict the behaviour of potential drug compounds before synthesis or *in vitro* testing. Using modern computer hardware, CADD software has the ability to predict drug behaviour on a massive set of compounds before undertaking any synthesis or *in vitro* testing. Depending on the methodology, CADD has the ability to save time and resources over traditional screening methods like high throughput screening (HTS). In the last two decades, CADD has developed into an essential tool to complement traditional drug discovery techniques. There are two major streams of CADD, structure based (SB) or ligand based (LB). SB techniques study and predict interactions between drugs and the target using crystal structures, homology models or NMR-derived models of the target. LB drug discovery makes predictions based on the known behaviour of a set of ligands in the absence of a reliable model of the target.

One of the most widely used and versatile drug discovery techniques is molecular docking.[79] In the context of drug discovery, docking entails the fitting of a small molecule into a designated binding site of a modelled target protein.[80] A model of the target can be derived from a crystal structure, a homology model or an NMR-derived structure. Crystal structures are the most accurate depictions of protein structure. Fortunately, the improvement of crystallisation techniques and the advent of high performance computers has dramatically reduced the time and cost required to grow and solve high resolution crystal structures. Protein crystallography is now the preferred method to study and visualise protein–small molecule interactions.

A docking algorithm operates by fitting a small molecule or ligand into a predefined space containing the binding site of the protein target. These algorithms consist of two main components, a search algorithm and a scoring function. The search algorithm explores the 6 degrees of freedom (3 degrees of translational and 3 degrees of rotational freedom) of the ligand to find an appropriate orientation. Once a satisfactory orientation has been found, the rotatable bonds are manipulated to determine the optimal ligand conformation. The simplest search algorithm would entail scanning all the possible conformations evaluating each conformation or pose after sequentially rotating each rotational bond by a small and finite increment. This approach becomes highly problematic as the molecules grow larger. The addition of each rotational bond adds an additional degree of freedom, thus exponentially increasing the conformational space available to the ligand. Such a calculation quickly becomes

prohibitively expensive due to combinatorial explosion as explained by equation 2.1. To work around this problem, many ingenious solutions have been devised to search a representative portion of the available conformational space.

$$N_{conformations} = \prod_{i=1}^{N} \prod_{j=1}^{n_{inc}} \frac{360}{\theta_{i,j}} \quad (2.1)$$

Docking was initially developed as a tool for virtual screening (VS). VS is the screening of a database of drug-like compounds for a novel chemical class of inhibitors against a target. The utility of docking protocols has since been expanded to many additional CADD applications such as *de novo* design, determining the starting geometry of ligands for molecular dynamics (MD) and binding energy calculations. This is thanks to the ability of docking algorithms to rapidly predict the bioactive conformation of a ligand within a protein binding site from a newly rendered 3D structure.

2.1.2 Search Algorithms

The role of a search algorithm is to explore the available conformational space while the scoring function evaluates the energetics of each suggested pose. The search algorithm and scoring function works synergistically to enable the docking programme to rapidly converge on a realistic prediction of the bioactive pose of small molecules within a protein binding site.

Docking simulations are usually performed in a predefined space inside the active site of the target in which the potential field is calculated. This predefined space in which docking simulations are performed is called a docking grid. A docking grid is typically defined as a rectangular box centred on a bound ligand in the respective binding site. While the grid may be defined with flat surfaces, the topology of the docking site is dictated by the surface of the protein. The atoms of the protein are modelled as hard spheres of a radius less than the van der Waals radii (Table 2.1)[81]. Sub van der Waal radii gives the ligand some space to move within the binding site. van der Waal radii violations do, however, lead to steric clashes and large energy penalties in the final pose. A compromise is usually found between allowing the ligand conformational freedom and applying energy penalties to favour accurate poses. This compromise is found by softening the potential and setting the atomic radii in this model to a fixed factor of the van der Waal radii. This topological consideration reduces the conformational space of the search.

***Table 2.1**: The van der Waals radii[81](r_w) of the common atoms found in biological systems*

Element	r_w (Å)
H	1.20
C	1.70
O	1.52
N	1.55

The vast majority of search algorithms fall into one of two classes, Monte Carlo[82] (MC) and Genetic Algorithms[83] (GA). MC algorithms apply a stochastic approach to generate random poses and quickly populate the conformational space with a representative set of poses. GAs apply a penalty-reward system mimicking the process of natural selection. Both methods have their advantages and disadvantages but the key to applying them effectively lies in their parameterisation and the refinement of the algorithm.

2.1.3 Scoring Functions

All good docking algorithms need a good scoring function to complement the search algorithm. The purpose of a scoring function is to cheaply evaluate or score the favourability of a generated pose. It is by no means an attempt to accurately quantify the binding energy between the ligand and protein but rather a tool to drive the search algorithm and rank the poses. There are three major classes of scoring functions employed in docking algorithms; force-field, empirical or knowledge-based. Force-field based functions evaluate the energetics of the pose arising from the placement of a ligand in the electrostatic potential field generated by the surrounding protein. Protein force-fields can be created using one of the major protein force-field generating algorithms such as CHARMM[84], AMBER[85] or OPLS[86]. Empirical scoring functions are based on the sum of uncorrelated terms to describe the interaction between a ligand and a binding site together with the change in solvent accessible surface area. The coefficients of these terms are weighted via a regression analysis of experimentally determined binding energies and X-ray crystal structure derived parameters such as dihedral potentials. Knowledge-based scoring functions are derived statistically by collecting data from large 3D databases of crystal interactions and creating an appropriate potential function to favour these observed contacts and poses.

Scoring functions can only rank poses relative to each other as docking scores correlate poorly with experimentally observed binding energies. The reason is that binding affinity is dependent on certain variables which cannot be accurately evaluated from an end point pose. Binding affinity is dependent on the change in energy (ΔG) of the protein-ligand-solvent system upon the ligand binding to the protein. Initially, the binding site of the unbound Apo protein will be solvated under biological conditions as will the ligand. Both the ligand and the binding site need to be desolvated before the ligand can bind. The binding of the ligand can therefore be described by equation 2.2. The two most popular algorithms for estimating ligand binding energy are the Molecular Mechanics Generalised Born Surface Area (MM-GBSA)[87] and Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA)[88] methods. These methods will be discussed in more detail at a later stage. Both the MM-GBSA and MM-PBSA methods are computationally exhaustive and performing these calculations on imprecise predicted ligand poses would be frivolous. Scoring functions are better suited to cheaply evaluate ligand poses.

$$\Delta G_{bind} = \Delta G_{Ligand-Protein} - \Delta G_{Ligand-desolvation} - \Delta G_{Protein-desolvated} \qquad (2.2)$$

Scoring functions are typically a sum of the interaction energies of a given pose. An example of a simple scoring function is ChemScore developed by Elridge et al[89] (equation 2.3). As shown, this function sums the energy terms of the different types of interactions. It has terms to describe the strength of the lipophilic interactions (lipo), the hydrogen bond interaction (hbond), the metal interactions (metal) and the enthalpy of a rotational bond (rotb). Each energy term carries a coefficient C which has been empirically determined to weight its corresponding interaction term accordingly.

$$\Delta G_{bind} = C_o + C_{lipo} \sum f(r_{ir}) + C_{hbond} \sum g(\Delta r)h(\Delta\alpha) + C_{metal} \sum f(r_{imetal}) + C_{rotb} H_{rotb} \qquad (2.3)$$

Weighting the lipophilic interaction term to reflect the realistic contributions made by lipophilic groups is a big challenge in endpoint scoring functions. (Table 2.2) shows the relative energies of each interaction type. Lipophilic interactions or dispersion forces are a full order of magnitude weaker than dipole-dipole interactions, which in turn are another order of magnitude weaker than hydrogen bond or ionic interactions. Nevertheless, it is well established that non-polar interactions can influence ligand binding to a similar extent as H-bonds. This observation is explained by the displacement of unfavourable water molecules trapped in hydrophobic pockets of the binding site. Displacing such molecules returns them into the bulk solvent system with a large binding energy reward. This desolvation of the binding cavity step depends on several variables which cannot be evaluated from the final pose. To estimate this multivariable binding energy, scoring functions have both weighted

coefficients and a lipophilic term incorporated into the function. This term takes the volumes of lipophilic cavities into account while considering the volume occupied by the hydrophobic groups.

***Table 2.2**: Relative interaction energies of the different types of intermolecular interactions*

Interaction type	Energy (kcal/mol)
Hydrogen	1–12
Dipole – Dipole	0.5–2
Lipophillic	< 1

2.1.4 Rigid Receptor

Arguably the biggest drawback of docking algorithms is the rigid receptor model. When a ligand binds to a receptor, small changes occur within the protein to accommodate the ligand. Most of these changes involve side-chain movements although backbone adjustments are also possible.[90] Accounting for full protein flexibility would render most docking exercises computationally unviable. Hence almost all docking is performed against a rigid receptor. The few exceptions impart flexibility on one or two side-chains. Wei et al[91] had some success in improving the predictions by docking against an ensemble of conformational variants of a single protein. However, fully flexible systems can only be modelled using MD simulations. There are a few examples of hybrid MD-docking protocols, which are attempts at a fully flexible docking system.[92] Most hybrid MD-docking protocols are still under development as only the latest computational hardware is capable of handling such simulations.

2.1.5 Docking Software

Having been first developed in the 1980s, docking algorithms have been steadily refined and diversified into the vast selection of docking algorithms available today. There are too many docking protocols available to list but some of the most popular examples today are Glide[93], Gold[94], ICM[95], AutoDock[96], MOE-Dock[97] and FlexX[98]. For this study, all docking simulations were performed using Glide for a few reasons. The first is that it is widely used in the field. The second is that there have been several studies comparing the different protocols on a variety of test cases.[99] In these studies, the metrics used for the comparison are the RMSD when reproducing the poses of ligands and the enrichment factor for identifying known active compounds from databases. Glide consistently performs favourably when measured using both metrics. It is not uncommon to perform docking studies on the same protein-ligand systems across multiple docking platforms and rank the docking scores via consensus scoring. While such methodologies have their merits, they also have the potential to be time consuming and computationally expensive without adding much value to the study.

Glide is one of the newer docking algorithms having been released in 2003. Its search algorithm is a hybrid of spatial analysis hierarchical filters and MC conformational sampling. The hierarchical filters help to quickly find the correct orientation and position within the docking grid, thus greatly reducing the conformational space to be scanned by the MC sampler. The result is a rapid convergence on the correct pose.

The Glide scoring function (GlideScore) is a derivative of the ChemScore scoring function (equation 2.4). The key modification is the splitting of the H-bond terms into three classes of H-bond interactions. These classes are H-bonds between charged-charged, charged-neutral and neutral-neutral H-bonding atom pairs. This sub classification of H-bond types provides an improved approximation of H-bonding energies. Another addition is the van der Waals and solvation terms, which help to better evaluate the effects of cavity desolvation in an end-point evaluation. This scoring function evaluates interaction energies using an OPLS force field, which was specially developed to handle side-chains and almost all the chemical species found in drug-like molecules.[86] The van der Waals terms are evaluated using the 12-6 Lennard-Jones potential (equation 2.5, figure 2.1). Potential softening by a factor of 0.8 is applied to the van der Waals radii. This enables ligands to access blocked pockets and better approximate the H-bonds, which are always shorter than the sum of two van der Waals radii. The Glide algorithm has three different precision settings, high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP). These levels of precision utilise progressively more integration points to evaluate their functions for more precise predictions. A greater number of integration points improves precision but at great computational cost.

$$\begin{aligned}\Delta G_{bind} = {} & C_o + C_{lipo-lipo}\sum f(r_{ir}) + C_{hbond-neut-neut}\sum g(\Delta r)h(\Delta\alpha) \\ & + C_{hbond-neut-charged}\sum g(\Delta r)h(\Delta\alpha) + C_{hbond-charged-charged}\sum g(\Delta r)h(\Delta\alpha) \\ & + C_{max-metal-ion}\sum f(r_{im}) + C_{rotb}H_{rotb} + C_{polar-phob}V_{polar-phob} + C_{coul}E_{coul} \\ & + C_{vdw}E_{vdw} + \text{Solvation terms}\end{aligned} \qquad (2.4)$$

$$E_{vdw} = \sum_{j=1}^{N} \sum_{i=i}^{N} 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \quad (2.5)$$

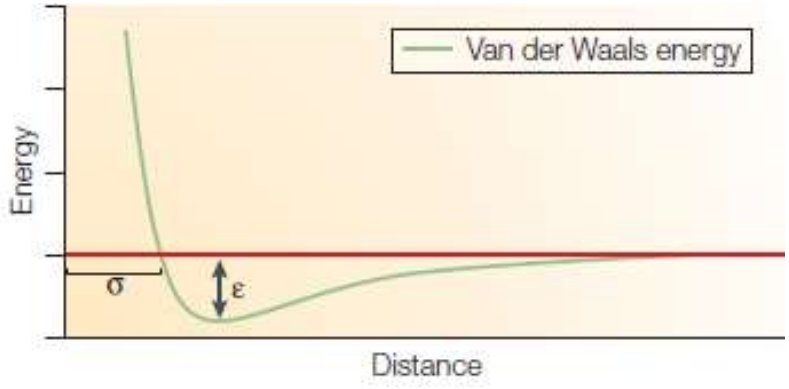


Figure 2.1*:The Lennard-Jones 6-12 potential as a function of a finite distance from an atom σ where a zero potential is experienced with an experimentally predetermined energy minima ε between each atom pair.*

2.1.6 Metalloprotein Docking

A big challenge associated with accurately modelling ACE is the central Zn atom. Zn is a transition metal with a preference for the 2^+ oxidation state. It has a propensity to coordinate with many lone pairs, all of which can affect its affinity for further coordination and the energy associated with such a coordination.

Accurately quantifying the interaction between the Zn and a chelating group requires very expensive quantum mechanical (QM) calculations. Such calculations are much too time consuming to incorporate into docking scoring functions. Glide has opted for an interaction pair approach to approximate metal coordination energies. Each metal-chelator pair has its interaction approximated from a template of pre-calculated values. These values are then applied to the output pose. This solution provides a rough approximation of an interaction, which is comparably inexpensive to calculate. This metal interaction term has improved the accuracy of GlideScore but it is still left with a few caveats. The first being that only ligands with the same metal chelators can be directly compared. The second is that ligands with more than one of the same chelating group are prone to having the incorrect group assigned to the metal for chelation. Glide is nevertheless sufficiently customisable to ensure the correct chelator-metal pair is always assigned.

There have been several previous attempts at applying docking techniques to find novel domain selective ACEis. A notable example is the one carried out by Cresset (unpublished) in 2010. In this study, 200 promising compounds from a database were docked into ACE using the LigandFit[100] docking algorithm. 38 Of the most promising compounds were selected for *in vitro* assays. None of the tested compounds showed more than 50% inhibition of the total ACE activity when tested against either

domain at a concentration of 500 µM. This poor enrichment factor was attributed to many weaknesses in their binding protocol.

LigandFit is a shape-based docking algorithm, which fits ligands to binding site cavities based on their shape complementarity. According to the developers, no special allowances or weightings are made to incorporate a metal binding term like those used in Glide. Another consideration is that the error associated with poor metal binding prediction propagates in larger ligands. The consequence is that errors caused by poor metal binding predictions will propagate in the overall binding prediction.

Considering the above-mentioned factors, it was important to develop a clear and focused docking strategy to address all the challenges presented by ACE. Glide is a better docking algorithm for drug targets containing metals. It also allows the user to set particular docking constraints forcing an interaction between a specific chelator and the metal. Another option to minimise the propagation of error is to dock smaller ligands with fewer rotatable bonds and therefore fewer degrees of freedom. Fragment docking is an attractive solution for the prevention of propagation of error in larger molecules.

2.1.7 Database Preparation

As previously mentioned, VS is the most commonly used application of docking algorithms. Screening exercises usually follow a strict protocol. This protocol typically starts with a selection of drug-like compounds from a database. There are many such freely available databases online. Some have been developed in-house by pharmaceutical companies and are derived from their own archives. Others are subscription based with access requiring costly subscriptions. Other databases are intended mainly for academic research and are widely available to academic institutions. These databases are a collection of decades of drug discovery research across academia and industry globally. As a result, these databases are of the order of a few million to tens of millions of compounds spanning a diverse range of chemical space, much of which will be irrelevant to any given target of interest. Most databases include some data relating to each hit such as biological assay information, physicochemical properties and publication references. The GVK[101] database is an example of a subscription database while ZINC[102] is freely available.

ZINC is the largest free to use database of drug-like compounds, containing contributions from the entire drug discovery community. The database is pre-divided into many subsets and the website contains many useful filtering tools to customise a set of compounds tailor-made for the target before downloading.

Handling and analysing these high-volume databases present their own challenges and has given rise to the field of Chemoinformatics. To analyse compounds in the databases, chemical structures for each molecule need to be entered in a format with low memory requirements that can easily be read. For this purpose the simplified molecular input line entry system (SMILES)[103] language was invented. SMILES was devised to describe molecules in their entirety with a string of standard input characters. For example, all the atoms are represented by their letter, but bonds are specified with numbers and stereochemistry is specified with the @ symbol. The advantage being that each molecule can be described with a one-line string which can easily fit on spreadsheets while certain motifs can be quickly identified from a string within the SMILES. This enables a wide variety of filtering and manipulation programs to be applied to the database reducing it to a more manageable size and streamlining it towards the relevant chemical space.

2.1.8 Database Filtering

Filtering often begins with the Lipinski 'Rule of 5,' a set of physicochemical parameters under which compounds are likely to make a good oral drug.[104] These parameters are a MW ≤ 500 Da, at most 10 H-bond acceptors, 5 H-bond donors and ClogP (measure of hydrophobicity) ≤ 5. While these are not strict rules to follow, they do serve as useful guidelines to filter databases for molecules with good drug-like properties.

Studying the target and its known ligands can help establish a set of secondary parameters. These may include the inclusion or exclusion of specific chemical groups, the size of a fused aromatic region or the number of rotatable bonds. Once filtered down to a manageable size, the database is docked against the target. The top poses are analysed and a cut-off docking score is decided upon based on the point at which the predicted poses become improbable. The next step is to further reduce the size of the hit compound set by visual inspections; removing ligands with an unfavourable docking score due to an improbable pose or conformation. Once the final set has been sufficiently reduced, a representative set is usually chosen for purchase and tested *in vitro* against the target.

2.1.9 Fragment Screening

A fragment screen is an adaptation of a traditional virtual screening protocol based on a few simple assumptions. Smaller molecules have fewer rotational degrees of freedom and therefore less conformational space to explore. The predicted poses from smaller molecules are therefore theoretically more accurate. A smaller molecule also leaves more room for synthetic optimisation at a later stage. Lastly, fragment screening exercises are computationally cheaper and therefore quicker to run. Instead of using the Lipinski 'rule of five' parameters, fragment screens use Congreve's rule of 3[105] as a guideline. These parameters are MW ≤ 300, CLogP ≤ 3 and at most 3 H-bond donors or

acceptors. Since the fragments are small, fragment screening sets will usually be selected based on a few defining chemical features. Once a set of fragments is selected for purchase, they are usually tested at higher concentrations with *in vitro* inhibition assays (500-1000 μM).

A fragment screening method would be well suited to finding a domain selective ACE inhibitor. The metal binding introduces a degree of uncertainty, which can be mitigated by smaller ligands. Absolute domain selectivity could be hard to come by with smaller ligands. A fragment may nevertheless provide useful information for selective drug design.

2.2 Aims and Objectives

2.2.1 Aim

The aim of this chapter was to screen a set of fragment compounds for potential leads on N-domain selective ACE inhibition.

2.2.2 Objectives

- Build a set of metal chelating drug fragments from the ZINC database.
- Perform constrained docking on this set against the N- and C-domains of ACE.
- Analyse and visually inspect the docking outputs for this fragment set against both domains.
- Screen a representative set of compounds for ACE inhibition *in vitro*.

2.3 Methods

2.3.1 Database Filtering

A chelating fragment database was created using the chelating fragment library (CFL) (Figure 2.2) devised by Agrawal et al[106] from the study of matrix metalloproteinase (MMP) inhibitors and the ZINC database. A substructure search was performed on a ZINC database subset of 1.2 million fragments titled “Fragment Leads Now” which by and large fall within the Congreve’s ‘rule of 3’ parameters. Using the ZINC substructure search function, substructure searches were performed using all 96 chelating fragments from the CFL against the Fragment Leads Now subset to find molecules containing these motifs.

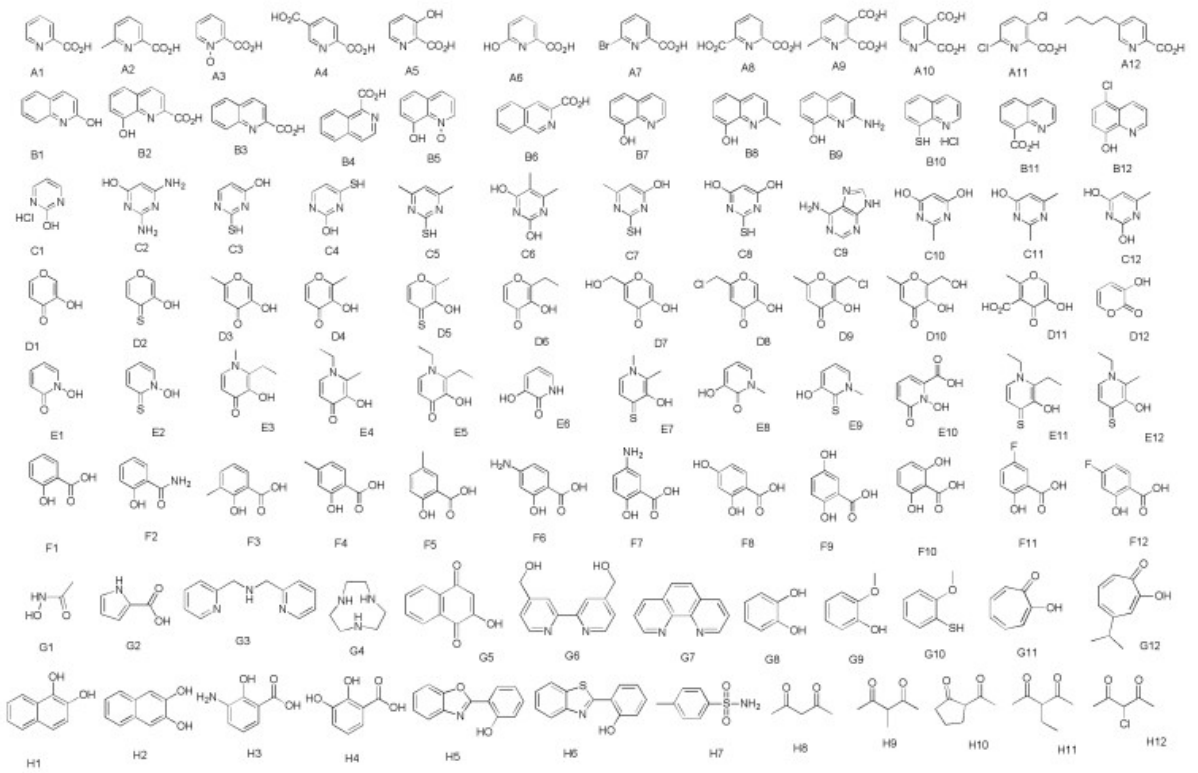


Figure 2.2: *The metal binding functional groups of the CFL.*

2.3.2 Protein Preparation

A high-resolution crystal structure for both the N- and C-domains of ACE was selected for docking simulations. The PDB crystal structures 3NXQ and 1O86 were selected to model the respective N- and C-domains. Each structure was prepared for docking according to the following protein preparation protocol using Maestro's protein preparation tool, PrepWizard (Schrödinger Suite 14.0) with the following settings applied. PyMOL 1.8 was used to generate all images containing protein ribbons from crystal and docked structures.

C-domain Steps:

1. Under PrepWizard's import and process tab, the Preprocess protocol was run on the 1O86 PDB file with the following boxes ticked; assign bond orders, add hydrogens, create zero order bonds to metals, fill in missing side-chains using prime and delete waters beyond 5 Å from a het group.
2. In the Review and modify tab, the Analyze Workspace button was selected and the structure was manually inspected. The structure of Lisinopril was manually corrected and the generate states function was run to generate the correct metal binding and ionisation state of the ligand.

3. Under the refine tab, H-bond assignment was run with just the sample water orientations box ticked. The remove waters option was run for waters with less than 3 H-bonds to non-waters. Lastly a restrained minimisation was run with the heavy atoms set to converge on the default 0.30 Å using the OPLS_2005 force field.

N-domain Steps:

1. The PDB file 3NXQ was imported and the Preprocess protocol was run in a manner identical to the C-domain.
2. The 3NXQ file contains two repeats of the N-domain structure in its asymmetric unit. Chain B, the less complete copy of the protein was deleted. As before the metal binding states were then generated.
3. As before the H-bonds were assigned, the waters removed and the structure minimised.

From these two prepared structures, docking grids were created. The docking grid was centred on the Zn. The size of the docking grid was adjusted to accommodate ligands of a maximum length of 12 Å, a length estimated to be the maximum theoretical length of a fragment. Lastly a docking constraint for a metal chelating interaction was set.

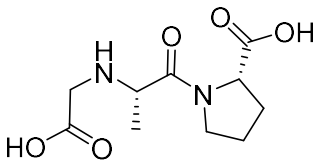
2.3.2 Ligand Preparations

The next step was to prepare ligands in the fragment database. The substructure filtered database was downloaded in the 2D sdf format. The Maestro LigPrep protocol (Schrödinger Suite 14.0) was run on this ligand set rendering the 3D structures for each ligand in the most likely ionisation and tautomerisation state under biological conditions. To achieve this, the pH was set at 7.0 ± 2.0 allowing for multiple tautomers. This step created a few duplicate entries but the acids were mostly deprotonated while the bases were mostly protonated.

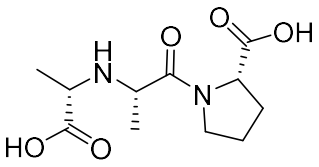
2.3.3 Validation and Execution of Docking Protocol

Before docking the fragment set, the protocol was validated against a set of fragments derived from known ACEis to demonstrate a correlation between the known ligand-enzyme interactions and the docking score. After an examination of the set of ACEis, a few fragments were created through the deconstruction of the Enalaprilat (E1, E2 and E3) and **RXP407** (R1) into fragments (Table 2.3). The dummy fragment (D1) was introduced to investigate the impact of the Zn binding disruption. These fragments were docked using the Glide (Schrödinger Suite 14.0) standard precision (SP) settings with default settings. After validation, the prepared set of compounds were docked against both targets with the settings unchanged.

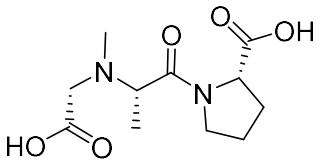
Table 2.3: *The four fragments used for benchmarking the fragment screening protocol.*



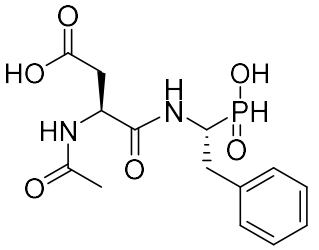
E1



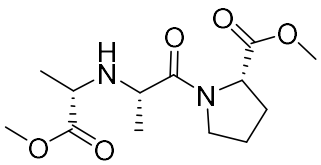
E2



E3



R1



D1

The docking results were then analysed. Due to the fragment nature of these ligands, ligand efficiency (LE), the docking score of the ligand divided by its heavy atom count (HAC), was chosen as the metric for ranking docking poses. A LE cut-off where the poses were deemed nonsensical was chosen. All ligands with a LE above the cut-off were discarded.

2.3.4 Visual Inspection

During the visual inspection, the set of remaining ligands was overlaid with the space known to be occupied by cocrystalised inhibitors of ACE. All compounds displaying unlikely poses or occupying a space within the active site where ligands are not known to bind were eliminated.

The second round of visual inspection entailed closer scrutiny of the space occupied by each fragment. All ligands found to occupy space outside the confines of the known binding site were eliminated. Special priority was given to ligands interacting favourably with Arg381 from the N-domain.

The remaining compounds were investigated for commercial availability. Often compounds are listed in these databases without regularly updating its availability status. This achieved another reduction in the set of remaining compounds. In order to reduce cost and maximise the chemical space of the

remaining set, the remaining compounds were clustered using an extended connectivity fingerprint (ECFP4) algorithm.[107] A chemically representative set of compounds spanning all the clusters was chosen and purchased for *in vitro* screening.

2.3.5 Inhibition Assays

The 60 compounds (Appendix 2.1) were tested *in vitro* in a fluorogenic competitive inhibition assay. For the C-domain the tACE Δ36NJ construct was used while the Ndom389 construct was used for the N-domain. The N- and C-domain samples were stored frozen in stock solutions of 14.0 µM and 7.3 µM respectively. The N- and C-domain stock solutions were diluted to 10 nM and 5 nM respectively, double the previously determined assay optimised enzyme concentrations. The 60 compounds were each dissolved in DMSO to make stock solutions of 50 mM. Aliquots of stock solutions were diluted to 10 mM with deionised H_2O followed by dilution into a phosphate buffer (100 mM $KHPO_4$, pH 8.3, 300 mM NaCl, 10 µM $ZnSO_4$, 1mg/ml albumin). All inhibitor-buffer solutions were calculated to have a DMSO concentration of no more than 0.5% when incubated with the prepared enzyme-buffer solutions.

To begin with, a broad inhibition screen at 500 µM for each of the compounds was performed. Each 10 mM stock solution was diluted down to 1 mM in buffer. 40 µL of the diluted compound was mixed with 40 µL of the enzyme in buffer halving the concentration of both solutions. This mixture was incubated at room temperature for 30 minutes then triplicate 20 µL aliquots were added into 3 individual wells of a 96 well plate. 30 µL of 1 mM Z-FHL substrate (Bachem Ltd., Bubenhof, Switzerland) was then added to each well. Each 96-well plate experiment contained a negative control with the inhibitor solution substituted for buffer only, a positive control with 10 µM of Lisinopril and a blank zero time (bzt) to measure the background using just buffer. The plate was then incubated at 40 ˚C for 15 minutes in a shaker. The reactions were then stopped and the product was derivatised with the addition of 190 µL of a base and *O*-pthaldehyde solution (0.28 M NaOH and 7 mM *O*-pthaldehyde) to each well. The plate was then incubated on a shaker at room temperature for 10 minutes. The wells were quenched with 25 µL of 3 M HCl and the plate was read in the fluorimeter (Varian Inc., Mulgrave, Victoria, Australia) with an excitation wavelength at 360 nm and an emission wavelength at 485 nm. Change in fluorescence correlated to nmol HL using linear regression analysis and activity expressed as nmol/ml/min or nmol/mg/min. Compounds displaying more than 50% inhibition at 500 µM were selected for a second round of inhibition assays at lower concentrations to approximate their IC_{50} values.

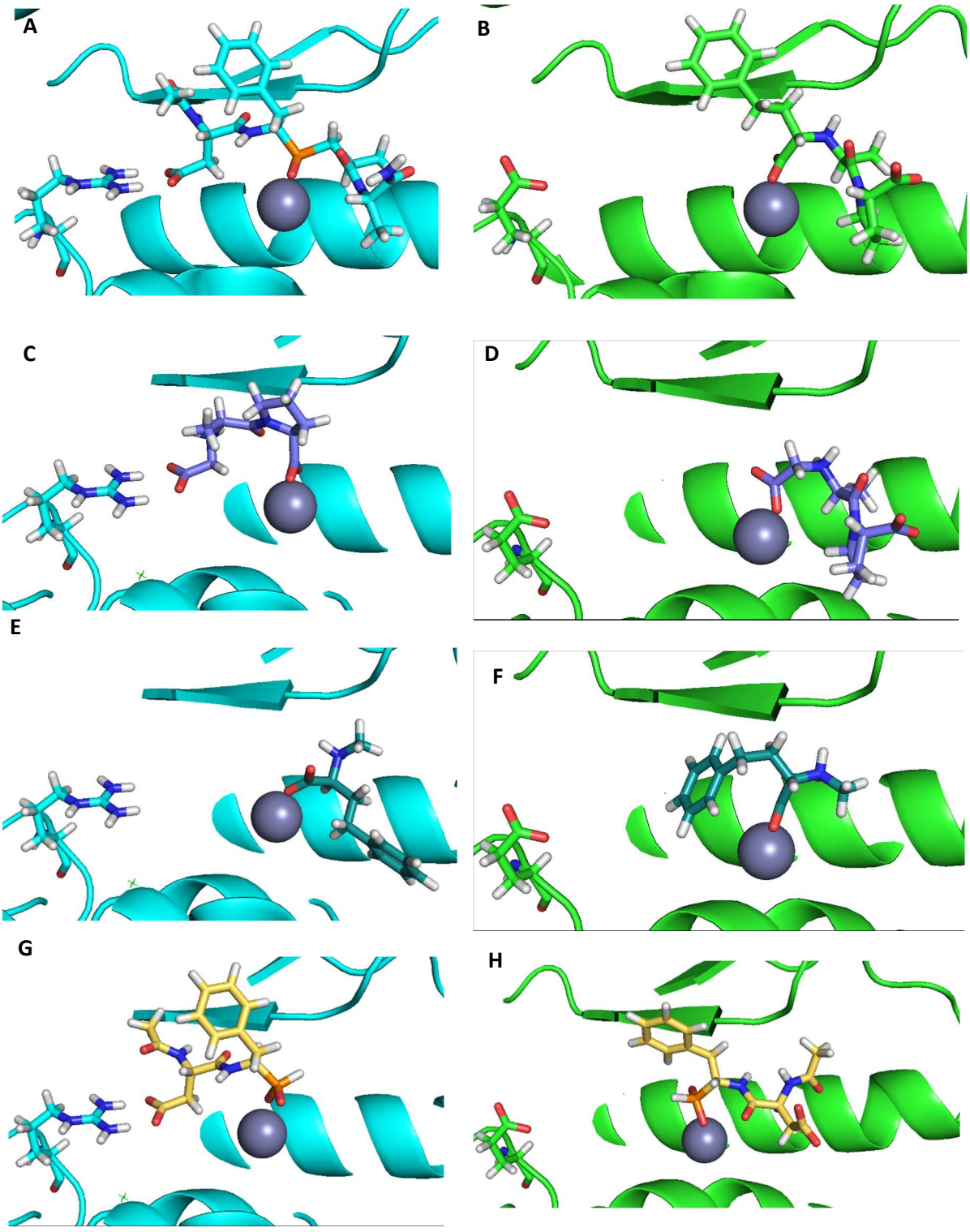
2.4 Results

2.4.1 Fragment Substructure Search

Substructure searches were run on each chelating fragment from the CFL against the Fragments Now subset of the ZINC database using the substructure tool on the ZINC website. Approximately 16 000 unique fragments were returned from these cumulative searches. The Initial inspection of this set showed a wide chemical diversity and a fair representation across the entire range of chelating groups listed in the CFL. Since 16 000 compounds is a manageable size for docking, no additional steps were necessary to filter this set before any further docking.

2.4.2 Fragment Docking Protocol Validation

All four ligands assigned the correct ZBG to the Zn metal (Figure 2.3) when docked with just a metal binding constraint. Fragments E1 and E2 were placed in the respective S_1' and S_2' subsites in an almost identical position to the pose of their parent crystal structure. The phenyl ring of E3 was placed in the correct S_1 pocket of the C-domain but was flipped around and placed in the S_1' pocket of the N-domain. Fragment R1 assigned the correct ZBG in both domains. The important salt bridge between its acid and Arg381 of the N-domain was also predicted. The fragment did, however, flip around in the C-domain when it encountered the Glu403. The dummy compound D1, which is E3 with methyl ester caps on the two acids managed to reproduce the poses observed for E1 in both domains despite the disruption of the Zn binding.


A
B
C
D
E
F
G
H

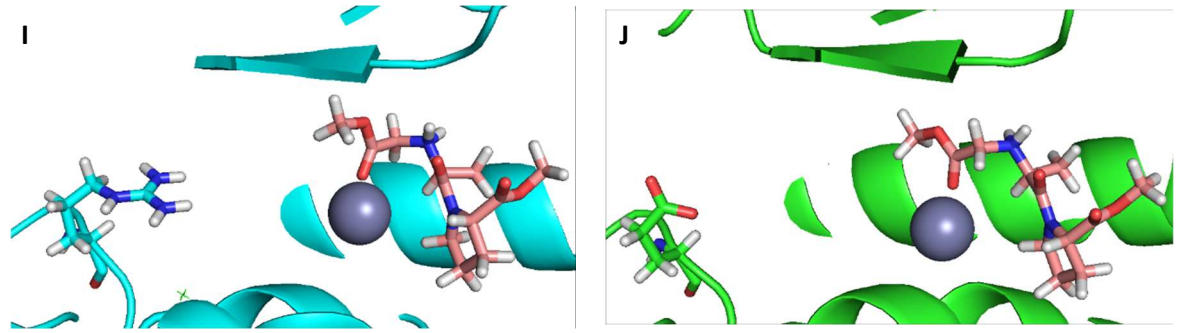


***Figure 2.3**: Docked poses of the benchmark fragments. 3A and 3B show the respective crystal poses of RXP407 and Enalaprilat, the parent compounds from which the benchmark fragments were derived in the respective N- and C-domains. 3C shows the fragment E1 docked in the N-domain. Here this Enalaprilat Ala-Pro scaffold coordinates with the Zn using its C-terminal carboxylic acid to facilitate a salt bridge with Arg381. In 3D this fragment returns to the more familiar prime side of the active site occupying a similar position to its parent Enalaprilat molecule. 3E shows fragment E3 docked into the N-domain with its phenyl group occupying the incorrect S_1' subsite unlike the C-domain (3F) where the fragment has swung around into correct S_1' subsite. 3E shows R1 docked into the N- domain with its phenyl group binding to the correct S_1' subsite to allow for the formation of a salt bridge with Arg381. In the C-domain (3H), the entire molecule swings around to keep its acid away from Glu403. Fragment D1 docked in almost an identical conformation in both the N- and C-domain (I and J respectively) despite the capping of its two acid groups with a methyl ester.*

The LE of these benchmark fragments is shown in Table 2.4. Fragments E1 and E2 fit snugly within the 2' side subsite poses similar to Enalapril and most ACEis. These poses are coupled with excellent LEs (<-0.5), thus validating the docking protocol against these fragments, which would be expected to these binding sites with a strong inhibtion given their small sizes. In the case of E3, the phenyl group finds the incorrect S_1' subsite of the N-domain while binding to the correct S_1 subsite of the C-domain. The failure of this fragment to find the correct binding pose in the absence of the Ala-Pro group emphasises its importance for ACE binding. The **RXP407** fragment made an important salt bridge interaction with Arg381 in the N-domain while flipping around in the C-domain to avoid Glu403. This pose correlated with a superior LE towards the N-domain. The effect of interrupting the Zn binding through an esterification of the carboxylic acid group was quantified with fragment D1. This fragment returned the correct ACE binding pose with LEs above -0.3 for both domains suggesting -0.3 to be a reasonable LE cut-off for this system.

Table 2.4: *LE of the benchmark ligands*

Ligand	N-domain LE	C-Domain LE
E1	-0.570	- 0. 611
E2	-0.504	-0.569
E3	-0.560	-0.634
R1	-0. 458	-0.350
D1	-0.221	-0.275

With all the predicted poses reproduced using these ACEi derived fragments, this fragment docking protocol was successfully validated. With the protocol validated and the domain selective fragment binding patterns illustrated, the entire fragment set was docked.

2.4.3 Fragment Set Docking

The metal binding moieties from the CFL were all represented in this set. These moieties were all easily recognised by Glide as a metal binding group. A set of 16 000 ligands can be docked using the Glide SP level of precision in a reasonable timeframe but a set this size is too large for Glide XP. Another point for consideration is that the improved accuracy of Glide XP on small ligands would be negligible.

While inspecting the docking output, the LE cut-off of -0.30 was chosen as an arbitrary point where the observed ligand poses became implausible. It was around this LE cut-off that the fragment poses became implausible and unrealistic in a similar manner to the benchmark fragments. 3500 Compounds remained after the cut-off was enforced, a set small enough to be manually inspected.

2.4.4 Visual Inspection

Two rounds of visual inspection were conducted. The first round of visual inspections focused on eliminating improbable poses and interactions. These poses predominantly contained improper metal binding and H-bonding interactions. Examples of these interactions include incorrect H-bond donor/acceptor pairings, polar groups in the hydrophobic regions and hydrophobic groups blocking protein donor/acceptor groups. As a bare minimum, each ligand needed to have a good metal chelating interaction and at least one H-bond with the target. This round of visual inspection reduced the set of remaining compounds down to 878 compounds.

The second round of visual inspection was more rigorous due to the smaller set while closely focusing on the space occupied by the cocrystallised ligands of ACE. All the ligands cocrystallised with ACE occupy a specific space or footprint. Priority was given to compounds with a favourable Arg381 interaction. Ligands in this set, which fell outside this observed crystal footprint were discarded. Figure 2.4 shows an example of both an accepted ligand and a rejected one. Figure 2.4A shows a ligand (yellow) where half the molecule fell outside the established ACE ligand footprint occupied by its native ligand, **RXP407** (cyan). In contrast, **SF37** (Figure 2.4B) passed this round of visual inspection due to its occupation of this footprint and its salt bridge with Arg381 of the N-domain.

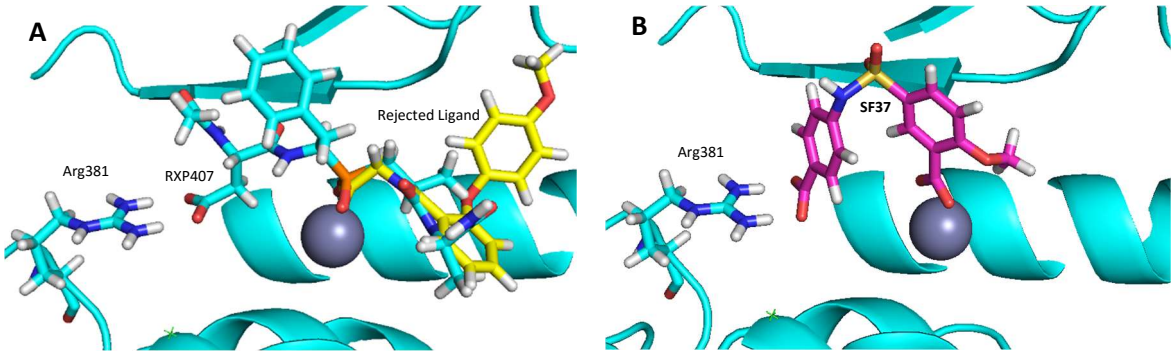


Figure 2.4: *An illustration of a rejected and an accepted docked fragment pose.* ***A****. An example of a fragment compound which was rejected during the visual inspection step of the virtual screening. This compound (yellow) was rejected on the basis that it occupied a space outside the confines of binding site defined by its native ligands like* ***RXP407****.* ***B****. An example of a ligand which passed the visual inspection (****SF37****) with both a prominent Zn chelator and a salt-bridge between the ligand and the N-domain selectivity inducing Arg381.*

Following two rounds of visual inspection, 350 compounds remained. Of the remaining 350 compounds, only 120 were available for purchase.

2.4.5 Clustering

To maximise the remaining 120 commercially available compounds, the ECPF4 clustering algorithm was applied. This sorted the 120 compounds into 24 clusters. A final set of 60 compounds (Appendix 2.1) was selected with each cluster represented. Special consideration was then given to compounds with a potential to interact with the Arg381/Glu403 residues.

2.4.6 Fragment Screen Overview

Figure 2.5 summarises the fragment screening protocol.

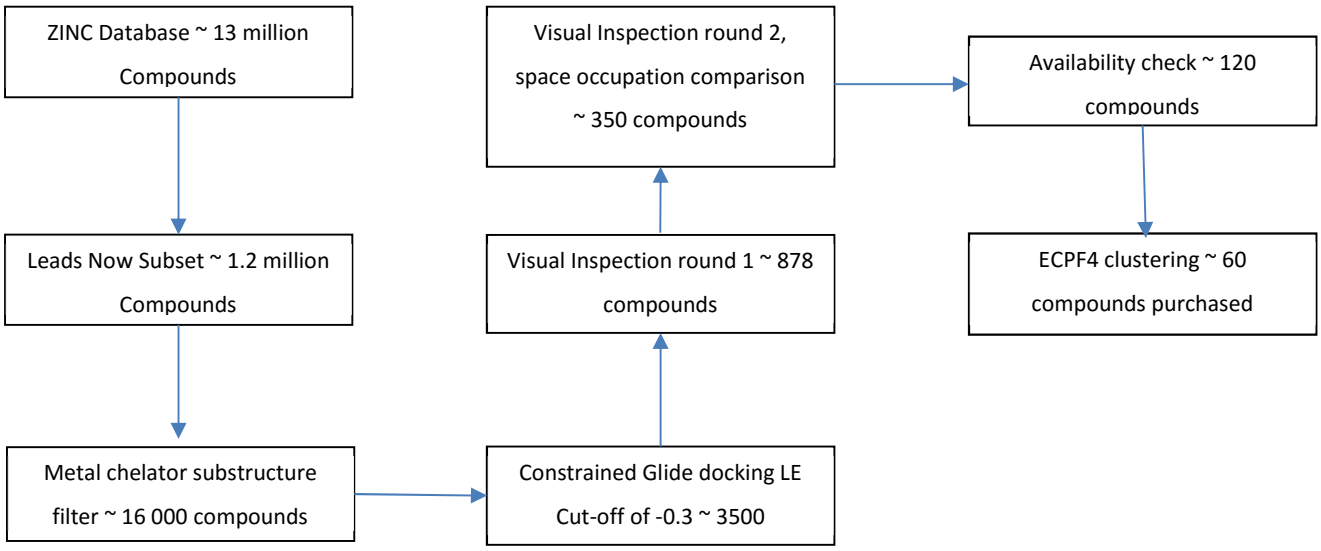
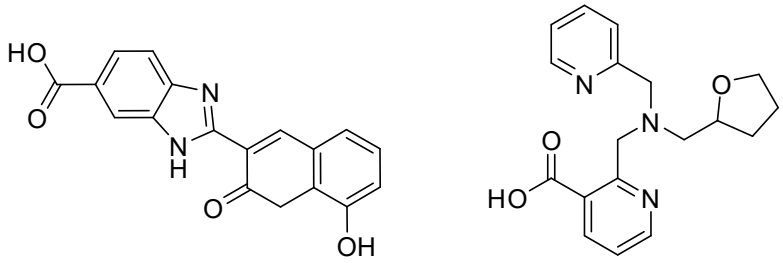


Figure 2.5: *Summary of the implemented docking procedure.*

2.4.7 Inhibition Assays

Of the 60 compounds tested (Appendix 2.1), only two displayed more than 50% inhibition of the total enzyme activity at 500 µM. These were compounds **SF43** and **SF49** (Figure 2.6).



SF43 **SF49**

Figure 2.6*: Structure of Fragments* ***SF43*** *and* ***SF49***

Compounds **SF43** and **SF49** were then tested at lower concentrations (Figure 2.7). When both inhibitor concentrations were reduced to 100 µM, the enzyme returned to full activity. **SF43** exhibits the typical dose response behaviour at these high concentrations with the almost full inhibition observed at 500 µM with the enzyme returning to full activity at 100 µM concentrations. **SF49** on the other hand displayed a slight inhibition of both domains at 500 µM with the enzyme returning to close to full activity at 100 µM. Both **SF43** and **SF49** contain carboxylic acid ZBGs and chemical groups novel to ACE inhibition while neither compound displayed any selectivity towards the N-domain.

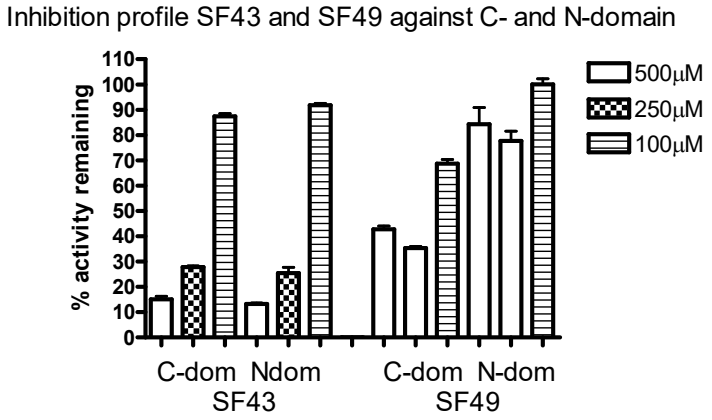


***Figure 2.7**: Inhibition of **SF43** and **SF49** against the N-and C-domains of ACE. These compounds were tested at concentrations of 500 μM, 250 μM and 100 μM. Both enzymes returned to full activity when inhibitor concentrations were dropped below 100 μM.*

2.5 Discussion

2.5.1 Hit Molecules

This fragment screen returned two hits out of 60 compounds screened. A hit rate of 3% for a fragment screen is about par for the course for this type of screen. It is also possible that screening at 1 mM may have brought more fragments to attention. However, with the existence of such potent ACEis, it would be misleading to classify such fragments as hits. Both the hit compounds share some chemical similarity with the current set of known ACEis such as a carboxylic acid ZBG and hydrophobic side-chain mimicking groups such as a the pyridyl group in **SF49**

While **SF43** and **SF49** may be potential leads for designing novel inhibitors from scratch, neither compound showed any domain selectivity with IC_{50}s 4-5 orders of magnitude weaker than the current set of ACEis. Despite returning two potential hit compounds, this exercise has not furthered the cause of finding a potential lead within the context of domain selective ACE inhibition.

The reason for the lack of domain selectivity among **SF43** and **SF49** is obvious when considering their docked poses (Figure 2.8). Neither compound is capable of forming a salt bridge with the Arg381 of the N-domain nor an unfavourable interaction with Glu403 of the C-domain. Adding a carboxylic acid to the biaryl system of **SF43** might have helped. While both these compounds contain similar core elements required for a strong ACEi, such as a metal chelator and peptide-like branching, they both lack sufficient H-bond donors and acceptors to rival established ACEis. Fragments are deliberately small with fewer chemical groups and interact with a small portion of the enzyme active site.

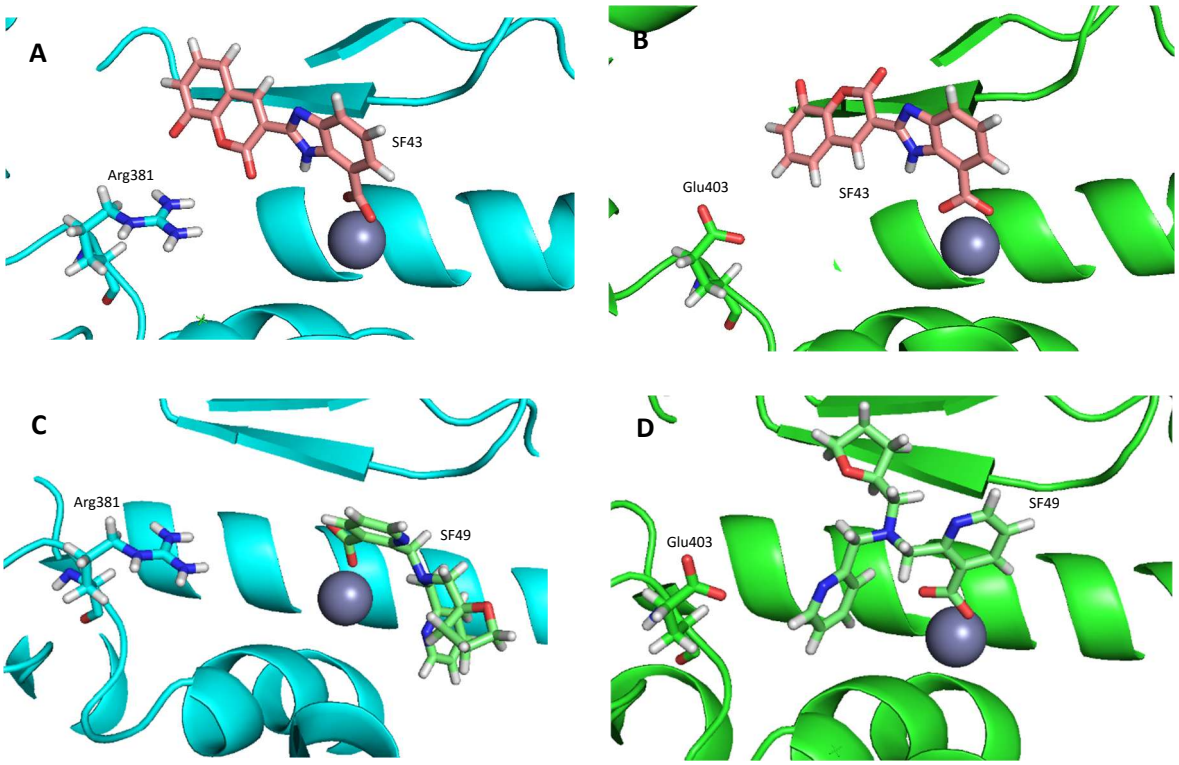


***Figure 2.8**: The docked poses of **SF43** (beige) and **SF49** (lime) in both the ACE N- (cyan) and C- domain (green). A. **SF43** docked in the ACE N-domain with its large heteroaromatic group occupying the S_2 pocket close to the region of the Arg381. B. **SF49** docked into the C-domain in an almost identical pose to the one seen in the N-domain. C. **SF49** docked in the ACE N-domain with its THF and pyridine groups sitting comfortably in the S_1' and S_2' pockets respectively. D. **SF49** docked into the ACE C-domain with the THF and pyridine groups occupying the S_1 and S_2' pockets respectively.*

2.5.2 Shortcomings

The failure of this approach to provide potential domain selective leads on domain selectivity highlights the difficulty in applying VS techniques to problems relating to target selectivity. VS relies on the brute force of numbers to help find a hit from a huge screening set. In the case of ACE, a potential domain selective lead hinges on one or two key interactions. VS by its very nature has a low hit rate. Typically, under 5% of compounds tested display activity against the target. Expecting one or two highly specific interactions from a hit was possibly too optimistic.

As shown during the benchmark docking, looking for a salt bridge between the ligand and the Arg381 in the N-domain is the only viable way to pick up on a possible N-domain selective compound using this model. Since the different domains represent distinct protein-binding systems, GlideScore and LE cannot be compared between the two domains for selectivity. Such an interaction requires a diacidic fragment with one acid binding to the Zn while the other remains free to interact with Arg381/Glu403. Diacids are rare in fragment screening databases as fragments usually focus on a few key functional groups while keeping the fragments within the bounds of strict physicochemical parameters. Too many charged groups on a fragment sized molecule would make it very polar resulting in poor cell permeability and a high LogP. For this reason, it is unlikely that many diacidic fragments would be

listed in a fragment screening database. Examining the final set of compounds screened, only two diacids were present (**SF2** and **SF37;** Figure 2.9). Two acids on each molecule should have been a condition on all compounds purchased.

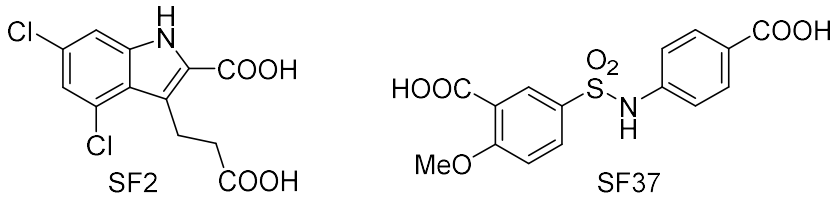


Figure 2.9**: Fragments **SF2** and **SF37

2.5.3 Conclusion

In retrospect, it appears that the conditions required for an N-domain selective inhibitor were improbable from the outset. This outcome changed the direction and focus of the project. The existing ACEis are much more potent than what could be expected from any fragment hit. At this point it became clear that a much more targeted approach than VS was required. A decision was made to pursue a *de novo* approach starting with existing ACEis rather than reinventing the wheel by searching for novel leads from a random set.

Chapter 3 – Enalaprilat Analogue *de novo* Design

3.1 Introduction

3.1.1 Background

Following the failure to find a novel N-domain selective inhibitor with a fragment screen in Chapter 2, a different approach was required to achieve this objective. As explained, the structural requirements of an N-domain selective ACE inhibitor appeared to be too specific to find in a set of 16 000 metal-chelating fragments. To improve upon the fragment screen, a new approach was required to find N-domain selective ACEis. This new approach focussed on the existing set of diverse ACEis which were not utilised during the fragment screen.

The SAR accumulated over four decades of ACEi development has fortuitously left behind a vast dataset of tens of thousands of compounds with varying degrees of ACE inhibition. The SAR accumulated over this period includes a large set of inhibitors displaying low to sub nanomolar inhibition. These inhibitors are several orders of magnitude stronger than can be expected from any fragment-based screening hit.

The extensive structural and inhibition data of ACE is ideally suited for *de novo* drug design methods. In a drug discovery context *de novo* design refers to the practice of starting from drug fragments or scaffolds and exploring the SAR required to build them into full sized drug molecules[108]. While the shortcomings of a fragment-based approach focused on a highly specific interaction were discussed in Chapter 2, some of the methodology can still be adapted to feed into certain *de novo* drug design protocols.

De novo drug design is a broad term encompassing a wide range of drug discovery techniques. These techniques are varied and have been developed to suit a variety of different systems. The most common *de novo* drug design techniques include methods based on (1) structural alignment, (2) molecular force-fields and docking, (3) fragment assembly, and (4) retrosynthesis and plausible reactions. As the names of these techniques suggest, optimising these fragments requires either knowledge of an inhibition pharmacophore or protein target as well as a basic understanding of the synthetic viability of the proposed structures. Like VS, *de novo* drug design is prone to the same problem of combinatorial explosion. It is therefore no surprise that many of tricks from VS are carried over into this discipline.

3.1.2 De Novo Drug Design Approaches

Alignment-based methods (Figure 3.1) require a large library of crystal structures of the target in complex with different ligands. Aligning the ligands within the same binding site can highlight the

space in the binding pocket not occupied by known inhibitors. It can also help link inhibitors to additional functional groups to expand its binding footprint within the binding site[109]. Alternatively, forcefield[110] and docking-based methods rely on creating combinatorial libraries of designed compounds which are then ranked based on docking experiments and other force-field-based energy calculations.

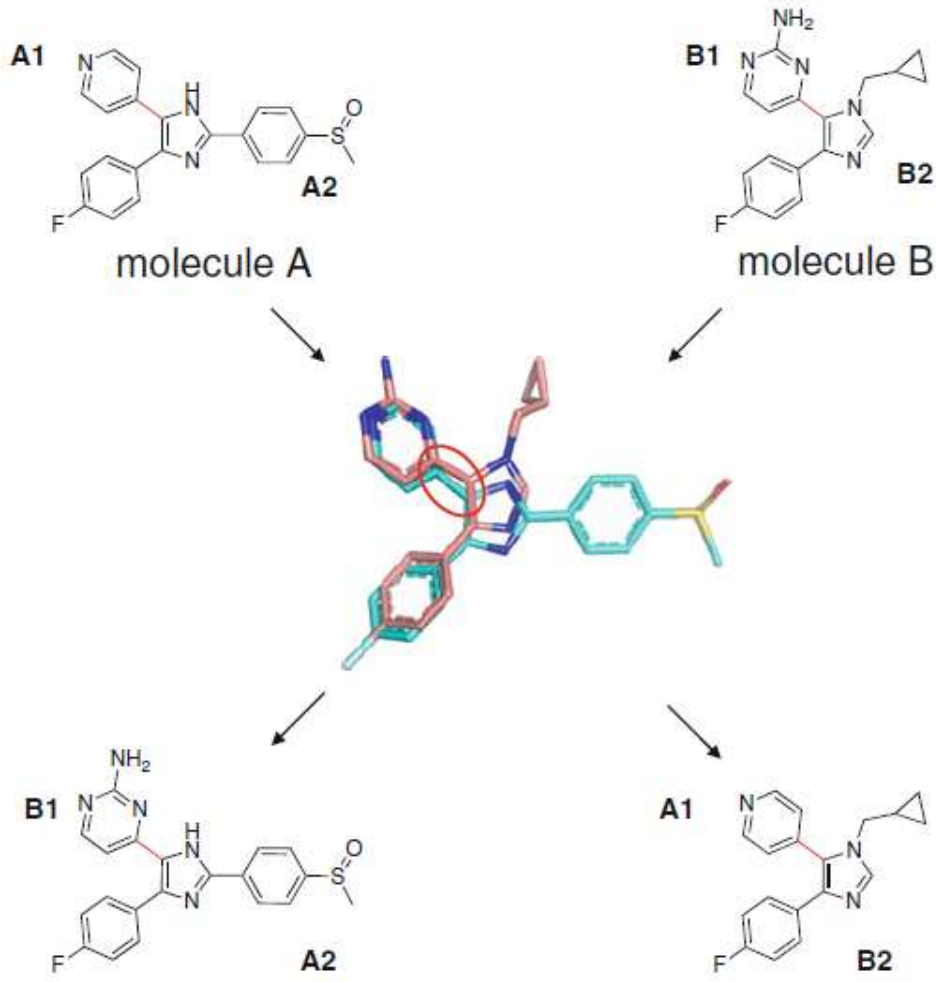


Figure 3.1*: An illustration of alignment-based de novo drug design. Two different inhibitors (molecule A and B) from the same target are aligned and their spatial occupation is compared.The key features of the two drugs are then combined around the circled bond to create two alignment combintions which maximise the target binding space.*

Fragment assembly methods are statistically-based algorithms, which cycle through large databases of fragments and connectivity options to generate vast libraries of designed molecules within a set of pre-specified parameters[111]. Retrosynthetic chemistry and reaction-based approaches involve the deconstruction of inhibitors into core fragments via retrosynthesis and identifying biologically active core compounds, which can be developed into novel potent inhibitors[112]. Such protocols allow one to remain mindful of the synthetic viability of proposed compounds.

3.1.3 Synthetic Chemistry Considerations

Theorising designed molecules is only helpful if one considers the synthetic viability of the suggested compounds. Examining the set of ACEis and their co-crystallised poses within the binding site of ACE, it can be deduced that Enalaprilat and Lisinopril (Figure 3.2) are the simplest ACEis that bind within the active site in reach of the P_2 Arg381/Glu403.

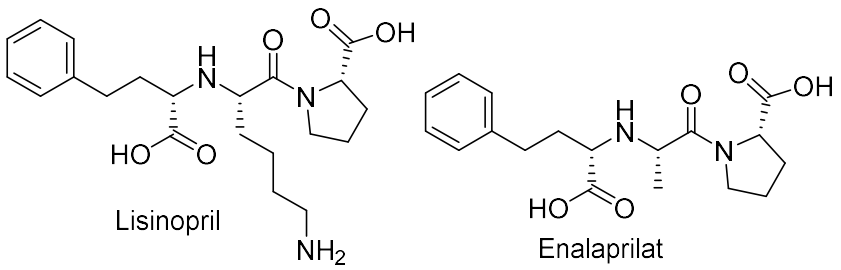


Figure 3.2: *The structures of Enalaprilat and Lisinopril*

Enalaprilat and Lisinopril were first synthesised via the reductive amination of a C-terminal proline dipeptide and an α-keto acid using a cyanoborohydride reducing agent (Figure 3.3). This reaction creates a third chiral centre in the molecule but is not diastereoselective. The two diastereomers were then separated via column chromatography. Inspecting this reaction helps to deconstruct these ACEis into the core scaffolds for *de novo* drug design. As previously discussed, the core scaffold of an ACEi consists of a central ZBG, a P_2' terminal carboxylic acid and a pseudopeptidic structure in the P_1' and P_2' positions. The most common P_1'-P_2' moieties are a pseudo Ala-Pro structure but other combinations were also investigated.

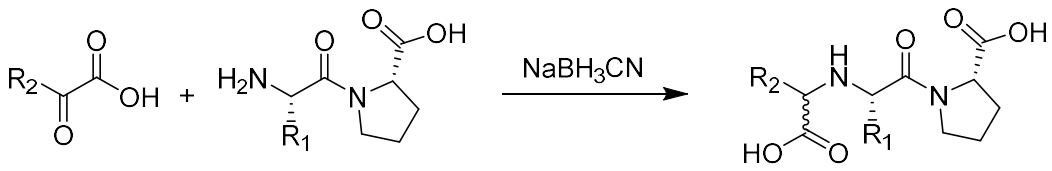
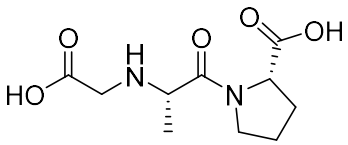
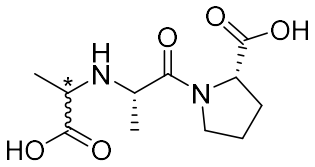
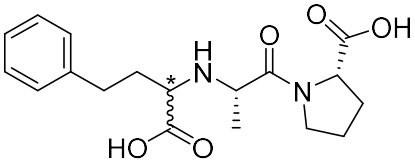
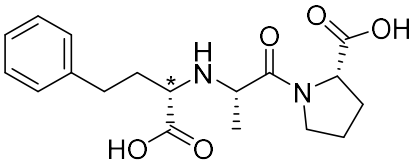


Figure 3.3*: The reaction used by Patchett et al[113] in ACEi synthesis to attach P_1 groups.*

Most of Patchett's work investigated the SAR around the P_1 substituent on the Ala-Pro scaffold where extensive testing was performed after introducing the chiral centre indicated in Table 3.1.[113] When comparing the merits of using Lisinopril and Enalaprilat as core scaffolds, the lysine moiety of Lisinopril greatly increases the MW of the compound while adding an amine which would need to be protected in many potential synthetic manipulations. The Ala-Pro backbone of Enalaprilat has a lower MW and the P_1' methyl group does not need any protection. This makes Enalaprilat a better candidate for synthetically exploring the SAR around the P_2 substituent

Table 3.1: *The SAR data of selected compounds tested by Patchett et al[113] examining the effect on inhibition caused by the introduction of a third chiral centre (*) and a P_1 phenyl group to the Ala-Pro backbone.*

Compound	Structure	sACE IC_{50} (nM)
1		2400
2		90
6		3.8
6a		1.2

3.1.4 Combinatorial Library Screening

Combinatorial library screening follows a similar methodology to VS. The major difference being that the library screened is of hypothetical molecules generated to deliberately probe specific SAR. Many docking programmes have a feature enabling the user to generate combinatorial libraries to feed into their docking protocol. In the case of Glide, there is an extension called CombiGlide that can be used to generate combinatorial libraries and feed them into Glide docking. It only requires input fragments and an input core structure. The points of attachment on both the core structure and additional fragment can then be specified with optional alkyl linkers of variable lengths. CombiGlide then generates all the possible permutations for the specified points of attachment and the alkyl linker length between the scaffold and the attached group. Once the library is generated, it is docked just like any database of real drug-like compounds.

Combinatorial library screening methodologies differ from traditional VS methodologies in the ranking step. Docking scores are designed to rank the viability of a pose above a certain cut-off. In this manner

they serve as an aid to help identify hit molecules from a diverse screening set. When investigating the virtual SAR (VSAR) in a particular region, one looks for subtle changes in binding energy. These subtle changes add a new quantitative aspect to docking protocols not seen in VS. To help quantify these subtle changes in binding energy, binding energy calculations have been developed.

3.1.5 Rescoring and Energy of Binding Calculations

Docking scores are used as a tool to qualitatively evaluate a predicted binding pose but they correlate poorly with experimentally determined ligand binding energies. For the purpose of predicting ligand binding energies with an improved experimental correlation, the more computationally demanding molecular mechanics Poisson-Boltzmann Surface Area[114] (MM-PBSA) and molecular mechanics Generalised Born Surface Area[115] (MM-GBSA) calculation algorithms were developed. Both these approximations begin with the same approach. The binding energy (ΔG_{bind}) can be approximated with equation 3.1. Here the enthalpy term is equated to the sum of molecular mechanical energy within the system (ΔE_{MM}) and a new term ΔG_{solv} represents the energy change created by adding solvent to the gas phase system under which the calculations are performed. ΔE_{MM} is the sum of the molecular mechanical changes to the system in gas phase. It is described using equation 3.2.

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (3.1)$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdw} \quad (3.2)$$

The change in energy brought about by solvating a gas phase system, ΔG_{solv} needs to be calculated by computing the difference in solvation energy between the free ligand and the protein target as described in equation 3.3. Each ΔG_{solv} term is then approximated with the PBSA or GBSA method (equation 3.4). This equation has two terms, one for the energy change originating from a change in polarisation energy ($\Delta G_{PB/GB}$) and the other term covering the non-polar contribution to the change in energy (ΔG_{SA}). These two methods use the same non-polar term but different polarisation terms.

$$\Delta G_{solv} = \Delta G_{solv}^{complex} - (\Delta G_{solv}^{Protein} + \Delta G_{solv}^{ligand}) \quad (3.3)$$

$$\Delta G_{PBSA/GBSA} = \Delta G_{PB/GB} + \Delta G_{SA} \quad (3.4)$$

Benchmarking comparisons between the two algorithms have shown that while the PBSA method is more computationally demanding than GBSA, it is also more accurate[116]. Despite this, GBSA energy calculations are still useful for comparing a homologous set of ligands. GBSA however, does not cope as well with highly charged or polar species. Both these methods have only been recently adopted in the field of CADD but their use is now widespread.

MM-GBSA energy predictions are included in Schrödinger's PRIME software suite. PRIME was originally developed for homology modelling, structure prediction and minimisation. A MM-GBSA algorithm was then added as an addition to its minimisation protocol. PRIME-MM-GBSA rescoring of GLIDE predictions are commonly used today in *de novo* studies as they are less concerned with whether the compound will bind and more interested in the effect that proposed VSAR modifications will have on ligand binding. Coupling an MM-GBSA prediction algorithm to the PRIME minimisation protocol (as provided with the Schrodinger software suite) generally produces a more accurate reflection of the ligand pose in the protein as small side-chain movements within the protein are permitted. This gives the MM-GBSA algorithm a more realistic structure to evaluate. Since minimisation protocols are easily caught in local minima, the minimisation and rescoring works on the assumption that the starting geometry predicted in the docking procedure is accurate and close to the minimised structure.

3.2 Aims and Objectives

3.2.1 Aim

The Aim of this chapter was to design a novel N-domain selective ACEis via a *de novo* approach.

3.2.2 Objectives

- Generate a combinatorial library via the modification of Enalaprilat exploring the P_2 SAR with the intention of forcing an Arg381/Glu403 interaction.
- Dock the combinatorial library into the two catalytic domains of ACE.
- Rescore the docked library poses using the MM-GBSA algorithm.
- Formulate a set of compounds for synthesis

3. 3 Methods

3.3.1 Constrained Docking Grid Generation

The first step in establishing a docking protocol for the N- and C-domain systems was to generate constrained docking grids. As in Chapter 2, the pdb structures 3NXQ and 1O86 were used to prepare docking grids for the respective N-and C-domains (Figure 3.4). The Grids for each domain were generated using the same settings as in Chapter 2 with the exception of grid length, which was set to the default 14 Å.

A new docking grid was created with the same metal binding constraint from Chapter 2 and two new H-bonding constraints both Lys511/489 and Gln281/Gln259. The grid centre was set to default centred on the native ligand. Figure 3.4 shows a representation of the native co-crystal ligands of **RXP407** and

Lisinopril from the PDB structures of 3NXQ and 1O86 respectively placed at the centre of the docking grid box. The space in which the metal chelator must be found is indicated with a red-sphere in each.

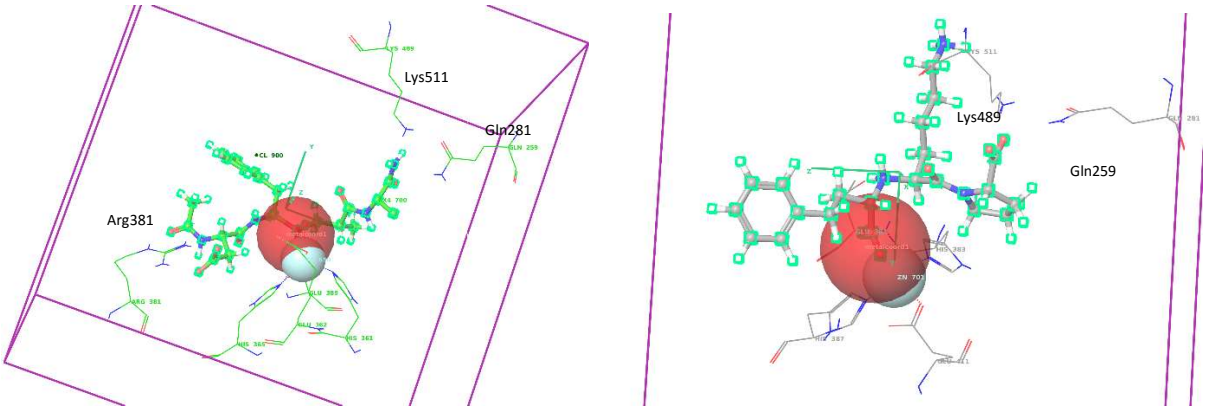


***Figure 3.4**: An illustration of the docking constraints implemented on the two domains of ACE. The N- (left) and C-domains (right) with their native ligands of Lisinopril and **RXP407** respectively. The Metal chelation constraint is illustrated with a red sphere, indicating the space in which a metal chelator must be found. The second constraint of a H-bond between either Lys511/489 or Gln281/259*

3.3.2 Docking Protocol

All docking performed in this chapter was done using the following settings:

- In the settings tab of Schrödinger GLIDE, the docking precision was set to extra precision (XP), while the remaining settings on this tab were left at default.
- In the ligands tab the ligand file was selected and the remaining settings remained at default.
- No core structure constraint was used as the constraints were deemed sufficient to fix the position of the core scaffold.
- In the constraints tab, all three constraints created for these grids were selected. A minimum of two of the three constraints was set.
- The remaining settings were all left as default.

3.3.3 System Validation

The docking systems of the N- and C-domain underwent a two-stage validation. Firstly, the set of 19 ACEis (Table 1.2) were docked into both prepared grids. A binding footprint for this set was established in both domains. The Enalaprilat ligands docked into the N and C-domain systems were then aligned with each other. The predicted poses of Enalaprilat were then overlaid with the C-domain crystal structure of the ligand of the PDB structure 1UZE.

The second stage of the validation was performed by aligning the Lisinopril ligand from the N-domain structure 2C6N with **RXP407** from the N-domain structure 3NXQ. Two hybrid ligands were created from the close alignment of the P_2 functional groups of Lisinopril and **RXP407**. These two hybrid ligands

were then docked into the prepared docking systems for both domains while the predicted poses were inspected.

3.3.4 Combinatorial Library Generation

After validating the protocol for these two systems, Enalaprilat was deconstructed into a core scaffold and then grown into the S_2 subsite of ACE by adding various negatively charged species. The core scaffold with the variable R group is depicted in Figure 3.5. A combinatorial library (Appendix 3.1) was then constructed using an assortment of carboxylic acids and acid bioisosteres reported by Ballatore et al.[117] Various alkyl chain lengths were tested as well as hydrophobic P_1 groups. A few non-acidic counter examples were included for benchmarking. This combinatorial library was created using CombiGlide's automated protocol. All the Fragments are listed in Appendix 3.1. A naming template for each ligand was established using the format FxLy with Fx indicating the fragment x attached to the core scaffold and Ly indicating an alkyl linker of y carbons inserted between fragment Fx and the core scaffold. The combinatorial library was then docked into both the N- and C-domains using the validated protocol

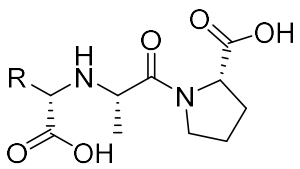


Figure 3.5: *R-group substitutions made to the Enalaprilat Scaffold. These substitutions probe the SAR in the S_1' and S_2' subsites of ACE.*

3.3.5 MM-GBSA Free Energy of Binding calculations and Rescoring

Prime MM-GBSA simulations were run on the docking outputs using the VSGB[118] solvent model. For the minimisation step, flexibility was tolerated for all protein atoms within a 15 Å radius of the ligand. The set was then reranked according to their ΔG_{GBSA} and a special note was made of the compounds with the most promising Arg381 interactions and the largest discrepancy between the N- and C-domain scores.

3.4 Results

3.4.1 ACEi Docking

To identify the space already occupied by the set of 19 clinically approved ACEis, the set was docked in the active sites of the N- and C-domains of ACE. When docked, the ACEis could all be easily overlaid in their analogous and neatly conserved poses within the S_1, S_1' and S_2' subsites. More importantly Enalaprilat was among this set. Docking this entire set showed that in addition to being the most

chemically suitable ACEi for combinatorial library design, its poses are no closer to the Arg381/Glu403 residues than the rest of the set.

The entire set followed the outline of an ACEi with a ZBG chelating the Zn atom and a P_2' carboxylic acid falling into the electropositive hole of the S_2' subsite. The majority of these ACEis have both a P_1 phenyl group and the P_1' methyl group. By and large, the docked poses of these ACEis show the ZBG, P_1' and P_2' in the same positions within the active sites. There is some movement seen in the resting poses of the P_1 phenyl group in the S_1 subsite, which is known to be large and cavernous affording the ligand a fair degree of flexibility in this region.

Figure 3.6 shows an alignment of the docked ACEi poses within each domain. Some ligands place their phenyl group within one or two carbon bonds of Arg381 suggesting an interaction with this residue is not present yet is attainable through synthetic additions to this ring.

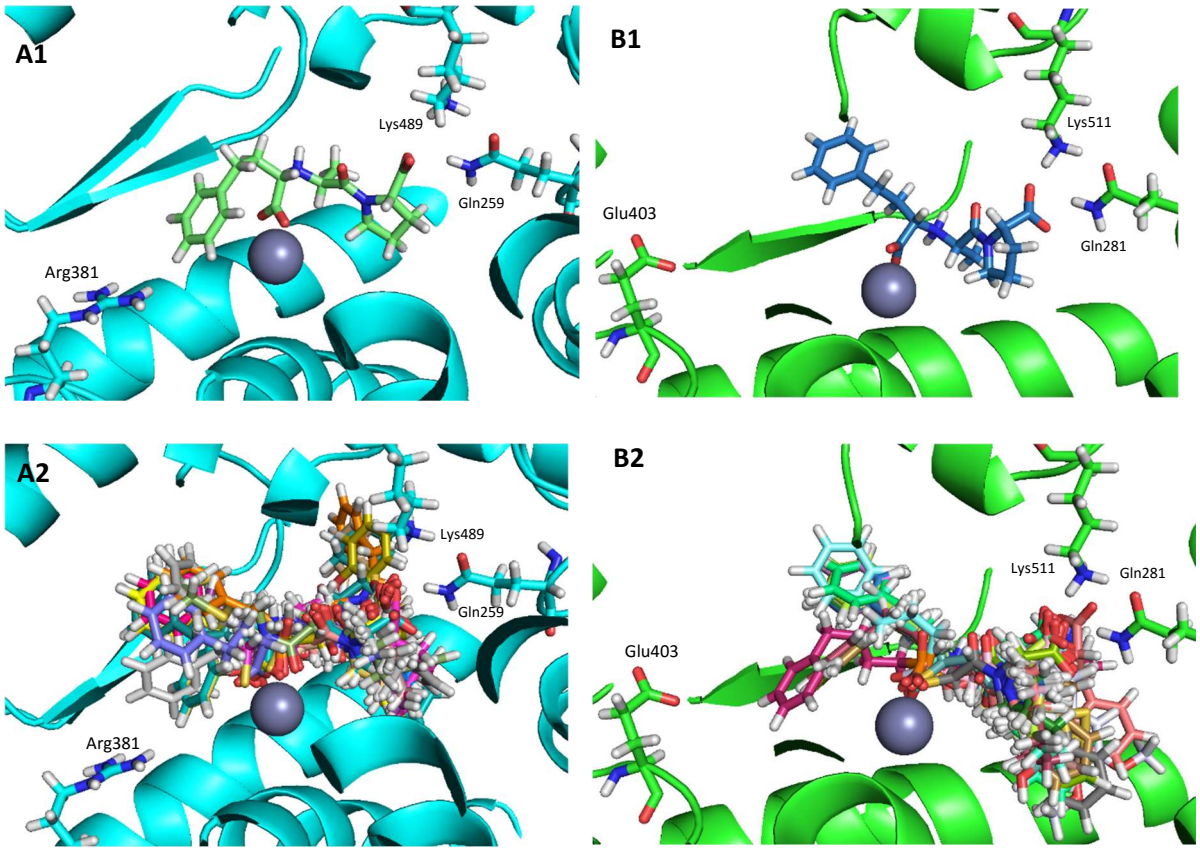


Figure 3.6: *Alignments of the set of ACEis within the N-domain (A) and C-domain (B) of ACE. A1 shows the pose of Enalaprilat in the N-domain with its Phenyl group just out of range of Arg381. The same is observed in B1 with Enalaprilat in the C-domain where the same phenyl group lies just out of range of Glu403. A2 and A3 show the docked structures of the rest of the ACEis set overlaid on top of Enalaprilat. All these ACEis have phenyl groups attached to an alkyl chain which are free to move around the lipophillic S_1 subsite but are out of range of Arg381/Glu403.*

3.4.2 Enalapril Docking Validation

Docking the ACEi set successfully validated the docking protocol for Enalaprilat, the parent compound of the combinatorial library dock. The docked poses were compared with the crystal pose of Enalaprilat from the PDB structure 1UZE. At first inspection, the crucial Zn chelation and H-bonds between the P_2' carboxylic acid and the respective Gln and Lys residues in each domain were reproduced. When aligned with the crystal pose of Enalaprilat from 1UZE, the only variation observed in this region was the conformation of the Phenyl group (Figure 3.7). An RMSD of 4.7 Å was observed between the crystallised Enalaprilat and the structure docked into this C-domain. 4.7 Å is high for two poses of a small molecule like Enalaprilat but all the variation appears to originate from the phenyl group forming a different rotamer to the crystal structure. The S_1 subsite is known to be a large and cavernous hydrophobic pocket. It is plausible that the phenyl group oscillates between these two poses under dynamic biological conditions. If the phenyl group is ignored, the RMSD is reduced to 0.38 Å providing a respectable validation of the docking algorithm using this system.

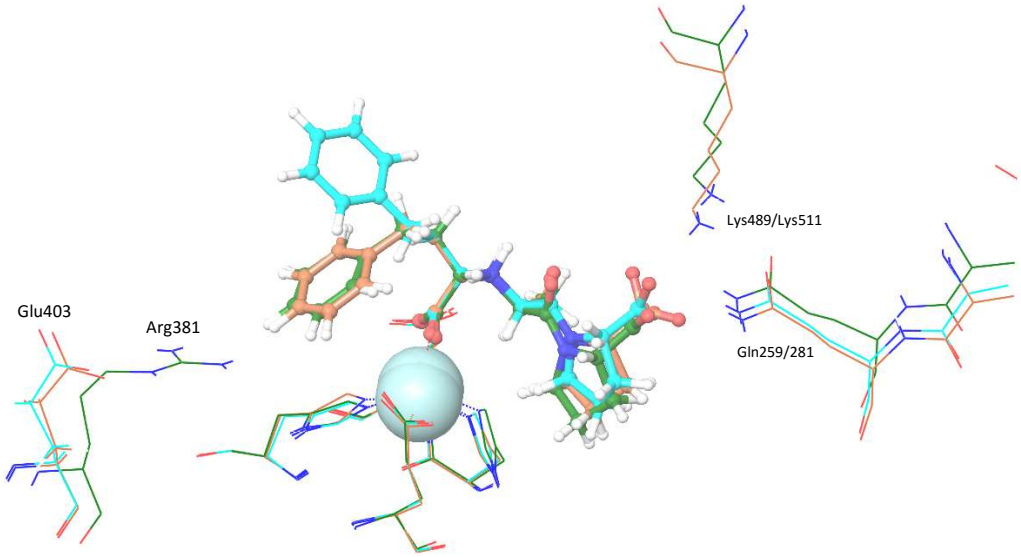


***Figure 3.7**: Overlays of the docking validation exercise. Enalaprilat is docked into the N-domain (green) the C-domain (orange) while being compared against the crystal structure of 1UZE (turquoise) with Enalaprilat co-crystalised with the C-domain. The docked structures of Enalaprilat superimpose exactly into each other. They also align closely in the P_2', and P_1'positions against the crystal pose. There is a big deviation in the P_1 position in the S_1 subsite indicating the presence of two competing local minima of torsional movement around the bond adjacent to the chelating acid.*

Glide generated a docking score for Enalaprilat of -12.78 and -11.13 for the C- and N-domains respectively. While nominal comparisons between Glide scores are meaningless, the score may give an indication as to whether strong binding between the protein and ligand exists. Glide scores below -10 are generally strong indicators of ligand binding.

3.4.3 Alignment and Recombination of RXP407 and Lisinopril

Lisinopril/Enalaprilat and **RXP407** are important ACE ligands as the former represents the chemical space of the ACEi set while **RXP407** represents the chemical space of the most N-domain selective ACEi. Although **RXP407** is a large peptide and an unsuitable drug molecule, its Arg381 binding motif can be adopted by more drug-like compounds. Aligning Lisinopril/Enalaprilat with **RXP407** is therefore useful for suggesting where to add a new acidic Arg binding functionality.

Aligning the N-domain crystal structures 3NXQ and 2C6N with their respective **RXP407** and Lisinopril ligands clearly illustrates the effect that a P_2 carboxylic acid has on the orientation of Arg381. **RXP407** extends into the S_2 subsite forcing Arg381 to swing in towards the binding site and form a salt bridge with the acid (Figure 3.8). The close alignment of these two ligands suggested two points of attachment where an acidic moiety may be attached to Enalaprilat. These points of attachment are a meta phenyl ring substitution and an alkyl carbon β to the phenyl group.

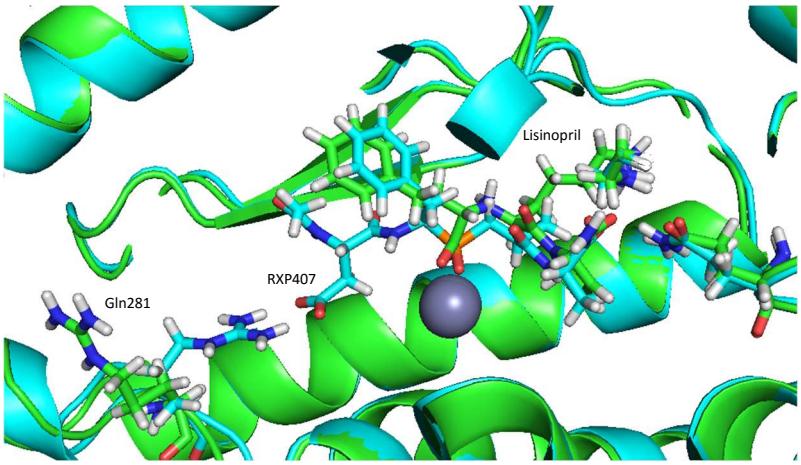


Figure 3.8: *The alignment of the crystal structures of **RXP407** (cyan) and Lisinopril (green) in the respective 3NXQ (cyan) and 2C6N (green) PDB structures. A small extension from the P_1 group of Lisinopril with a carboxylic acid causes Arg381 to swing in to face the ligand forming a salt bridge interaction with the carboxylic acid of **RXP407**.*

Since Enalaprilat has been observed to bind in a structurally identical pose to Lisonopril, it is an ideal scaffold for the design of N-domain selective inhibitors via the use of a combinatorial library. The observed crystal pose of Enalaprilat within the active site of both domains suggests only minor synthetic additions will be sufficient to create an interaction with Arg381 and Glu403 in the S_2 subsite of the respective N- and C-domains.

The alignment and recombination of **RXP407** and Lisinopril led to the design of two theoretical molecules (Figure 3.9). Docking these two molecules (Figure 3.10) returned the predicted interactions of a salt bridge with Arg381 in the N-domain and an unfavourable interaction with Glu403 of the C-domain, thus justifying the exploration of P_2 VSAR from the standard ACEi scaffold.

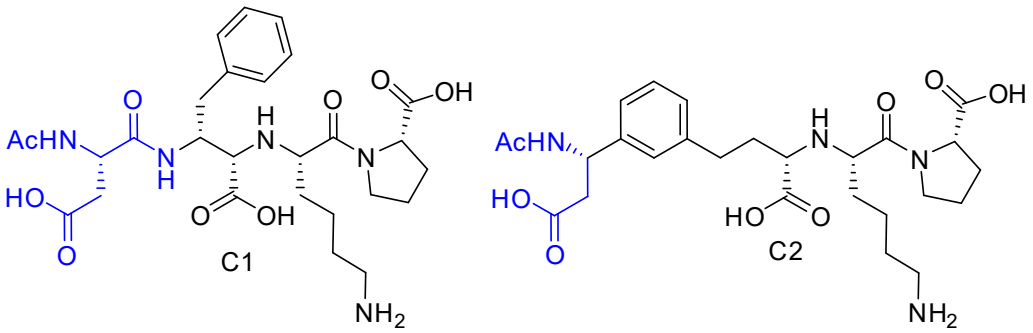


Figure 3.9: *Two theoretical molecules created by aligning the crystal poses of Lisinopril and* ***RXP407***

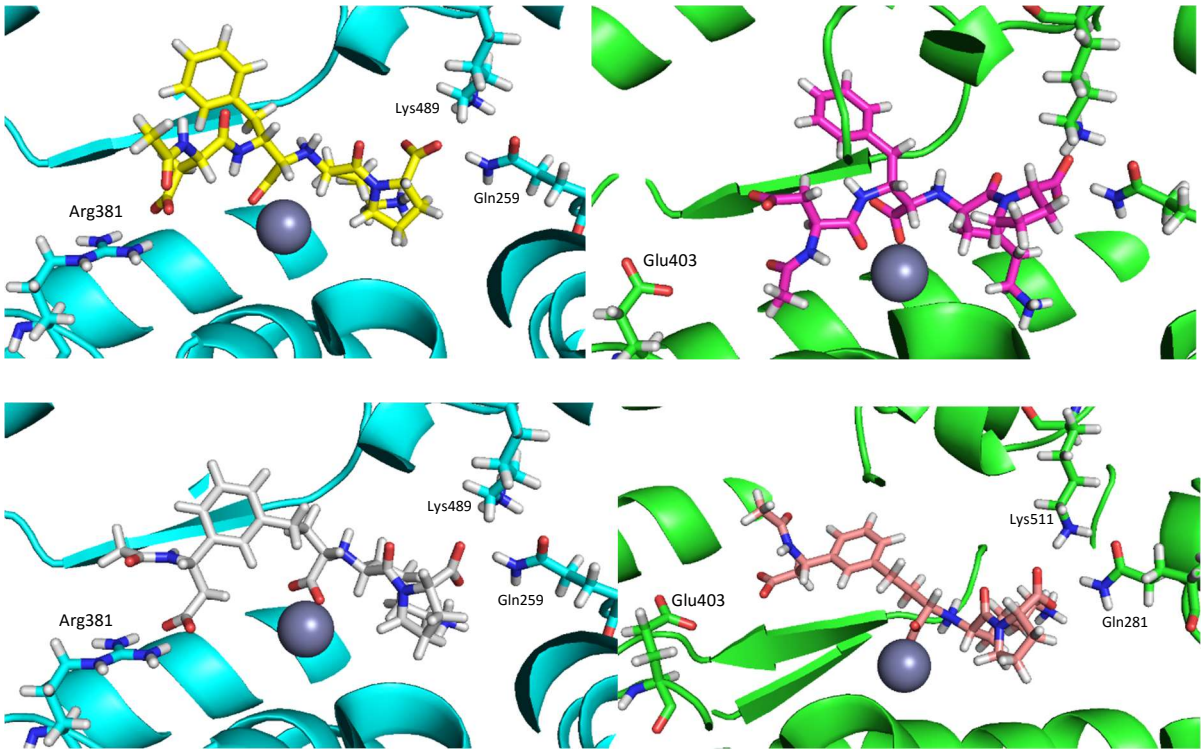


Figure 3.10*: Docked poses of the two theoretically combined molecules in the N-domain (left) and the C-domain (right). Both these molecules form a salt bridge interaction with Arg381 (left) and an unfavourable interaction with Glu403 (right).*

3.4.4 Combinatorial Library Preparation

The combinatorial library was generated using the 197 introduced fragments and attaching them with a linker of 0-4 alkyl carbons in length. Generating the library using these conditions created 985 compounds. The fragments which were too long to fit in the binding site were eliminated, reducing the set to 415 compounds (Appendix 3.1).

3.4.5 Docking and Visual Inspection of Binding Poses

After validation, the revised library of 415 compounds was docked into both domains. A thorough visual inspection of each compound in both domains found the core scaffold pose to be retained in each instance thus achieving the key objective of designing a system where flexibility was only allowed in the introduced R group. Many compounds displayed the desired salt bridge between Arg381 and

the P_2 acidic moiety. Conversely, when placed within close proximity of Glu403, negatively charged moieties were strongly repelled and thus assumed highly strained conformations. Such strained conformations were not as prominent with positively charged moieties in the N-domain. This is likely due to these moieties having more space to shield themselves from the electrostatic repulsion. (Figure 3.11).

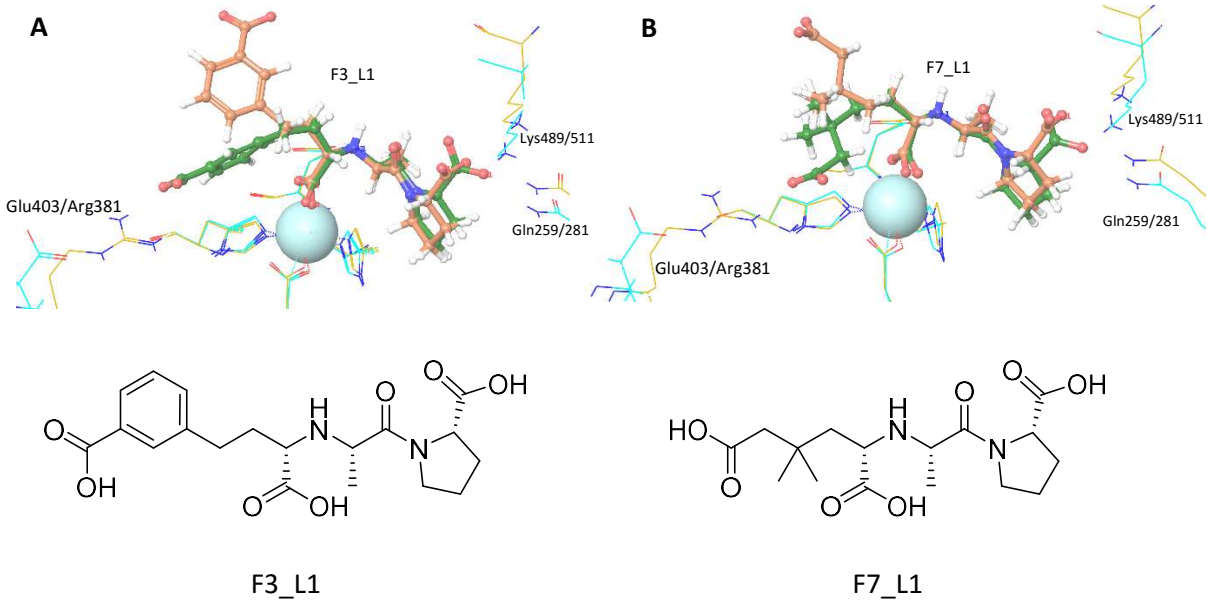


***Figure 3.11**: A comparison of poses of two potential N-domain selective ligands F3_L1 (A) and F7_L1. The green ligand is the structure docked into the N-domain and the orange ligand is the structure docked into the C-domain. C-domain residues are coloured turquoise and N-domain ones are coloured yellow. These two ligands form comfortable salt bridges with Arg381 and are strongly repelled by Glu403.*

3.4.6 MM-GBSA Rescoring

The Prime MM-GBSA free energy of binding calculations rescored and re-ranked the compounds from the Glide-score metrics (Appendix 3.2). Comparing the calculated energy of binding (ΔG_{GBSA}), there was no discernible correlation between GlideScore and ΔG_{GBSA} (Figure 3.12). The plots of Glide Score vs ΔG_{GBSA} are extremely scattered showing R^2 values below 0.1 in each case. Figure 3.12 therefore reinforces the fact that docking scores cannot be used to quantitatively predict the strength with which a compound binds to a target.

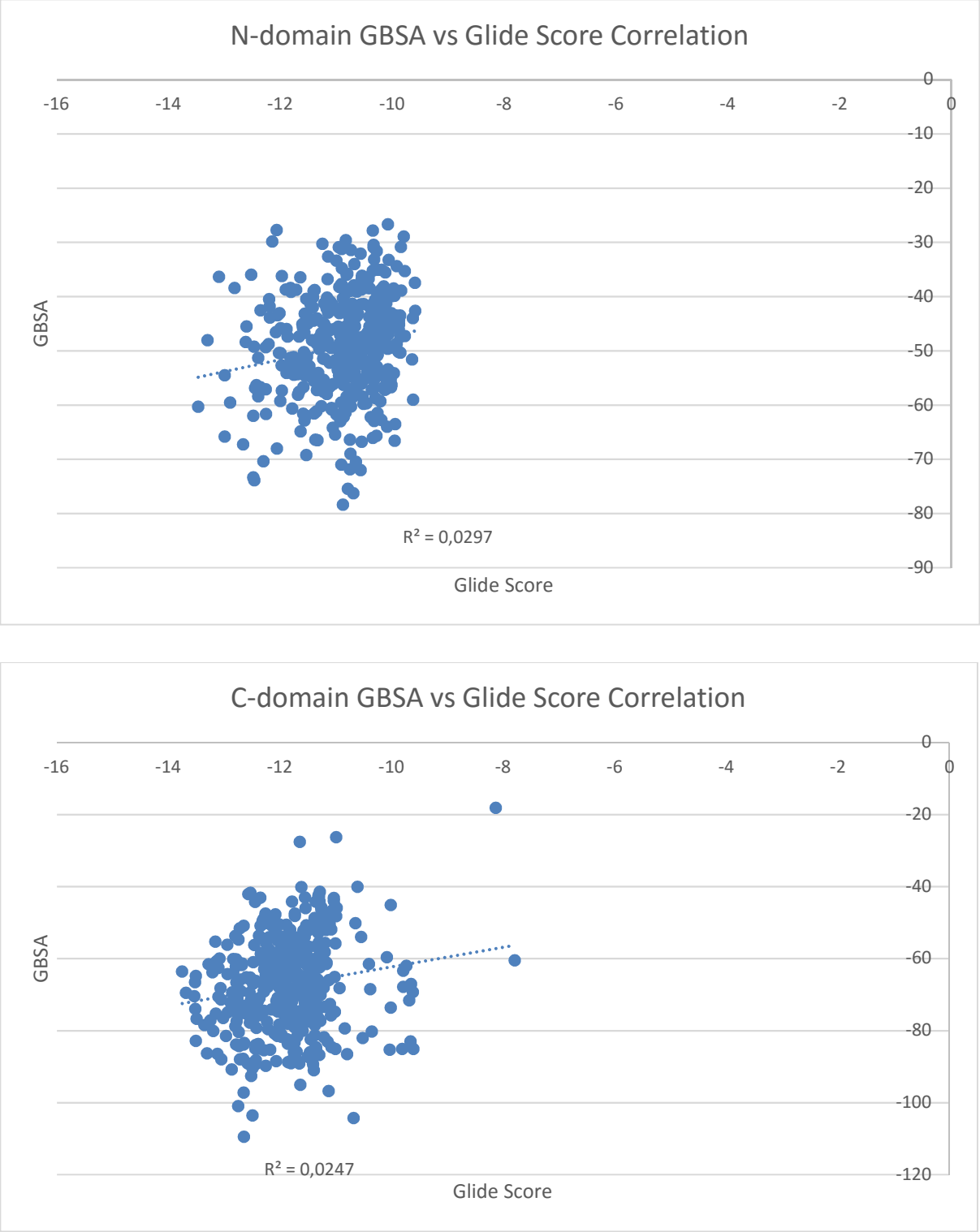


***Figure 3.12**: Scatter Plots of the Predicted ΔG_{GBSA} and Glide Score for each compound in the N-(Top) and C-domain (bottom).*

3.4.7 Data Summary and Analysis

The docking scores and ΔG_{GBSA} for each compound in the individual domains were compared. Since the docking scores and calculated binding energies between different systems cannot be compared, a new metric was devised. This metric was termed the selectivity factor. It arises from the assumption

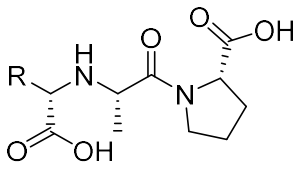
that the inhibition of the N- and C-domain by Enalaprilat are equivalent. The K_i of Enalaprilat has been measured at 2.60 and 0.63 nM for the N- and C-domains respectively[3b]. A four-fold difference in K_i exceeds the precision of the MM-GBSA method, hence this assumption is fair. Given the similar K_i values for Enalaprilat binding to the two domains they were approximated to be equivalent for the sake of the predictions. The ΔG_{GBSA}s were then normalised to 1.00 against the calculated ΔG_{GBSA} of Enalaprilat in that specific system by dividing each ΔG_{GBSA} by the ΔG_{GBSA} of Enalaprilat. Selectivity factors were then calculated by dividing the normalised N-domain binding energy by the normalised C-domain binding energy.

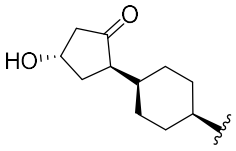
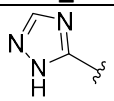
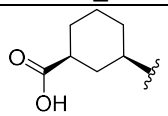
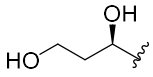
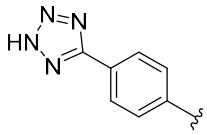
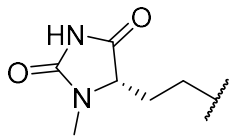
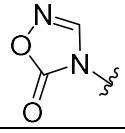
Table 3.2 shows a representative selection of the docking and MM-GBSA results for a selection of compounds from the combinatorial library (see Appendix 3.2 for the full results). With the library sorted by the selectivity factor, a few examples from the top, the middle and bottom sections are displayed. Acidic compounds were found throughout the range and the highest selectivity factors were assigned to the compounds where the acidic groups were able to make the desired Arg381 salt bridge interaction.

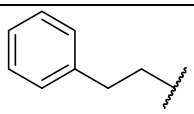
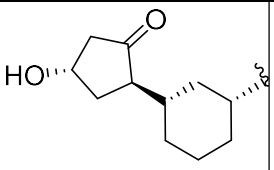
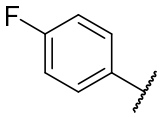
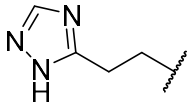
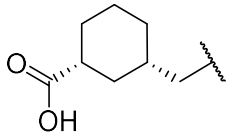
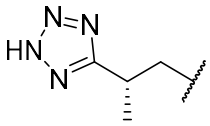
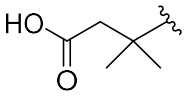
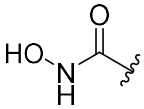
Compounds containing proximal acidic groups showed less discrimination in terms of the selectivity factor were often constrained by their geometry. It is also worth noting that the discrimination in favour of N-domain selectivity is significantly greater as the highest selectivity factor was 2.6 while the lowest was 0.6. None of the compounds can therefore be considered C-domain selective while there is strong evidence to suggest N-domain selectivity.

According to this selectivity factor, the most N-domain selective compound was F146-L0 with a selectivity factor of 2.6. It then follows that the most C-domain selective compound was F63-L2 with a selectivity factor of 0.60. This is encouraging as out of the 415 compounds predicted, 235 proved to be N-domain selective with selectivity factors over 1.00 while another 100 were in the marginal range with factors of 0.9 – 1.0. On the top end of the scale, factors of over 2.0 were observed while nothing below 0.5 on the bottom end was observed. These results are consistent with the hypothesis that the S_2 subsite can only induce N-domain selectivity.

Table 3.2: *A summary of the docking data of the combinatorial library against the two domains of ACE. The docking score, ΔG_{GBSA} and the normalised (norm) are given for each ligand. The selectivity factor is calculated as N ΔG_{GBSA} Norm/ C ΔG_{GBSA} Norm. Selectivity factors over 1.0 indicate N-domain selectivity and factors below 1.0 indicate C-domain selectivity.*



		C-Domain			N-Domain			Selectivity
Top 5								
Rank	**R-Group**	**Glide Score**	**ΔG_{GBSA}**	**ΔG_{GBSA} Norm**	**Glide Score**	**ΔG_{GBSA}**	**ΔG_{GBSA} Norm**	**N_{norm}/C_{norm}**
1	 **F46_L0**	-11.283	-41.465	0.622	-10.791	-75.425	1.628	2.614
2	 **F187_L1**	-10.989	-26.263	0.394	-10.051	-39.480	0.852	2.160
3	 **F12_L1**	-11.640	-27.544	0.413	-10.400	-38.464	0.830	2.007
4	 **F36_L2**	-12.347	-50.858	0.763	-11.534	-69.203	1.493	1.956
5	 **F61-1_L1**	-12.442	-44.240	0.664	-10.364	-57.132	1.233	1.856
Middle 5								
232	 **F194_L2**	-12.241	-59.384	1.201	-11.460	-41.343	0.928	0.772
233		-12.538	-77.816	1.168	-11.512	-54.148	1.168	1.000

	F95_L1							
234	 **F43_L2 (Enalaprilat)**	**-12.777**	**-66.596**	**1**	**-11.131**	**-46.326**	**1**	**1**
235	 **F165_L0**	-11.972	-81.025	1.216	-10.868	-56.288	1.215	0.998
236	 **F59_L0**	-11.567	-64.950	1.314	-9.875	-45.119	1.013	0.770
Bottom 5								
411	 **F187_L2**	-11.890	-70.177	1.420	-11.145	-32.641	0.732	0.516
412	 **F11_L1**	-11.941	-65.623	1.328	-10.331	-30.483	0.684	0.515
413	 **F87_L1**	-11.789	-64.765	1.310	-10.827	-29.586	0.664	0.506
414	 **F7_L0**	-11.479	-65.231	0.979	-10.344	-27.838	0.600	0.6135
415	 **F63_L2**	-11.595	-74.851	1.123	-10.275	-31.616	0.682	0.607

3.5 Discussion

3.5.1 System Validation

The validation of the molecular docking protocol in this chapter was multi-faceted as many aspects of this approach needed to be tested. The first aspect of the validation was to show that S_2 interactions

do not occur in the clinically available ACEis and that these interactions lie outside the S_1, ZBG, S_1' and S_2' footprint in which these ACEis have been observed to operate. The ACEis in question (table 1.2) are structurally homologous to the compounds that have been crystallised while most of the variation was introduced in the P_1' group. It is therefore a fair assumption that the ACEis all follow the binding pattern observed with the cocrystallised Enalaprilat, Lisinopril and Captopril ligands. The purpose of docking the set of ACEis was to test the docking protocol against the cocrystalised ligands in these two distinct catalytic domains of ACE. The close reproduction of the crystal poses of Captopril, Lisinopril and Enalaprilat demonstrated that the protocol with its implemented constraints was adequate to accurately predict binding poses in this system. The further docking of the uncrystallised ACEis confirmed these homologous drugs bind within the known footprint further justifying the VSAR exploration of S_2 interactions.

The alignment and recombination suggest that small synthetic additions to the established ACEi drug scaffold can create an interaction with the Arg381/Glu403 residues. Native peptides have a branched structure while a new chiral centre is introduced with each residue. Molecules with more than three chiral centres can be potentially challenging to synthesise while excessive branching can raise the MW unnecessarily high. Aligning the cocrystal **RXP407** and Lisinopril structures in the N-domain of ACE suggested two points of attachment for adding the P_2 groups of **RXP407** to a Lisinopril-like ACEi. The first position was the alkyl carbon α to the ZBG carboxylic acid and the second was a meta substitution on the phenyl ring (Figure 3.8).

Before devoting considerable resources to docking an entire combinatorial library, it was necessary to verify whether ACEi attachments proposed by the alignment had a realistic chance of producing the desired Arg381/Glu403 interactions. Since the docking protocol had been verified against known crystal ligands, docking the compounds created in Figure 3.8 would predict an accurate pose in both domains. Fortunately docking these two compounds produced the desired Arg381 salt bridge and Glu403 repulsion (Figure 3.9).

These two theoretical molecules introduce an additional H-bond donor, acceptor and chiral centre to an already polar molecule with three chiral centres. More polar groups would have a detrimental effect on the already poor permeability of ACEis while an additional chiral centre can make the synthesis prohibitively challenging. The purpose of the combinatorial library is to mimic some of these introduced features while keeping the ACEi within acceptable physicochemical parameters and the realm of synthetic viability.

3.5.2 Library Generation

Combi-Glide simplified the task of generating a combinatorial library from the input set of chemical groups. More than half of the ligands did not properly fit in the binding site due to the large variation in size of the attached groups. Most of the entries with large introduced groups and a long alkyl chain linker needed to be manually eliminated as they became too large to fit into the binding pocket. A molecular length constraint filter would have prevented these oversized ligands from being generated.

Many of the groups introduced a third additional chiral centre. Seeing as ACE is a protease, each subsite binds selectively to a side-chain of the native peptide while each additional residue adds another chiral centre. The absolute stereochemistry at a given stereogenic centre of each compound is therefore extremely important for this class of inhibitor. With the introduction of each stereocenter, many diastereomers needed to be generated. This created a fair amount of redundancy but all diastereomers needed to be considered when docking against proteases.

The primary focus of the combinatorial library was to attach a negatively charged group, which could interact with Arg381/Glu403. While this was the primary objective, the lipophilic nature of the S_1 subsite could not be ignored. A wide array of both aromatic and non-aromatic lipophilic groups were therefore inserted between the carboxylic acids and the scaffold in order to maintain strong S_1 binding. The library also contained a few basic negatively charged, pure lipophilic and non-ionisable polar groups. These are virtual negative controls to reinforce the hypothesis that no domain selectivity would be observed in the absence of a strong electrostatic repulsive interactions with Glu403.

3.5.3 Enalaprilat Benchmarking

In order to develop a scoring system capable of discriminating between C- and N-domain selective compounds, it was important to Benchmark the docking protocol against Enalaprilat, the library's parent ligand. The first objective of successfully finding the correct backbone orientation was achieved. The top five poses for each domain all placed the correct acid in the position chelating the Zn and the C-terminal acid was in the electropositive pocket of the S_2' subsite. The only variation brought about by these constraints was amongst the phenyl group. This variation is acceptable as the S_1 subsite is large and cavernous and can therefore accommodate the movement. The docking algorithm seems to have found another local minima and it is conceivable that in a dynamic system the phenyl group regularly switches between the two positions.

The two docked poses returned excellent docking scores both below -10.0. While it is impossible to compare docking scores between two different systems, these scores indicate highly probable ligand

binding poses. These poses are shared by Enalaprilat in the PDB structure 1UZE. This excellent correlation between the docked and the observed crystal pose of Enalaprilat successfully validated the docking protocol.

3.5.4 Library docking

A final library of 415 compounds was easily docked at the GLIDE XP level of precision. When inspected, each compound conformed to the expected pose for the Enalaprilat scaffold. The only variation in ligand pose was in the P_2 position where the ligand was free to move by design. The majority of the ligands had acidic or negatively charged P_2 groups while most of the variation was introduced with different linkers. Not every introduced acidic group was capable of forming the desired interaction with Arg381/Glu403 given the spatial constraints in place. Every example inspected showed that the desired Arg381 salt bridge also displayed an unfavourable repulsion with Glu403.

Despite this interaction being the assumed cause of N-domain selectivity, the docking score did not suffer much of a penalty as a result of this perceived unfavourable interaction. A good example of this is compound F7_L1 (Figure 3.11B). While the desired C-domain repulsion is visible, the docking scores for the C- and N-domains were -11.418 and -11.254 respectively. These similar docking scores gave no indication of the selective binding alluded to by the docked pose. This is a possible consequence of Glide not assigning a low penalty to the Glu403 electrostatic repulsion.

These poses seem to suggest that N-domain selectivity hinges on a poor Glu403 interaction more than any other N-domain interaction. Glide has been shown to be a powerful tool for predicting the correct geometry of a ligand within a binding site. Conversely Glide docking scores are useful tools for discriminating against ligands, which do or do not bind to a given target yet they provide a poor comparison of relative binding strengths.[89] In theory, rescoring using MM-GBSA predictions is better suited for the task of predicting relative binding energies.

3.5.5 MM-GBSA Rescoring

Rescoring is an essential exercise for the target binding of a chemically similar set of molecules such as this combinatorial library. Running the MM-GBSA prediction on this set appeared to provide reasonable values for the binding of these ligands (Table 3.2). For both domains there appears to be no correlation between the ΔG_{GBSA} and the docking scores assigned to these ligands by Glide. Most importantly, the poor interaction between a negatively charged group and Glu403 appears to have incurred a binding penalty.

While the binding penalty reflected in ΔG_{GBSA} suggests this method is capable of accurately predicting binding energies relative to other ligands in the same protein system, a problem still exists when

comparing the ΔG_{GBSA} between two different protein systems such as the N- and C-domains of ACE. To devise a metric, which can be compared across the two systems, the ΔG_{GBSA}s for each ligand needed to be normalised against a common ligand. The only ligand in the set with known biological activity is Enalaprilat.[3b] Theoretical binding energy, the quantity calculated by the MM-GBSA protocol is proportional to the log of K_i ($\Delta G_{Bind} = -RT\ln K_i$). This works out to a binding energy of -50 and -54 kJ/mol for the N- and C-domains respectively at a biological temperature of 310 K.

When visually inspecting docking results, the non-polar and basic side-chains were mainly concentrated in the middle of the table with selectivity factors close to 1. The flexibility of Arg381 during the minimisation step seems to make the N-domain completely tolerant of positively charged groups in the S_2 subsite with no discernible penalty attached to the placement of such groups in this position.

3.5.6 Strengths and Shortcomings

There are two major shortcomings to this approach. Quantitative binding energy predictions for a set of molecules this size usually form part of a quantitative SAR (QSAR) study. While there is a vast wealth of binding data for sACE on many open access databases, the drug inhibition data for the individual domains of ACE is almost entirely restricted to a single study conducted by Wei et al over 20 years ago, a few phosphinic peptide studies[56,58] as well as the work performed on keto-ACE analogues[60, 66]. In total, there are less than 20 molecules with which to build a QSAR training set. These molecules vary from phosphinic peptides to drugs with a carboxylic acid or thiol ZBG. An effective QSAR training set would need to be more homologous than this set. In the absence of an adequate training set, an effective QSAR model is not possible.

The other shortcoming of this approach is that the constraints in place forced many of the ligands into an unnatural position within the C-domain binding site. This implies that many of the C-domain poses are inaccurate making MM-GBSA calculations for these poses meaningless. The selectivity factor therefore needs to be considered with a fair degree of scepticism as it is unlikely that much correlation would be observed between the C-domain ΔG_{GBSA} and experimental K_is. These results therefore need a more qualitative interpretation as scoring penalties do not accurately reflect the impact these interactions have on binding.

Identifying ligands with strong Arg381 interactions and pronounced Glu403 repulsions via a visual inspection is therefore the best use of this method. The selectivity factor may help emphasise some promising compounds but it is unlikely to show a strong experimental correlation with any measured binding.

Another point of contention is the synthetic viability of some of these compounds, especially ligands that introduced multiple chiral centres. While it is true that many of these compounds are unlikely to be synthetically viable, they can become synthetically viable with some simple modifications. It is easy to rerun these predictions on slightly modified ligands from this set if necessary. Even if these compounds are not synthetically viable, they can still elucidate new chemical space which can be explored via more synthetically viable analogues.

3.5.7 Concluding Remarks

Combinatorial library docking coupled with MM-GBSA rescoring is an attractive method for designing new ligands and exploring targeted VSAR around a particular ligand. Large libraries can be generated and docked at a relatively low computational cost. Challenges, however, arise when comparing two similar homologous targets. In the absence of extensive enzyme binding data for the individual domains, it is impossible to create a training set to test this prediction model. This prediction model therefore has provided some ideas for synthesis when considered on a purely qualitative level. On the other hand, these predicted binding energies and the selectivity factors need to be treated with a touch of scepticism before it has been correlated with any *in vitro* data.

Chapter 4 – Enalaprilat Analogue Synthesis

4.1 Introduction

4.1.1 Enalaprilat Scaffold

In Chapter 3 the VSAR of the ACE S_2 subsite was undertaken to explore the possibility of adding an acidic P_2 group to an Enalaprilat scaffold. In order to test the hypothetical molecules conceived via VSAR studies, a viable synthetic pathway allowing for small changes at the P_1 position needed to be established. The scheme in which Patchett et al[113] synthesised Enalaprilat was centred around the reductive amination of an α-keto acid and Ala-Pro using sodium cyanoborohydride (Figure 4.1)

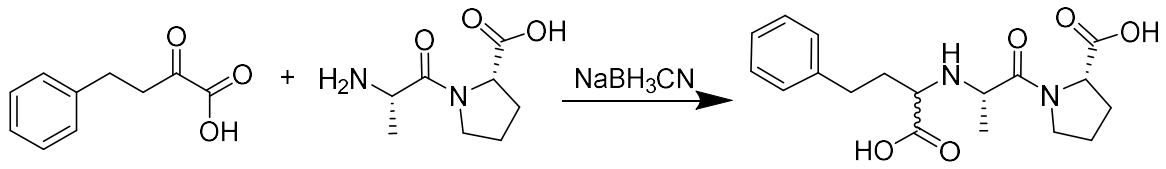


***Figure 4.1**:Patchett's synthesis of Enalaprilat[113]*

While this scheme produced both Lisinopril and Enalaprilat, it has one major drawback. The reductive amination step is non-diastereoselective resulting in an equimolar ratio of both the *SSS* and *RSS* diastereomers. As a protease, ACE is a highly stereospecific target showing a strong preference towards the *SSS* diastereomer of Enalaprilat (*SSS* IC_{50} 1.2 nM vs *RSS* IC_{50} 820 nM)[113]. Patchett et al synthesised many P_1 analogues of Enalaprilat while performing inhibition assays on diastereomeric mixtures, only separating the diastereomers once strong hits had been identified.

Since the synthesis of this pioneering series, a handful of diastereoselective synthetic routes have been published.[119] These routes centre around nucleophilic substitution of a secondary bromo group. The nucleophilic substitution of secondary halogens progresses either via S_N1 or S_N2 mechanisms depending on the stability of the dehalogenated carbocation intermediate. An S_N2 substitution entails a stereocentre inversion and is therefore diastereoselective while an S_N1 mechanism proceeds via a carbocation intermediate producing both diastereomers with the degree of selectivity depending on stereo control features in the vicinity of the carbocation intermediate.

***Figure 4.2**: A generalised nucleophillic substitution reaction to synthesise P_2 analogues of Enalaprilat.*

In the context of Enalaprilat P_1 analogues (Figure 4.2), a nucleophilic substitution would entail the substitution of a secondary Br group from a 2-bromopropanoic acid proline. Since the Br in 2-bromopropanoic acid is on the chiral carbon adjacent to an electron withdrawing carbonyl group, the expected stabilisation of a carbocation intermediate may not be feasible. Thus, a diastereoselective S_N2 mechanism would be expected as the favoured mechanism. If stabilisation of the carbocation intermediate were to occur, it would likely constitute the minor pathway producing a small proportion of the other diastereomer via a hybrid S_N1/S_N2 mechanism. The net result would be a partially diastereoselective substitution.

4.1.2 Accessing the P_2 VSAR series

The VSAR study undertaken in Chapter 3 delivered a large variety of Enalaprilat variants. Replacing Patchett's reductive amination with a diastereoselective S_N2 nucleophilic substitution provides an alternative. A scheme centred around a substitution reaction would require the protection of acids to facilitate such a reaction.

A diastereoselective synthesis of Enalaprilat analogues would therefore be highly advantageous if it could provide a simpler synthetic route to P_2 Enalaprilat analogues. There are a few published examples of a nucleophilic Br substitution applied specifically to the synthesis of ACE inhibitors (Figure 4.3)[119c]. In such reactions, the substitution was highly diastereoselective making a strong case for its adoption in the synthesis of Enalaprilat derivatives.

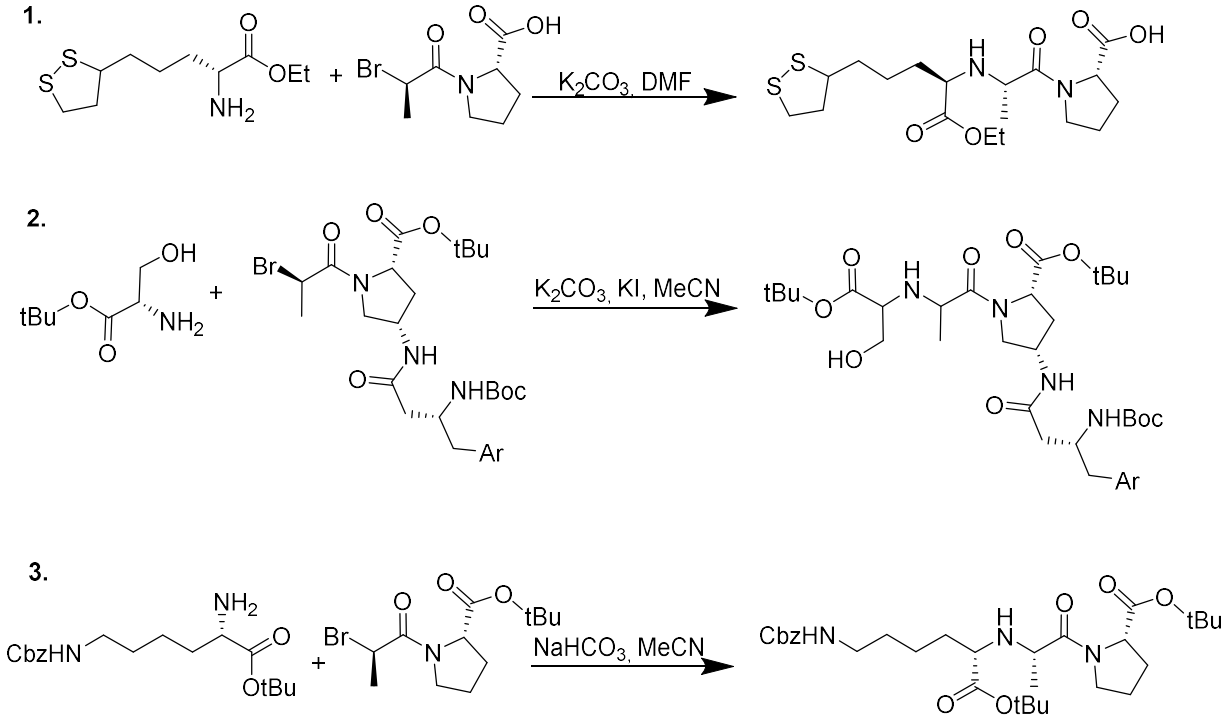


Figure 4.3: *Published examples of a nucleophilic substitution successfully employed in the diastereoselective synthesis of ACEis.*[119]

In contrast to the key amination step of Patchett et al, the nucleophilic substitution reactions require the protection of acid and amine groups to ensure the correct regioselectivity. Selecting a protecting group for both acid and amine groups has the potential to play an important role in the final yield and purity of the compound. Boc is the most commonly used amine protecting group while a variety of acid protecting groups have been employed including methyl, ethyl and t-Butyl esters.

4.1.3 Novel ACE N-domain selective SAR

Until now the existing SAR on the individual ACE domains has been restricted to a set of bradykinin potentiating peptides,[120] Ang-I metabolites,[52] phosphinopeptides[56, 58] and peptidomimetic ACEis.[3b] These studies have explored P_2' SAR in search of C-domain selectivity. While the core contribution of P_2' P_1', ZBG and P_1 moieties of Enalaprilat towards binding in the two domains is well understood[62], the vacant P_2 position is a novel site with which to exploit interactions in the S_2 subsite.

The effect of the Arg381/Glu403 mutation on domain selective ACE inhibition has only been explored using phosphinic peptides.[56, 72] Targeting these S_2 residues via modifications to the Enalaprilat scaffold would be the first attempt designing N-domain selective drug-like ACE inhibitors.

4.2 Aims and Objectives

4.2.1 Aim

The aim of this chapter was to explore a novel diastereoselective synthetic route for the synthesis of Enalaprilat analogues. This route was intended to provide easy access to the set of Enalaprilat analogues modelled in Chapter 3.

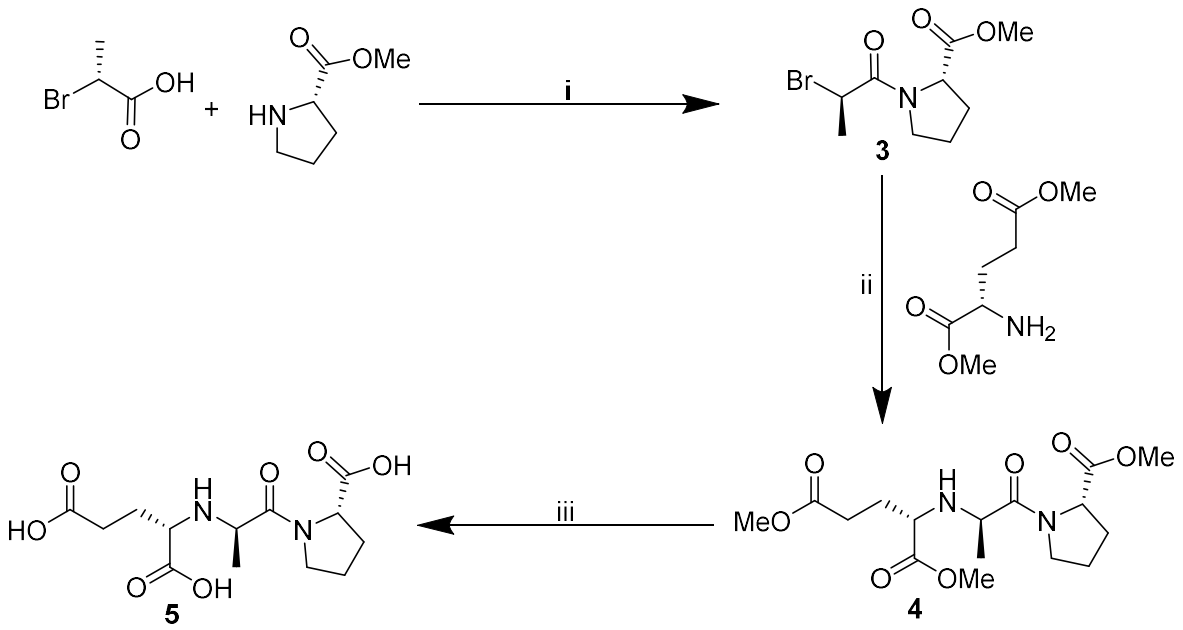
4.2.2 Objectives

In pursuit of this novel synthetic route, the following objectives were laid out:

1. Establish a new diastereoselective synthetic route to Enalaprilat analogues.
2. Synthesise Enalaprilat analogues for SAR studies.
3. Test the analogues in ACE competitive inhibition assays to determine their potency and domain-selectivity.

4.3 Methods

4.3.1 Enalaprilat Analogue Synthesis



Scheme 4.1: *Synthetic route used to synthesise an Enalaprilat P_1 analogue. i. EDC/HOBt, N-methyl morpholine, DCM 0 °C 2 hrs, 25 °C 12 hrs (43%). ii. $NaHCO_3$, MeCN 80 °C 36 hrs (73%). iii NaOH, H_2O, 25 °C, 2hr (39%).*

An Enalaprilat analogue was synthesised according to Scheme 4.1. For reaction i 2-(R)-bromo propanoic acid, *N*-hydroxy benzatriazole (HOBt) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) were dissolved in chilled dichloromethane (DCM) with stirring for 30 mins. Proline methyl ester was then added together with *N*-methylmorpholine (NMM) after which the mixture was allowed to warm to 25 °C and stirred for 12 hours. In reaction ii, compound **3** and an amino acid methyl ester

were added to MeCN together with $NaHCO_3$ under inert conditions. The reaction was then refluxed at 80 °C for 24 hours. After purification, product **4** was then deprotected by dissolution in a 1 M aqueous NaOH at 25 °C with stirring for 2 hours.

4.3.2 *in vitro* Competitive ACE Inhibition Assay

Once synthesised the compound was tested *in vitro* for ACE inhibition. The ACE assays followed the same methods described in Chapter 2, section 2.3.5. After the initial screening of the compounds, IC_{50} values were determined for compounds displaying sub-micromolar inhibition.

4.4 Results

4.4.1 Peptide Coupling

The first conditions attempted utilised EDC and HOBt in DMF without an organic base.[121] These conditions produced a poor yield of 31% with a by-product formed in equal mass quantities. In an attempt to improve the yield of this reaction, a variety of different reagents and solvents were tested. The best yields were obtained using the combination of EDC and HOBt in DCM with NMM. Compound **3** (**SF05**) formed in a roughly 1:1 mass ratio with a by-product. A yellow-brown residue was purified over silica (10:90 EtOAc/Hexane) to yield **SF05**, a white crystalline powder (mp 107-110 °C), in a yield of 43%.

4.4.2 Nucleophilic Bromine Substitution

The most popular conditions for this reaction, are a weak inorganic base such as K_2CO_3 or $NaHCO_3$ with an aprotic solvent such as MeCN or DMF.[119a, 119c] For this reaction $NaHCO_3$ was utilised with the mixture refluxed in MeCN for 36 hours at 80 °C.[122]

The first amino acid chosen was Glutamic acid dimethyl ester based on availability and modelling results described in Chapter 3. The newly formed amine bond from reaction ii was easily visualised using a ninhydrin TLC stain. The product was purified over silica (2 : 98, MeOH : DCM) and obtained as a yellow-brown oily residue with a yield of 73%.

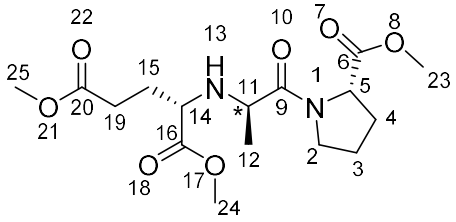


Figure 4.4: SF06

The diastereomeric purity of **SF06** (Figure 4.4) was then evaluated. With a mixture of two diastereomers around a single chiral centre, ^1H NMR peak duplication would be expected for the

protons attached to either the chiral carbon or diastereotopic protons as these protons are shared by two analogous peaks in slightly different environments, hence the slightly different chemical shifts. The **SF06** proton H-11 is attached to the chiral carbon C* while the H-12 protons belong to a methyl group attached to this chiral carbon. In the presence of a diastereomeric mixture, the H-11 and H-12 protons would be found in slightly different environments for each diastereomer.

Considering the crude ^{1}H NMR of **SF06** (Figure 4.5), a clear peak duplication was observered for H-12 in its ^{1}H NMR spectrum. The H-12 peak is characterised by a doublet (J = 6.82 Hz) with a shift in the 1.0 ppm region. In Figure 4.5, two doublet peaks of almost identical J-values (J = 6.82, 6.60) but different heights and integrations were observed. The large peak labelled 12M was assigned to H-12 from the major **SF06** diastereomer while the smaller peak labelled 12m was assigned to the minor diastereomer. The relative integrations of the major and minor peaks can now be compared to determine the diastereomeric excess of this substitution reaction.

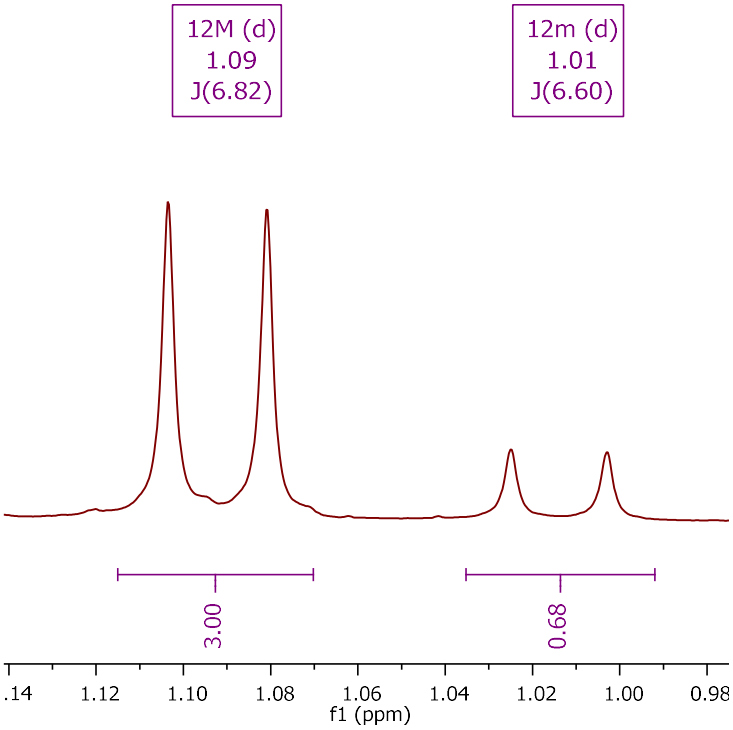


Figure 4.5: *An extract from the ^{1}H NMR spectrum of a crude **SF06** sample showing the duplication of the doublet peak for the H-12 methyl group.*

The 12M major peak integrated for 3 protons when normalised relative to single proton peaks (Figure 4.5). With this major peak normalised to 3 protons, the integration of the minor peak was measured at 0.68 protons. The ratio of areas for the minor and the major peak is therefore 1:4.41 translating to a diastereomeric excess (de) of 63%.

This reaction is therefore moderately diastereoselective following a presumed hybrid S_N1/S_N2 mechanism. A de in excess of 50% suggests the S_N2 mechanism is the dominant substitution pathway while the S_N1 mechanism still makes an important contribution. Assigning these two diastereomer peaks would require a diastereomer separation and 2D NMR methods. Based on precedence, these diastereomers can be assigned using ACE inhibition data, therefore no attempt was made to separate the diastereomers at this stage.

4.4.3 Methyl ester Deprotection

The three methyl esters present in **SF06** were hydrolysed by stirring the compound in aqueous NaOH for two hours. As this is an aqueous phase reaction, it was monitored via LC-MS instead of TLC. Initially 0.25 M NaOH was used but, at this concentration the only mass seen was the product of **SF06** with just two hydrolysed esters. In an effort to hydrolyse the third methyl ester, the NaOH concentration was increased in increments of 0.05 until the mass of the triple hydrolysed ester product was observed via LC-MS using a NaOH concentration of 1.00 M.

After the workup involving HCl neutralisation, the resultant product was obtained as a salt mixture. In a similar reaction by Greenlee et al,[123] the salt was removed by running the sample through a Dowex ion exchange resin but in this case, the sample was loaded directly onto the Prep-HPLC. **SF07** was isolated as a sticky transparent-white paste with a post HPLC recovery yield of 39% and a purity of 97%.

4.4.4 Racemisation

An unforeseen consequence of raising the NaOH concentration to achieve a triple ester hydrolysis was the racemisation of chiral centres adjacent to the ester groups. Base-catalysed ester hydrolysis has been documented to cause racemisation via a reversible side reaction (Figure 4.6)[124]. Despite the diastereoselectivity around the methyl group introduced in reaction ii, **SF06** contains two methyl ester groups attached to chiral carbons. The base-catalysed ester hydrolysis reaction has a significantly lower activation energy than the racemisation reaction, hence racemisation is minimal at low base concentrations. The presence of a third methyl ester in **SF06** required the NaOH concentration to be increased to a point where the racemising side reaction became a significant factor.

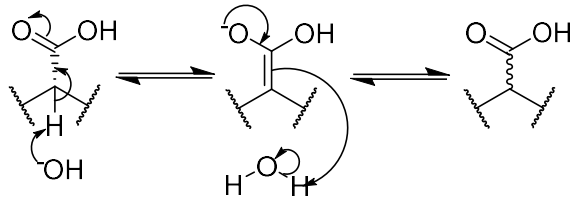


Figure 4.6: *The reversible base catalysed racemisation of chiral acid groups.*

Two of the esters in **SF06** are adjacent to chiral centres, which can undergo base-catalysed racemisation during the ester hydrolysis. As a mixture of two diastereomers around the methyl group, racemisation at the two chiral methyl esters creates four diastereomer permutations for each of the two existing diastereomers, bringing the total number of diastereomer permutations to eight. All eight diastereomer permutations from the three chiral centres of **SF07** are now possible. With four charged groups, **SF07** is a highly polar compound in solution. This high polarity results in short HPLC retention times. The short retention times made diastereomer separation difficult. The diastereomerism was clearly seen in the ^{1}H NMR spectrum with multiple duplicates of the methyl peak H-12 (Figure 4.7).

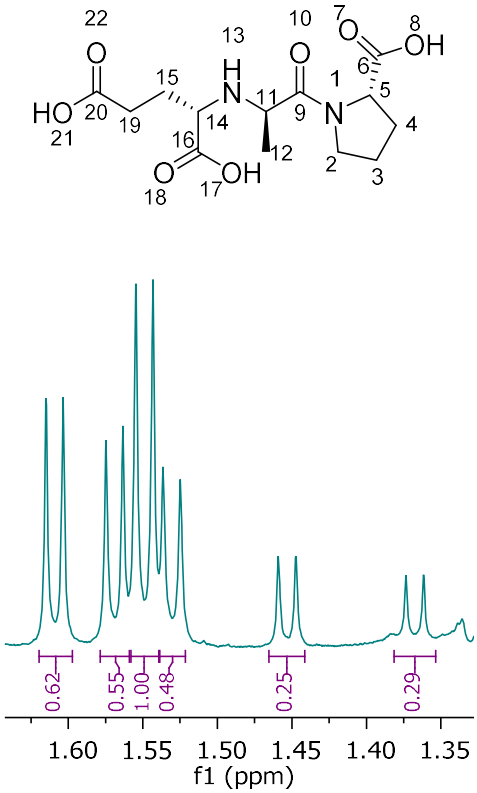


Figure 4.7: *An extract from the ^{1}H NMR spectrum of **SF07** showing the peak duplication of several doublets for the methyl H-12 protons indicating several distinct diastereomers.*

Racemisation is often the reason the base concentration is kept low during ester hydrolysis. In the case of **SF07**, racemisation was an unavoidable consequence of the higher base concentration required for the triple ester hydrolysis. Scheme 4.1 was devised to introduce diastereoselectivity in the reaction. However, the triple ester hydrolysis unfortunately resulted in up to eight diastereomers instead of the original two. The final HPLC purified sample of **SF07** showed a high purity despite the high degree of diastereomerisation present.

4.4.5 Characterisation of SF07

The high level of diastereomerism in **SF07** meant its ^{1}H NMR peaks needed to be compared directly with those of **SF06** for the correct assignment. The assignment of the ^{1}H NMR of **SF06** (Figure 4.8) begins with a quartet (*J* = 4.08 Hz) at 4.31 ppm that can be assigned to H-11 with an adjacent methyl group. As there is only one proton in the H-11 position, the integration of this peak can be normalised to 1.0. The doublet (*J* = 6.77 Hz) at 1.09 ppm can be assigned accordingly to the methyl at H-12. The broad multiplet at 3.61 ppm integrates for two protons corresponding with the two deshielded cycloalkyl protons at H-2 with an adjacent N. The three methyl ester peaks of H-23, 24 and 25 appear as singlets at 3.61, 3.60 and 3.58 ppm respectively. Two peaks each integrating for one proton are seen immediately upfield from the methyl ester peaks. The peak at 3.67 ppm is a sharp multiplet corresponding to the chiral proton H-14. A broad multiplet integrating for another single proton immediately upfield from the sharper multiplet corresponds to the proton at position H-5 on the pyrrolidine ring adjacent to both a N and an ester. The two protons at H-19 can be assigned to the multiplet at 2.32 ppm. These protons are slightly deshielded by the adjacent ester while its expected quartet has been distorted by the diastereomerism of the compound creating two interfering peaks for H-19 resulting in the multiplet. The only peak remaining is a large broad multiplet at 1.87 ppm integrating for six protons. These six protons match the six remaining unassigned downfield protons of H-3, H-4 and H-15.

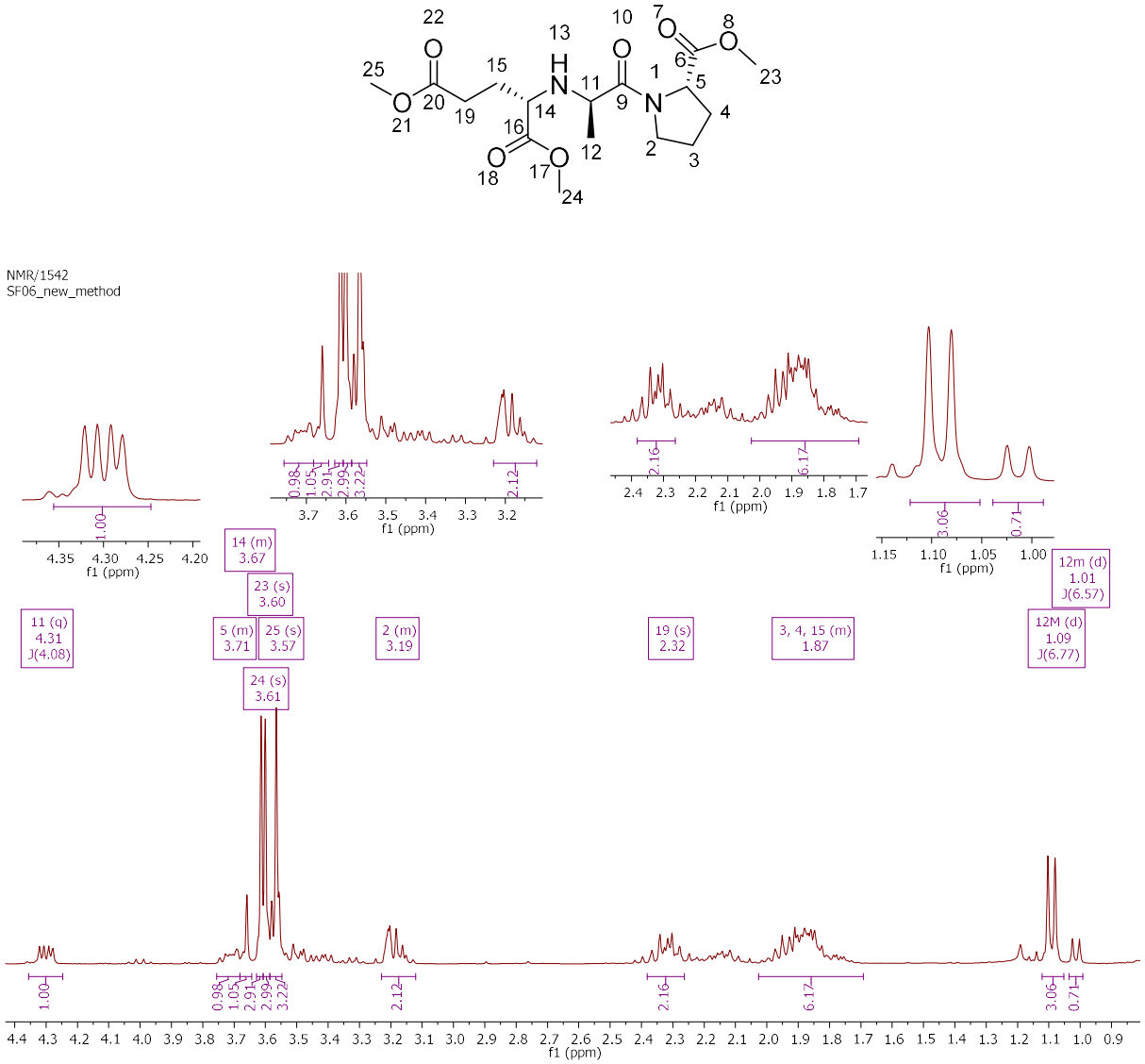


Figure 4.8: *^{1}H NMR spectrum of **SF06** at 300 MHz*

With the ^{1}H NMR spectrum of **SF06** assigned, it was stacked onto the spectrum of **SF07** (Figure 4.9) as this spectrum was harder to assign due to peak duplication on account of a mixture of eight possible diastereomers. Protons attached to stereocentres can show a pattern of distinct yet repeated peaks while broader peaks tend to remain together with their multiplicity merged into multiplets. Figure 4.9 shows the stacking of the spectra of **SF06** and **SF07**. As expected, the defining difference between these two spectra is the absence of the three methyl ester peaks (3.68-3.57 ppm) following the deprotection. In the absence of these peaks, the multiplets for the cycloalkyl protons H-2 and the single chiral proton H-14 peaks are observed in this region.

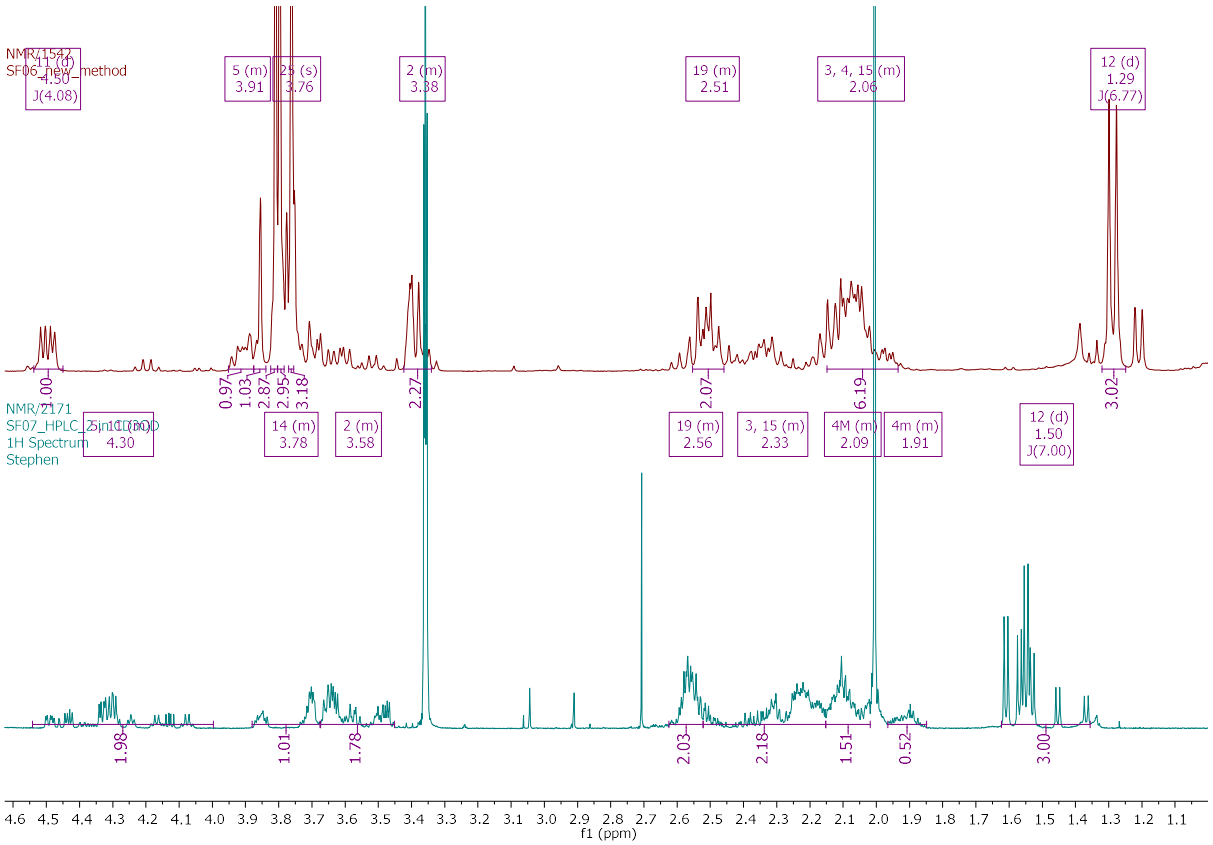


Figure 4.9: *Stacked* ***SF06*** *(red) and* ***SF07*** *(cyan)* ^{1}*H NMR spectra*

With the peaks of **SF06** mapped to the ^{1}H NMR spectrum of the diastereomerically diverse **SF07**, its ^{1}H NMR spectrum (Figure 4.10) could now be assigned. To normalise the integration, the total area of the six doublet peaks (J = 6.95 Hz) between 1.35 and 1.60 ppm was set to 3 protons on the assumption that they cumulatively add up to the three methyl H-12 protons with contributions from each diastereomer present. Peaks in the 4.05 to 4.50 ppm range appeared to be a mixture of quartets (J= 6.94 Hz) and triplets (J = 5.66 Hz). These peaks collectively integrated for two protons matching the individual protons at H-14 (quartet) and H-5 (triplet). These two peaks are duplicated several times with different areas as these protons are both attached to chiral carbons whose environment varies significantly with each diastereomer. The peak for the chiral proton H-14 is split between multiplets at 3.84 and 3.71 ppm which collectively sum to a full proton. The broad multiplets found between 3.46 and 3.66 ppm integrate to a total of two protons matching the expected shift pattern of the deprotected cycloalkyl protons at H-2. The multiplet at 2.58 ppm integrates for two protons coinciding with the expected shift of the two protons found at H-19. The broad repeated quintet or multiplet expected for the two diastereotopic protons at H-15 is seen between 2.3 and 2.5 ppm. Multiple diastereomers and overlapping chemical shifts have caused some overlap and interference, hence the pattern of this peak is messy. The two remaining unassigned proton pairs at H-3 and H-4 can now be assigned to the multiplets at 2.28 and 2.09 ppm respectively. The multiplet of H-4 is duplicated with a

minor peak at 1.91 ppm. This peak is split due to the different environments arising from different through space interactions with the different configurations of the acid at C-5. Collectively these two peaks integrate for two protons.

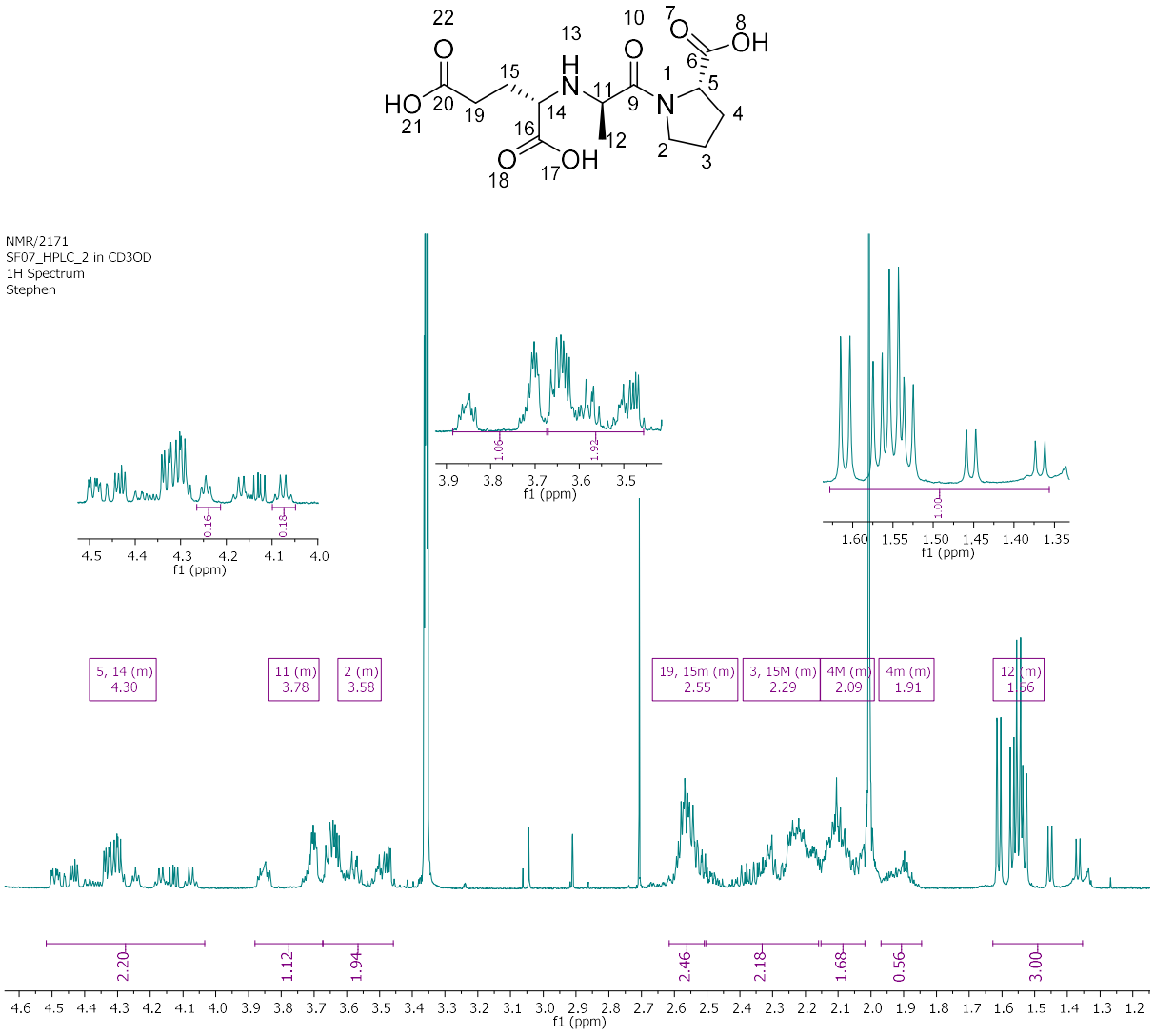


Figure 4.10: *^{1}H NMR spectrum of* ***SF07*** *at 600 MHz*

4.4.6 ACE Competitive Inhibition Assay

SF07 as a diastereomeric mixture, was sufficiently pure for a competitive inhibition assay against the N- and C-domains of ACE. The **SF07** mixture was initially screened at 100 and 1 µM (Figure 4.11). At 100 µM nearly 70% of the C-domain activity remained while the N-domain was almost completely inhibited. At 1 µM, there was no inhibition of the C-domain while the N-domain still showed 82% of the total enzyme activity remaining. This preliminary screen suggests the IC_{50} for the N-domain lies in the low µM range while the IC_{50} for the C-domain is over 100 µM giving a selectivity factor of roughly two orders of magnitude.

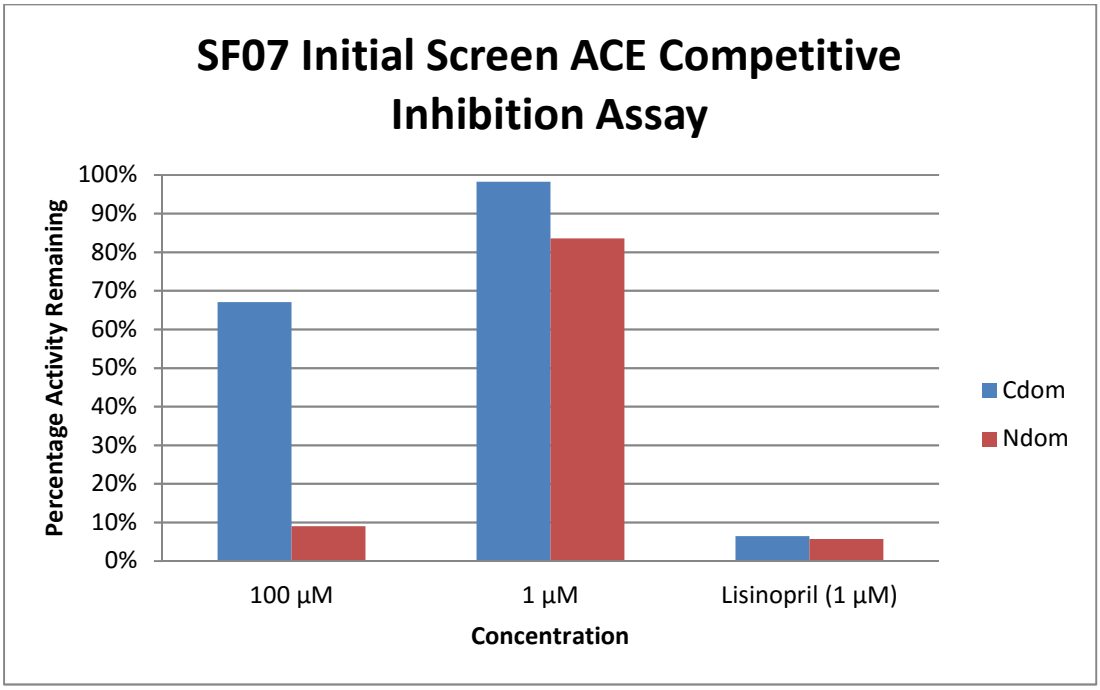


Figure 4.11: *The enzymatic activity remaining after the two ACE domains were screened with* ***SF07****. A 1 µM Lisinopril was used as a negative control.*

With an IC_{50} > 100 µM for the C-domain, only the N-domain IC_{50} of the diastereomeric **SF07** mixture was determined (Figure 4.12).

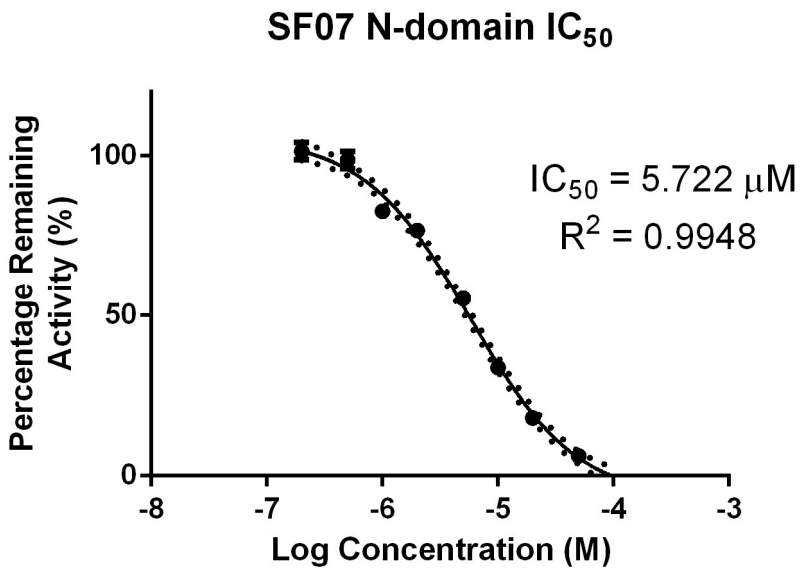


Figure 4.12: *Dose response curve measuring the IC_{50} of* ***SF07*** *against the ACE N-domain. The IC_{50} was measured at 5.722 µM with an R^2 of 0.9948*

This low micromolar inhibition of the N-domain and only 33% inhibition at 100 µM for the C-domain unambiguously confirms the N-domain diastereomeric selectivity of the **SF07** mixture. Based on the structural similarity between Enalaprilat and **SF07**, the IC_{50} of the *SSS* diastereomer against the N-

domain was expected to be in the low nanomolar range. However, the IC_{50} of **SF07** was measured to be in the low micromolar range, 2-3 orders of magnitude higher than what was expected. Low micromolar inhibition is consistent with the wrong diastereomer or a mixture of different diastereomers. Patchett et al measured the IC_{50} of the *RSS* Enalaprilat diastereomer to inhibit sACE with an IC_{50} of 820 nM as opposed to 1.2 nM for the *SSS* diastereomer while the mixture of the two diastereomers inhibited sACE with an IC_{50} of 3.8 nM.[113] The 6-8 diastereomers present in this **SF07** sample adequately explains the lower than expected potency observed for the N-domain. Based on these preliminary results, **SF07** is a potentially promising N-domain selective ACE inhibitor. The synthesis, however, does need to be refined to avoid racemisation during deprotection.

4.4.7 Rationalisation of Domain Selectivity

To explain the ACE domain selectivity behaviour of **SF07**, its predicted binding pose was analysed in greater detail. **SF07** was named compound F1_L1 in Chapter 3. Based on MM-GBSA binding scores, F1_L1 was predicted to bind strongly to both catalytic domains of ACE. **SF07** (F1_L1) was the only compound from the modelled set, which was synthesised hence it deserves a closer look. Apart from the strong MM-GBSA binding scores towards both ACE domains, the predicted binding pose between the **SF07** and the two domains was now analysed.

The distance between an oxygen the **SF07** P_1 acid and the nearest oxygen atom of Glu403 of the C-domain is 12.1 Å while the distance between an oxygen atom of this acid and the nearest nitrogen of Arg381 in the N-domain is 6.3 Å (Figure 4.13). At a distance of 12.09 Å, it is safe to say that Glu403 has no influence on the binding of **SF07** to the C-domain. At a distance of 6.0 Å from **SF07**, it is plausible that Arg381 is involved in a water mediated H-bond with **SF07.** However, it must be noted that this interaction is not essential for strong binding between **SF07** and ACE.

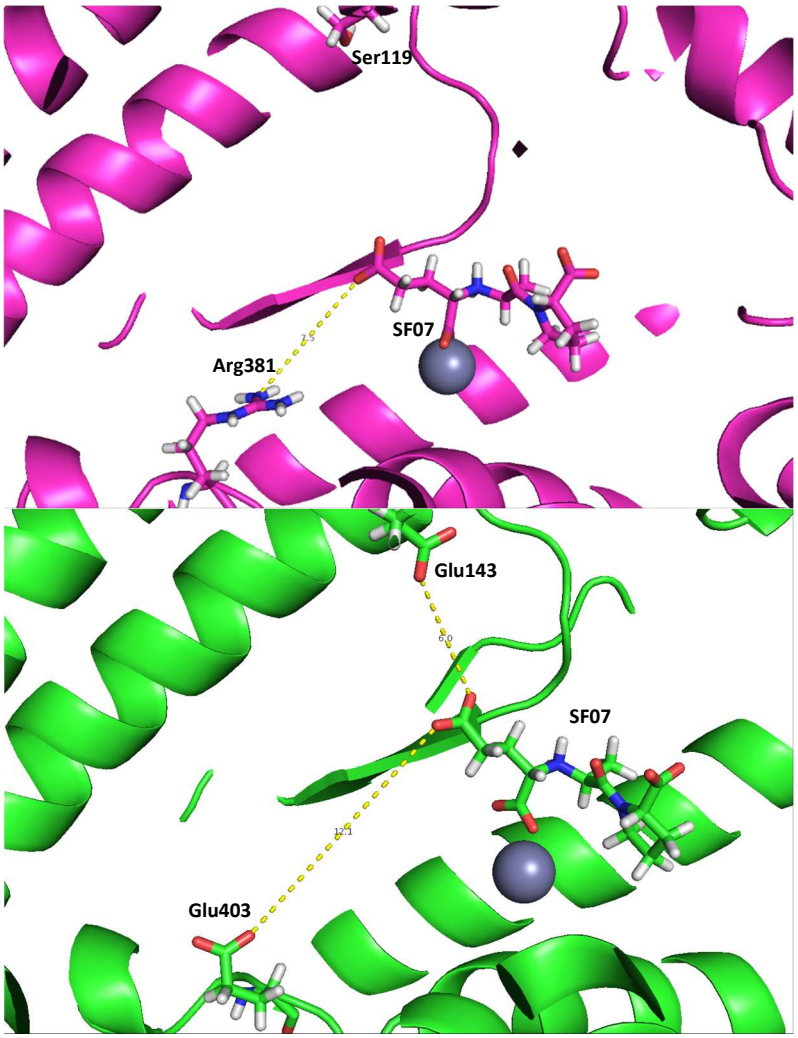


***Figure 4.13**: The docked and MM-GBSA minimised poses of **SF07** in the ACE N- (magenta) and C- (green) domains. The measured distance between the **SF07** P_1 acid and Arg381 in the N-domain was measured at 6.27 Å while the distance between this acid and Glu403 in the C-domain was measured at 12.09 Å.*

With no discernible interaction between **SF07** and Glu403, the interaction responsible for the poor C-domain interaction must lie elsewhere. Judging from the predicted pose, the most likely candidate would be Glu143 in the C-domain, which is replaced with Ser119 in the N-domain. Glu143 lies 5.99 Å from the P_1 acid in the C-domain. This negatively charged residue is now close enough to influence the binding of **SF07** to the C-domain while replacing this residue with a polar-neutral Ser119 in an analogous N-domain position provides a possible explanation for the stronger N-domain binding of **SF07**.

4.5 Discussion

4.5.1 Diastereoselectivity of Substitution Reaction

Despite the lack of any diastereomeric purity in the deprotected **SF07**, the substitution reaction ii appeared to be largely diastereoselective. As already indicated, the Br leaving group of secondary carbons can be substituted via an S_N1 or S_N2 pathway depending on the extent to which a carbocation

intermediate is stabilised by the groups attached to this carbon. In the case of **SF05**, the chiral carbon is bonded to both a methyl and a carbonyl group. While a methyl group has a slight stabilising effect on a carbocation intermediate, the electron-withdrawing carbonyl group destabilises it. Given this combination of groups attached to this secondary chiral carbon, minimal stabilisation of the carbocation intermediate was predicted favouring the S_N2 mechanism. Despite this prediction, a significant proportion of the reaction progressed via an S_N1 pathway to create a hybrid S_N1/S_N2 mechanism. A 50:50 mixture of two ACEi diastereomers will theoretically increase the IC_{50} by roughly a factor of two as observed by Patchett et al during the synthesis and evaluation of Enalaprilat.

Given the apparent high diastereomeric excess of this reaction, a methyl ester deprotection free of racemisation may, by all indications, have yielded the preferred *SSS* diastereomer without changing the diastereomer composition from **SF06**. The expected factor of two order of magnitude drop in inhibition present in a 50:50 diastereomeric mixture around a single diastereomer is not too significant to mask ACE inhibition while constituting 50% of the mass, as demonstrated by Patchett et al. While this substitution reaction has proven to be a diastereoselective alternative to Patchett's reductive amination, the subsequent deprotection negated all efforts of reaction ii to impart diastereoselectivity into the synthesis of this Enalaprilat analogue.

4.5.2 Ester Deprotection

Generic base catalysed ester hydrolysis conditions were used for the deprotection. In an analogous reaction by Greenlee et al,[123] only double methyl esters were deprotected using 0.25 M NaOH. LC-MS reaction monitoring of the **SF06** deprotection at this concentration only showed the presence of the double hydrolysed product. The required triple hydrolysis necessitated incremental increases in NaOH concentration resulting in the observed racemisation.

Replacing the methyl ester protecting groups with a Bn protecting group could provide a viable alternative that avoids racemisation during deprotection. Bn protecting groups have been used successfully in a diastereoselective synthesis of Captopril by Fisher et al.[125] Adapting Scheme 4.1 to use Bn protection groups would require different starting materials.

Patchett et al observed a two-fold decrease in potency when measuring the inhibition of a mixture of *RSS* and SSS enalaprilat diastereomers as opposed to the pure *SSS* diastereomer. Crystal data would suggest a uniform drop in potency across both ACE domains when tested against a diastereomeric mixture. These predictions were confirmed during the inhibition assays.

4.5.3 ACE Competitive Inhibition Assay

The IC_{50} of **SF07** towards the N-domain was predicted to be in the low nanomolar range for a diastereomerically pure SSS stereoisomer. However, the IC_{50} of the given diastereomeric mixture was measured at 5.722 μM, two orders of magnitude greater than the expected inhibition. This is consistent with a mixed diastereomer sample containing 6-8 diastereomers with the SSS diastereomer constituting only a small fraction of the mixture. It was impossible to assign the different diastereomers evident in the ^{1}H NMR without separating them individually. Separating six different diastereomers of a 36 mg sample would be a challenging task given the short HPLC retention time of **SF07**.

The **SF07** diastereomeric mixture inhibited the C-domain with an IC_{50} above 100 μM suggesting a possible two order of magnitude N-domain selectivity. While a diastereomeric mixture cannot be assumed to uniformly weaken the binding towards both domains by the same factor, the specific ligand-protein interactions around these stereocentres need to be considered. There are no structural differences between the two domains present around the three stereocentres in the ACE binding footprint of **SF07** making a uniform drop in binding strength plausible.

4.5.4 Conclusion

These results provide some motivation to rework the synthetic route to avoid racemisation. There is room to expand upon this work by synthesising and testing a SAR series to further refine the structural motifs and chemical space responsible for ACE N-domain selectivity.

SF07 serves as further validation of Dive et al's observation that an interaction with an acid in the S_2 subsite is responsible for N-domain selective ACE inhibition.[56] The longer distance predicted between the P_1 acid of **SF07** and both Arg381 and Glu403 meant the observed N-domain selectivity was unexpected. It was thought that a minimum of one additional alkyl carbon in length was required to effect an interaction with Arg381/Glu403. The observation of N-domain selectivity with the first and only compound tested in this series holds encouraging prospects for future SAR work on this series.

SF07 is potentially the first documented Enalaprilat analogue compound to display ACE N-domain selectivity via an interaction with an acidic S_1 group. A thorough evaluation of the predicted binding pose of this Enalaprilat derivative would require a cocrystal of **SF07** with the two ACE catalytic domains. This would clarify the role that the C-domain Glu143 plays in disrupting the binding of **SF07**.

The predicted distances between the P_1 acid of **SF07** and Glu403/Arg381 suggest that these residues have little to no involvement in the binding of **SF07** to either domain of ACE. Interactions with the

oppositely charged Glu403 and Arg381 residues cannot rationalise the N-domain selectivity observed in the inhibition assay of **SF07**.

A closer inspection of the binding suggested that Glu143 of the C-domain could play a more pivotal role in the disruption of **SF07** binding. Glu143 is replaced with Ser119 in an analogous N-domain position. This substitution has the potential to expand upon the current hypothesis of Arg381/Glu403 as the sole source of N-domain selectivity. The S_1 subsite could therefore also contribute to N-domain selectivity in certain compounds with appropriate P_1 groups.

A P_1 acid would experience no adverse effect on its binding from an interaction with a polar-neutral Ser residue while a Glu in that position has the potential to cause an electrostatic clash with a P_1 acid. The 5.99 Å distance between the P_1 acid of the docked **SF07** and Glu403 is not generally associated with an interaction but is close enough to suggest this acid has been forced into an unfavourable position in order to avoid this residue. It was therefore surprising, in light of the existing ACE domain selectivity data, that **SF07** displayed any N-domain selectivity.

To date, P_1 SAR contributions towards N-domain selectivity have not been documented. **SF07** is an encouraging example justifying a further investigation. Acidic P_1 attachments to ACE inhibitors are better suited in the context of drug-like N-domain selective ACE inhibition as it would allow for compounds of the same length as traditional ACEis. While it is impossible to draw any conclusions from just one example, the **SF07** inhibition data warrants a deeper P_1 ACEi SAR investigation.

Chapter 5 – Database Mining for Old ACEis

5.1 Introduction

5.1.1 Chemical Databases for Drug Discovery

Drug screening and SAR optimisation against drug targets has been in practise for well over a century and while millions of compounds have been patented, the number of compounds synthesised and tested in the process far exceeds this. This research has left behind vast archives of drug-like compounds with their accompanying biological data. The advent of chemoinformatics has necessitated the need to convert these archives into digital chemical databases. These databases have been adapted to suit many different types of software and applications. A noteworthy development from these databases is the invention of the simplified molecular-input line-entry system (SMILES),[103] which uniquely specifies a molecule in a one line string of characters. This one-line entry system has created a low-data method for specifying molecules allowing for the construction of databases with millions of entries.

As already discussed, virtual screening is one of the most common uses of drug discovery databases. Many software applications have been developed to analyse and manipulate these databases according to the requirements of the user. This has led to the development of a new discipline in the field of chemoinformatics termed database mining. Modern software and processors can sort, filter, analyse and run simulations on these databases at a rate many orders of magnitude faster than what can be achieved manually. Databases can now be mined on an unprecedented scale to find specific target-binding patterns.

Database mining may be undertaken using a variety of computational tools. Substructure searching and pharmacophore screening are two of the most popular tools. While these tools help facilitate the processing of immense datasets, they are not a complete substitute for human intuition. Effective database mining protocols usually combine computational tools with manual inspections to effectively utilise the database.

The organisation and indexing of chemical databases can vary depending on the application. As previously mentioned, ZINC is a free database of diverse commercially available compounds for virtual screening.[102] The entries in the ZINC database are indexed with their physicochemical properties to enable the filtering of databases based on a custom set of parameters. Other databases are better suited to mining as each compound is indexed against measured biological data from the original publications in which they were reported. The most common examples of such free databases are

ChEMBL[126] and BindingDB[127]. More extensive subscription databases are available but these charge a large fee for sharing this IP.

5.1.2 ACE and the Drug Discovery Databases

Database mining should prove useful in the search for domain selective ACE inhibitors due to the target's extensive cumulative SAR data. ChEMBL and BindingDB, the open access databases of drug-like compounds indexed with biological data both contain entries for human ACE. BindingDB contains ACE inhibition data on 2252 compounds while ChEMBL contains ACE data on 1154 compounds. The subscription database GVK contains 24 958 entries with some form of biological data pertaining to ACE.

An important aspect of this pre-existing ACE inhibition data is the nature of the assays that were performed. The pioneering competitive inhibition studies were performed on sACE using the HHL substrate.[128] ACE inhibition was evaluated on the vast majority of entries in these databases using this assay. Since the C-domain is responsible for over 90% of the HHL hydrolysis,[3a] it follows that the sACE inhibition of the HHL hydrolysis can be roughly equated to C-domain inhibition data. N-domain selective inhibitors would have likely been missed by the assays since N-domain selective inhibition would leave the C-domain free to hydrolyse the substrate.

This chapter deals with the efforts to search the available chemical databases of compounds containing ACE inhibition data for potential ACE N-domain inhibitors. It is possible that such N-domain selective compounds have remained hidden for decades due to the lack of structural information available in the early days of ACE drug discovery. If an N-domain selective ACEi is hidden in any database, the ACE entries in the GVK database would be an ideal place to look.

5.2 Aims and Objectives

5.2.1 Aim

The Aim of this chapter was to search for potential N-domain selective ACEis among the vast databases of existing sACE inhibitors.

5.2.2 Objectives

- Analyse the set of database compounds, which have been tested on ACE.
- Apply filters to the set of compounds to keep them within the confines of strict chemical parameters conducive to ACE inhibition.
- Run docking simulations on the most promising set of compounds.
- Recommend compounds for purchase/synthesis and *in vitro* competitive inhibition assays.

5.3 Methods

5.3.1 Database Selection

Access to the GVK database was granted through the Novartis Institute of Biomedical Research (NIBR). All the compounds containing biological data on ACE were extracted. The entries pertaining to ACE were compared to those in the open access databases of ChEMBL and BindingDB. GVK was chosen due to its greater size, diversity of compounds and an almost complete overlap with both ChEMBL and BindingDB.

5.3.2 Inspection and Filtering

Once all the entries with biological data pertaining to ACE had been selected, a filter was applied to reduce the database to compounds with a maximum of 15 rotatable bonds. This is a common filter to remove larger biopolymers like peptides to make the set more drug-like. All compounds remaining in this set had been tested on human ACE with inhibition reported in the IC_{50} form. The set was then imported into Schrodinger's Maestro GUI.

The remaining set was separated according to their ZBG using the Maestro filtering tools. Only drugs containing carboxylic acid ZBGs were selected for further scrutiny. The remaining set was visually inspected for esterified prodrugs and peptides. All the esterified prodrugs were manually converted into their bioactive form while peptides which had passed the 15 rotatable bond filter were removed. Duplicate compounds created by de-esterifying the prodrugs were removed using Maestro's remove duplicate tool. Only compounds with polar groups that could potentially interact with the S_2 subsite were selected.

This smaller reduced set was further scrutinised and subjected to a second round of visual inspection. The second round of visual inspection identified core binding motifs in each compound and considered the reported biological data. Only compounds with IC_{50}s in the nanomolar range and a P_2 polar group passed this round of inspections.

5.3.3 Ligand Docking

Ligands which passed the two rounds of visual inspection were all docked into the constrained ligand docking system created in Chapter 3. This set of docking constraints was defined as set C1. Ligands which failed to find a suitable pose using the C1 set of constraints were docked into either one of two newly created constraint systems defined as C2 or C3.

Conditions C2 and C3 were created by introducing three optional positional constraints to the docking grid based on the crystal pose of Lisinopril in each domain (Figure 5.1). The first positional constraint was for an alkyl C or an amine to be placed in a sphere centred on the alkyl carbon of Lisinopril in a

positon adjacent to the ZBG on the non-prime side. This constraint was named α. An analogous positional constraint was added to the system centred on the secondary N of Lisinopril in the other position adjacent to the ZBG but on the prime side. This constraint was named α'. A third positional constraint was placed on the amide N of Lisinopril. This constraint was named ε'. Each constraint was set to only allow poses with either an N or C atom within a 1 Å radius sphere centred on these Lisinopril atoms.

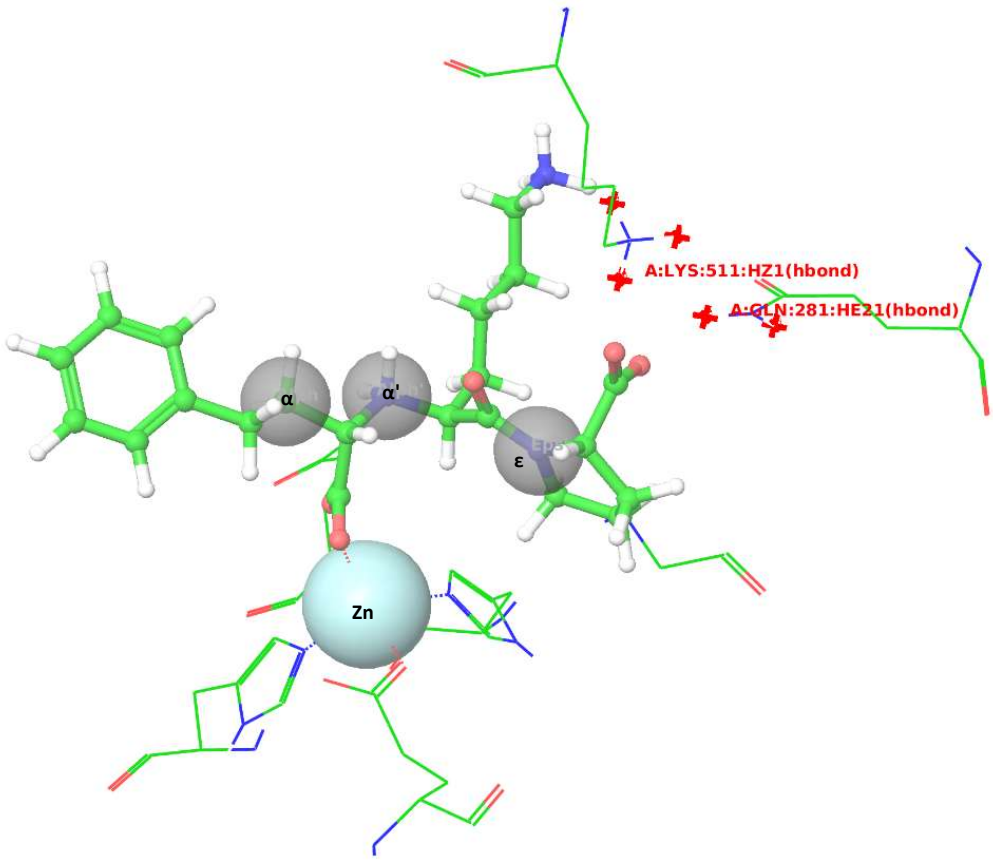


***Figure 5.1**: An illustration of the docking constraints employed against the two domains of ACE. These consisted of three possible positional constraints α, α' and ε', the metal binding constraint and the two H-bond constraints overlaid on the native Lisinopril ligand.*

From these three new positional constraints, two additional sets of docking constraints were defined (Table 5.1). The first constraint condition C1, can be defined as the set of constraints used in Chapter 3. The second condition C2 was defined as a metal-ZBG chelation with the ε' positional constraint sphere containing an amide/amine N. The third constraint condition was defined as the same two constraints as C2 with the addition of positional constraint α, which must contain either an alkyl C or an amine.

***Table 5.1**: A summary of the three docking constraint conditions defined as C1, C2 and C3.*

Condition	Constraints
C1	Carboxylic acid – Zn chelation, H-bond with Lys511/489 or Gln281/Gln259
C2	Carboxylic acid – Zn chelation, amide/amine N in ε' constraint sphere
C3	Carboxylic acid – Zn chelation, amide/amine N in ε' constraint sphere, amine N of alkyl C in α constraint sphere

Once plausible binding poses had been created for each entry on the list, the potential of each compound to interact with Arg381/Glu403 was assessed. Compounds from the same chemical series were grouped together. The series that were most likely to interact with Arg381/Glu403 were then considered for *in vitro* testing in competitive ACE inhibition assays.

5.4 Results

5.4.1 Database Mining Overview

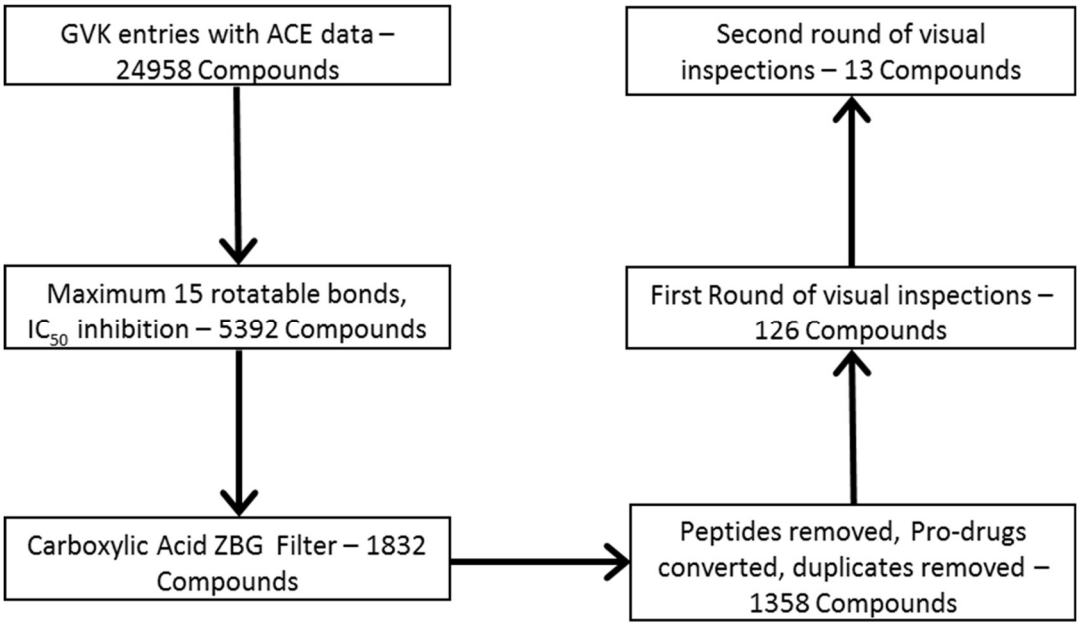


***Figure 5.2:** A schematic outlining the database mining protocol.*

Figure 5.2 illustrates an overview of the procedure used to filter this database and identify the most promising N-domain selective ACEis. To begin with, a total of 24958 compounds in the GVK database were found to have been tested on ACE, many of which were peptides. This set was reduced to 5392 compounds after applying a rotatable bond limit of 15 and limiting the selection to only include

compounds with IC_{50} inhibition data. The set was then reduced to 1832 compounds after it was restricted to compounds containing only a carboxylic acid ZBG. The remaining set contained many esterified prodrugs and small peptides, which had escaped the rotatable bond filter. After removing these small peptides, de-esterifying the prodrugs and removing the duplicates, 1358 compounds remained. The first round of visual inspection reduced it to 126 compounds of which only 13 passed the second round of inspection for docking.

5.4.2 GVK Database Sorting of ACE inhibitors

A total of 5392 compounds in the GVK database had been tested on human sACE in a competitive inhibition assay with their IC_{50} values reported. The majority of these IC_{50} values were obtained via competitive inhibition against a HHL substrate. However, different substrates were used in some of the more recent studies. There was a much smaller subset of compounds with reported K_i values but these were later all found to be duplicates of compounds in the IC_{50} subset. This set of 5392 compounds was investigated further.

Upon rendering the SMILES into 2D structures, it was evident that the three main classes of ACE ZBG (carboxylic acid, thiol and phosphinic acid) were well represented. Many peptides were still present in addition to the traditional ACEis. The majority of the peptides were removed by the maximum 15 rotatable bonds filter. Since thiol ZBGs have been shown to be prone to toxicity issues[129] and phosphinic acids suffer from poor permeability, only compounds with carboxylic acid ZBGs were retained. This reduced the set down to 1832 compounds.

A set of 1832 compounds is usually small enough to dock. However, this was not possible using ACE as a target. No blanket set of constraints could find plausible poses for a diverse set of compounds when docked against ACE. Test runs were performed and on average, 45% of the set could be docked plausibly with any one set of constraints. A low docking accuracy is unacceptable for evaluating possible domain selectivity within this set of known ACEis.

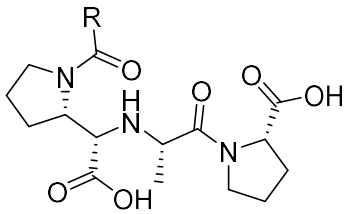
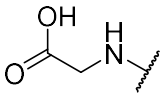
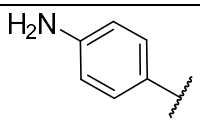
5.4.3 Visual Inspection and Manual Filtering

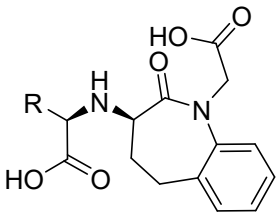
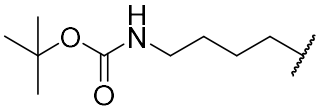
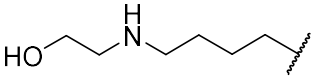
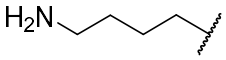
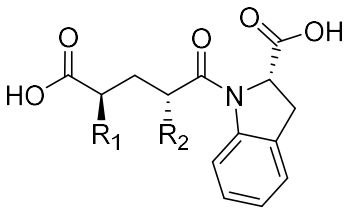
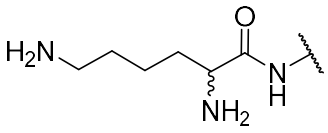
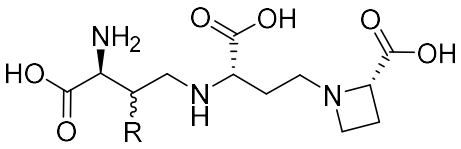
Upon first inspection of the remaining set of 1832 compounds, it was apparent that many small di- and tripeptides had escaped the 15 rotatable bond filter. Another concern was the presence of many prodrugs in the set with at least one ester group. These remaining peptides were removed and the prodrugs were de-esterified. The de-esterification created a few duplicates, which were easily removed using Maestro's remove duplicate tool. This step reduced the set to 1358 compounds before the visual inspection could officially commence.

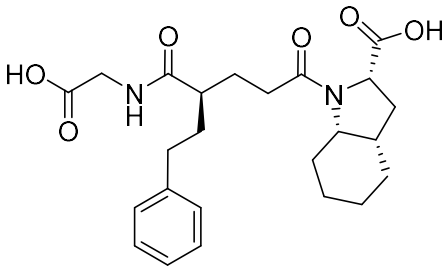
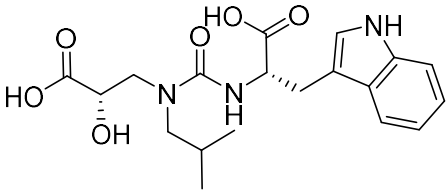
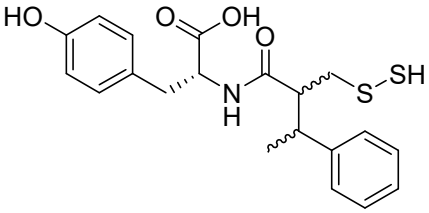
While a pharmacophore filter with an Arg381 interaction pattern was considered, the dataset was now small enough for visual inspection. To ensure adequate ACE binding, compounds with a prominent P_2' carboxylic acid and a known P_2' moiety of Pro, Trp or Phe analogous structure were prioritised. With an ACE binding core established, compounds containing a polar group with the potential for a polar interaction with Arg381/Glu403 were selected from the set.

Drug-like compounds with the most promising P_2 moieties were selected for docking. The set considered for docking consisted of 13 compounds. Of the 13 compounds, 10 compounds belonged to one of four series (Table 5.2): Series 1 – Diprolyl; Series 2 – Benazeprilat analogues; Series 3 – Quinaprilat analogues; Series 4 – Nicotianamine natural product derivative. Three of the series (Series 1-3) were related to existing ACEis while the fourth consisted of two natural products isolated from buckwheat. These 13 compounds were then docked against both domains of ACE.

***Table 5.2**: The compounds of interest which were selected from the GVK database for docking studies.*

Entry	R	sACE IC_{50} (nM)
	Series 1[123]	
		
DM1		6
DM2		6
DM3		5.4
	Series 2[130]	

		
DM4		4.0
DM5		7.0
DM6		40
Series 3[131]		
		
DM7	$R_1 = CH_3$ $R_2 = CH_3$	10
DM8	 $R_1 =$ $R_2 = H$	188
Series 4[132]		
		
DM9	OH	80
DM10	H	85
Ungrouped		

DM11		10
DM12		103
DM13		52

5.4.4 Molecular Docking of Selected Compounds into the N-domain

Thanks to an extensive collection of ACEis co-crystalised with both catalytic domains, the most important ligand-target interactions are well understood. The P_1' group is always an amino acid or pseudo amino acid residue with a carboxylic acid. The most common P_2' moiety is a Pro residue but Trp, Phe and many artificial peptidomimetic non-polar rings and fused aromatics have been documented to bind to this subsite. This outline for ACEi binding gives a good indication of what docking poses to expect, making it easy to identify when the docking protocol has found a plausible pose.

Unlike the Enalaprilat-derived combinatorial library docked in chapter three, these compounds of interest are not structurally homologous. As expected, no one set of constraints was found to be capable of docking all 13 of these ligands in plausible poses. To work around this problem, three different constraint conditions (C1-C3) were tested to find convincing or reasonable poses for the diverse set of ligands.

In the N-domain, plausible poses for compounds DM1-9 were obtained using constraint conditions C1 while plausible poses for DM9 and DM10 were obtained using constraint conditions C2. Acceptable poses for DM11 and DM12 could not be obtained using any of the three constraint conditions. This is

likely due to the bulky P_2' indole or saturated bicyclic moiety. Such bicyclic P_2' groups have been shown to cause selective inhibition of the C-domain.[58]

When docked in the C-domain, acceptable poses were obtained for all compounds except DM6 using one of the three constraint conditions (Table 5.3). The poor docking results for DM6 were probably due to the presence of a lysine moiety. Long alkyl chains push the limits of docking algorithms due to the many degrees of freedom and the raised entropy of the system with the addition of each rotatable bond in an alkyl chain.

Once each compound was docked, the distance between Arg381/Glu403 and the nearest heavy atom on each ligand was measured. If the predicted pose is plausible and the ligand falls within 5 Å of these residues, an interaction with either of these residues can be expected to make a significant contribution towards ligand binding.

***Table 5.3**: A table cataloguing the docking constraint conditions required to help each mined ligand find its correct binding pose. The Proximity between the ligand and either Arg381 in the N-domain and Glu403 in the C-domain in each correct pose is given.*

Entry	N-domain	Distance from Arg381 (Å)	C-domain	Distance from Glu403 (Å)
DM1	C1	2.85 (Salt-bridge)	C3	5.93
DM2	C1	5.62	C1	7.59
DM3	C1	4.05	C1	5.09
DM4	C1	9.10	C1	6.69
DM5	C1	9.74	C2	12.47
DM6	C1	8.19	-	
DM7	C1	10.18	C1	12.25
DM8	C1	8.75	C2	9.52
DM9	C2	6.01	C2	9.47
DM10	C2	6.01	C2	9.47
DM11	-	-	C2	7.64
DM12	-	-	C1	9.86

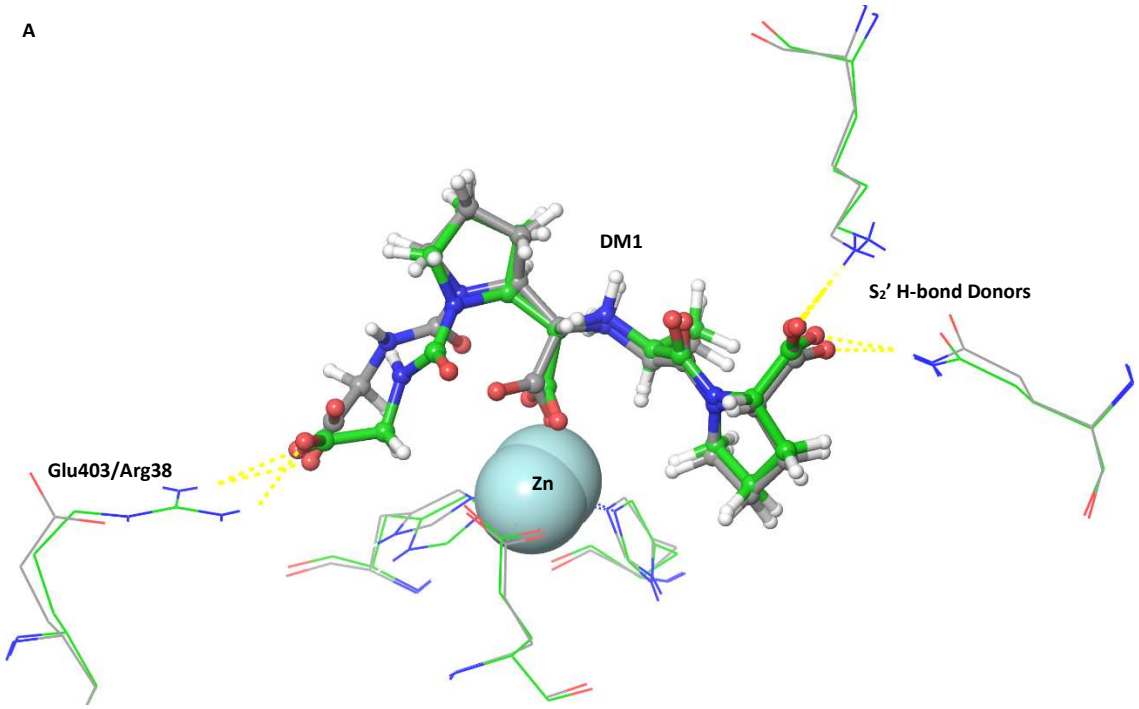
DM13	C1	6.54	C1	9.99

5.4.5 Docking Poses

Figure 5.3 illustrates a few docking poses achieved for some of the ligands from the set of 13 ACEis. In Figure 5.3A, **DM1** found its core H-bond interactions in the S_2' subsite. In the S_2 subsite, the terminal carboxylic acid forms a salt bridge with Arg381 in the N-domain while in the C-domain it appears that this acid has been held in place by the constraints have forced it into raised energy conformation to avoid an interaction with Glu403.

In Figure 5.3B, the same Zn chelation and S_2' H-bonding interactions are observed for **DM10**. Its terminal carboxylic acid lies 5.10 Å from Arg381, too far away for an interaction. Lengthening the compound by one or two alkyl atoms could probably help it reach this residue.

Figure 5.3C shows an example of a compound which is too large and cumbersome to accurately dock. Long alkyl chains add too many degrees of freedom to the simulation, hence the docking algorithm struggles to correctly place the carboxylic acid in the S_2' subsite. It also shows that this Benazepril-like core is only capable of getting close to the Arg381/Glu403 residues at the end of a long alkyl chain whose conformation cannot be predicted by Glide.



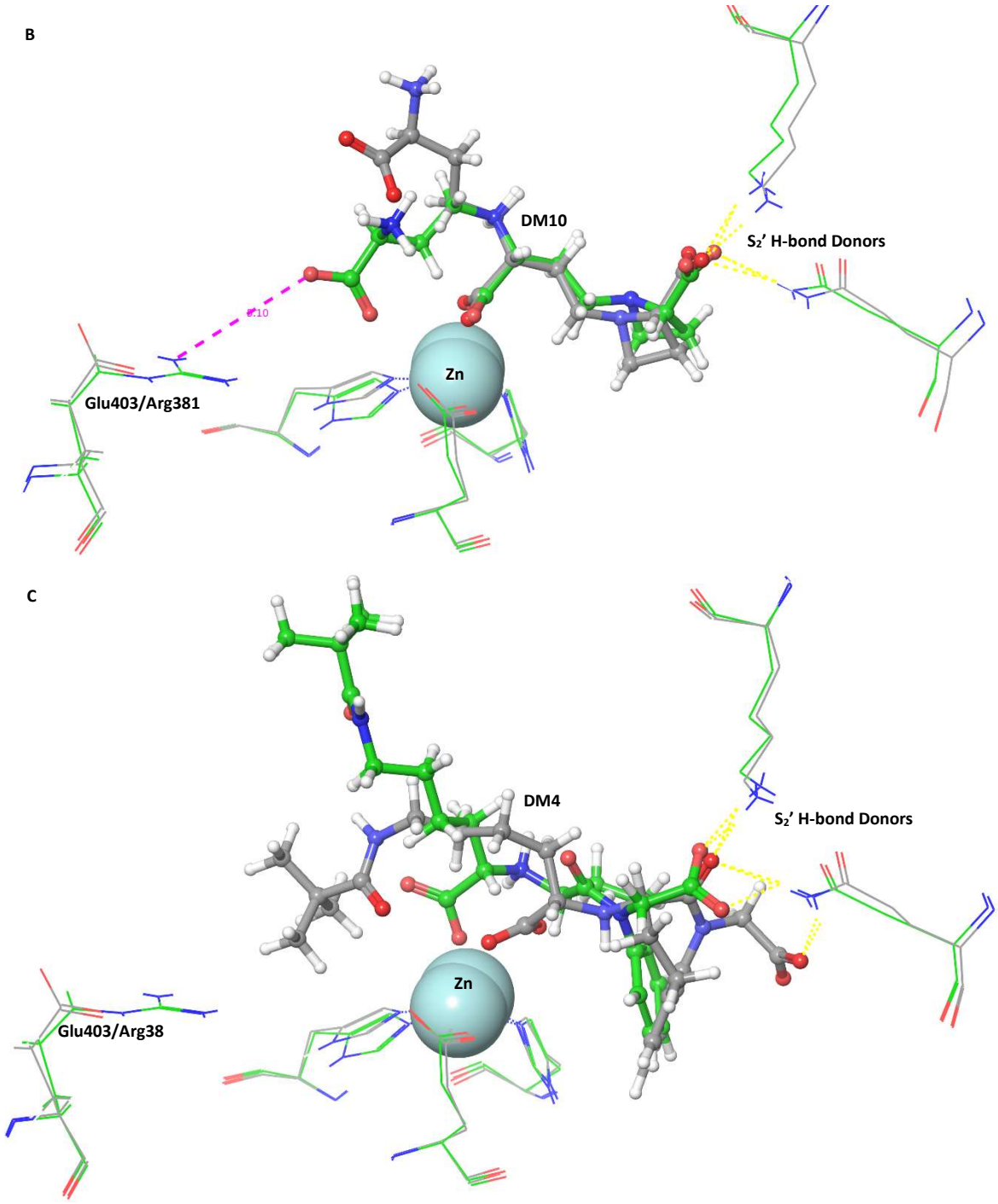


***Figure 5.3**: Sample poses of **DM1** (A), **DM10** (B) and **DM4** (C) docked into the N-(green) C-domains (grey). Each image depicts an overlay of selected N- and C-domain residues around the central Zn atom. The two key features are a carboxylic acid Zn chelation and a second carboxylic acid acting as an H-bond acceptor in the S_2' subsite. All H-bond interactions are indicated with a yellow dotted line. The P_2 group of each ligand is then allowed to freely interact with Arg381/Glu403 in the S_2 subsite. Figure A shows a near perfect example of a N-domain selective pose with an ionic interaction between the charged species of a carboxylic acid and Arg381. The ionic interactions between this acid and Glu403 in the C-domain creates an unfavourable torsion to avoid this residue. Figure B shows a ligand which almost makes the desired interactions with Arg381/Glu403 but ends up 5.10 Å short of these residues. Figure C illustrates the issues associated with docking large compounds with long alkyl chains. This ligand almost makes interactions with Arg381/Glu403 but the entropy of the Boc-Lys is too high to accurately predict the binding pose.*

5.4.6 ACEi Series with Potential for N-domain Selectivity

Considering the docked poses of all the selected compounds, some compounds stood out as promising domain selective inhibitors. The most promising series by a considerable margin is Series 1 (Table 5.1) reported by Greenlee et al.[123] Series 1 is an ideal scaffold with which to probe Interactions in the S_2 subsite. Pro-ZBG-Ala-Pro is known to bind tightly with S_1, Zn, S_1' and S_2' subsites from peptide studies.

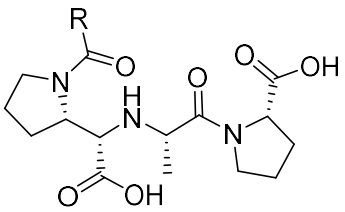


***Figure 5.4**: Series 1 scaffold*

The best compound from this set (**DM1**) easily fitted into the binding pocket of the N-domain forming a salt-bridge with Arg381 (Figure 5.3A). Since the R-group in this series only varies around a peptide bond, a large range of fragments can be attached at this position. Of the three compounds selected from the database, **DM1** contained an acid in this position.

The second most promising series was series 4 (Figure 5.1). This series was derived from a natural product extraction performed by Aoyagi et al.[132] They found two compounds with a cyclobutyl analogue of the terminal Pro observed in some ACEis. These two compounds were selected due to their strong ACE inhibition and the crucial third acidic group which could interact with Arg381/Glu403. Unfortunately, docking simulations showed these two compounds to be too short to interact with Arg381/Glu403 as well as the Zn and the S_2' group. Another point of consideration was that these two compounds were derived from natural products for which a synthetic pathway has not been established.

Series 2 and 3 seemed interesting as the long alkyl chains emanating from their core structures seemed capable of reaching Arg381/Glu403. Upon closer inspection, it became apparent that only a lys attachment gave the ligand the ability to reach these residues. Deconstructing these compounds back to that point of attachment leaves a distance of over 10 Å between the scaffold and the Arg381/Glu403 residues. This distance is too great to access with the simple synthetic introduction of a new chemical group.

5.5 Discussion

5.5.1 Mining Protocol

There are many approaches to follow in a database mining exercise. Database mining protocols are developed on the merits of the target and the available ligands. Finding N-domain selective ACE inhibitors posed a unique challenge due to the wealth of existing sACE inhibition data. Since crystal structures have almost conclusively shown the Arg381/Glu403 mutation to be one of the origins of N-domain selectivity, an effective N-domain inhibitor needs to retain its core ZBG and P_2' carboxylic acid before it can interact with Arg381/Glu403. These stringent requirements left a very narrow window of permissible structural features. The reduction of the set of compounds down to 1832 after the appropriate filters left a set which was suitable for manual inspection. While much consideration was given to constructing a pharmacophore model comprising of the core ACE binding motifs, a set of 1832 compounds was small enough to inspect manually. A visual inspection also ensured that none of the interesting compounds would slip through undetected due to a flawed automated screen.

The two thorough rounds of visual inspection returned 13 interesting compounds from this set. Not all of them had an acid in the P_2 position but they all showed some potential for a polar interaction with the S_2 subsite. Faced with the twin challenges of poor Zn binding predictions and a set with no common scaffold, finding plausible binding poses via docking required some innovation. The three constraint conditions were devised to test a range of ligand-binding motifs observed in crystal structures with the hope that at least one combination of constraints could help guide these ligands into a plausible pose. It was also important to give Glide enough freedom to ensure that the poses it returned were indeed minimised. As expected, docking with different constraints yielded varying degrees of success against the two domains of ACE.

5.5.2 Docking Overview

The C1 set of constraints found a plausible pose for 9/13 inhibitors in the N-domain and 6/13 inhibitors in the C-domain. The ligands for which a plausible pose could not be found in the N-domain belonged to either series 4 or were ungrouped. The docking algorithm likely struggled to find the correct poses for compounds in series 4 under these conditions because the distance between the ZBG and the cyclobutyl is one bonded C less than the analogous length in Enalaprilat. Its predicted poses consequently appear to be more strained. Series 4 is of a fundamentally different shape to the other ACEis and peptides but its ACE inhibition is reported to be in the low nM range. It is likely that the protein experiences small side-chain shifts when binding to this ligand compared to the peptides and peptidomimetics. It is therefore possible that the final pose of series 4 is compromised as the side-chains around the binding site are optimised to fit a peptidomimetic compound.

Compounds **DM11** and **DM12** likely experienced difficulty docking into the N-domain active site on account of their bicyclic P_2' groups in a manner similar to **RXPA380**. A more polar S_2' subsite in the N-domain is a possible explanation for the failure to find the pose for these compounds using any of the tested combination of constraints. The difficulties in finding the correct pose using the docking protocols can be interpreted as evidence of possible C-domain selectivity.

The constraint conditions, which helped a ligand find the correct pose in the N-domain did not always work for the same ligand in the C-domain. A prime example of this is ligand **DM1**. Judging by its salt bridge with Arg381 in the N-domain, its Glu403 interaction was expected to have a detrimental effect on C-domain binding. A consequence of this poor interaction was that Glide failed to find a plausible pose with this orientation in the C-domain using conditions C1 or C2. Only in the most constrained system C3, was a plausible pose found but with unfavourable torsions around the P_2 group. The unfavourable torsions seem to indicate that in a relaxed system, the protein conformation would be slightly altered to better accommodate the clash between the P_2 acid and the Glu403. The ease with which Glide found the correct pose for this ligand in the N-domain compared with the C-domain is further evidence to suggest favourable N-domain binding. The other two compounds from series 1 (**DM2** and **DM3**) easily found the correct pose in the C-domain using the C1 constraint conditions on account of the more agreeable Gly and 4-amino P_2 groups.

There were also some difficulties docking Series 2 into the C-domain. These difficulties can be attributed almost entirely to the existence of a long alkyl chain challenging the limits of Glide's simulation due to excessive rotatable bonds in a high entropy alkyl chain. Docking series 4 into the C-domain and N-domain yielded similar poses. Its P_1 acid was not capable of getting close enough to Glu403 to make a meaningful interaction. Compounds **DM11** and **DM12** were able to dock into the C-domain with either C1 or C2 constraint system, which was not possible with the N-domain. This again suggests that P_2' bicyclic groups bind favourably towards the C-domain.

5.5.3 Compounds Overview

After a thorough analysis of each docking pose for this selected set of potentially domain selective ACEis, the most promising compound is unequivocally compound **DM1**. This compound has also revealed a very important chemical series as it supplies a versatile molecular framework with which to explore the P_2 SAR of an ACEi. The second most promising series was series 4 but as already mentioned these compounds were unable to interact with Arg381/Glu403. At the same time natural products limit the synthetic options for probing P_2 SAR.

Upon docking series 2 and 3, it is clear that neither series had much potential to interact with the S_2 subsite. Apart from their inability to access the S_2 subsite, their bicyclic P_2' groups may be detrimental to N-domain selectivity ironically resulting in C-domain selectivity. As for the ungrouped molecules, **DM11** and **DM12** also contain bicyclic P_2' groups and **DM13** contains a disulphide group with no novel leads regarding S_2 interactions.

Examining Table 5.3, the docking protocol appears to treat each ligand in one of three ways. A ligand can easily find a plausible pose in both domains, the N-domain only or the C-domain only. **DM1**, the compound with the most overtly N-domain selective motif easily found a plausible pose in the N-domain but not the C-domain. The converse was observed with compound **DM11**, which carries a known C-domain selective P_2' bicyclic moiety. While this is not conclusive evidence of domain selectivity, it is an indication. The docking protocol ran into difficulty evaluating the poses of compounds with long entropically disordered chains (series 2) and compounds, which did not fit the binding pattern embedded into the crystal pose of the protein by native peptides or peptidomimetic ligands (Series 4).

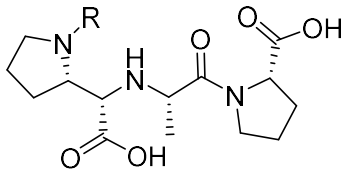
5.5.4 Potential for N-Domain Selectivity

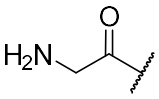
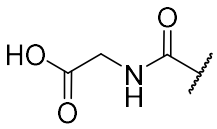
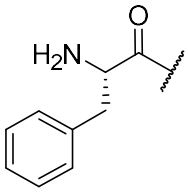
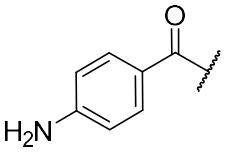
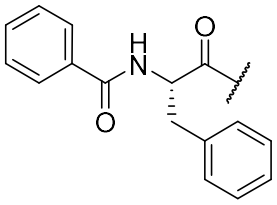
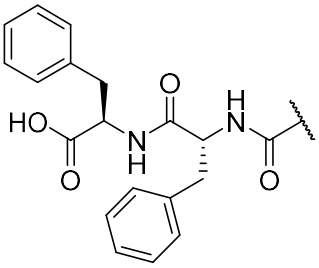
Since series 1 was overwhelmingly the most promising series, it warranted a closer look. Greenlee et al synthesised two series of ACEis using the Enalaprilat prime side scaffold of ZBG-NH-Ala-Pro-COOH scaffold. Series 1 follows a R–Pro-ZBG-NH-Ala-Pro-COOH format and seven of these compounds are described in Appendix 5.1[123]. The groups attached were an acetyl, Benzyl, Gly, Phe, BnPhe and pGlu-Lys. The seven compounds from this series did not match the 11 compounds reported to belong to this series in the GVK database. The four unaccounted compounds contained IC_{50} data matching compounds within this set. Upon closer inspection of the reported structures and a reconstruction of the synthetic scheme, it became clear that these four compounds were duplicates of compounds from the set of seven entered into the database using incorrect SMILES containing small errors.

Comparing Greenlee's reported structures and the rendered structures, it was easy to determine which structures where incorrectly reported (Table 5.4). Unfortunately, mined compounds **DM1** and **DM3** correspond to incorrect renderings of compounds **34** and **36** respectively from the Greenlee publication. Even more disconcerting is the fact that the most promising mined compound (**DM1**) is not a published compound as confirmed by SciFinder. The only additional mention of this compound was found in the BindingDB database again incorrectly recorded and referencing the same Greenlee paper.

Table 5.4: *An illustration comparing the correct entries as reported by Greenlee et al[123] and the incorrect renderings of these compounds found in the GVK database. The numbering scheme used is the same as in the cited paper. The incorrect entries*

column compares the compounds synthesised (left) vs incorrect renderings of these compounds (right) found in the GVK database.



Compound Number	Correct Structure	Incorrect Rendering
34		
36		
38		

5.5.5 The Improvisation of Series 1

Despite the most promising hit from Series 1 not being a known compound, the R-Pro-ZBG-NH-Ala-Pro-COOH still appears to be an excellent platform for probing S_2 SAR of ACE. A simple peptide coupling can add a multitude of carboxylic acid containing groups including the acidic amino acids Glu and Asp. Coupling Glu and Asp to this scaffold would place a carboxylic acid in the ideal position to interact with Arg381/Glu403. Despite the failure of this database mining exercise to find a promising N-domain selective ACEi, it has found a promising new drug-like scaffold with which to probe the very SAR under investigation.

Series 1 is an ideal example of a chemical series synthesised in the early days of ACEi development using an sACE competitive inhibition assay against the ZFHL substrate to measure ACE inhibition. Had crystal structures of the two domains been known back then, it is possible that a more comprehensive

SAR study in the S_2 subsite would have been undertaken. The seven variants of Series 1 explored a very limited chemical space using P_2 substitutions.

The P_2 SAR of this series is now understood to be key to probing the key interactions responsible for inducing N-domain selectivity. With just three compounds from this series docked, **DM1** made a polar salt bridge interaction with Arg381 while showing a strong repulsion towards Glu403. **DM2** appeared to form a favourable π-stacking interaction with Arg381 while the amine group of **DM3** was positioned close enough to these residues to interact with these residues (Figure 5.5).

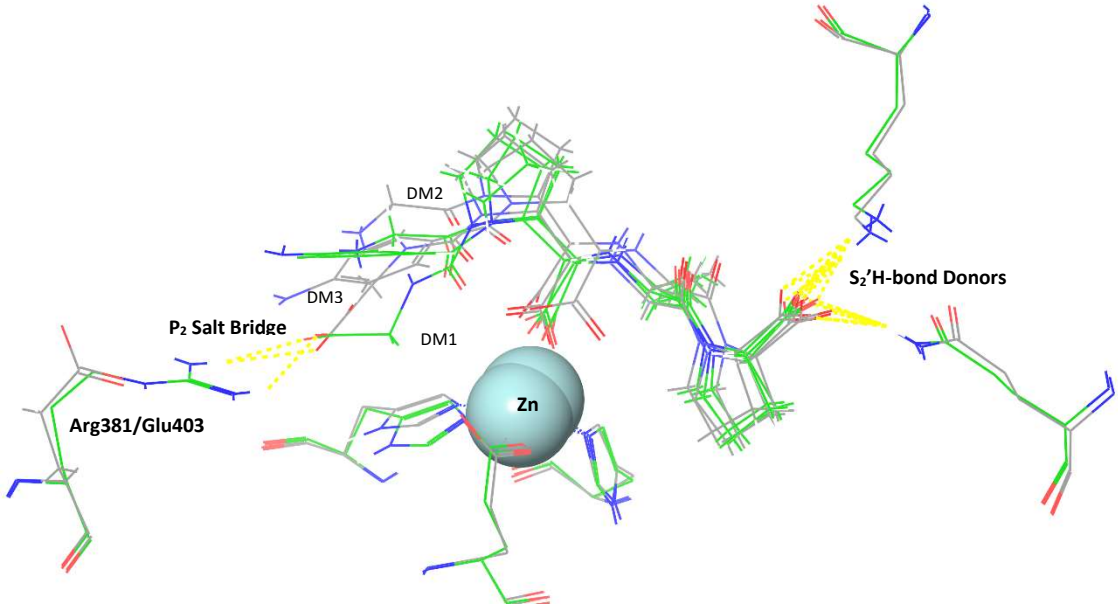


***Figure 5.5**: An overlay of all the series 1 compounds docked into the N- (green) and C- domains (grey). The common backbone for each compound is held firmly in place with the central carboxylic acid ZBG and the P_2' carboxylic acid H-bonding to the S_2' subsite. All observed variation occurs in the P_2 position where each group interacts differently with Arg381/Glu403.*

5.5.6 Concluding Remarks

While finding an N-domain selective ACEi mining the GVK database was never guaranteed, it has provided an opportunity to rediscover older ACEi binding patterns. ACE is a special case since a vast database of inhibition data exists from decades of drug discovery research. Revisiting these archives in search of N-domain-selective inhibitors added much more stringent chemical parameters to the requirements of the inhibitor. A thorough search for the chemical signature responsible for N-domain selectivity was conducted and turned up a meagre 13 compounds with the potential to make polar contacts in the S_2 subsite. Out of the 13, only one chemical series showed promise.

Unfortunately, the best compound from series 1 (**DM1**) turned out to be an undocumented compound rendered as a result of an incorrectly archived entry in the GVK database. **DM1** is structurally similar to two compounds which can be synthesised via simple modifications to the published synthetic scheme and will be covered in Chapter 6.

Chapter 6 – Synthesis of the Diprolyl Inhibitor Series

6.1 Introduction

6.1.1 Diproline Series

The database mining protocol followed in Chapter 5 failed to identify potential N-domain selective ACE inhibitors. Despite the failure to find a potent N-domain-selective inhibitor, Greenlee's diprolyl series (Figure 6.1) did show promise as reported.[123] The original kinetic data and docking studies showed the diprolyl backbone to interact strongly with the S_1, Zn, S_1' and S_2' subsites of both ACE domains. The last step in the synthesis of this series attaches a unique P_2 group to the scaffold. Greenlee et al attached seven different groups to this position, none of which contained the requisite P_2 acidic group required to selectively compromise C-domain binding. Following Greenlee's synthetic approach (scheme 6.1), it seemed feasible to attach an acidic group at the P_2 position in the last step of the synthesis.

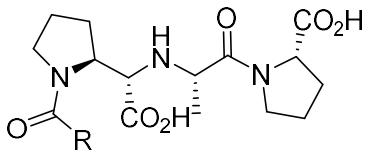


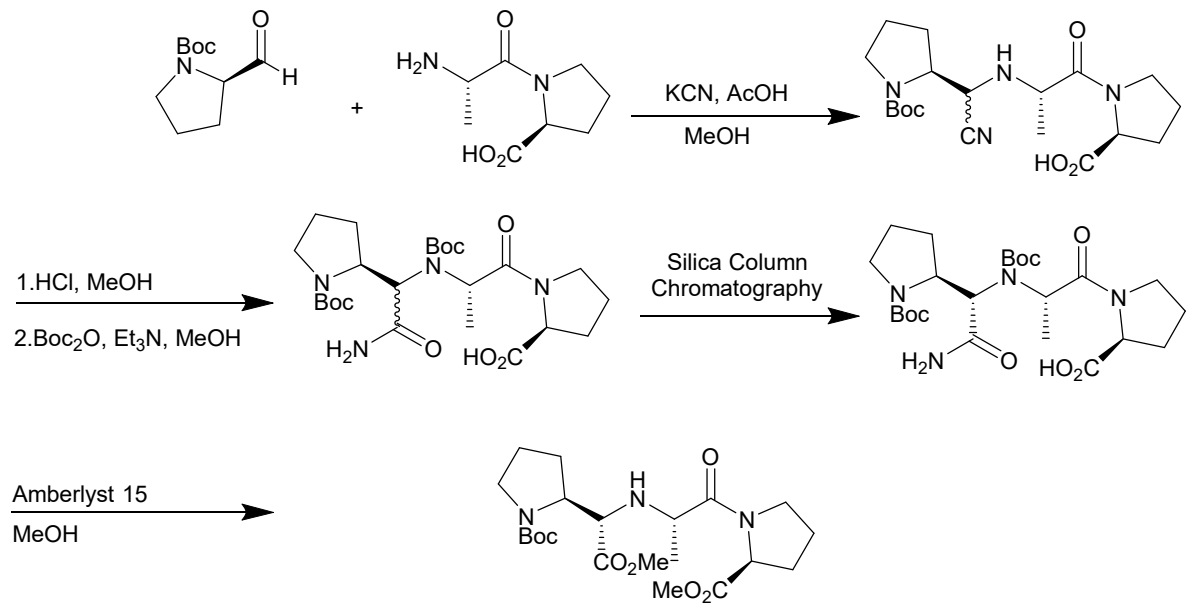
***Figure 6.1**: Greenlee et al's diprolyl scaffold, an ideal molecular scaffold for exploring ACE S_2 SAR.[123]*

The seven compounds in this series were tested *in vitro* using competitive ACE inhibition assays (

Table 6.**1**). It was from this series that the most promising compound **DM1** was identified. As already mentioned, **DM1** was shown to be the result of an erroneous entry of compound **34** in the GVK database. Despite the error, the diprolyl series appeared to be an ideal platform with which to probe SAR in and around the S_2 subsite of ACE. This series provides a promising drug-like scaffold with which to exploit the Arg381/Glu403 mutation elucidated by Kroger et al[77] in the crystal structure 3NXQ.

6.1.2 Synthetic Scheme

As discussed in Chapter 3, Enalaprilat is synthesised via the reductive amination of an α-keto acid and Ala-Pro using sodium cyanoborohydride. The diprolyl series was synthesised via an analogous scheme (Scheme 6.1). Scheme 6.1 modifies the reductive amination step used by Patchett et al[113] coupling a Boc protected pyrolidine aldehyde to Ala-Pro in the presence of KCN. This step creates a new chiral centre around the introduced cyano group yielding equimolar quantities of the two diastereomers around the new chiral centre. The cyano group is then converted into an amide after which the unprotected amine is Boc protected. An acidic Amberlyst 15 ion exchange resin then converts the amide and the acid groups into methyl esters. The resulting product is the diproline scaffold with methyl ester protected acids. Seven different groups were added to the diprolyl amine via an amide coupling. The amide was formed using either *N*-hydroxy succinimide (HOSu) or *N*,*N*'-dicylcohexylcarbodiimide (DCC) reagents.



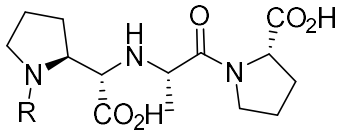
Scheme 6.1: *An outline of the synthetic scheme by which Greenlee et al[123] synthesised the diprolyl Series 1*

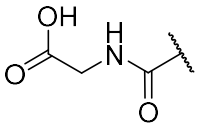
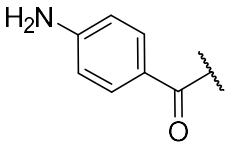
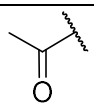
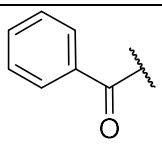
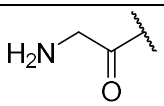
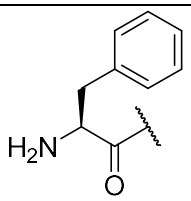
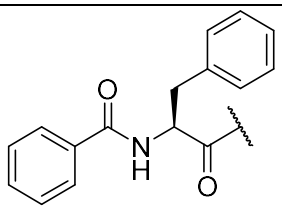
6.1.3 Revisiting the Series

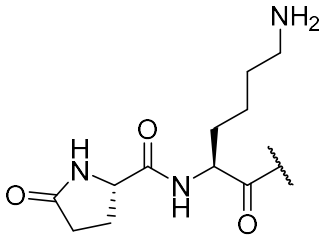
Of the seven compounds published in this series (Table 6.1), not all were relevant for studying Arg381/Glu403 interactions. In this set only compounds **32**, **34** and **36** are of a suitable size to probe

the S_2 subsite. Three compounds from this series were docked in Chapter 5, of which only **DM2** was a published compound corresponding to compound **34**. **DM1** contains a urea group which is a chemical deviation from the published synthesis. **DM3** can be synthesised using this scheme via the amide coupling of 4-amino benzoic acid to the scaffold. Docking simulations on these three compounds suggested Glu and Asp attachments to the scaffold could create a highly favourable interaction with Arg381/Glu403 residues.

Table 6.1: *The three compounds extracted from the GVK database mining exercise (prefixed with DM) and the seven compounds in the Diprolyl series synthesised by Greenlee et al with the reported sACE IC_{50} and the original numbering from the paper.*



Compound	R	sACE IC_{50} (nM)
DM1		6
DM3		5.4
28	H	700
30		9.4
32		2.9
34, **DM2**		6.0
36		5.4
38		2.9

41		8.5

While providing a strong platform for investigating S_2 interactions, this scaffold is not particularly drug-like. Two amide bonds together with an acid and amine terminus make this scaffold too peptidic. A strong hit from this series can always be modified at a later stage to better adapt into a true peptidomimetic. Another point of contention is the high level of chirality with up to five chiral centres present in these compounds. This level of chirality can create synthetic complications when it comes to obtaining absolute diastereomeric purity. Such chirality is characteristic of peptides and protease targets. As such, mimicking the chirality of this natural peptide is crucial for strong ACE inhibition.

6.1.4 Competitive Inhibition Assay

ACE inhibition assays have been performed since the 60s[133]. Since then, both protein expression and inhibition assays have seen considerable advancements. The first spectrophotometric assays measuring the rate of ACE hydrolysis on synthetic peptides were developed independently in the early 70s by both Roth et al[133] and Cushman et al[128]. Roth's assay made use of a Cbz-Phe-His-Leu (ZFHL) substrate while Cushman's assay used Hip-His-Leu (HHL). Cushman later went on to develop Captopril with Ondetti[17] in 1977 hence the greater prevalence of HHL assays in the early days of ACEi research.

Both ZFHL and HHL are N-terminal capped peptides with the same HL dipeptide product of ACE hydrolysis. The rate of hydrolysis for both substrates can be determined by quenching the reaction with NaOH after a fixed time interval and derivatising the product with O-pthaldehyde. The uncapped N-terminus of HL forms a fluorogenic complex under basic conditions while the concentration of this complex can be measured fluorometrically. K_m values for HHL and ZFHL sACE hydrolysis are 2.6 mM and 40 µM respectively which are both suitable for competitive inhibition assays.[133]

Since the elucidation of the two domains of ACE, the K_m of HHL has been shown to be roughly equivalent in both domains while for ZFHL the K_m of the N-domain is roughly five times that of the C-domain at 0.93 and 0.18 mM respectively[134]. With the two domains assayed separately, no competitive hydrolysis is experienced between the two domains, hence the choice of substrate for assaying the two domains individually is irrelevant.

An alternative to these quenched fluorometric assays is a continuous fluorogenic assay developed by Araujo et al[135] using the internally quenched Abz-FRK(Dnp)P-OH substrate. This substrate displays similar kinetic properties towards the two catalytic domains of ACE. Despite the benefits of a real-time assay, Abz-FRK(Dnp)P-OH is considerably more expensive than both ZFHL and HHL making them more appropriate for broad based inhibitor screening assays.

6.2 Aims and Objectives

6.2.1 Aim

The aim of this chapter was to probe the P_2 SAR of the Greenlee et al's diprolyl series for N-domain selective ACE inhibition.

6.2.2 Objectives

- Design a series of diprolyl ACE inhibitors with different P_2 groups.
- Carry out molecular docking of the diprolyl series against the N- and C-domains of ACE.
- Synthesise a small library of diprolyl derivatives using the approach of Greenlee et al.
- Determine the IC_{50} values of all the compounds for the N- and C-domains of ACE.

6.3 Methods

6.3.1 Redocking and Expanding Dataset

Compounds **32**, **34** and **36** from the Greenlee publication were docked into the two catalytic domains of ACE using Glide XP with the C1 set of constraints and the docking grids created (Chapter 5; section 5.3.2). Once satisfactory binding poses were determined, a search for chemical monomers with a potential for peptide coupling to the diprolyl scaffold were investigated. The availability, price and likelihood to induce a strong Arg381/Glu403 interaction were considered. A small VSAR set of compounds was constructed using some of these building blocks. New variations to this scaffold were then proposed, docked and evaluated.

6.3.2 MM-GBSA Binding Energy Predictions

Once satisfactory docking poses had been determined for each compound in the extended set, the binding energy of each pose was evaluated using the Prime MM-GBSA binding energy prediction protocol described in Chapter 3 with the calculation preceded by a minimisation of all residues within a 10 Å radius of the ligand.

Using the predicted binding energy and final pose as a guide, binding hypotheses were made to predict the domain selectivity of each compound in this extended set. Considering these hypotheses, starting

material availability and chemical space representation, a final set of compounds was decided upon for synthesis.

6.3.3 Diprolyl Series Synthesis

The contract research organisation (CRO) Syngene International Ltd was contracted to synthesise a selection of compounds from the published diprolyl series and the additional modelled compounds. The compounds were synthesised via Scheme 6.1 as published by Greenlee et al. Each compound was separated into their respective diastereomers with each diastereomer pair unassigned for *in vitro* ACE competitive inhibition assays.

6.3.4 Competitive Inhibition Assays

ACE N- and C-domain competitive inhibition assays were performed using the ZFHL substrate to determine IC_{50}s for each compound. The same assay described in Chapter 2 (section 2.3.5) was followed but with different inhibitor concentrations.

Initially, each compound was screened against the two catalytic domains of ACE at concentrations of 100 µM and 1 µM. Compounds displaying sub micromolar inhibitions were then screened in a 10-fold serial dilution series from 1 µM down to 1 nM. These serial dilutions identified the concentration range in which the IC_{50} for each compound could be found.

Once an estimate of the IC_{50} had been made, a 2-fold serial dilution series of nine inhibitor concentrations was constructed spanning the range of the IC_{50} estimates. The remaining activity was plotted against the logarithmic inhibitor concentration and sigmoidal dose-response curves were fitted using the least squares algorithm with the removal of statistical outliers. Once the data was fit to a dose response curve within acceptable R^2 parameters ($R^2 > 0.95$), the IC_{50} was read off the curve (GraphPad Prism, v6.0).

6.4 Results

6.4.1 P_2 Monomer Considerations

Only three compounds published in the diprolyl series appeared capable of probing the S_2 subsite of ACE. A search for additional RCO_2H groups was conducted to find similar interactions. Some of the cheapest and most readily available R groups suitable for attachment to the scaffold were Boc and methyl ester protected Glu and Asp monomers. 4-Amino benzoic acid, the building block required to make **DM2** was also readily available. Many different options for adding an acidic substitution to a P_2 benzyl ring were explored. 13 Compounds (Table 6.2) in the diprolyl series whose R-group were readily available were selected for further analysis. These compounds tested both standard amino acid monomers and non-peptidic moieties focussing on phenyl ring substitutions and varying chain lengths.

This set was kept small and probed only a few specific P_2 variations to stay within reasonable time and cost constraints.

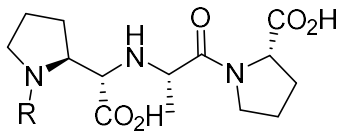
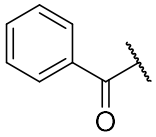
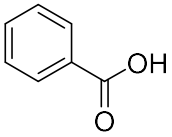
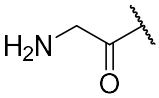
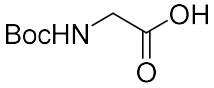
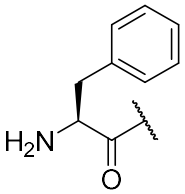
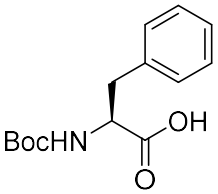
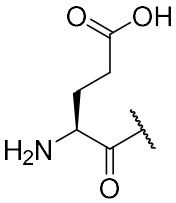
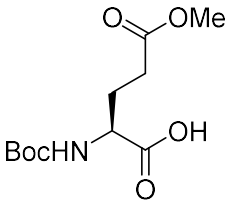
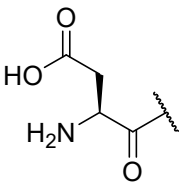
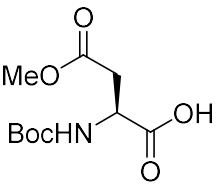
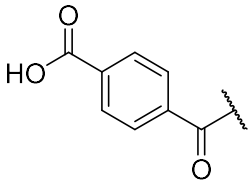
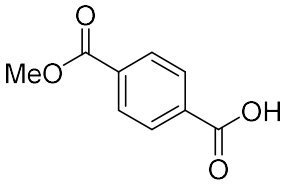
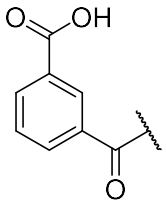
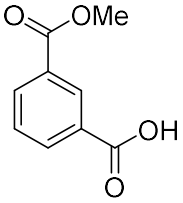
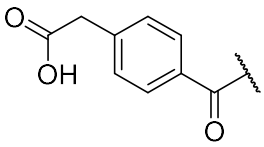
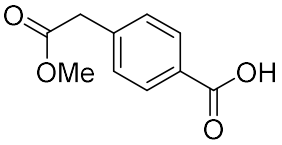
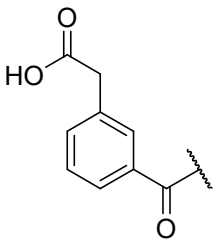
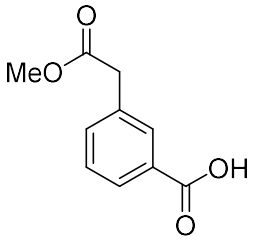
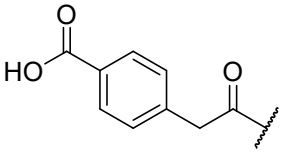
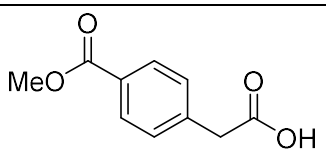
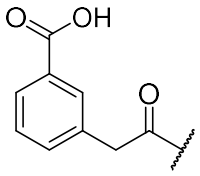
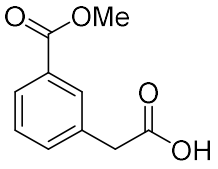
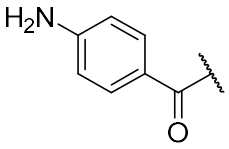
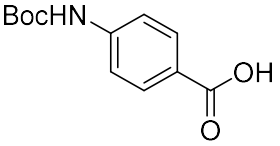
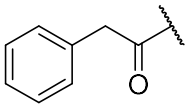
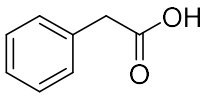


***Table 6.2**: Proposed R groups and the corresponding protected RCO_2H starting material required to attach to the diprolyl scaffold*

Name	R	Protected RCO_2H Monomer
32		 Benzoic Acid
34		 N-Boc-Glycine
36		 N-Boc-Phenylalanine
DP-Glu		 N-Boc-Glutamate ester
DP-Asp		

		N-Boc-Aspartate ester
DPI1		 Terepthalate methyl ester
DPI2		 1,3-Beneze dicarboxylic acid methyl ester
DPI3		 4-Carboxy Benzylic acid methyl ester
DPI4		 3-Carboxy Benzylic acid methyl ester
DPI5		 4-methyl ester Benzylic acid

DPI6		 2-(3-phenyl methyl ester) Acetic acid
DPI7		 3-Boc-amine benzoic acid
DPI8		 Benzylic acid

6.4.2 Molecular Docking of Diprolyl Compounds into the N- and C-Domains

The set of 13 compounds in the diprolyl series were docked into the N- and C-domains of ACE. Docking constraints C1 (Chapter 5; section 5.4.4) only found a plausible pose for 5/13 compounds in the C-domain and 9/13 compounds in the N-domain. To improve consistency, a core docking constraint was employed. When docked using the C1 constraints system defined in Chapter 5, **DPI5** displayed an idealised pose for the core scaffold in the C-domain while **DPI1** displayed this idealised pose in the N-domain. The docking was then repeated by setting the core scaffold in each compound to be held within a 1 Å RMSD tolerance of the core atoms of the chosen reference ligands. This restriction only permitted significant movement for the P_2 region of each ligand as desired.

Many of the compounds with P_2 acids interacted with Arg381 at awkward or unnatural dihedral angles to which Glide assigned no docking score penalties. This is a clear work-around built into the rigid receptor model specifically designed to be lenient with interactions involving highly flexible residues like arginine and lysine. Minimisation can make small adjustments to the side-chains giving the ligand a better fit. The minimised complex is therefore a more suitable model for evaluating binding energy.

6.4.3 MM-GBSA Analysis and Interactions Between the Diprolyl Series and the Two Domains of ACE

Prime MM-GBSA minimisation and binding energy predictions were performed to further guide the selection of compounds for synthesis. Minimised poses were evaluated and the binding energies were calculated to provide an additional metric with which to compare differential binding and establish a selection threshold to further reduce the set.

Predicting differential binding between two targets is not a function for which docking software is explicitly designed. A docking algorithm is a minimisation protocol which converges on the most favourable pose for a ligand within a given system. Evaluating the binding energy of such poses only gives comparable predictions once the ligand has been relaxed and minimised within its binding site. Ligands with poor binding avoid certain interactions resulting in poses being hard to predict. If the pose cannot be predicted accurately, binding energy predictions hold no merit. This problem is exacerbated when the constraints used for docking force the compound into a position where it makes unfavourable and highly unlikely interactions with the protein, as was observed for some compounds docked into the C-domain. Such forced binding poses need to be evaluated qualitatively as quantitative binding energy predictions of such poses are meaningless.

With the predictive power of these energy calculations for some of these systems in doubt, a qualitative selectivity hypothesis was introduced. The selectivity hypothesis predicts the domain selectivity of each compound based upon an analysis of the predicted pose for each ligand in their respective domain. The compounds were hypothesised to be either non domain-selective (NS) or N-domain selective (N). C-domain selectivity was not predicted for any of the compounds.

Some compounds are flagged as strained in the Arg381 distance and hypothesis columns. This flag indicates ligand strain when interacting with Arg381. In reality, this residue would likely reposition itself to a greater extent than what energy minimisations are capable of predicting. These ligands have had to force themselves into improbable and strained conformations disrupting the Arg381 salt bridge and weakening the overall N-domain binding. Conversely, the C-domain inhibition would probably receive a greater penalty as some of these compounds bring an acid even closer to Glu403. These compounds would therefore have a poor interaction with the C-domain while the introduced strain would also compromise the N-domain inhibition to a degree.

Table 6.3 summarises the minimised docking poses, the MM-GBSA binding energy and selectivity hypothesis for each compound. The distance between each ligand and either Arg381 or Glu403 was measured to assess the likelihood of an interaction with these two residues. The strain of each pose

was evaluated and then compared with the predicted ΔG of binding for each domain. The combination of predicted binding poses and binding energy helped to formulate the binding hypothesis. Compounds **32**, **34** and **36** were all 3.5 Å or more away from Arg381 or Glu403, too far for a direct interaction. None of the docked poses for these three compounds were strained hence the predicted binding energies were valid. The binding energies of these three compounds were generally of the order of -50 kJ/mol or lower suggesting strong ligand binding against both domains.

The compounds with P_2 acid moieties (**DP-Asp, DP-Glu, DPI1, DPI2, and DPI3**) all docked with strained poses in the C-domain but showed relaxed poses forming salt-bridges with Arg381 in the N-domain. The ΔG of these five compounds were all calculated to be strong below -50 kJ/mol while the ΔG of its strained C-domain pose was often significantly higher than the N-domain ΔG. Due to the strained C-domain poses of these compounds, the ΔG values cannot be considered accurate. The strong N-domain binding predicted for these compounds coupled with a strained C-domain pose strongly suggested N-domain selectivity. Compounds **DPI4**, **DPI5** and **DPI6** produced strained docking poses in both domains as a consequence of their larger size. This suggested weaker binding in both domains of ACE.

***Table 6.3**: A summary of the docking and MM-GBSA predictions for the expanded diprolyl series. The measured distance between the nearest heavy atom of the ligand and a N from Arg381 in the N-domain or an O from Glu403 in the C-domain is given. Some of these interactions are described as either in plane with the Arg381 or forming a salt bridge while unnatural torsions are described as strained, crowded or cramped. Such descriptors indicate an implausible pose was created due to the strict docking constraints. The last column provides a selectivity hypothesis. Compounds are classed as either non-selective towards either domain (NS), N-domain selective (N), or N-domain selective but strained leading to an overall weaker interaction with both domains despite the N-domain selectivity. Binding energy values marked with an asterix (*) are those run using input ligand poses which were already implausible.*

Compound	Arg381 distance	Glu403 distance	C ΔG (kJ/mol)	N ΔG (kJ/mol)	Hypothesis
32	5.63	8.78	-47.14	-59.24	NS
34	6.85	10.02	-52.14	-53.50	NS
36	3.50 (In plane)	7.86	-82.72	-56.41	NS
DP-Asp	2.64 (salt Bridge)	6.85	-51.57*	-71.76	N
DP-Glu	2.94 (Salt Bridge)	7.90 (Strained)	-52.10*	-70.32	N
DPI1	2.71 (Salt Bridge)	4.85	-54.82*	-51.62	N
DPI2	2.82 (Salt bridge)	9.96 (Strained)	-43.86*	-62.44	N
DPI3	2.89 (strained salt bridge)	6.50 (crowded)	-39.55*	-65.84	N
DPI4	2.82 (Salt Bridge)	5.28 (crowded)	-48.63*	-70.40*	N (strained)
DPI5	2.89 (strained salt bridge)	6.85 (cramped)	-53.82*	-60.10*	N (strained)
DPI6	2.90 (Improbable salt bridge)	9.68 (cramped)	-34.90*	-58.15*	N (strained)
DPI7	5.43	7.01	-62.94	-73.82	NS
DPI8	6.37 (In plane)	6.88	-53.46	-53.88	NS

Some of the poses described in Table 6.3 are illustrated in Figure 6.2. It is important to bear in mind that all N-domain docking simulations were performed using the 3NXQ crystal structure where Arg381 is bent inwards to face the ligand and make a salt bridge contact with the said ligand. Ligands with no P_2 acid lack the ability to draw this residue inwards. Since Glide does not allow protein flexibility, all compounds were therefore docked with Arg381 in this orientation.

Figure 6.2A and 6.2B show the P_2 phenyl group of compound **34** sitting snugly in the S_2 pocket of both domains. The slight changes in orientation of this group between the domains is acceptable as there is sufficient space to move in this region allowing for the possibility of two conformers of similar energies to exist in this region. For compound **DP_Glu** docked into the N-domain (Figure 6.2C), this ligand easily makes contact with Arg381 while it is forced into an energetically strained position to avoid a contact with Glu403. In Figure 6.2E, the acid substituted phenyl of **DP12** makes an ideal contact with Arg381 but is forced into a highly unfavourable dihedral angle around the indicated bond. Figures

6.2G and 6.2H show the ligand **DPI4**, which can make the required contact with Arg381 but only under a fair amount of steric strain on the benzylic acid group. Figure 6.2H shows that **DPI4** makes another unfavourable interaction with Glu403 as it is forced into a strained conformation to avoid this residue (Figure 6.2H).

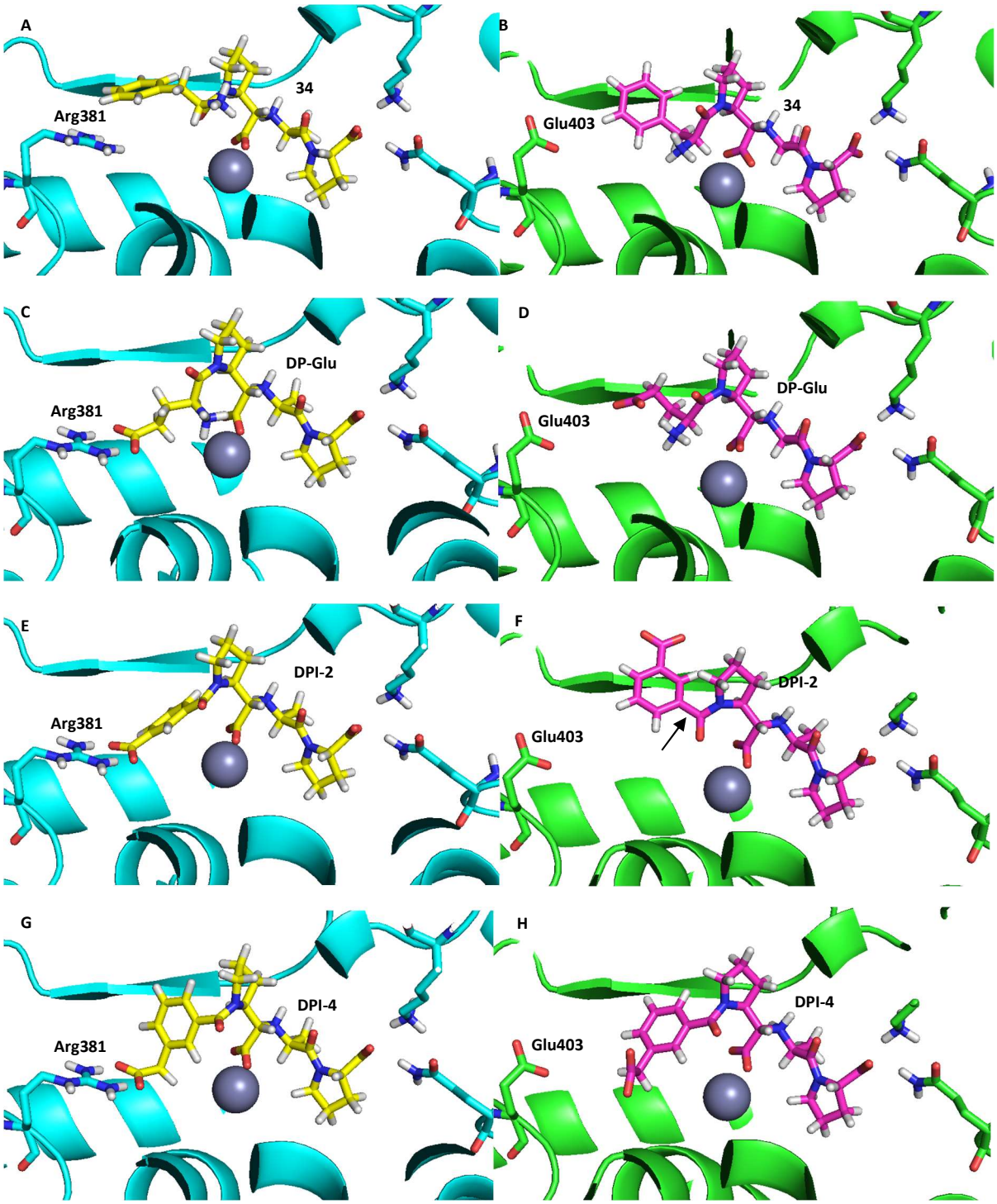


Figure 6.2: *Representative samples of ligand poses from the extended set of Diprolyl compounds. A and B show compound* ***36*** *in the N (cyan) and C (lime) domains respectively. C and D show* ***DP-Glu*** *docked in the respective N and C-domains. E and F show* ***DPI2*** *in the respective E and F domains. In figure F the strained bond between the benzoic acid and prolyl moieties is indicated with an arrow. G and H show* ***DPI4*** *in the respective N and C-domains.*

Table 6.1 and Figure 6.2 reinforce the notion that phenyl groups bind strongly to the S_2 subsite. Coupling acidic residues Asp and Glu to the diprolyl scaffold creates favourable interactions with Arg381 and unfavourable interactions with Glu403. Of the substituted phenyls, it was shown that the introduction of an acid directly onto the phenyl group created a ligand of the ideal length to interact with both the Arg381 and Glu403. An extra C linker in this region as is seen in **DPI4** appears to make the ligands too long resulting in a strained binding pose. Glu and Asp additions are good substituents to test the effect of different chain lengths interacting with the S_2 subsite.

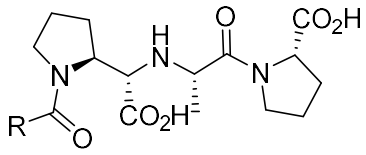
6.4.4 Selection Criteria

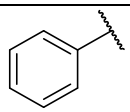
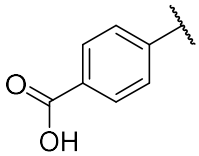
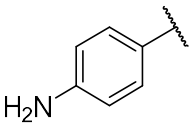
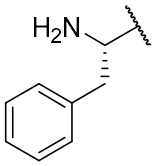
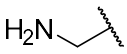
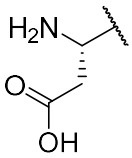
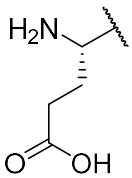
With insights from molecular modelling, a set of compounds most likely to yield an effective proof-of-concept was selected for synthesis. Out of the 13 compounds in this set, only three have been previously documented. Two of the compounds require the addition of natural amino acid substituents while the remaining compounds require the attachment of unnatural amino acids.

Table 6.4 shows the seven compounds selected for synthesis. It was crucial to resynthesise the published compounds (**32**, **34**, and **36**) as they could be compared against the published sACE inhibition. These compounds will also help test the domain selectivity hypotheses that have been made. **DP-Glu** and **DP-Asp** showed promising poses as a natural extension of the set attaching two additional standard amino acid residues. The substituted phenyls were also very important as they began to depart from the peptidic characteristics of the scaffold in the search for more drug-like compounds.

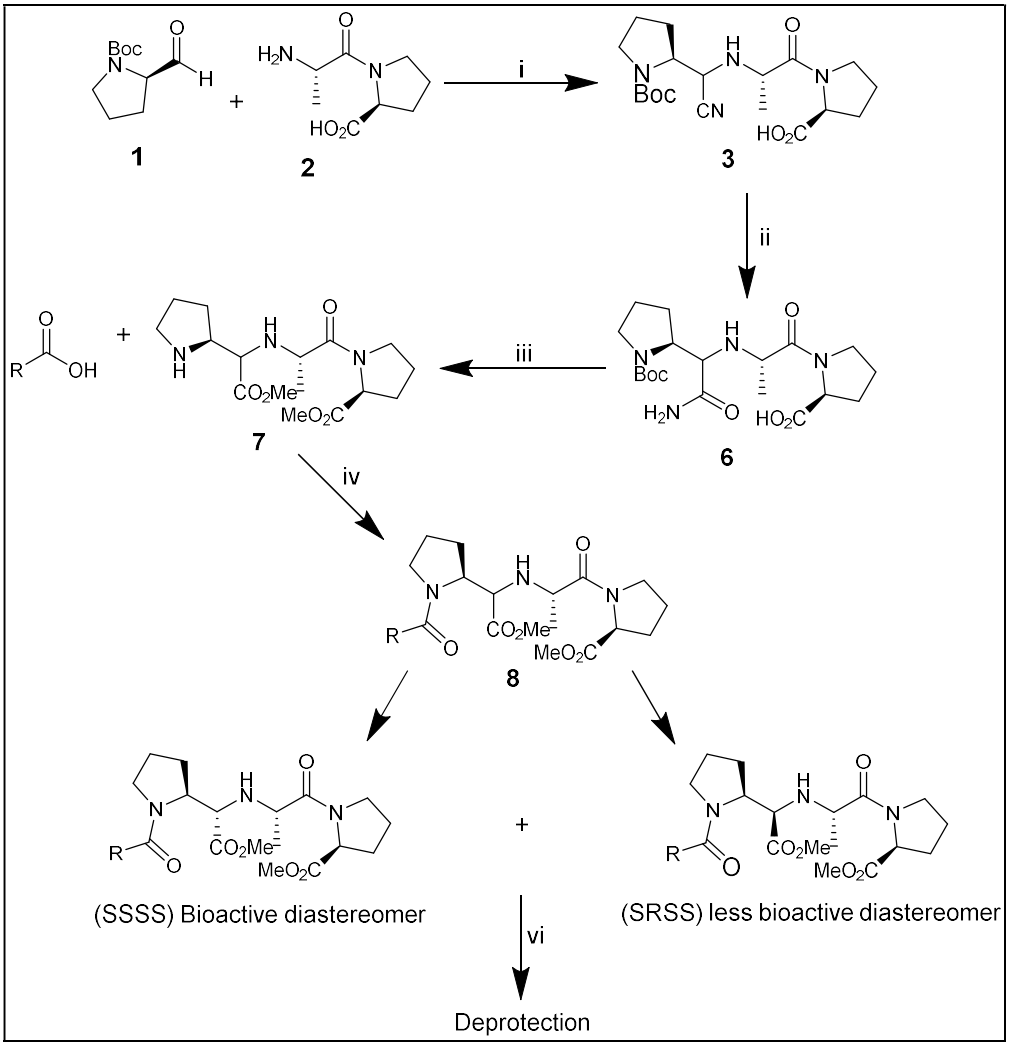
Rather than synthesise all 13 compounds, the initial focus was testing the hypothesis around compounds in Table 6.3 with a representative sample size. Priority was therefore given to compounds **32**, **34**, **36**, **DP-Asp** and **DP-Glu**. Two substitutions to the phenyl ring of compound **32** were selected to begin investigating this chemical space. **DPI1** and **DPI7** were chosen to test the opposing amino- and carboxy-4-phenyl substitutions. The final series chosen is shown in Table 6.4. For ease of reference, each selected compound was assigned a code **SG1-SG7** for the purpose of the synthesis.

***Table 6.4**: A summary of the final chosen series and its renaming of the final series before its synthesis*



R	Docking Code	Synthesis Code
	**32**	**SG1**
	**DPI1**	**SG2**
	**DPI4**	**SG3**
	**36**	**SG4**
	**34**	**SG5**
	**DP-Asp**	**SG6**
	**DP-Glu**	**SG7**

6.4.5 Diprolyl Synthesis



Scheme 6.2: *i*. TMSCN, MeOH/$NH_4^+Cl^-$, 25 °C 36h, 36% *ii. a. HCl in MeOH, 25 °C, 2d. b. $(Boc)_2O$, Et3N/MeOH, 72% iii. a. Amberlyst 15, MeOH, 60°C, 10 days b. HCl/Dioxane 25 °C, 12hr, 40% iv. a. RCO_2H, T3P, Et3N, DCM, 0 °C – 25 °C b. LiOH, THF/H_2O, (c. HCl/Dioxane)*

Seven compounds were synthesised via Scheme 6.2. Reaction i was first attempted using KCN in a AcOH/MeOH solvent as published by Greenlee. Replacing KCN with TMSCN in a NH_4Cl/MeOH solvent returned superior yields and quicker reaction times. This reaction delivered a diastereomeric mixture around the newly introduced chiral carbon to which the cyano group is attached. Reaction ii converted the cyano group to an amido group after reacting compound **3** in saturated HCl/MeOH at 25 °C for two days. The crude product **6** formed in a 4:1 ratio with a side-product. This side-product was removed using silica chromatography. In reaction iii, compound **6** was loaded onto the acidic amberlyst-15 resin in MeOH for 10 days. The compound was then eluted from the resin with 0.25 M Et_3N in MeOH. Once eluted, compound **7** was deprotected using HCl in dioxane at 25 °C for 12 hours to yield compound **7**.

Compound **7** branched off into seven different compounds. The RCOOH groups (Table 6.4) were each coupled to the core scaffold using a propylphosphonic anhydride (T3P) coupling reagent. Diastereomers were separated via Chiral Prep-HPLC. The diastereomer separation was performed before the deprotection at the compound **8** stage of the synthesis. This was due to the final compounds all being extremely polar giving them inadequately short retention times on a C-18 stationary phase. Once separated, these diastereomers were deprotected first using LiOH in THF/MeOH/H_2O to hydrolyse the esters then HCl in dioxane if the R-COOH group contained a Boc protected amine. These final products were each purified via Prep-HPLC.

6.4.6 Diastereomer Separation

Successful diastereomer separations were achieved using HPLC with a reverse phase C-18 chiral column. The separation was performed at the penultimate stage of the scheme where the compounds had the ideal retention time for an effective separation in a water/MeCN mobile phase.

Theoretically, the diastereomerism created in reaction i should yield both diastereomers in equal quantities but the ratios were difficult to gauge with such close peaks. Figure 6.3 shows the HPLC readout from the protected form of compound **SG1** at the compound **8** stage of scheme 6.1. It shows a clear example of an HPLC chromatogram before and after diastereomer separation. The unseparated mixture (Figure 6.3a) shows the two diastereomers of the compounds as a split peak with retention times of 8.661 and 8.814 minutes respectively. There is a fair amount of overlap at the base of these peaks. Figure 6.3B and 6.3C show the purified diastereomers with their own sharp peaks and similar but slightly different retention times of 7.957 and 8.269 minutes respectively.

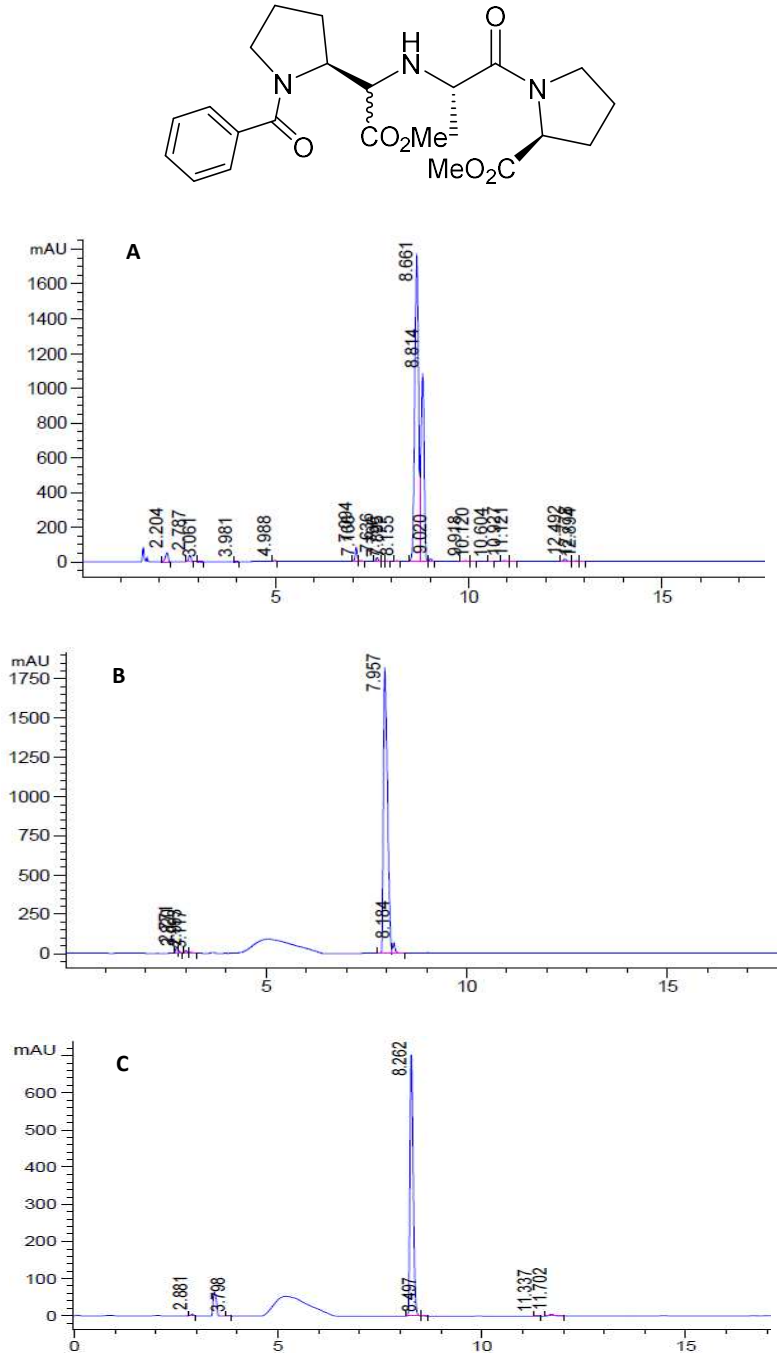


***Figure 6.3**:The HPLC chromatograms for the separation of the two diastereomers of **SG1**. Before the separation (A), two sharp peaks are visibly overlapping on account of the diastereomeric mixture. The subsequent HPLC runs show the separated disastereomers run on the same method with individual peaks and retention times comparable to the mixture in A.*

While a separated diastereomer may appear clean on HPLC, ^{1}H NMR is a more sensitive method for evaluating the diastereomeric purity of a compound. The methyl peak from this series often showed a clear duplication with a smaller copy of itself. Figure 6.4 shows an example of a peak duplication. The diastereomeric purity of each peak can be determined by comparing the relative areas. In the case

of Figure 6.4, the minor peak constitutes 13% of the total combined area with the remainder belonging to the other diastereomer.

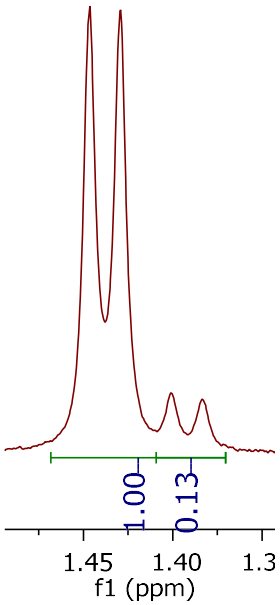


Figure 6.4: *A duplicated methyl ^{1}H NMR doublet peak from the ^{1}H NMR spectrum of SG4_1.*

Table 6.5 summarises the yields and diasteremeric purity of each compound at the compound **8** stage of the synthesis where the diastereomer separation was performed. It is important to note that the diastereomers were not assigned and both the SRSS and SSSS diastereomers were later deprotected and purified. The final yields varied greatly depending on the success of the separation. The highest yielding separation was achieved in the case of **SG6** with 38% of the original mass recovered in diastereomer 1 and 50% recovered in diastereomer 2. Cumulatively, 88% of the original mixture was recovered in this case. **SG2** separated poorly with 17% of the total mixture recovered in diastereomer 1 and 22% recovered in diastereomer 2. In this case only 39% of the original mixture was recovered. The protected form of each compound is referred to as **SGX-8** to indicate the protected compound **8** of this scheme.

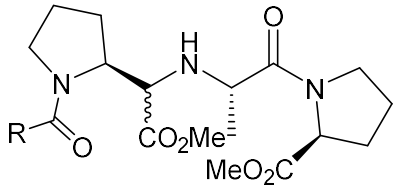
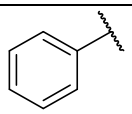
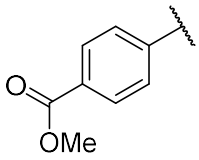
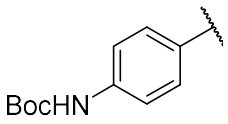
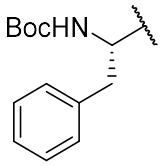
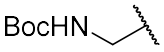
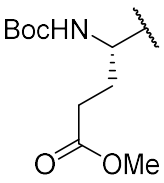
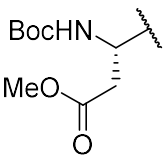


***Table 6.5**: Summary of yields for the compound 8 stage of each compound and the corresponding percentage of total mass recovered with each diastereomer.*

Name	R	Diastereomer 1 Recovered	Diastereomer 2 Recovered
SG1-8		15%	50%
SG2-8		17%	22%
SG3-8		20%	20%
SG4-8		56%	12%
SG5-8		45%	10%
SG6-8		38%	58%
SG7-8		35%	40%

From here on, the compounds were named as either **SGX_1** or **SGX_2** indicating the chronological order in which they eluted during the separation. This assignment is arbitrary as the two diastereomers were not yet assigned. According to Greenlee's study, the SSSS diastereomer displays an IC_{50} 2-3 orders of magnitude lower than the SRSS diastereomer in each instance. The ACE inhibition assays were therefore performed on both diastereomers allowing for them to be assigned according to their differential ACE inhibition. Comparing the activity between the two diastereomers, a vastly superior ACE inhibition by one diastereomer would be expected over the other.

6.4.7 Characterisation of Diprolyl Series

Post separation, the **SGX-8** compounds were deprotected then purified on the Prep-HPLC. The purity of these compounds was then analysed via HPLC, LC-MS and ^{1}H-NMR. The diprolyl series is highly analogous hence all the spectra shared the same core ^{1}H NMR peaks. All compounds with the exception of **SG3_1** were recovered with a purity above 95% with their mass verified via LC-MS. For the purpose of analysing the ^{1}H NMR spectra of the series, **SG4** has been selected as a representative example.

The ^{1}H shifts (Figure 6.5) for this series are characterised by a doublet peak (J = 7.0 Hz) at a shift of 1.40-1.50 ppm corresponding to methyl protons H-16. This doublet is duplicated with a much smaller peak. This smaller peak appears to belong to the other diastereomer which is present in small quantities after the separation. 86% of the combined area of the two peaks belongs to the larger major peak. Three broad multiplets are found with shifts in the 1.75-2.25 ppm range. These broad multiplets are consistent with constrained alkyl protons found in the two pyrolidine rings. These three peaks integrate for 8 protons cumulatively and correspond with protons H-2, H-3, H-8 and H-9. Compounds with a full amino acid residue in the P_2 position introduced a chiral centre with two diastereotopic protons on the β-carbon of this residue. In the case of **SG4**, protons H-12a and H-12b are diastereotopic and correspond to a symmetrical pair of doublets of doublets (J = 14.7, 5.5 Hz) and (J = 14.8, 7.8 Hz) at the respective shifts of 3.06 and 3.23 ppm. Further downfield two additional broad multiplets can be seen at 3.47 and 3.64 ppm corresponding to the deshielded cycloalkyl protons of H-10 and H-1 respectively, both found on C atoms adjacent to the N atoms. A doublet at 3.82 ppm (J = 4.4 Hz) corresponds to the proton H-6. This proton is attached to the chiral carbon from which the two diastereomers in this system originate. In the **SG4_2** diastereomer, the coupling value changes from 4.4 Hz to 5.8 Hz. These two coupling constants suggest different through space environments for H-6 and H-7 as expected in the two diastereomers.

Two of the remaining H-4, and H-7 protons are bonded to chiral carbons on similar alkyl rings in almost identical environments. These protons therefore share a broad multiplet peak overlapping with the expected quartet of H-5. The multiplet at 4.36 ppm therefore corresponds to the overlapping peaks of these 3 protons. The triplet (J = 7.6 Hz) at 4.51 ppm corresponds to proton H-11. The two downfield multiplets at 7.25 and 7.32 ppm integrate for two and three protons respectively. These five aromatic protons can be unambiguously assigned to the aromatic phenyl group.

Despite slight variations in the peak shifts, all compounds showed the same set of core peaks corresponding to the shared core structure. Chiral P_2 groups all showed the distinctive pair of doublets of doublets while 4 or 5 protons were observed in the aromatic region depending on the aromatic group. A crucial feature of these spectra are the chiral peaks seen at H-6 with contrasting J-values.

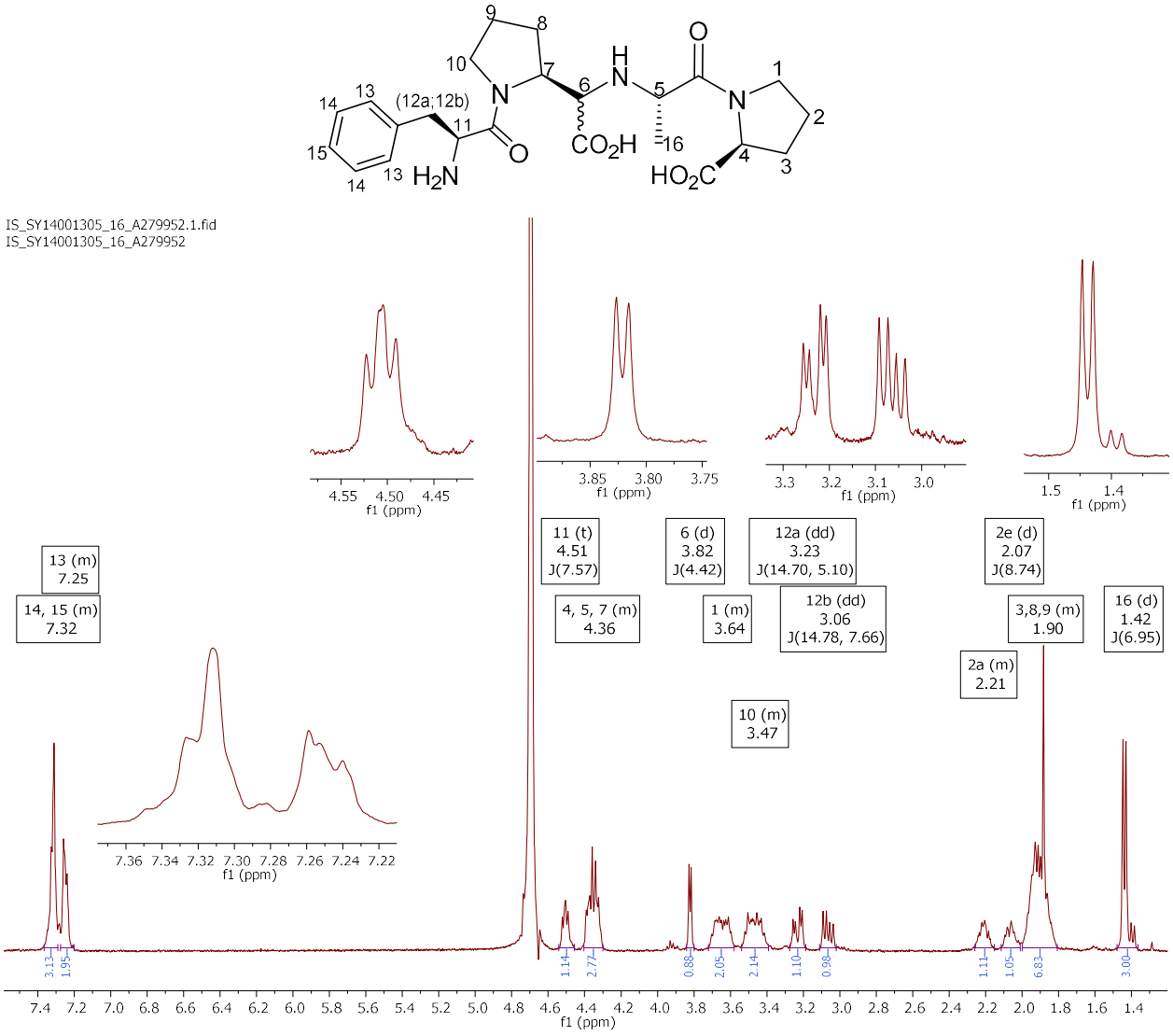


***Figure 6.5**: ^{1}H NMR spectrum of SG4_1*

6.4.8 Determination of Binding Affinities for the Diprolyl Derivatives

An initial inhibition screen was performed for each compound at concentrations of 100 µM and 1 µM where all but one compound displayed strong to moderate inhibition of the two domains of ACE at 1 µM. **SG6_1** showed no inhibition of the C-domain at 1 µM and only moderate inhibition of the N-domain at this concentration. **SG6** was immediately flagged as a potentially N-domain selective compound. Sub-micromolar inhibition for almost all compounds in the initial screen served as a secondary validation of both the assay conditions and the high purity of the compounds.

IC_{50} values were determined in a 2-fold serial dilution series of the inhibitor spanning the activity range. Each dataset was successfully fitted to a dose response curve with $R^2 > 0.95$ using Graphpad Prism. Dose-response curves of all inhibitors against both ACE domains are given in Appendix 6.1. Figure 6.6 shows an example of a dose-response curve from the set.

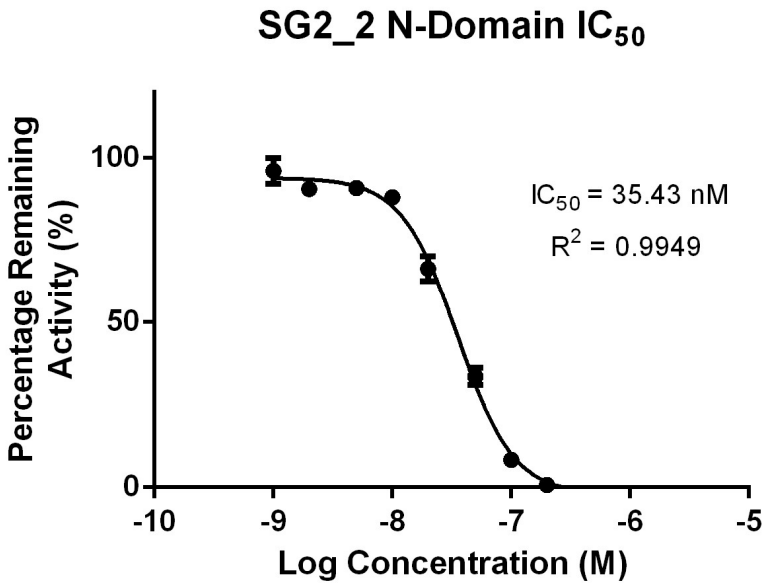


Figure 6.6: *The dose response curve used to determine the IC_{50} of a compound. Normalised activity is plotted against the log of the inhibitor concentration. These data-points are fitted against a sigmoidal dose-response curve and the IC_{50} is read as the point of inflection where the curve passes the midway point between the top and bottom plateaux.*

Table 6.6 lists the IC_{50} values measured for each compound against the two catalytic domains of ACE. In the absence of any NMR-based diastereomer assignment, both diastereomers were tested with each compound. In the case of **SG3,** only the second diastereomer was tested as it was the only diastereomer recovered. Competitive ACE inhibition assays are a reliable method for assigning these two diastereomers. Greenlee et al[123] observed a roughly 200-fold difference in ACE inhibition activity between between (S)SSSS and (S)SRSS diastereomers of each compound in this series using the same ZFHL substrate but with the full sACE enzyme. In this study, a roughly two order of magnitude difference in IC_{50} values was observed between each diastereomeric pair agreeing with the values

cited by Greenlee et al. This allowed the diastereomers to be unambiguously assigned. In the case of **SG3**, the only diastereomer tested appeared to be the active one.

The K_is were calculated from the IC_{50}s using equation 6.1. [S] is the substrate concentration which was set at a constant 1 mM for all assays. The K_m of ZFHL was measured at 0.93 and 0.18 mM for the N- and C-domain respectively.[134] These K_i*s* are directly related to the strength of the protein-ligand interaction, hence the K_i of each compound determined for the individual domain was calculated and compared with the selectivity factor defined as the C-domain K_i/N-domain K_i.

$$K_i = \frac{IC_{50}}{1 - \frac{[S]}{K_m}} \quad (6.1)$$

Table 6.6: *A comparison of the measured IC_{50} values for each diastereomer of the seven compounds synthesised in the SG series. The IC_{50} for each compound against the two catalytic domains of ACE is listed as well as an N-domain selectivity factor (C-domain K_i/N-domain K_i)*

Compound	N-Domain IC_{50}	N-Domain K_i	C-domain IC_{50}	C-domain K_i	N-Domain Selectivity Factor	Diastereomer Assignment
SG1_1	9 nM	4.34 nM	2 nM	0.3 nM	0.07	**SSSS**
SG1_2	294 nM	142 nM	133 nM	20 nM	0.14	SRSS
SG2_1	336 nM	162 nM	634 nM	97 nM	0.60	SRSS
SG2_2	35 nM	19 nM	15 nM	2 nM	0.14	**SSSS**
SG3_2	7 nM	3 nM	1 nM	0.15 nM	0.05	**SSSS**
SG4_1	580 nM	279 nM	311 nM	47 nM	0.17	**SSSS**
SG4_2	1.1 µM	530 nM	831 nM	127 nM	0.24	SSRSS
SG5_1	146 nM	70 nM	1.1 µM	168 nM	2.39	**SSSS**
SG5_2	816 nM	393 nM	1.5 µM	229 nM	0.58	SSRSS
SG6_1	3.3 µM	1.59 µM	> 100 µM	-	-	SSRSS
SG6_2	24 nM	12 nM	6.3 µM	961 nM	83.10	**SSSS**
SG7_1	1.7 µM	819 nM	3.0 µM	458 nM	0.56	SSRSS
SG7_2	20 nM	10 nM	114 nM	17 nM	1.80	**SSSS**

6.4.9 Domain Selectivity

The selectivity factors based on the measured *Ki*s can now be compared with the selectivity hypotheses. This comparison is important for validating the interpretation of the docking model. By

and large, these predictions held true with a few exceptions. Compounds **SG1**, **SG3**, **SG4** and **SG5** were predicted to be non-domain selective (Table 6.7) and judging by the measured selectivity factors ranging from 0.05 – 2.39, this appears to be largely accurate. K_i is an exponential quantity and is best compared on a log scale, hence selectivity factors of 0.1-10 are trivial. Compounds with selectivity factors in this range are therefore not domain-selective. **SG1_1** and **SG3** showed selectivity factors of 0.07 and 0.05 respectively, these would appear to be marginally C-domain selective but this selectivity may be deemed negligible due to experimental uncertainty.

SG2 was predicted to be N-domain selective yet, surprisingly, **SG2_2** proved a potent inhibitor of both domains without any significant difference in inhibition between the two domains (K_i = 19 and 2 nM for the N- and C-domains respectively). Both diastereomers of **SG6** were strongly N-domain selective. The selectivity factor of **SG6_2** (SSSSS) was measured at 83 while that of **SG6_1** (SSRSS) could not be measured as no C-domain inhibition was observed at concentrations below 100 µM. **SG7** was expected to have a similar N-domain selectivity to **SG6** but none was observed.

Table 6.7: *Summary of hypothesised domain selectivity vs the observed domain selectivity (NS = non-selective).*

Compound	Selectivity Hypothesis	Observed Selectivity
SG1	NS, strong inhibition	NS, strong inhibition
SG2	N-selective	NS, strong inhibition
SG3	NS, strong inhibition	NS, strong inhibition
SG4	NS, strong inhibition	NS, moderate inhibition
SG5	NS, strong inhibition	NS, moderate inhibition
SG6	N-Selective	Strongly N-Selective
SG7	N-Selective	NS

6.5 Discussion

6.5.1 Series Prioritisation

The database mining exercise proved more complicated than expected. Instead of finding promising hits for *in vitro* testing, the best compound appeared to originate from an incorrect database entry. Despite the error, the parent diprolyl series appeared to hold great promise for exploring the P_2 SAR of ACE inhibitors. A closer examination of the parent diprolyl series was performed leading to an

alternative approach involving the attachment of acidic amino acid residues onto the P_2 position of the scaffold. The docking results for both **DP-Glu** and **DP-Asp** were promising as the acid interacted with Arg381 in the N-domain, but compromised C-domain binding. **DP-Glu** and **DP-Asp** dock in a manner reminiscent of **RXP407** (Figure 6.7).

RXP407 shares a P_2 acidic group with **DP-Glu** and **DP-Asp** but its terminal amine is acetylated. **RXP407** was synthesised alongside a non-acetylated equivalent compound III[56]. The acetylation weakened the C-domain K_i 30-fold from 800 nM to 25 µM while the N-domain K_i only experienced a 5-fold drop from 5 nM to 25 nM. It is therefore plausible that the positively charged amine lowers the charge separation in that region of the molecule negating the clash of two negatively charged species to a degree. Acetylating this amine prevents a positive charge from forming, concentrating the negative charge and polarity at this end of the molecule. Greater charge polarisation would create a harsher binding penalty from a stronger electrostatic clash. The amine group of **DP-Asp** is in the same position as **RXP407** with the common Asp making a salt bridge with Arg381. **DP-Glu** still makes these two interactions but with a different orientation around the P_2 C-α. The effect of this different P_2 conformation is not yet clear.

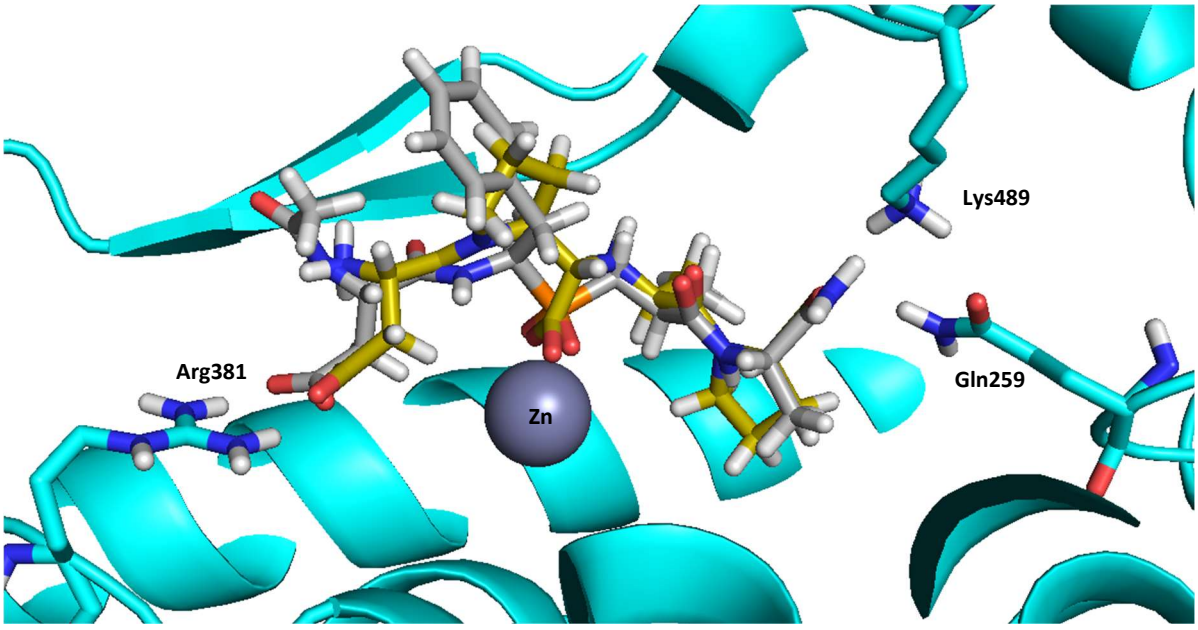


Figure 6.7: *An overlay of the docked poses of* ***DP-Asp*** *and the crystal pose of* ***RXP407*** *in the N-domain.*

As promising as **DP-Asp** and **DP-Glu** appear to be, they are still peptidic in nature. A major objective of this project was to design more drug-like N-domain selective ACE inhibitors than **RXP407**. These two compounds have a significantly reduced MW over **RXP407**, fewer bulky side-chains and lack the phosphinic acid, a major obstacle to cell permeability. These two compounds however, still contain two peptide bonds and four amino acid residues. Substituting the P_2 amino acid residue with

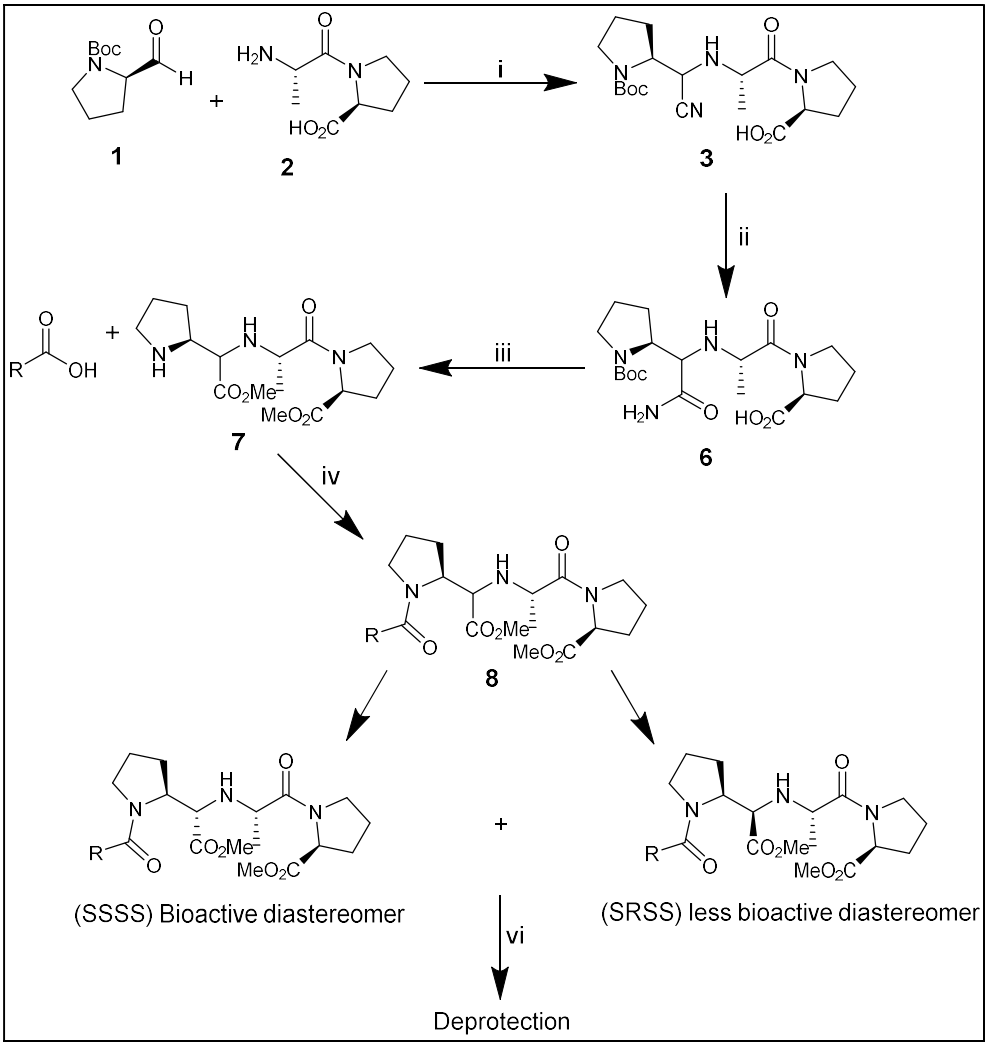
something more artificial and better customised to complement the unique S_2 subsite of ACE would be preferable.

Substitutions on the P_2 phenyl ring of compound **32** (Table 6.4) could be the first step in exploring a more drug-like chemical space with this SAR study. Docking studies suggested a carboxylic acid substitution directly onto the phenyl ring of **32** would have a high chance of making a salt bridge with Arg381. The rigid receptor dock was promising for both the *meta* and *para* substituted phenyl group but a choice had to be made and the *para* position was chosen due to a more readily available starting material.

6.5.2 Diprolyl Synthesis

The synthetic scheme reported by Greenlee et al was challenging as it required a difficult diastereomer separation. For the most part this published scheme was followed. A few minor changes were made as some of the reactions have been updated over the years (Scheme 6.2). Reaction i is a good example as Greenlee et al introduced a cyano group using KCN in an acetic acid and MeOH solvent. This reaction was greatly improved by replacing KCN with TMSCN in NH_4Cl/MeOH.

Reactions ii and iii were easily reproduced but a different amide coupling reaction was implemented for reaction iv. In the reference, the R-group amino acids were first activated using *N*-hydroxy succinimide and then mixed with the diprolyl scaffold in an Et_3N/DCM solution to complete the amide coupling. T3P was chosen as the preferred coupling reagent in this reaction and in many reactions the yields showed a marked improvement over the referenced conditions.



***Scheme 6.2**: i*. TMSCN, MeOH/$NH_4^+Cl^-$, RT 36h, 36% *ii. a. HCl in MeOH, 25 °C, 2d. b. $(Boc)_2O$, Et3N/MeOH, 72% iii. a. Amberlyst 15, MeOH, 60°C, 10 days b. HCl/Dioxane 25 °C, 12hr, 40% iv. a. RCO_2H, T3P, Et3N, DCM, 0 °C – 25 °C b. LiOH, THF/H_2O, (c. HCl/Dioxane)*

6.5.3 Diastereomer separation

The most challenging aspect of this synthesis was the diastereomer separation. Each compound from the diprolyl series contained either four or five chiral centres. The first two chiral centres originate from the natural optically pure amino acids, $_L$Ala and $_L$Pro. Reaction i introduces a third chiral centre producing both the R and S configurations in equimolar quantities. Studies on analogous Enalaprilat and Lisinopril ACEis have demonstrated that the S configuration around this centre is preferable for ACE ihibition.[113] The fourth chiral centre is then introduced when the R group is an amino acid in the natural S configuration.

The two diastereomers predicted were observed with compound **3** as observed by its twin peaks on the HPLC chromatogram. Assigning the two diastereomers spectroscopically is more challenging than identifying their presence. Greenlee et al assigned them by converting compound **3** into bicyclic ureas and comparing the ^{1}H NMR *J*-values between H_a and H_b (Figure 6.8). Today improved NMR technology

with 2D NOESY techniques and QM predictions are often used alternatively. Fortunately, there is a third method mentioned by Greenlee et al.

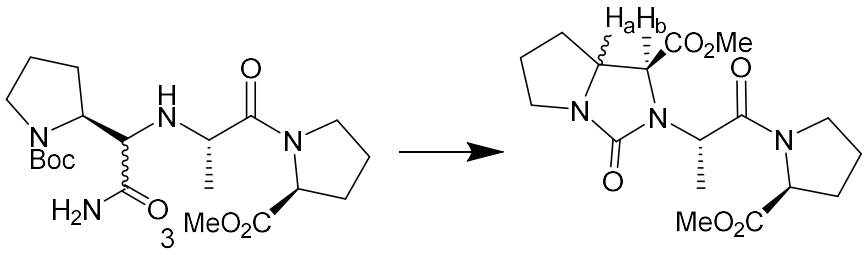


Figure 6.8: *Bicyclic ureas synthesised by Greenlee et al for the purpose of diastereomer assignment via ^{1}H NMR.*

Greenlee et al observed an average 200-fold decrease in ACE inhibition between the SSSS and the SRSS diastereomers during the development of the analogous Lisinopril and Enalaprilat ACEis. For practical reasons the enzyme assay method was chosen as the favoured method for assigning the two diastereomers. Relying on the enzyme inhibition assays to assign the diastereomers was not just the easiest method for assigning the diastereomers but also removed the possibility of omit incorrectly assigning the compound.

Another point of deviation from the scheme was the stage at which the diastereomers were separated. Greenlee et al separated the diastereomers at the compound **6** stage of the synthesis using normal phase medium pressure liquid chromatography (MPLC).

Diastereomer separation was first attempted at the compound **6** stage using a chiral HPLC column. Although the retention time of compound **6** was good, the peak separation was poor. The compound **8** stage of the synthesis proved favourable over compound **6** for diastereomer separations. Superior separations were observed with larger molecules and longer retention times.

6.5.4 Domain Selectivity of the Diprolyl Series

The reported 200-fold difference in IC_{50} between (S)SSSS and (S)SRSS diastereomers is consistent with the K_i values reported in Table 6.6. Since two different enzyme constructs were used for the two domains, it is more appropriate to compare N- and C-domain inhibition using the K_i metric. Despite the exceptions of **SG4** and **SG5**, a roughly two order of magnitude difference in K_i was observed across this set of inhibitors between the two stereoisomers separated for each compound. **SG1_1**, **SG4_1** and **SG5_1** correspond with compounds **32**, **36** and **34** from the Greenlee et al paper respectively. The K_i values of **SG1_1** were calculated at 4 and 0.3 nM for the N- and C-domain respectively. These values corroborate with the low nanomolar IC_{50} reported by Greenlee. The K_i values of **SG4_1** were a bit higher than what would be expected from the Greenlee's values but not alarmingly so.

Of the compounds added to this set, **SG3** was predicted to be a strong non-selective inhibitor of both domains. The 4-amino benzyl moiety served as a counterexample of the acid dependent selectivity hypothesis. It was predicted that when presented with this 4-amino benzyl moiety, Arg381 would swing away from the ligand allowing it to settle comfortably. Despite only recovering one diastereomer of this compound, the recovered diastereomer appears to be the favoured one. As predicted, **SG3** proved to be a potent inhibitor of both ACE domains with a 20-fold selectivity towards the C-domain while the N-domain K_i is still potent at 3 nM.

SG2 was predicted to be N-domain selective. Docking simulations showed the 4-carboxyl phenyl moiety easily formed a strong interaction with Arg381 while in the C-domain it appeared trapped in a position proximal to Glu403. The lack of observed N-domain selectivity suggests a shortcoming in this docking model. A possible explanation is the shielding of an unfavourable interaction between the P_2 acid and Glu403 by an unexpected residue. Another possibility is the ligand finding a lower energy conformation violating the constraints. Such a conformation could shield the compound from an unfavourable interaction between these two acids.

In contrast to the lack of domain selectivity observed with **SG2** and **SG7**, **SG6** behaved exactly as predicted. **SG6** binds strongly to the N-domain while unfavourable interactions between the P_2 acid and Glu403 compromises the C-domain binding with a high 83-fold K_i penalty in line with the predictions.

The respective Asp and Glu groups of **SG6** and **SG7** differ from each other by a single alkyl carbon. **RXP407** contains an Asp residue in the same position as **SG6** and is of a similar length. Comparing the docked poses of **SG6** and **SG7**, both compounds make a salt-bridge with Arg381 while a strong N-domain interaction was observed. When it comes to disrupting the unfavourable interaction with Glu403, an extra alkyl carbon adds an additional torsional degree of freedom to the ligand. This extra degree of freedom could possibly allow the ligand to find a new conformation shielding this acid from Glu403.

Figure 6.9 provides a direct comparison between the predicted poses of **SG6** and **SG7** docked into the C-domain. The additional strain experienced by **SG7** to avoid a contact with Glu403 was thought to contribute to the unfavourable interaction with Glu403. In reality, the extra carbon in the chain appears to have enabled the ligand to find a new pose shielding it from this unfavourable interaction. The failure to predict this behaviour is a shortcoming of the rigid receptor model. A minimisation of a

greater portion of the protein or dynamics simulation would have improved the chances of predicting such behaviour.

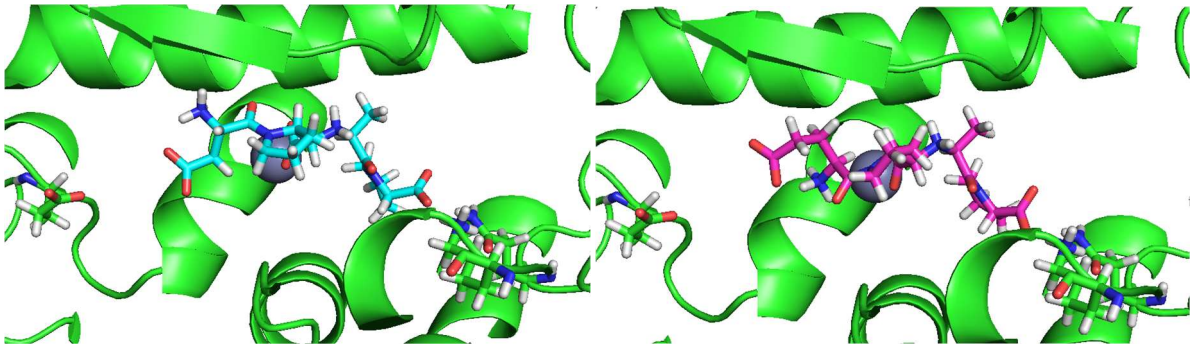


Figure 6.9: *A side by side comparison of SG6 (left) and SG7 (right) docked into the C-domain of ACE. The raised energy pose experienced in the P_2 region is a result of an unfavourable interaction with Glu403.*

The strong inhibition displayed by this set of compounds and the discovery of a novel N-domain selective ACEi makes a compelling case for the expansion of this set to include all the ligands docked against the two ACE domains. Expanding the SAR series would improve the approximation of the chemical space in which N-domain selectivity can be expected. In addition to defining this chemical space, crystal structures would help explain the interactions responsible for the N-domain selectivity of **SG6** and the lack thereof observed in **SG2** and **SG7**.

Despite only **SG6** showing the desired domain selectivity, a direct comparison between the inhibition of **SG6** and **SG7** provides invaluable SAR data. These two compounds have elucidated an important structural aspect of N-domain selective ACE inhibition. Cocrystalising these two ligands with the C-domain would confirm these assertions and improve the ability of this model to predict domain selectivity.

The database mining and subsequent docking, synthesis and *in vitro* assays produced one success in the form of **SG6**. While **SG6** is certainly less peptidic than **RXP407**, it still contains two peptide bonds and three carboxylic acids. In a similar manner to most ACEis, one or two of these acids can be esterified into a prodrug. The P_1 pyrolidine and amine bond can be altered without too much difficulty. Aromatising this pyrolidine would remove a stereo centre and could facilitate the replacement of the second amide bond. While **SG6** does not appear to be drug-like, it is a step closer to a peptidomimetic compound than **RXP407** and provides a suitable platform from which a more drug-like compound can be derived.

6.5.5 Concluding Remarks

This chapter demonstrates an example of the disconnect, which still exists between the predictive power of the ACE docking model and what one can expect to observe *in vitro*. While docking can be a

great aid in lead optimisation and SAR studies, when applied to N-domain selective ACE inhibition there were some caveats. The dual-domain and off-target docking was able to correctly predict one N-domain selective inhibitor but incorrectly suggested two additional compounds. These results reinforce the notion that N-domain selective ACE inhibition exists within a small and precise chemical space. The comparative docking model can help approximate this space but not pinpoint it. A more extensive SAR series introducing more variety in this P_2 region would help to better define this chemical space while this diprolyl series has been shown to be an ideal scaffold with which to probe this SAR.

Chapter 7 – Summary, Conclusions and Recommendations for Future Work

7.1 Premise

A well-established drug target such as ACE is an unlikely novel therapeutic target. The elucidation of the dual-domain structure of ACE has revealed a more intricate synergistic relationship between the two catalytic domains of ACE.

The collection of high resolution crystal structures for the two domains of ACE and the availability of powerful molecular modelling software provided a unique opportunity to develop drug-like N-domain selective ACE inhibitors using CADD software.

7.2 Summary

7.2.1 Docking Constraints

Since docking simulations form the foundation of most CADD protocols, it was important to find docking conditions capable of accurately reproducing plausible ligand poses given this challenging metalloprotease binding site. The two most important ACEi interactions have been identified as the Zinc chelation and a polar hydrogen bonding interaction between the P_2' acid and S_2' donors. Introducing docking constraints to fix these two interactions allowed for the recreation of the ACEi poses observed in crystal structures, thus demonstrating the docking protocol to be capable of predicting the binding pose of an ACEi.

With a constrained docking system against the two catalytic domains successfully validated, a CADD approach was devised. Constrained docking formed the basis of three CADD approaches that were devised to maximise the chances of finding an N-domain selective ACEi.

7.2.2 Fragment-Based Screening

A fragment-based approach was attempted in order to find novel inhibitors beyond the known ACEi chemical space. This set of fragments was restricted to compounds containing metal-binding groups to maximise its chances of finding hits.

Following a standard fragment docking procedure, two potential hits were identified with IC_{50} values in the 500 µM range. Despite the inhibition, these two compounds produced no lead on domain selectivity. This lack of selectivity was attributed to the highly homologous domains. The interactions responsible for domain selectivity were too specific to be found in a set of 16 000 compounds. Two borderline inhibitors were identified from a set of 60 that were tested. This is usually an acceptable outcome for a fragment screen where weak fragment inhibitors are built into stronger drug-sized

inhibitors. The absence of a lead on domain selectivity coupled with the availability of many potent sACE inhibitors meant closer attention was paid to the structural features of existing ACEis.

7.2.3 Combinatorial Library Screening

In response to the failure of the fragment screen to find an N-domain selective ACEi, a *de novo* approach was devised to better utilise the existing clinical ACEis, some of which are highly potent. Enalaprilat was deemed the simplest ACEi from which Arg381 in the S_2 subsite is accessible. The non-prime side of Enalaprilat was deconstructed to replace the phenyl group with a large range of acidic moieties and alkyl linkers of variable length extending into the S_2 subsite. A library of Enalaprilat analogues was generated and docked against both domains. Many compounds from the library showed promise following docking and binding energy calculations. These results prompted the progression from modelling these Enalaprilat analogues to synthesising them.

Despite difficulties with diastereomerism and racemisation, one Enalaprilat analogue, **SF07** was synthesised and found to be N-domain selective. This is an encouraging result and warrants further investigation into the structural causes of this observed N-domain selectivity.

7.2.4 ACEi Database Mining

The search for novel N-domain selective ACEi provided an opportunity to explore the extensive archives of sACE inhibitors. Almost 25 000 ACEis were extracted from the GVK database and investigated for potential selectivity towards either domain. Several structure filters were applied to the database to reduce the dataset to a smaller set of drug-like compounds. After applying these filters and manually inspecting the best 1800 compounds, one chemical series in particular stood out. This series of diprolyl compounds synthesised by Greenlee et al contains a carboxylic acid ZBG coupled with an Ala-Pro in the P_1'-P_2' position and a pseudo-pro P_1 moiety. Attached to this scaffold is a variable P_2 group, the exact position being probed for N-domain selectivity.

This series was modelled using the same constrained docking protocol developed for the fragment-based docking. Docking studies suggested this series to be an excellent platform from which to probe interactions with the S_2 subsite. The synthesis of this series was deconstructed and four additional P_2 groups were added to the series and modelled accordingly. Seven of the best compounds were then synthesised and tested *in vitro*. One of the seven compounds (**SG6**) was shown to have an 88-fold K_i selectivity towards the N-domain of ACE, thus validating the predictive power of this docking protocol.

7.3 Conclusion

This project has developed a constrained docking model capable of accurately reproducing the poses of ACE ligands inside the binding site. Applying constraints to fix the position of metal-chelator pairs

helped to overcome previously reported problems originating from poorly approximated metal chelation interactions. Docking constraints fixing a specified chelating group in a position proximal to the Zn atom proved capable of reproducing the natural pose. However, each set of docking constraints needed to be optimised for a given chemical series leaving them unsuitable for screening a large database of varied ligands. This constrained docking protocol formed the foundation of all three CADD techniques attempted in this project.

Docking played a vital role in predicting the poses for each CADD method attempted. In the fragment screen, docking provided plausible poses for every compound examined. Constrained docking then provided a simple yet elegant method for examining the docked poses of the Enalaprilat. It helped to identify ligands which were likely to bind to ACE selectively compared to those which failed to fit into the binding site. Lastly, constrained docking helped to predict the binding pose of many archival ACEis. A potent series of ACEis with a variable P_2 group was identified as a promising scaffold with which to explore N-domain selectivity.

Validation of the CADD methodology was only possible with the *in vitro* inhibition assays. While the fragment-based screening procedure proved capable of finding weak ACEi fragments, it failed to find a domain selective inhibitor. However, it is also likely that such a fragment was never present in the screening database. Many N-domain selective Enalaprilat analogues were predicted but only one such analogue was successfully synthesisedand shown to be N-domain selective. A few selective P_2 SAR manipulations to the diprolyl series identified during GVK database mining yielded another N-domain selective drug-like ACEi.

This project has achieved its primary objective of designing N-domain selective drug-like ACE inhibitors with **SF07** and **SG6**. These two compounds are similar in structure and are the products of two different CADD approaches centred around constrained docking. The success of these two techniques emphasise the initial assertion that high resolution crystal structures and modern docking software can make accurate ligand-binding predictions, which have been utilised to achieve the goal of designing novel N-domain selective drug-like ACEis.

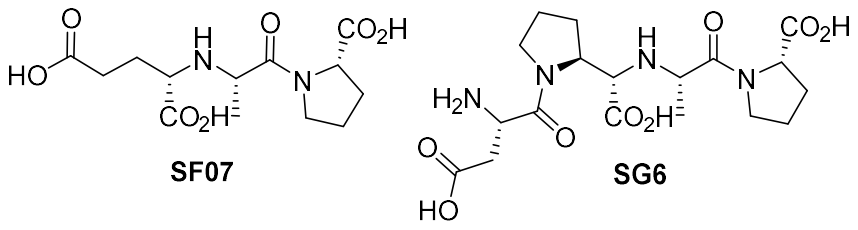


Figure 7.1: SF07** and **SG6

7.4 Future Work

The discovery of an N-domain selective ACEi from the synthesis of just one Enalaprilat analogue emphasises the need to explore both the P_1 and P_2 SAR of Enalaprilat analogues. Scheme 4.1 will need to be modified to avoid the hydrolysis of the methyl ester and avoid the racemisation in the final step. As suggested, a Bn ester protecting group deprotected through hydrogenation is a possibility. The diprolyl series also returned an N-domain selective ACEi. The N-domain selectivity of **SG6** also justifies the synthesis of a more comprehensive SAR series.

Apart from a full SAR investigation of these two similar drug-like ACEi scaffolds, there are structural questions about the interactions between these ligands and the receptor which need to be answered. Crystals structures of both **SF07** and **SG6** cocrystallised against the N-domain of ACE need to be grown. These crystal structures will confirm the accuracy of the predicted binding poses and whether the two postulated residues are responsible for N-domain selective ACE inhibiton.

A complete SAR study on both the Enalaprilat analogue and Diprolyl series will suggest many novel drug-like N-domain selective ACEis. Cocrystal structures of **SG6** and **SF07** and similar compounds bound to the N-domain will help to further narrow the chemical space in which N-domain selectivity is found, which in turn may drive future SAR series. These future studies would lay the groundwork for the future development of N-domain selective ACEi drug development.

Chapter 8 – Experimental

8.1 Chemistry

8.1.1 Reagents and Solvents

All chemical reagents and anhydrous solvents used during the synthesis of the Enalaprilat analogue were purchased from Sigma Aldrich, South Africa. Analytical Reagent (AR) grade solvents ethyl acetate, hexane, DCM and acetone were purchased from Kimix Chemicals. HPLC grade solvents were purchased from Sigma Aldrich (ammonium acetate, triflouroacetic acid and DMSO), Merck (glacial acetic acid) and Microsep (Acetonitrile and Methanol) for Chromatography, Mass Spec and HPLC.

The diprolyl series was synthesised by Syngene International in Bangalore India. Their chemical suppliers were withheld but all spectra provided for their compounds were verified locally.

8.1.2 Chromatography

Thin layer chromatography (TLC) reaction monitoring was performed using Merck F_{254} aluminium-backed silica gel 60 plates. Spots were visualised with either ultra violet (UV) light (254/366 nm), anisaldehyde or ninhydrin stains. Reaction products were all purified via column chromatrography using Merck kieselgel 60:70-230 mesh via gravitational column chromatography and flash chromatography.

8.1.3 Physical and Spectroscopic Characterisation

Melting points were determined using a Reichert-Jung Thermovar hot-stage microscope.

The reported compounds were characterised using ^{1}H NMR while novel compounds were characterised via ^{1}H-NMR, ^{13}C-NMR and LC-MS.

HPLC: Peak purity of the Enalaprilat analogue was determined locally using preparatory HPLC with a thermo separation system comprising of a Spectra Series P200 pump, an AS100 automated sampler and a UV 100 variable wave detector. The UV detector was set to monitor peak absorption at 214 nm. A Waters® X-bridge C18 5.0 µm (4.6 x 150 mm) (Phenomenex, Torrance, CA) column stationary phase used was fitted to a Supelguard ® Ascentis ™ C18 guardcartridge (2cm x 40 mm, 3 µm) (Supelco Analytical, Bellefonto, PA).

The Enalaprilat analogue was purified and analysed using a mixture of mobile phase A; 10 mM ammonium acetate in water and mobile phase B; 10 mM ammonium acetate in MeCN running the method described in Table 8.1 at flow rate of 1.2 mL/min.

Table 8.1: *HPLC method used to test the purity of Enalaprilat analogues.*

Time	% A	% B
Initial	75	25
9.00	0	100
14.00	0	100
14.10	75	25
20.00	75	25

Diastereomer separation of the diprolyl series was performed using a thermo separation system comprising of an Analytical Technologies® P2230 HPLC pump with an automated sampler and Analytical Technologies® UV 2230 variable wavelength detector set to detect absorbance at 214 nm. The stationary phase column used was a Gemini® NX – C18 3.5 µm (4.6 x 50 mm) (Phenomenex, Torrance, CA).

This HPLC configuration was run with a mixture of mobile phase A; 10 mM ammonium acetate in water and mobile phase B; 10 mM ammonium acetate in MeCN using the method described in Table 8.2 at a flow rate of 1.0 mL/min.

Table 8.2*: HPLC method used to separate the two diastereomers of each diprolyl compound.*

Time	% A	% B
Initial	90	10
15.00	0	100
20.00	0	100
23.00	90	10
28	90	10

Peak purity of the diprolyl series was analysed with the same equipment used for the diastereomer separation but with an Atlantis® dC18 5 µm (4.6 x 250 mm) column stationary phase.

The HPLC was run in this configuration with mixture of mobile phase A; 0.1 % TFA in water and mobile phase B; 0.1% TFA in MeCN using the method described in Table 8.3 at a flow rate of 1.5 mL/min.

Table 8.3: *HPLC method used to test the purity of diprolyl compounds.*

Time	% A	% B
Initial	90	10
15.00	0	100
20.00	0	100
23.00	90	10
30.00	90	10

LC-MS: Liquid chromatography mass spectrometer (LC-MS) analysis on the Enalaprilat analogues was performed using an Agilent® 1260 Infinity Binary Pump, Agilent® 1260 Infinity Diode Array Detector (DAD), Agilent® 1290 Infinity Column Compartment, Agilent® 1260 Infinity Standard autosampler and an Agilent® 6120 Quadruple (Single) mass spectrometer equipped with an APCI and ESI multimode ionisation source. Compound purity was determined using an Agilent® LC-MS with a Kinetex Core C18 2.6 µm column (50 x 3 mm).

A mixture of mobile phase A; 0.4% acetic acid in 10 mM ammonium acetate and mobile Phase B; 0.4% acetic acid, 10 mM ammonium acetate in a 9:1 ratio was run at a flow rate of 0.9 mL/min.

Table 8.4*: LC-MS method used to analyse Enalaprilat analogue compounds with the described mobile phases A and B.*

Time	% A	% B
Initial	75	25
1.00	75	25
3.00	0	100
4.50	0	100
5.20	75	25
6.00	75	25

The diprolyl series was monitored via LC-MS using the same Agilent 1200 series equipment with an XBridge® C8 column (50 mm x 4.6 mm, 3.5 µM) stationary phase. This LC-MS configuration was run using a mixture of mobile phase A; 0.1% TFA in H_2O and mobile phase B; 0.1% TFA in MeCN was run at a flow rate 0.6 mL/min using the method described in Table 8.5 at a flow right of 0.6 mL/min.

***Table 8.5**: LC-MS method most commonly used for analysis of diprolyl series.*

Time	% A	% B
Initial	95	5
8.00	0	100
8.50	95	5
10.00	95	0

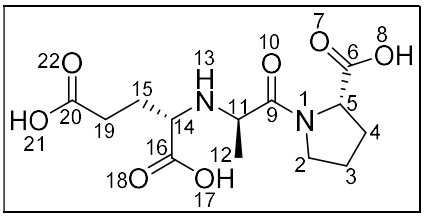
NMR: All NMR spectra were recorded on either a Brucker Ultrashield-Plus Spectrometer (^{1}H-300 MHz; ^{13}C-75 MHz), a Brucker Ultrashield-Plus Spectrometer (^{1}H-400 MHz; ^{13}C-100 MHz) or a Brucker Ultrashield-Plus Spectrometer (^{1}H-600 MHz; ^{13}C-150 MHz) with compounds dissolved in either deuterated Methanol (MeOD-d_4) or deuterium oxide (D_2O) solvents. Chemical shifts (δ) are recorded in ppm and coupling constants (*J*) are recorded in hertz (Hz). Abbreviations used in assigning ^{1}H NMR signals are: br (broad), d (doublet or doublets), m (multiplets), q (quartet), s (singlet), t (triplet), dd (doublet of doublets), ddd (doublet of doublet of doublets) and td (triplet of doublets).

8.2 Characterisation

8.2.1 Enalaprilat Analogue

SF07

105 mg of **SF06** was dissolved in an aqueous 1 M NaOH solution at 25 °C and stirred for 2 hours. The reaction was monitored via LC-MS. The solution was then neutralised with the dropwise addition of HCl. The solution was stirred for an additional 30 minutes after the water was evapourated yielding a mixture of a white paste and NaCl. The mixture of the compound and NaCl was loaded onto the prep-HPLC and purified (t_r = 3.86 min) to yield 36.6 mg (39%) of an opaque white paste.



^{1}H NMR (600 MHz, Methanol-d_4): δ 4.34 (q, *J* = 4.36 Hz, 1H, H-11), 4.30 (t, *J* = 5.4 Hz, 1H, H-5) 3.70 (m, 1H, H-14), 3.64(m, 2H, H-2), 2.62 – 2.51 (m, 2H, H-19), 2.50 – 2.27 (m, 2H, H-15), 2.26 – 2.16 (m, 2H, H-3), 2.15-2.07 (m, 2H, H-4), 1.55 (d, *J* = 6.9 Hz, 3H, H-12); ^{13}C NMR (150 MHz, CD_3OD): δ 180.08 175.58, 168.45 (2C), 62.18, 61.62, 60.58, 53.39, 46.84, 33.53, 29.47, 25.19, 24.13, 20.53; LC-ESI-MS (+ve mode): m/z 317.1 $[M+H]^+$ purity 97.0% (t_r = 0.26 min)

8.2.2 Diprolyl Series

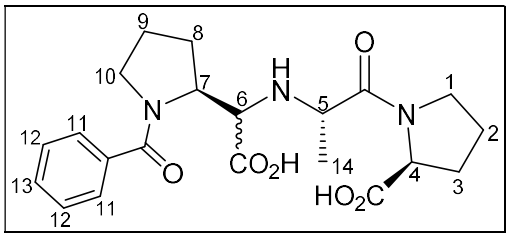
General ester deprotection method:

5 equivalents of a 4.0 M aqueous solution of LiOH was added to a 0.1 M solution of the methyl ester protected form of each diprolyl compound in a THF/MeOH/H_2O (3:2:1) solution. The mixture was stirred for 2 hours at room temperature. The mixture was then diluted with 5 mL of water then neutralised by adding 1 N HCl dropwise. The compound was either dried and purified via Prep-HPLC or the crude Product was Boc-deprotected.

General Boc Deprotection method:

Crude Boc protected compound was dissolved in chilled 4 M HCl/Dioxane under N_2 at 0 °C. The mixture was stirred for 30 min at 0 °C and monitored via TLC. Upon completion the solvent was evaporated and the residue was washed with diethyl ether. The final compound was then purified over prep HPLC.

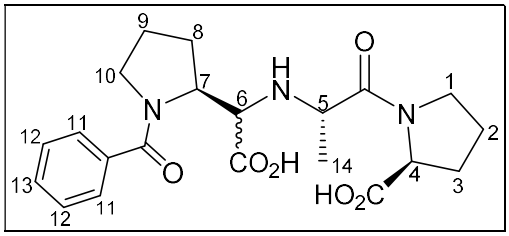
SG1_1



The general ester deprotection was performed on 45 mg SG1_1-8 and then purified via prep HPLC to obtain 26 mg (62%) of a sticky white-yellow solid.

^{1}H NMR (400 MHz, Methanol-*d*$_4$) δ 7.62 – 7.55 (m, 2H, H-11), 7.55 – 7.44 (m, 3H, H-12, H-13), 4.72 (t, *J* = 8.7 Hz, 1H, H-4), 4.52 – 4.47 (m, 1H, H-7), 4.37, (d, J = 1.9 Hz, 1H, H-6), 4.23 (q, J = 6.8 Hz, 1H, H-5), 3.70 (m, 1H, H-1e), 3.60 (m, 2H, H-1a, H-10e), 3.50 (d, 1H, H-10a), 2.44 – 2.25 (m, 3H, H-2, H-9a), 2.12 – 1.93 (m, 4H, H-3, H-8a, H-9e), 1.87-1.73 (m, 1H, H-8e), 1.64 (d, *J* = 7.0 Hz, 3H, H-14). ^{13}C (100 MHz, Methanol-d$_4$): δ 175.9, 175.1, 173.2, 168.2, 132.3, 129.4, 127.9 (2C), 126.5 (2C), 66.3, 64.2, 59.5, 58.3, 50.5, 49.3, 27.2, 25.3, 23.4, 22.7, 22.1 LC-ESI-MS (+ve ion mode): 418 $[M+H]^+$ HPLC purity 98%, t_r = 8.26 min

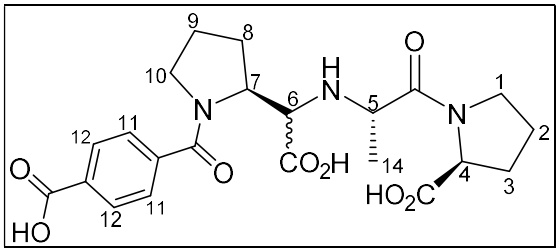
SG1_2



The general ester deprotection was performed on 75 mg SG1_2-8 and then purified via prep HPLC to obtain 55 mg (78%) of a sticky white-yellow solid.

^{1}H NMR (400 MHz, Methanol-*d*$_4$) δ 7.60 (m, 2H, H-11), 7.55 – 7.45 (m, 3H, H-12, H-13), 4.67 (q, *J* = 7.1 Hz, 1H, H-5), 4.63 – 4.56 (m, 1H, H-7), 4.33-4.23 (m, 1H, H-4), 4.13 (d, *J* = 6.6 Hz, 1H, H-6), 3.75 – 3.58 (m, 2H, H-1), 3.53 (m, 2H, H-10), 2.41-2.28 (m, 1H, H-2a), 2.27 – 2.14 (m, 2H, H-2e, H-9a), 2.13 – 1.94 (m, 4H, H-3, H-8a, H-9e), 1.91 – 1.76 (m, 1H, H-8e), 1.61 (d, *J* = 6.8 Hz, 3H, H-14). ^{13}C (100 MHz, Methanol-d$_4$): δ 176.3, 174.7, 173.8, 167.9, 133.2, 129.1, 128.2 (2C), 125.9 (2C), 66.5, 65.2, 58.5, 57.2, 50.1, 48.8, 26.9, 25.1, 23.8, 22.0, 21.5. LC-ESI-MS (+ve ion mode): 418 $[M+H]^+$ HPLC purity 98%, t_r = 7.96 min

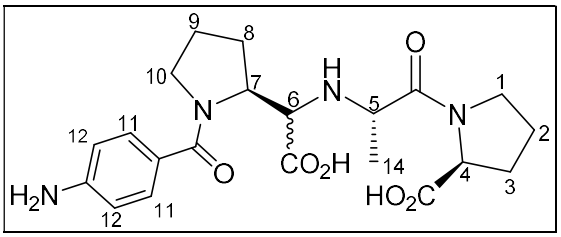
SG-2_2



The general ester deprotection was performed on 30 mg of SG2_2-8 and then purified via prep HPLC to obtain 23 mg (62%) of a sticky white-yellow solid.

^{1}H NMR (400 MHz, Methanol-*d*$_4$) δ 8.07 (d, *J* = 7.8 Hz, 2H, H-12), 7.65 (dd, *J* = 8.2, 2.9 Hz, 2H, H-11), 4.71 – 4.61 (t, J = 8.2 Hz, H-6), 4.40 (q, *J* = 6.9 Hz, 1H, H-5), 4.12 (m, 1H, H-4), 3.96 (m, 1H, H-7), 3.62 (m, 2H, H-1), 3.43 (m, 2H, H-10), 2.32 (m, 3H, H-2, H-8a), 1.98 (m, 3H, H-3, H-8e), 1.86 (m, 1H, H-9a), 1.74 (m, 1H, H-9e), 1.58 (d, J = 6.4 Hz, 3H, H-14). ^{13}C (100 MHz, Methanol-d$_4$): δ 175.9, 175.1, 174.5, 171.3, 167.2, 134.3, 128.8, 128.1 (2C), 126.3 (2C), 67.1, 66.2, 59.3, 57.9, 51.5, 49.1, 27.3, 25.8, 23.1, 22.5, 21.3. LC-ESI-MS (+ve ion mode): 462 $[M+H]^+$ HPLC purity 97%, t_r = 7.31 min

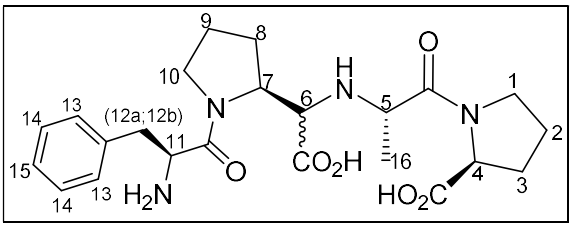
SG3_2



The general ester deprotection was performed on 50 mg of SG3_2-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified via Prep HPLC to yield 30 mg (78%) of SG3_2 as a sticky yellow solid.

^{1}H NMR (400 MHz, D_2O) δ 7.61 – 7.51 (m, 2H, H-12), 7.27 (dd, *J* = 8.7, 2.9 Hz, 2H, H-11), 4.51 (t, *J* = 8.5 Hz, 1H, H-7), 4.33 (q, J = 6.9 Hz, 1H, H-5), 4.29 (m, 1H, H-4), 3.96 (m, 1H, H-6), 3.62 – 3.36 (m, 4H, H-1, H-10), 2.21 (m, 2H, H-2), 2.13 – 2.00 (m, 1H, H-9a), 1.98 – 1.79 (m, 4H, H-2, H-8a, H-9e), 1.66 (m, 1H, H-8e), 1.50 (d, *J* = 6.9 Hz, 3H, H-14). ^{13}C (100 MHz, D_2O): δ 176.3, 174.8, 174.1, 168.3, 133.9, 128.1, 127.7 (2C), 126.8 (2C), 67.5, 66.1, 60.7, 58.2, 52.7, 48.9, 28.1, 26.4, 23.7, 22.1, 21.1. LC-ESI-MS (+ve ion mode): 433 $[M+H]^+$ HPLC purity 93.79%, t_r = 4.95 min

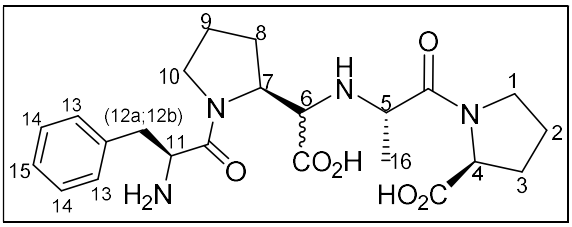
SG4_1



The general ester deprotection was performed on 50 mg of SG4_1-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified vi prep-HPLC to yield 38 mg (97%) of SG4_1 as a sticky yellow solid.

^{1}H NMR (400 MHz, D_2O) δ 7.32 (m, 3H, H-14, H-15), 7.25 (m, 2H, H-13), 4.51 (t, J = 7.6 Hz, 1H, H-11), 4.40-4.30 (m, 3H, H-4, H-5, H-7), 3.82 (d, *J* = 4.4 Hz, 1H, H-6), 3.64 (d, 2H, H-1), 3.54 – 3.39 (m, 2H, H-10), 3.23 (dd, *J* = 14.7, 5.1 Hz, 1H, H-12a), 3.06 (dd, *J* = 14.8, 7.7 Hz, 1H, H-12b), 2.21 (m, 1H, H-2a), 2.07 (m, 1H, H-2e), 2.00 – 1.81 (m, 6H, H-3, H-8, H-9), 1.48 – 1.36 (m, 3H, H-14). ^{13}C (100 MHz, D_2O): δ 177.1, 175.3, 174.8, 171.2, 136.7, 129.1 (2C), 128.5 (2C), 127.2, 68.3, 67.5, 62.1, 57.3, 53.6, 51.9, 49.2, 40.1, 29.3, 25.7, 24.1, 22.8, 22.1 LC-ESI-MS (+ve ion mode): 461 $[M+H]^+$ HPLC purity 97.4%, t_r = 5.52 min

SG4_2

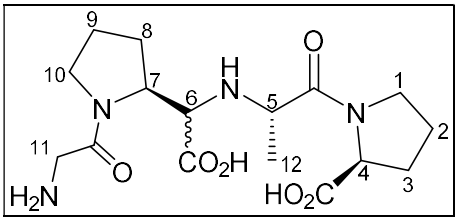


The general ester deprotection was performed on 20 mg of SG4_2-8 after which the crude residue was Boc deprotected via the general Boc deprotection after which it was purified via prep-HPLC to yield 12 mg (77%) of SG4_2 as a sticky yellow solid.

^{1}H NMR (400 MHz, D_2O) δ 7.31 (d, *J* = 7.7 Hz, 3H, H-14, H-15), 7.21 (d, *J* = 7.1 Hz, 2H, H-13), 4.43 (t, *J* = 7.6 Hz, 1H, H-11), 4.37 (m, 1H, H-4), 4.23 (m, 1H, H-7), 4.13 (q, *J* = 6.9 Hz, 1H, H-5), 3.71 (d, *J* = 5.8 Hz, 1H, H-6), 3.51 (m, 2H, H-10), 3.30 (m, 1H, H-1), 3.13 (dd, *J* = 14.6, 7.8 Hz, 1H, 12a), 3.07 (dd, J = 15.4, 6.4, 12b), 2.19 (m, 2H, H-2), 1.87 (m, 4H, H-3, H-9), 1.67 (m, 2H, H-8), 1.47 (d *J* = 6.9 Hz, 3H). ^{13}C (100 MHz, D_2O): δ 177.8, 176.2, 174.4, 172.1, 137.1, 130.3 (2C), 127.8 (2C), 127.1, 69.5, 68.4, 63.5, 56.8, 54.1, 52.1, 50.1, 42.2, 30.5, 26.7, 25.5, 23.8, 22.9 LC-ESI-MS (+ve ion mode): 461 $[M+H]^+$ HPLC purity 95.3%, t_r = 6.69 min

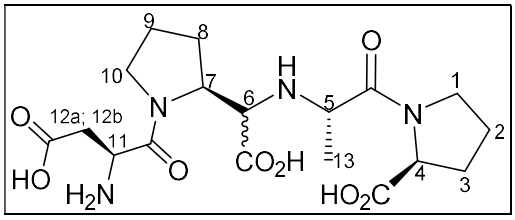
SG5_1

The general ester deprotection was performed on 90 mg of SG5_2-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified via HPLC to yield 65 mg (97%) of SG5_2 as a sticky yellow solid.



^{1}H NMR (400 MHz, D_2O) δ 4.42 (m, 2H, H-4, H-7), 4.29 (q, *J* = 6.9 Hz, 1H, H-5), 3.95 – 3.84 (m, 3H, H-6, H-10), 3.54 (m, 2H, H-1), 3.48 – 3.33 (m, 2H, H-11), 2.25 (s, 1H, H-2a), 2.15 (m, 1H, H-2e), 2.04 (m, 1H, H-9a), 1.95 (m, 4H, H-3, H-8a, H-9e), 1.82 (m, 1H, H-8e), 1.50 (d, *J* = 6.9 Hz, 3H, H-12). ^{13}C (100 MHz, D_2O): δ 177.1, 176.2, 174.7, 170.9, 70.5, 63.1, 56.3, 55.5, 52.3, 51.5, 42.9, 31.8, 26.9, 25.9, 24.1, 23.5 LC-ESI-MS (+ve ion mode): 371 $[M+H]^+$ HPLC purity 93.6%, t_r = 5.66 min

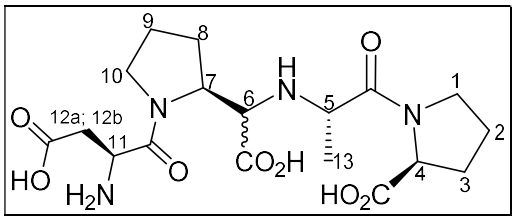
SG6_1



The general ester deprotection was performed on 150 mg of SG6_1-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified via prep-HPLC to yield 100 mg (89%) of SG6_1 as a sticky yellow-brown solid.

^{1}H NMR (400 MHz, D_2O) δ 4.53 (m, 1H, H-11), 4.47 – 4.38 (m, 1H, H-5), 4.30 (m, 1H, H-4), 4.04 (m, 1H, H-7), 3.98 (d, *J* = 5.6 Hz, 1H, H-6), 3.59 (m, 4H, H-1, H-10), 3.06 – 2.71 (m, 2H, H-12a, H-12b), 2.28 (m, 2H, H-2), 2.11 – 1.82 (m, 6H, H-3, H-8, H-9), 1.47 (d, *J* = 6.8 Hz, 3H, H-13). ^{13}C (100 MHz, D_2O): δ 177.9, 177.2, 175.4, 173.5 171.2, 72.3, 67.4, 59.2, 57.5, 52.3, 51.5, 49.4, 43.2, 30.5, 27.5, 26.3, 24.8, 23.1, LC-ESI-MS (+ve ion mode): 429 $[M+H]^+$ HPLC purity 93.83%, t_r = 4.73 min

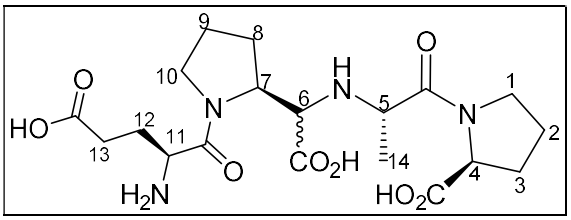
SG6_2



The general ester deprotection was performed on 230 mg of SG6_2-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified via prep-HPLC to yield 85 mg (49%) of SG6_2 as a sticky yellow-brown solid.

^{1}H NMR (400 MHz, D_2O) δ 4.60-4.50 (m, 1H, H-11), 4.48-4.35 (m, 2H, H-4, H-5), 4.33-4.24 (m, 1H, H-7), 3.85 (d, *J* = 2.1 Hz, 1H, H-6), 3.70-3.48 (m, 3H, H-1, H-10a), 3.47-3.35 (m, 1H, H-10e), 3.03-2.91 (m, 1H, H-12a), 2.84 – 2.71 (m, 1H, H-12b), 2.23-2.14 (m, 2H, H-2), 2.13 – 1.75 (m, 6H, H-3, H-8, H-9), 1.59 – 1.35 (d, *J* = 6.1 Hz, 3H, H-13). ^{13}C (100 MHz, D_2O): δ 177.1, 176.5, 175.9, 174.2 170.6, 70.9, 68.5, 59.5, 58.2, 53.8, 50.2, 49.1, 41.9, 31.1, 28.3, 27.1, 24.1, 23.5. LC-ESI-MS (+ve ion mode): 429 $[M+H]^+$ HPLC purity 97.55%, t_r = 4.76 min

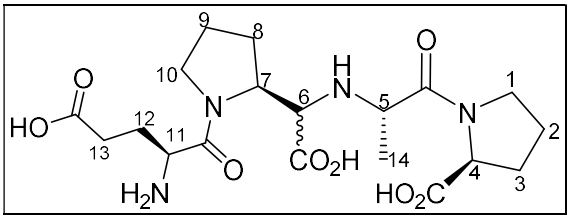
SG7_1



The general ester deprotection was performed on 70 mg of SG7_1-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified to yield 4 mg (8%) of SG7_1 as a sticky yellow-brown solid

^{1}H NMR (400 MHz, D_2O) δ 4.41 (t, J = 4.4 Hz, 1H, H-11), 4.32 (q, J = 5.8 Hz, 1H, H-7), 4.15 (q, *J* = 6.9 Hz, 1H, H-5), 3.78 (d, *J* = 5.7 Hz, 1H, H-6), 3.66 – 3.47 (m, 4H, H-1, H-10), 2.59 – 2.44 (m, 2H, H-13), 2.24 (m, 1H, H-2a), 2.09 (m, 2H, H-2e, H-12a), 1.98 (m, 2H, H-9a, H-12b), 1.93 (m, 5H, H-3, H-10, H-9e), 1.48 (d, *J* = 7.1 Hz, 3H, H-14). ^{13}C (100 MHz, D_2O): δ 177.5, 177.1, 176.4, 174.8 172.4, 71.5, 69.1, 60.3, 57.5, 52.4, 51.3, 49.8, 42.1, 32.5, 29.1, 27.7, 25.6, 23.9, 23.3. LC-ESI-MS (+ve ion mode): 443 $[M+H]^+$

SG7_2

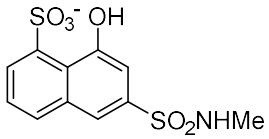


The general ester deprotection was performed on 80 mg of SG7_2-8 after which the crude residue was Boc deprotected via the general Boc deprotection method after which it was purified to yield 6 mg (10%) of SG7_2 as a sticky yellow-brown solid.

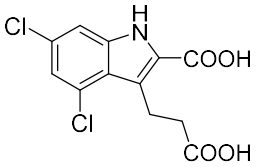
^{1}H NMR (400 MHz, D_2O) δ 4.40 – 4.23 (m, 4H, H-4, H-5, H-7, H-11), 3.76 (d, *J* = 2.0 Hz, 1H, H-6), 3.65 (m, 1H, H-1a), 3.53 (m, 2H, H-1e, H10a), 3.39 (m, 1H, H-10e), 2.54 (t, *J* = 7.5 Hz, 2H, H-13), 2.22 (m, 2H, H-2), 2.13 (m 2H, H-9a, H-12a), 1.95 (m, 5H, H-3, H-8a, H-9e H-12b), 1.77 (q, 1H, H-9a), 1.49 (d, *J* = 6.9 Hz, 3H). ^{13}C (100 MHz, D_2O): δ 178.2, 177.4, 175.5, 174.9 171.4, 72.3, 68.9, 61.5, 58.1, 51.9, 51.2, 48.3, 41.9, 31.1, 29.7, 28.1, 26.4, 24.2, 22.9. LC-ESI-MS (+ve ion mode): 443 $[M+H]^+$ HPLC purity 98.95%, t_r = 4.76 min

Appendices

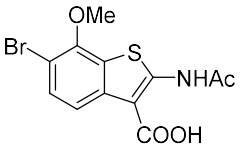
Appendix 2.1 - Fragment Compounds Screened against ACE



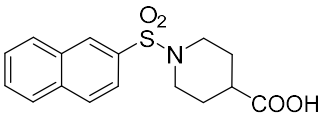
SF1



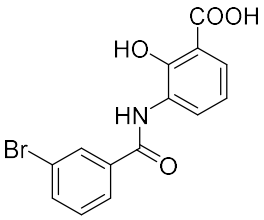
SF2



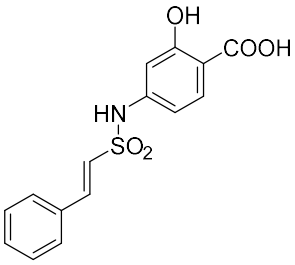
SF3



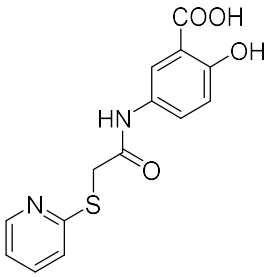
SF4



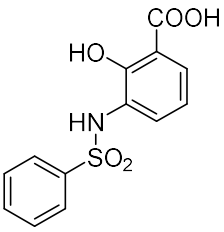
SF5



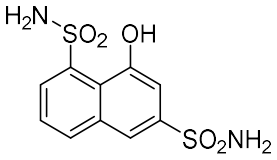
SF6



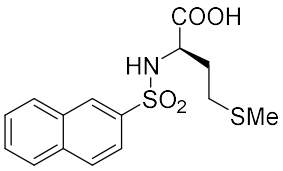
SF7



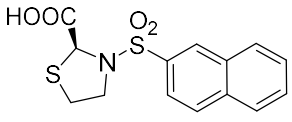
SF8



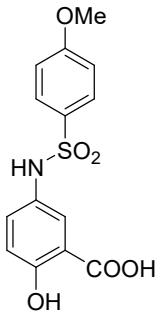
SF9



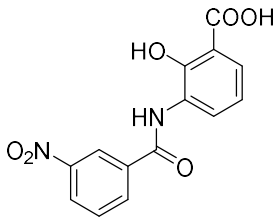
SF10



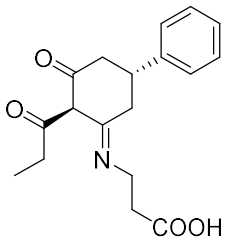
SF11



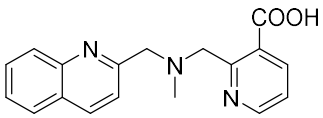
SF12



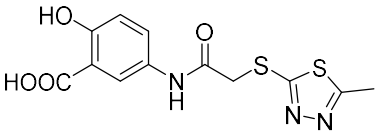
SF13



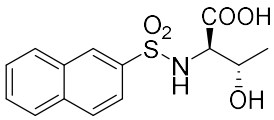
SF14



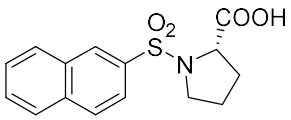
SF15



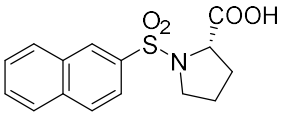
SF16



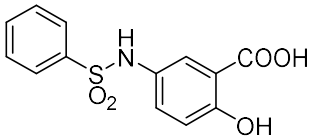
SF17



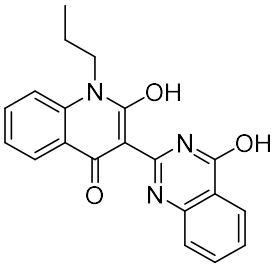
SF18



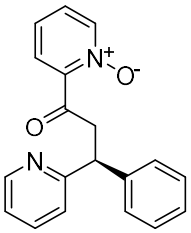
SF19



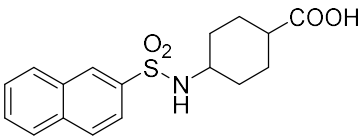
SF20



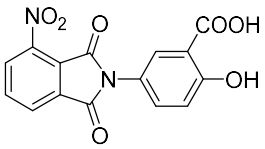
SF21



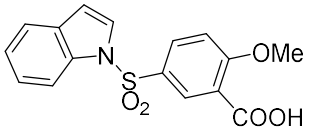
SF22



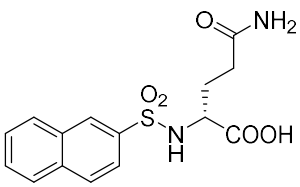
SF23



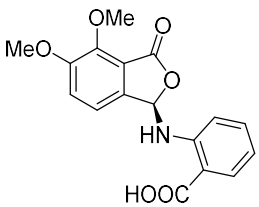
SF24



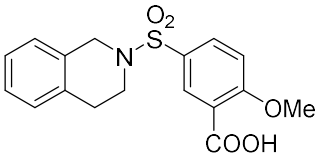
SF25



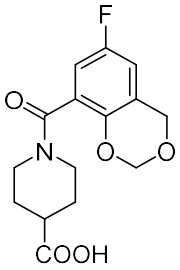
SF26



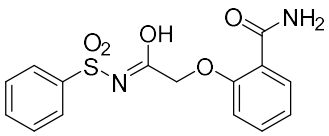
SF27



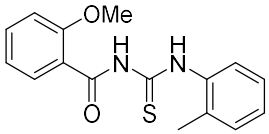
SF28



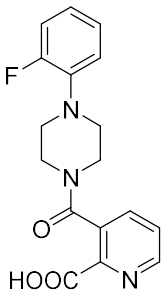
SF29



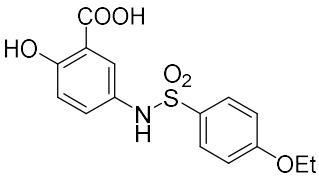
SF30



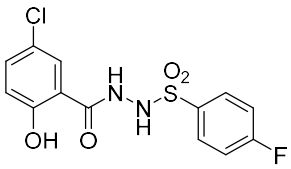
SF31



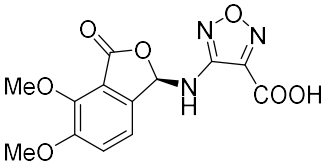
SF32



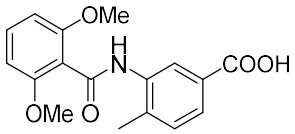
SF33



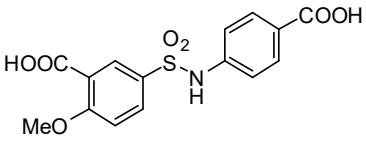
SF34



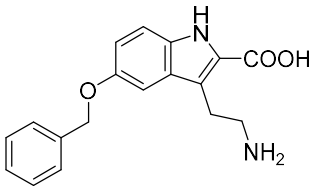
SF35



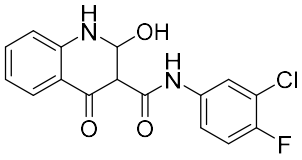
SF36



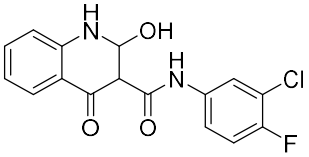
SF37



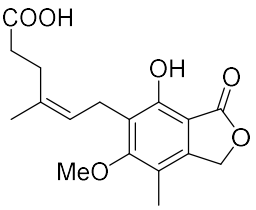
SF38



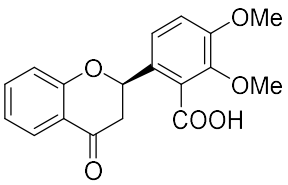
SF39



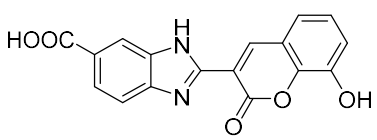
SF40



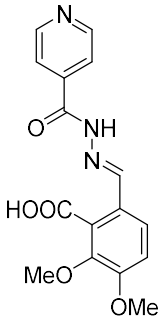
SF41



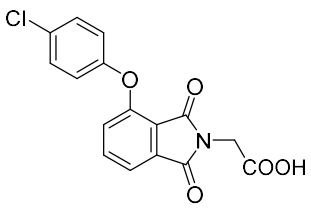
SF42



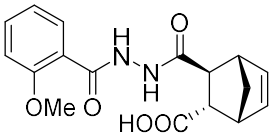
SF43



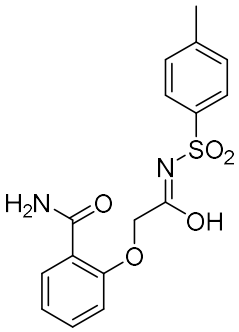
SF44



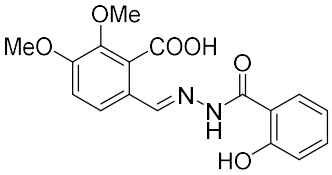
SF45



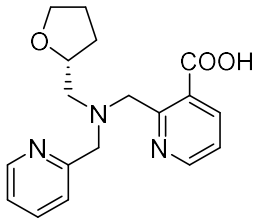
SF46



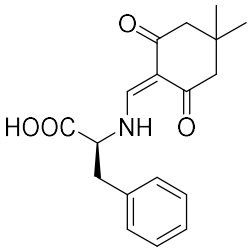
SF47



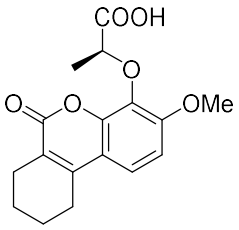
SF48



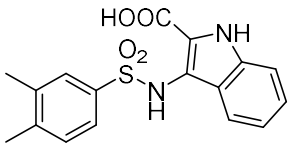
SF49



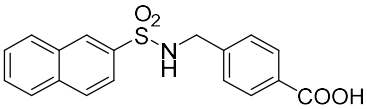
SF50



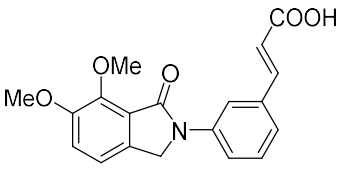
SF51



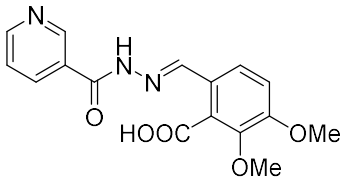
SF52



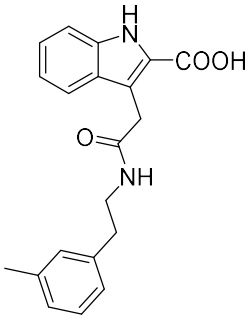
SF53



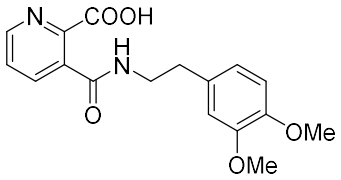
SF54



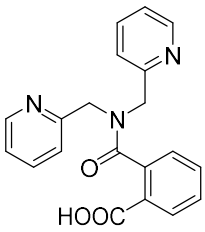
SF55



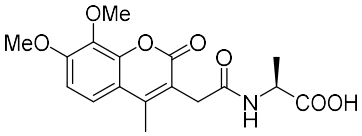
SF56



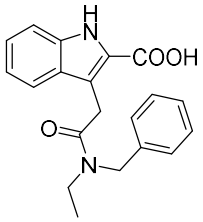
SF57



SF58

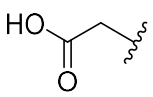
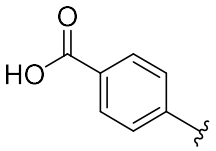
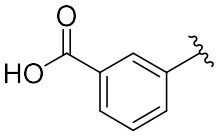
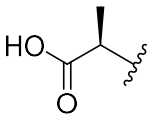
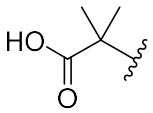
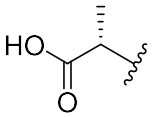
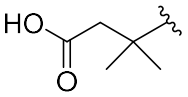
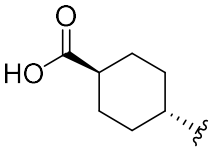
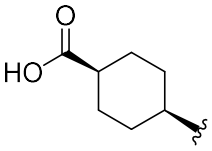
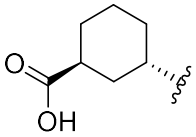
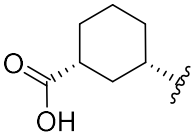
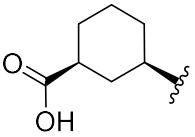
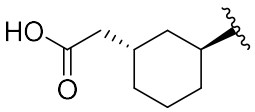
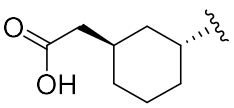
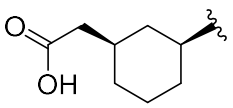
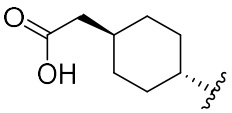
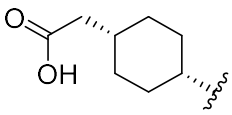
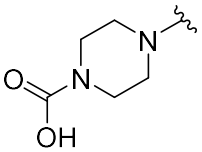
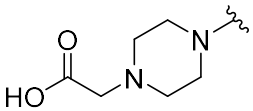
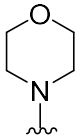
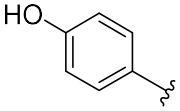
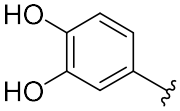
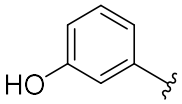
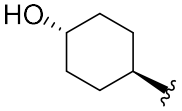


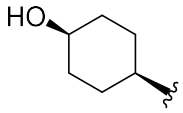
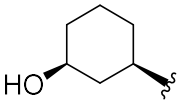
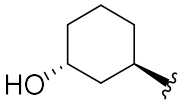
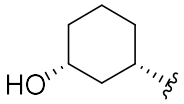
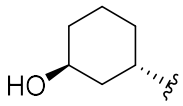
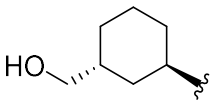
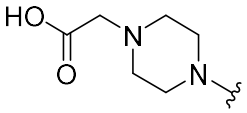
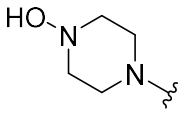
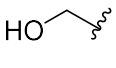
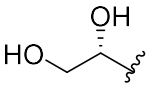
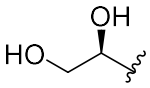
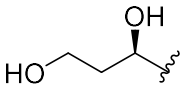
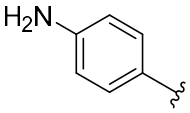
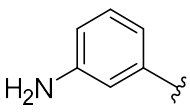
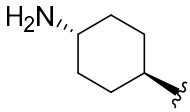
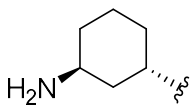
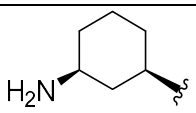
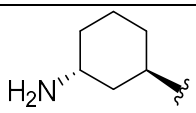
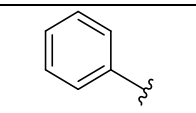
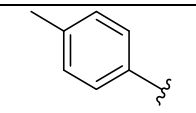
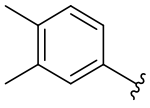
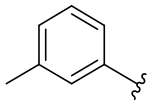
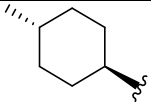
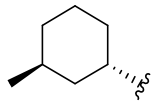
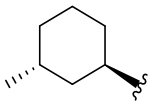
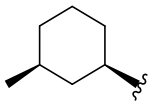
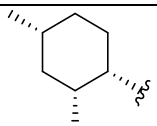
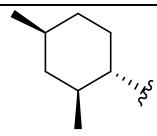
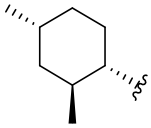
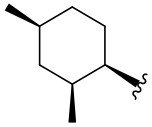
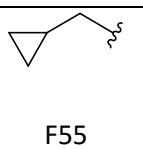
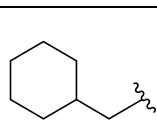
SF59

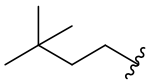
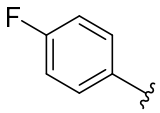
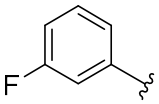
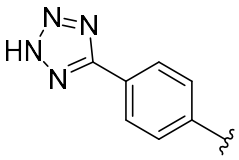
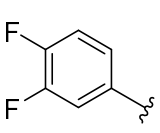
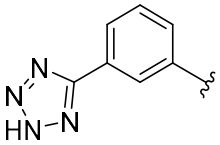
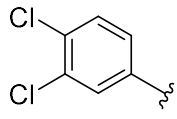
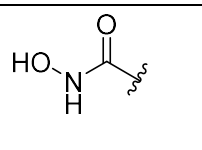
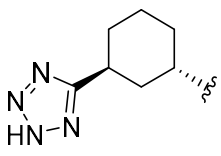
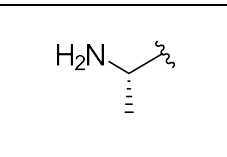
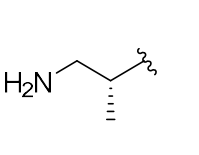


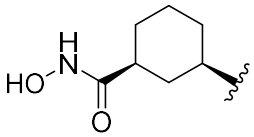
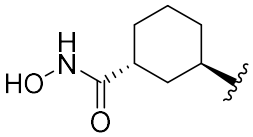
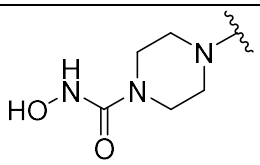
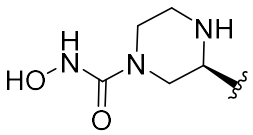
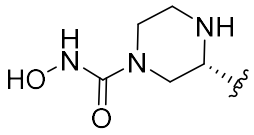
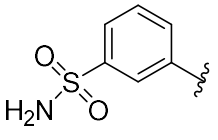
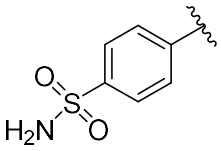
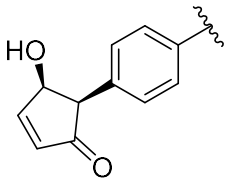
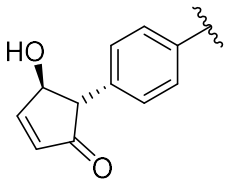
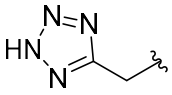
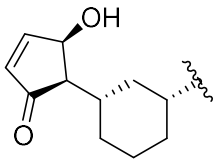
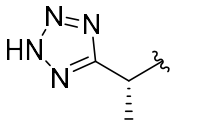
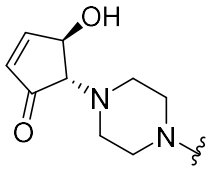
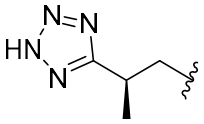
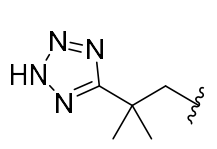
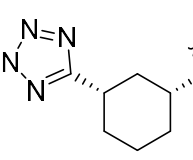
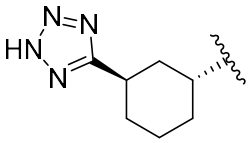
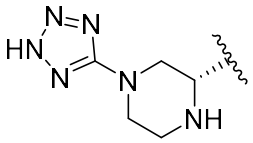
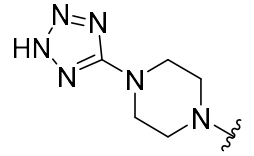
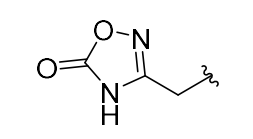
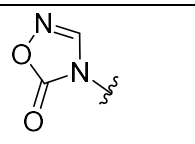
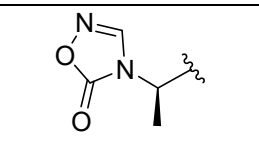
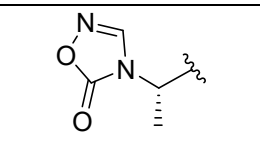
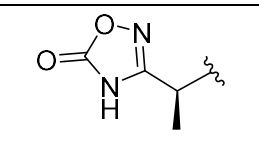
SF60

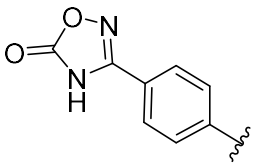
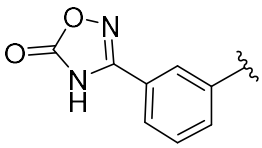
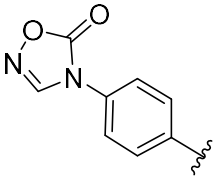
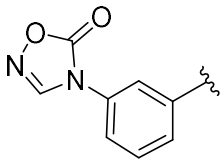
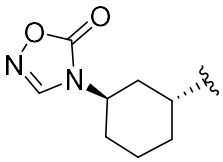
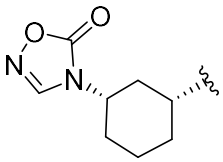
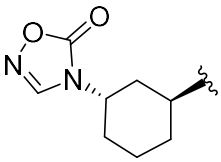
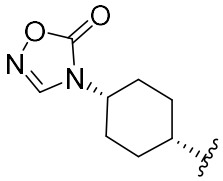
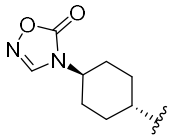
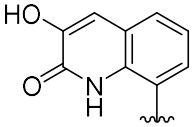
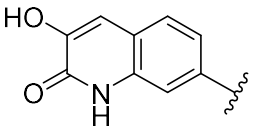
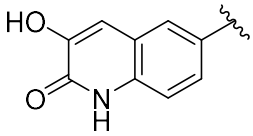
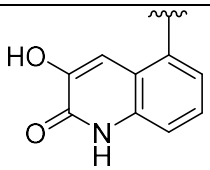
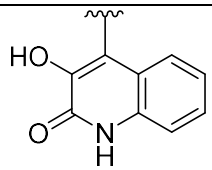
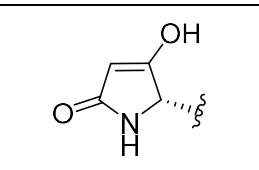
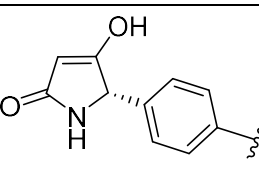
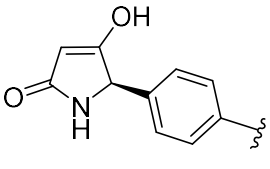
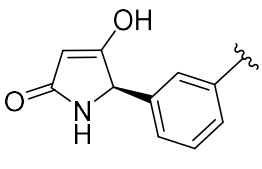
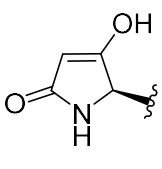
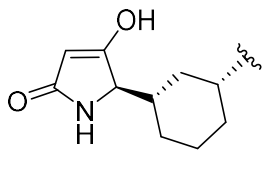
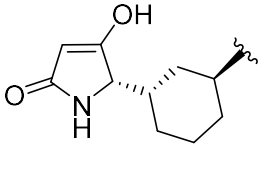
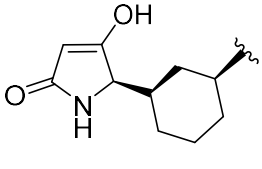
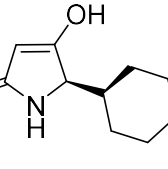
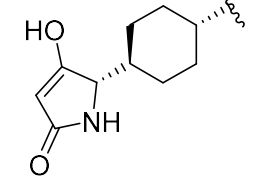
Appendix 3.1- Combinatorial Fragment Groups

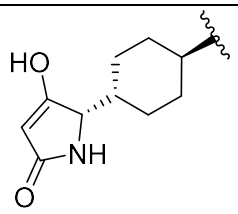
 F1	 F2	 F3	 F4
 F5	 F6	 F7	 F8
 F9	 F10	 F11	 F12
 F13	 F14	 F15	 F16
 F17	 F18	 F19	 F20
 F21	 F22	 F23	 F24

			
F25	F26	F27	F28
			
F29	F30	F31	F32
			
F33	F34	F35	F36
			
F37	F38	F39	F40
			
F41	F42	F43	F44
			
F45	F46	F47	F48
			
F49	F50	F51	F52
			
		F55	

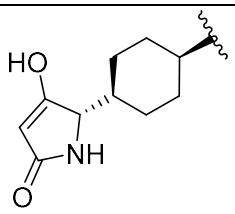
F53	F54		F56
			
F57	F58	F59	F60
			
F61-1	F61	F62-1	F62
			
F63	F64	F65	F66
			
F67-1	F67	F68	F69
			
F70	F71	F72	F73
			
F74	F75	F76	F77

			
F78	F79	F80	F81
			
F82	F83	F84	F85-1
			
F86-1	F86	F87-1	F87
			
F88-1	F88	F89	F90
			
F91	F92	F93	F94
			
F95	F96	F97	F98

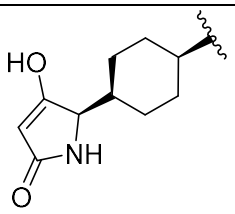
			
F99	F100	F101	F102
			
F103	F104	F105	F106
			
F107	F108	F109	F110
			
F111	F112	113	114
			
F115	F116	117	118
			
F119	F120	121	122



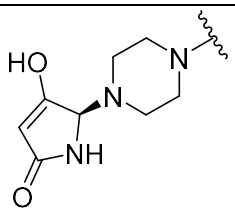
F123



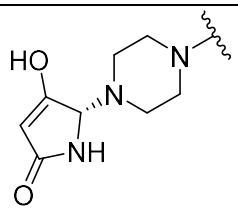
F124



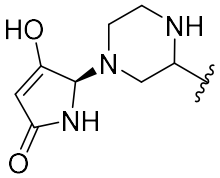
125



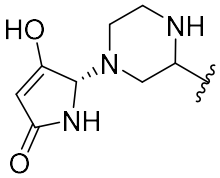
126



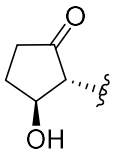
F127



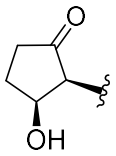
F128



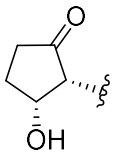
129



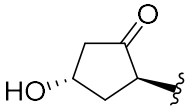
130



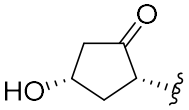
F131



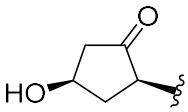
F132



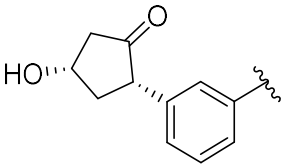
133



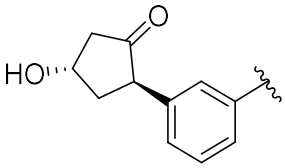
134



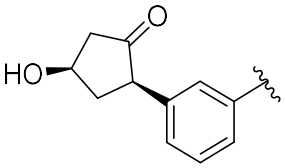
F135



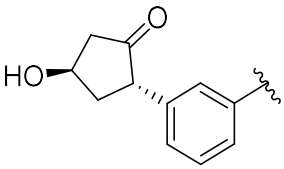
F136



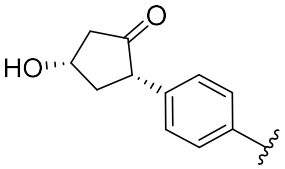
F137



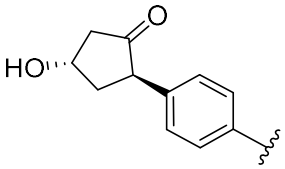
F138



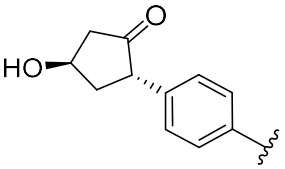
F139



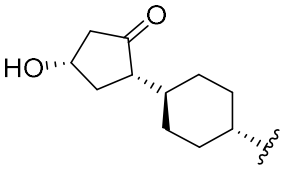
F140



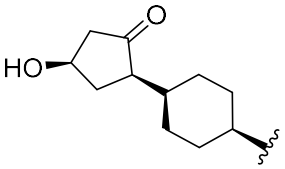
F141



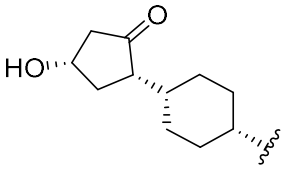
F142



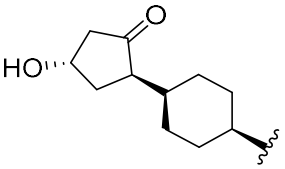
F143



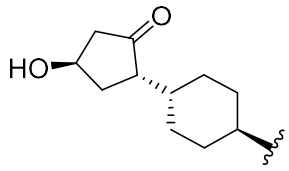
F144



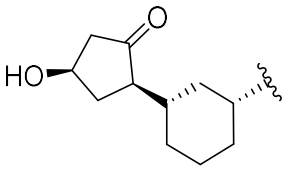
F145



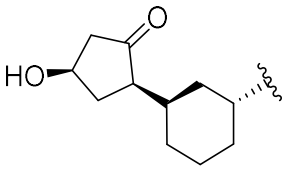
F146



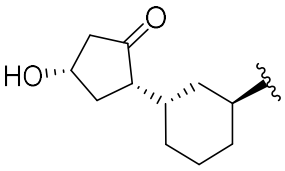
147



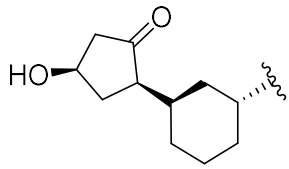
F154



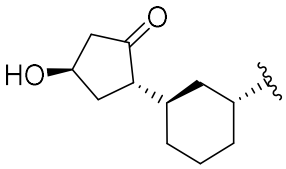
F155



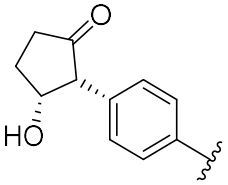
F156



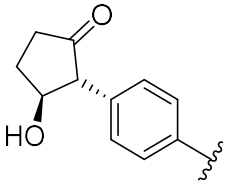
157



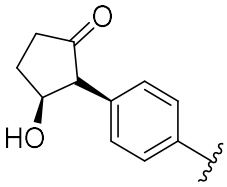
F158



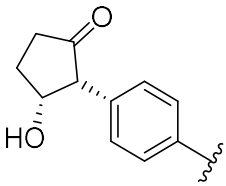
F159



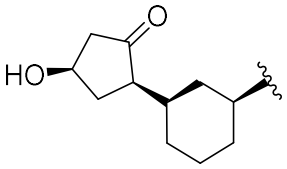
F160



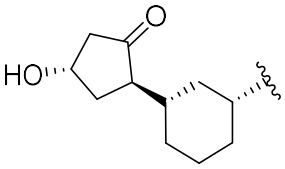
161



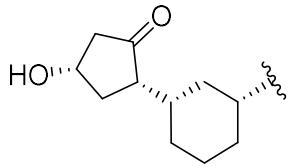
F162



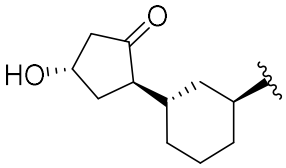
F163



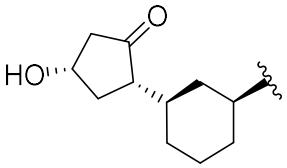
164



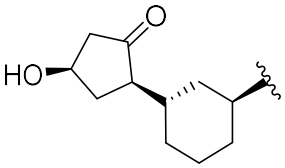
F165



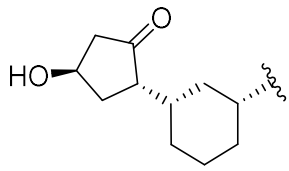
F166



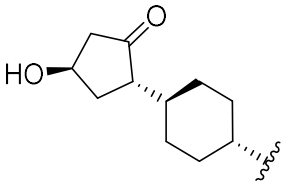
F167



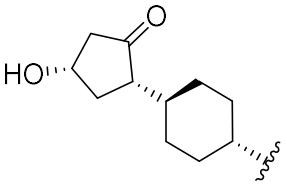
F168



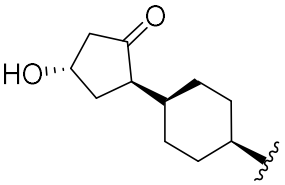
F169



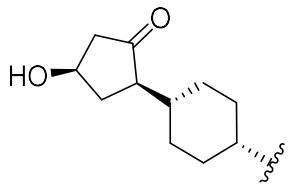
F171



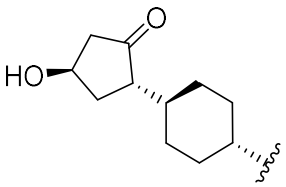
F172



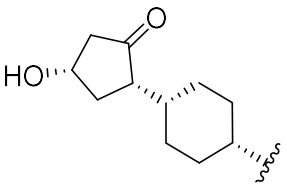
F173



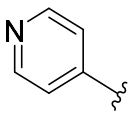
F174



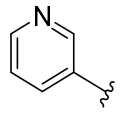
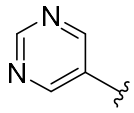
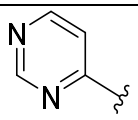
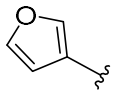
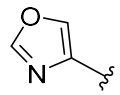
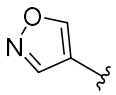
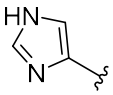
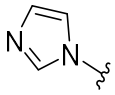
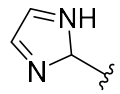
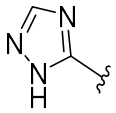
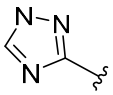
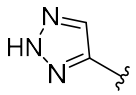
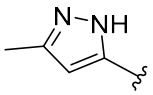
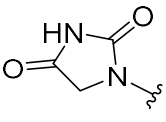
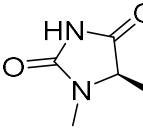
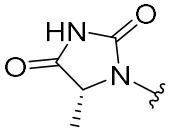
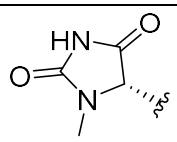
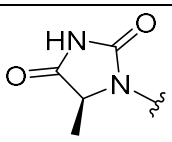
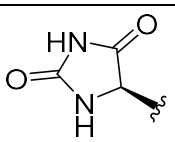
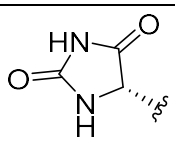
F175



F176



F177

			
F178	F179	F180	F181
			
F182	F183	F184	F185
			
F186	F187	F188	F189
			
F190	F191	F192	193
			
F194	F195	F198	F197

Appendix 3.2 - Full Table of Ligand Docking Scores and MM-GBSA Calculated Binding Energies

	C-Domain			N-Domain			
Ligand	**Docking score**	**ΔG**	**C Norm**	**Docking score**	**ΔG**	**N Norm**	**N/C**
F61-1_L0	-8.129	-18.148	0.367	-9.852	-50.344	1.130	3.078
F146_L0	-11.283	-41.466	0.839	-10.791	-75.425	1.694	2.018
F187_L1	-10.990	-26.264	0.532	-10.052	-39.481	0.887	1.668
F12_L1	-11.641	-27.544	0.557	-10.400	-38.465	0.864	1.549
F36_L2	-12.347	-50.859	1.029	-11.534	-69.204	1.554	1.510
F61-1_L1	-12.443	-44.240	0.895	-10.365	-57.132	1.283	1.433
F76_L0	-11.033	-43.127	0.873	-10.077	-53.393	1.199	1.374
F141_L0	-13.102	-62.547	1.266	-10.696	-76.269	1.713	1.353
F161_L0	-12.460	-56.189	1.137	-11.335	-66.489	1.493	1.313
F21_L0	-11.576	-53.394	1.081	-12.486	-61.978	1.392	1.288
F119_L0	-10.011	-45.101	0.913	-10.752	-50.834	1.141	1.251
F21_L1	-12.797	-53.683	1.086	-13.472	-60.314	1.354	1.247
F64_L0	-11.296	-42.692	0.864	-9.774	-47.255	1.061	1.228
F32_L0	-11.233	-49.514	1.002	-11.891	-54.074	1.214	1.212
F38_L1	-12.786	-60.189	1.218	-11.639	-64.852	1.456	1.195
F171_L0	-12.058	-73.675	1.491	-10.878	-78.354	1.759	1.180
F75_L1	-11.278	-60.142	1.217	-9.944	-63.530	1.427	1.172
F160_L0	-12.014	-67.902	1.374	-10.911	-70.982	1.594	1.160
F191_L3	-12.349	-43.102	0.872	-11.511	-44.966	1.010	1.158
F179_L1	-11.250	-46.727	0.946	-10.938	-48.589	1.091	1.154
F35_L0	-12.256	-47.501	0.961	-12.470	-49.251	1.106	1.150
F22_L1	-13.753	-63.554	1.286	-12.998	-65.831	1.478	1.149
F123_L0	-11.177	-51.818	1.049	-10.824	-53.670	1.205	1.149
F50_L1	-11.687	-54.150	1.096	-11.236	-55.635	1.249	1.140
F4_L2	-11.309	-42.407	0.858	-11.519	-43.553	0.978	1.140
F191_L2	-11.549	-43.020	0.871	-12.188	-43.877	0.985	1.132
F188_L3	-12.080	-47.700	0.965	-11.354	-48.630	1.092	1.131
F72_L2	-12.045	-59.072	1.195	-10.461	-59.666	1.340	1.121
F19_L0	-9.787	-63.348	1.282	-10.088	-63.974	1.437	1.121
F100_L0	-11.997	-61.529	1.245	-11.591	-61.578	1.383	1.110
F35_L3	-12.838	-60.101	1.216	-12.004	-59.274	1.331	1.094
F107_L0	-10.010	-73.580	1.489	-10.756	-71.855	1.613	1.084
F5_L2	-11.538	-45.976	0.930	-10.118	-44.390	0.997	1.071
F196_L1	-11.615	-40.158	0.813	-11.902	-38.701	0.869	1.069
F69_L1	-11.527	-56.983	1.153	-10.039	-54.884	1.232	1.069
F143_L0	-12.077	-64.275	1.301	-10.995	-61.760	1.387	1.066
F111_L0	-12.301	-60.864	1.232	-12.399	-58.382	1.311	1.064
F142_L0	-12.854	-69.340	1.403	-10.347	-66.060	1.483	1.057
F191_L1	-11.732	-48.108	0.974	-10.894	-45.800	1.028	1.056

F181_L1	-12.719	-51.636	1.045	-11.263	-49.042	1.101	1.054
F87_L0	-12.321	-54.321	1.099	-9.641	-51.574	1.158	1.053
F70_L0	-11.333	-59.760	1.209	-10.018	-56.706	1.273	1.053
F170_L0	-11.906	-76.073	1.540	-10.565	-71.983	1.616	1.050
F130_L1	-12.565	-65.536	1.326	-11.547	-61.976	1.392	1.049
F104_L0	-11.482	-65.863	1.333	-10.261	-61.460	1.380	1.035
F109_L0	-13.158	-55.298	1.119	-11.821	-51.318	1.152	1.030
F125_L0	-11.727	-62.069	1.256	-11.228	-57.514	1.291	1.028
F72_L3	-12.745	-72.395	1.465	-10.546	-66.768	1.499	1.023
F135_L2	-11.932	-72.017	1.457	-11.372	-66.384	1.491	1.023
F62-1_L0	-11.419	-50.686	1.026	-10.928	-46.623	1.047	1.021
F17_L0	-10.607	-40.025	0.810	-10.420	-36.705	0.824	1.017
F182_L1	-12.257	-60.611	1.227	-11.222	-55.375	1.243	1.014
F2_L1	-12.567	-42.045	0.851	-12.818	-38.408	0.862	1.014
F44_L0	-11.694	-57.508	1.164	-10.474	-52.436	1.177	1.012
F86_L3	-12.072	-51.729	1.047	-10.076	-47.106	1.058	1.010
F197_L2	-12.263	-50.141	1.015	-12.607	-45.521	1.022	1.007
F181_L2	-12.010	-50.547	1.023	-11.998	-45.862	1.030	1.007
F155_L0	-12.541	-69.485	1.406	-10.926	-62.993	1.414	1.006
F116_L0	-11.780	-55.418	1.122	-10.870	-50.083	1.125	1.003
F189_L2 (Enalprilat)	**-12.128**	**-49.413**	**1.000**	**-9.964**	**-44.535**	**1.000**	**1.000**
F58_L3	-11.418	-61.708	1.249	-10.873	-55.580	1.248	0.999
F189_L0	-11.528	-50.716	1.026	-9.936	-45.278	1.017	0.991
F84_L0	-11.531	-66.542	1.347	-10.209	-59.287	1.331	0.989
F61_L0	-11.873	-57.337	1.160	-10.266	-50.848	1.142	0.984
F186_L3	-12.254	-64.496	1.305	-11.160	-57.146	1.283	0.983
F63_L0	-11.173	-52.027	1.053	-9.887	-46.068	1.034	0.982
F30_L0	-11.862	-53.398	1.081	-10.931	-47.048	1.056	0.978
F23_L1	-13.508	-64.817	1.312	-12.264	-57.090	1.282	0.977
F63_L1	-11.255	-45.768	0.926	-10.882	-40.289	0.905	0.977
F108_L0	-11.834	-65.125	1.318	-11.973	-57.327	1.287	0.977
F187_L0	-11.056	-46.578	0.943	-11.106	-40.964	0.920	0.976
F56_L3	-12.216	-58.305	1.180	-10.790	-51.258	1.151	0.975
F67-1_L0	-11.296	-55.401	1.121	-10.107	-48.691	1.093	0.975
F37_L2	-13.018	-76.517	1.549	-12.665	-67.231	1.510	0.975
F172_L0	-10.350	-80.270	1.624	-10.649	-70.448	1.582	0.974
F28_L0	-11.778	-57.647	1.167	-10.660	-50.455	1.133	0.971
F22_L0	-12.193	-61.927	1.253	-11.757	-54.178	1.217	0.971
F86-1_L0	-13.066	-68.189	1.380	-10.915	-59.547	1.337	0.969
F18_L0	-9.732	-62.027	1.255	-9.968	-54.105	1.215	0.968
F117_L1	-11.783	-44.140	0.893	-11.822	-38.404	0.862	0.965
F31_L0	-11.953	-60.893	1.232	-11.804	-52.758	1.185	0.961
F115_L0	-11.885	-57.348	1.161	-11.305	-49.658	1.115	0.961
F180_L1	-12.309	-66.942	1.355	-11.170	-57.953	1.301	0.961

F189_L1	-11.376	-48.729	0.986	-11.175	-42.167	0.947	0.960
F55_L1	-12.056	-53.670	1.086	-10.171	-46.380	1.041	0.959
F30_L1	-12.371	-75.715	1.532	-11.019	-65.413	1.469	0.959
F56_L0	-11.387	-66.906	1.354	-10.643	-57.765	1.297	0.958
F139_L0	-11.995	-78.903	1.597	-12.061	-68.037	1.528	0.957
F45_L0	-10.082	-59.631	1.207	-10.911	-51.298	1.152	0.954
F86_L1	-11.034	-43.948	0.889	-10.536	-37.547	0.843	0.948
F47_L1	-11.652	-68.390	1.384	-10.525	-58.428	1.312	0.948
F194_L1	-11.732	-47.457	0.960	-11.534	-40.426	0.908	0.945
F190_L2	-12.528	-41.744	0.845	-10.123	-35.548	0.798	0.945
F86_L0	-11.080	-51.916	1.051	-10.250	-44.195	0.992	0.945
F126_L0	-11.355	-56.596	1.145	-11.218	-48.088	1.080	0.943
F13_L0	-11.920	-61.695	1.249	-10.514	-52.247	1.173	0.940
F183_L3	-11.921	-65.639	1.328	-11.397	-55.566	1.248	0.939
F55_L2	-12.005	-63.796	1.291	-11.512	-54.000	1.213	0.939
F5-1_L0	-11.696	-60.834	1.231	-11.233	-51.438	1.155	0.938
F186_L1	-12.139	-50.944	1.031	-10.940	-43.038	0.966	0.937
F4-1_L0	-11.578	-52.181	1.056	-11.348	-44.036	0.989	0.936
F102_L0	-11.718	-67.838	1.373	-10.130	-57.207	1.285	0.936
F101_L0	-12.434	-66.192	1.340	-10.528	-55.557	1.248	0.931
F94_L1	-12.059	-67.051	1.357	-10.013	-56.261	1.263	0.931
F34_L2	-12.499	-73.498	1.487	-12.256	-61.648	1.384	0.931
F190_L3	-11.981	-63.742	1.290	-10.309	-53.131	1.193	0.925
F60_L1	-12.683	-61.811	1.251	-11.207	-51.513	1.157	0.925
F178_L0	-11.913	-61.055	1.236	-10.556	-50.810	1.141	0.923
F192_L1	-11.651	-56.052	1.134	-10.437	-46.622	1.047	0.923
F177_L2	-12.741	-54.704	1.107	-10.407	-45.494	1.022	0.923
F66_L2	-12.340	-58.710	1.188	-12.212	-48.759	1.095	0.921
F69_L4	-12.658	-73.142	1.480	-11.789	-60.655	1.362	0.920
F60_L0	-11.941	-61.953	1.254	-10.341	-51.173	1.149	0.916
F98_L0	-12.884	-72.197	1.461	-12.901	-59.547	1.337	0.915
F128_L0	-11.002	-55.785	1.129	-11.018	-45.959	1.032	0.914
F88_L0	-12.091	-49.486	1.001	-10.247	-40.732	0.915	0.913
F134_L1	-11.733	-75.537	1.529	-10.868	-62.170	1.396	0.913
F18_L1	-11.437	-69.287	1.402	-10.333	-56.978	1.279	0.912
F33_L2	-11.588	-57.690	1.168	-11.873	-47.369	1.064	0.911
F69_L2	-11.322	-65.787	1.331	-10.259	-53.957	1.212	0.910
F6_L2	-11.887	-50.563	1.023	-10.711	-41.340	0.928	0.907
F58_L0	-11.824	-54.228	1.097	-10.000	-44.270	0.994	0.906
F1_L1	-10.979	-45.920	0.929	-9.591	-37.435	0.841	0.905
F180_L0	-11.637	-58.410	1.182	-9.852	-47.358	1.063	0.900
F94_L2	-12.834	-63.990	1.295	-11.740	-51.854	1.164	0.899
F118_L0	-10.377	-68.468	1.386	-10.328	-55.410	1.244	0.898
F34_L1	-11.973	-55.730	1.128	-11.608	-45.091	1.012	0.898

F49_L1	-11.960	-71.557	1.448	-10.650	-57.752	1.297	0.895
F184_L1	-12.377	-53.651	1.086	-10.344	-43.246	0.971	0.894
F57_L1	-11.674	-59.641	1.207	-10.462	-47.920	1.076	0.891
F57_L3	-11.816	-71.372	1.444	-11.345	-57.264	1.286	0.890
F180_L2	-13.085	-59.972	1.214	-11.435	-48.039	1.079	0.889
F195_L0	-11.816	-51.939	1.051	-10.358	-41.502	0.932	0.887
F87-1_L0	-13.209	-63.786	1.291	-10.626	-50.927	1.144	0.886
F58_L1	-11.786	-57.276	1.159	-11.598	-45.708	1.026	0.885
F59_L2	-12.943	-56.170	1.137	-11.252	-44.724	1.004	0.883
F10_L1	-11.290	-44.066	0.892	-10.187	-35.074	0.788	0.883
F188_L0	-11.191	-48.296	0.977	-10.433	-38.402	0.862	0.882
F190_L1	-11.637	-59.768	1.210	-10.239	-47.485	1.066	0.881
F73_L2	-11.540	-61.207	1.239	-10.450	-48.589	1.091	0.881
F177_L1	-12.532	-65.230	1.320	-11.026	-51.757	1.162	0.880
F167_L0	-12.231	-77.514	1.569	-10.861	-61.207	1.374	0.876
F196_L0	-11.106	-50.481	1.022	-9.960	-39.832	0.894	0.875
F183_L0	-11.678	-64.149	1.298	-10.256	-50.608	1.136	0.875
F195_L1	-12.007	-57.745	1.169	-10.965	-45.545	1.023	0.875
F15_L0	-11.723	-54.764	1.108	-10.608	-43.158	0.969	0.874
F99_L0	-11.703	-77.992	1.578	-10.835	-61.401	1.379	0.874
F130_L0	-11.796	-59.878	1.212	-10.383	-47.038	1.056	0.872
F93_L0	-9.605	-85.116	1.723	-9.957	-66.591	1.495	0.868
F181_L3	-11.662	-66.507	1.346	-10.693	-52.026	1.168	0.868
F188_L1	-10.990	-48.210	0.976	-10.909	-37.677	0.846	0.867
F65_L1	-11.683	-75.567	1.529	-9.625	-59.002	1.325	0.866
F189_L3	-12.307	-58.409	1.182	-10.669	-45.475	1.021	0.864
F196_L2	-11.553	-53.895	1.091	-11.419	-41.959	0.942	0.864
F46_L1	-12.707	-65.792	1.331	-11.765	-51.153	1.149	0.863
F32_L1	-11.862	-72.671	1.471	-10.611	-56.439	1.267	0.862
F163_L0	-12.285	-85.462	1.730	-10.757	-66.372	1.490	0.862
F65_L0	-11.321	-77.243	1.563	-10.512	-59.790	1.343	0.859
F83_L0	-12.286	-69.125	1.399	-10.974	-53.498	1.201	0.859
F34_L3	-13.100	-70.553	1.428	-12.995	-54.502	1.224	0.857
F42_L1	-12.031	-71.029	1.437	-10.135	-54.814	1.231	0.856
F177_L0	-12.280	-62.916	1.273	-10.349	-48.445	1.088	0.854
F185_L1	-12.478	-60.966	1.234	-10.733	-46.868	1.052	0.853
F58_L2	-11.796	-63.966	1.295	-10.272	-49.058	1.102	0.851
F162_L0	-12.649	-83.881	1.698	-11.062	-64.200	1.442	0.849
F28_L1	-12.233	-67.482	1.366	-10.886	-51.637	1.159	0.849
F157_L0	-12.493	-90.429	1.830	-10.744	-68.988	1.549	0.846
F37_L0	-11.560	-71.733	1.452	-10.088	-54.671	1.228	0.846
F174_L0	-12.648	-97.195	1.967	-12.467	-73.880	1.659	0.843
F40_L1	-11.209	-81.859	1.657	-10.382	-62.203	1.397	0.843
F66_L3	-12.287	-74.506	1.508	-11.114	-56.576	1.270	0.843

F117_L3	-12.145	-63.370	1.282	-10.805	-48.052	1.079	0.841
F62_L0	-12.812	-69.646	1.409	-10.908	-52.808	1.186	0.841
F66-1_L0	-11.191	-55.630	1.126	-10.899	-42.170	0.947	0.841
F183_L1	-12.221	-61.971	1.254	-10.834	-46.859	1.052	0.839
F8_L0	-11.247	-58.571	1.185	-10.016	-44.213	0.993	0.838
F127_L0	-11.517	-67.793	1.372	-10.622	-51.113	1.148	0.837
F182_L0	-11.737	-56.743	1.148	-10.745	-42.672	0.958	0.834
F86_L2	-12.148	-58.543	1.185	-9.631	-43.966	0.987	0.833
F38_L0	-11.549	-72.452	1.466	-10.257	-54.039	1.213	0.828
F178_L1	-12.697	-69.086	1.398	-12.398	-51.309	1.152	0.824
F183_L2	-12.454	-74.173	1.501	-11.361	-55.048	1.236	0.823
F185_L3	-12.676	-61.324	1.241	-10.820	-45.452	1.021	0.822
F53_L1	-11.664	-65.130	1.318	-10.653	-48.252	1.083	0.822
F20_L0	-9.805	-85.077	1.722	-10.321	-62.929	1.413	0.821
F50_L2	-12.041	-70.861	1.434	-10.574	-52.367	1.176	0.820
F133_L1	-11.656	-78.755	1.594	-11.684	-58.099	1.305	0.819
F136_L0	-12.060	-71.719	1.451	-11.673	-52.847	1.187	0.818
F178_L3	-12.937	-64.323	1.302	-10.514	-47.386	1.064	0.817
F66_L0	-11.202	-60.761	1.230	-10.684	-44.722	1.004	0.817
F33_L1	-11.519	-59.050	1.195	-12.054	-43.434	0.975	0.816
F80_L0	-10.031	-85.336	1.727	-10.193	-62.734	1.409	0.816
F26_L0	-11.900	-69.023	1.397	-10.931	-50.643	1.137	0.814
F54_L1	-11.855	-82.488	1.669	-10.737	-60.234	1.353	0.810
F94_L0	-11.269	-70.105	1.419	-11.521	-51.072	1.147	0.808
F23_L2	-13.522	-66.516	1.346	-12.618	-48.362	1.086	0.807
F35_L1	-11.570	-67.883	1.374	-12.264	-49.326	1.108	0.806
F193_L2	-12.308	-49.544	1.003	-12.522	-35.957	0.807	0.805
F179_L0	-12.116	-62.340	1.262	-10.301	-45.159	1.014	0.804
F88_L2	-11.345	-67.235	1.361	-10.818	-48.654	1.093	0.803
F188_L2	-11.982	-53.856	1.090	-9.839	-38.912	0.874	0.802
F20_L1	-11.390	-90.979	1.841	-10.280	-65.681	1.475	0.801
F85-1_L0	-12.933	-74.926	1.516	-11.323	-54.082	1.214	0.801
F36_L3	-13.147	-75.254	1.523	-11.696	-54.305	1.219	0.801
F62-1_L1	-11.019	-65.078	1.317	-10.716	-46.914	1.053	0.800
F43_L0	-11.645	-56.140	1.136	-10.160	-40.468	0.909	0.800
F134_L2	-12.442	-85.001	1.720	-11.355	-61.180	1.374	0.799
F106_L0	-11.976	-81.665	1.653	-10.299	-58.736	1.319	0.798
F4_L3	-11.856	-60.669	1.228	-10.631	-43.595	0.979	0.797
F131_L2	-12.968	-81.459	1.649	-10.814	-58.457	1.313	0.796
F124_L0	-11.968	-65.772	1.331	-10.288	-46.986	1.055	0.793
F71_L3	-13.052	-87.995	1.781	-11.570	-62.823	1.411	0.792
F64_L1	-10.929	-68.221	1.381	-9.919	-48.642	1.092	0.791
F89_L0	-11.511	-73.515	1.488	-10.565	-52.400	1.177	0.791
F33_L4	-11.015	-74.790	1.514	-10.924	-52.985	1.190	0.786

F59_L1	-12.776	-77.390	1.566	-11.452	-54.821	1.231	0.786
F194_L0	-11.360	-44.142	0.893	-10.904	-31.237	0.701	0.785
F34_L0	-12.094	-65.987	1.335	-11.571	-46.692	1.048	0.785
F132_L1	-12.039	-74.660	1.511	-11.481	-52.810	1.186	0.785
F43_L3	-13.494	-76.741	1.553	-11.597	-54.262	1.218	0.785
F8_L1	-11.807	-54.069	1.094	-10.878	-38.203	0.858	0.784
F16_L0	-11.304	-66.505	1.346	-10.381	-46.696	1.049	0.779
F110_L0	-12.837	-71.850	1.454	-11.997	-50.442	1.133	0.779
F46_L0	-11.935	-59.742	1.209	-10.831	-41.934	0.942	0.779
F181_L0	-11.728	-61.871	1.252	-10.224	-43.396	0.974	0.778
F23_L0	-11.974	-61.130	1.237	-10.663	-42.856	0.962	0.778
F24_L0	-11.338	-75.214	1.522	-11.980	-52.711	1.184	0.778
F192_L2	-12.220	-67.948	1.375	-11.311	-47.616	1.069	0.778
F62_L1	-13.246	-77.146	1.561	-10.672	-54.018	1.213	0.777
F193_L0	-11.290	-62.434	1.264	-9.943	-43.694	0.981	0.777
F74_L1	-11.687	-81.613	1.652	-10.098	-56.931	1.278	0.774
F74_L3	-12.747	-100.95	2.043	-12.303	-70.366	1.580	0.773
F3_L1	-12.190	-63.436	1.284	-11.495	-44.165	0.992	0.772
F194_L2	-12.242	-59.384	1.202	-11.460	-41.343	0.928	0.772
F95_L1	-12.539	-77.817	1.575	-11.512	-54.149	1.216	0.772
F43_L1 (Enalaprilat)	**-12.778**	**-66.597**	**1.348**	**-11.132**	**-46.326**	**1.040**	0.772
F164_L0	-11.973	-81.025	1.640	-10.869	-56.289	1.264	0.771
F59_L0	-11.568	-64.950	1.314	-9.875	-45.119	1.013	0.771
F192_L0	-11.468	-56.715	1.148	-10.267	-39.355	0.884	0.770
F114_L0	-12.650	-50.848	1.029	-10.344	-35.219	0.791	0.769
F140_L0	-12.522	-71.702	1.451	-10.781	-49.572	1.113	0.767
F81_L0	-11.371	-78.490	1.588	-10.448	-54.242	1.218	0.767
F4_L1	-11.518	-61.816	1.251	-9.589	-42.630	0.957	0.765
F186_L2	-12.733	-68.773	1.392	-11.669	-47.394	1.064	0.765
F90_L0	-11.391	-72.061	1.458	-10.081	-49.642	1.115	0.764
F48_L1	-11.402	-89.356	1.808	-11.396	-61.548	1.382	0.764
F175_L0	-11.450	-87.636	1.774	-10.775	-60.315	1.354	0.764
F120_L0	-11.634	-65.699	1.330	-10.271	-45.211	1.015	0.764
F112_L0	-10.836	-79.414	1.607	-10.950	-54.558	1.225	0.762
F63_L3	-11.591	-80.048	1.620	-10.666	-54.908	1.233	0.761
F184_L2	-12.398	-71.855	1.454	-10.520	-49.257	1.106	0.761
F55_L0	-11.871	-68.763	1.392	-10.179	-47.110	1.058	0.760
F75_L4	-12.068	-88.471	1.790	-11.083	-60.567	1.360	0.760
F82_L0	-11.276	-77.308	1.565	-10.418	-52.908	1.188	0.759
F130_L2	-12.731	-84.288	1.706	-10.759	-57.673	1.295	0.759
F25_L1	-12.149	-74.149	1.501	-10.920	-50.649	1.137	0.758
F44_L1	-13.141	-60.745	1.229	-10.561	-41.439	0.930	0.757
F53_L0	-11.817	-71.309	1.443	-10.732	-48.567	1.091	0.756
F73_L3	-12.431	-88.193	1.785	-10.843	-60.058	1.349	0.756

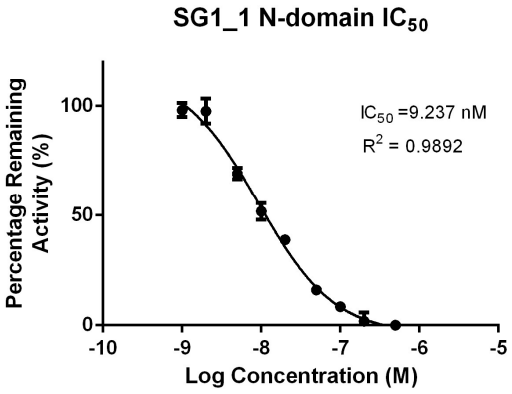
F42_L0	-9.788	-67.829	1.373	-9.878	-46.120	1.036	0.754
F132_L2	-12.377	-83.852	1.697	-10.731	-56.942	1.279	0.753
F185_L2	-13.517	-74.007	1.498	-11.578	-50.227	1.128	0.753
F11_L0	-11.479	-63.886	1.293	-10.128	-43.337	0.973	0.753
F35_L2	-12.387	-83.684	1.694	-12.361	-56.701	1.273	0.752
F56_L1	-12.013	-78.283	1.584	-10.270	-52.944	1.189	0.750
F195_L2	-10.543	-53.987	1.093	-11.643	-36.473	0.819	0.750
F48_L0	-11.354	-71.424	1.445	-10.198	-48.176	1.082	0.748
F9_L1	-11.513	-77.475	1.568	-10.848	-52.229	1.173	0.748
F70_L2	-11.842	-73.493	1.487	-10.547	-49.517	1.112	0.748
F169_L0	-10.514	-82.034	1.660	-10.863	-55.259	1.241	0.747
F72_L0	-11.373	-81.270	1.645	-10.901	-54.618	1.226	0.746
F47_L2	-12.210	-79.614	1.611	-10.356	-53.499	1.201	0.746
F145_L0	-12.042	-81.471	1.649	-10.748	-54.721	1.229	0.745
F56_L2	-11.856	-83.647	1.693	-10.549	-56.109	1.260	0.744
F24_L1	-12.251	-89.772	1.817	-11.264	-60.210	1.352	0.744
F46_L2	-13.280	-61.534	1.245	-11.217	-41.227	0.926	0.743
F154_L0	-12.647	-109.47	2.215	-12.487	-73.343	1.647	0.743
F131_L1	-12.363	-70.908	1.435	-10.874	-47.480	1.066	0.743
F69_L0	-11.119	-65.952	1.335	-9.904	-44.161	0.992	0.743
F40_L0	-11.373	-74.987	1.518	-10.347	-50.043	1.124	0.740
F182_L2	-12.739	-80.431	1.628	-10.400	-53.577	1.203	0.739
F117_L0	-11.195	-58.191	1.178	-10.198	-38.700	0.869	0.738
F92_L0	-11.081	-75.746	1.533	-10.930	-50.212	1.127	0.736
F191_L0	-11.581	-57.282	1.159	-10.683	-37.936	0.852	0.735
F113_L1	-11.634	-54.621	1.105	-10.542	-36.169	0.812	0.735
F80_L1	-11.102	-72.570	1.469	-10.878	-47.999	1.078	0.734
F4_L0	-10.647	-50.168	1.015	-10.322	-33.159	0.745	0.733
F50_L0	-11.387	-72.673	1.471	-10.836	-48.024	1.078	0.733
F179_L2	-13.509	-82.820	1.676	-11.553	-54.614	1.226	0.732
F71_L0	-11.523	-74.978	1.517	-9.962	-49.365	1.108	0.731
F79_L0	-9.613	-69.320	1.403	-10.072	-45.632	1.025	0.730
F185_L0	-11.325	-66.029	1.336	-10.686	-43.440	0.975	0.730
F38_L2	-13.119	-86.453	1.750	-12.451	-56.863	1.277	0.730
F5_L1	-11.254	-46.068	0.932	-11.246	-30.281	0.680	0.729
F21_L2	-13.355	-78.439	1.587	-11.550	-51.517	1.157	0.729
F70_L3	-12.573	-89.131	1.804	-10.698	-58.498	1.314	0.728
F138_L0	-10.792	-86.542	1.751	-11.589	-56.687	1.273	0.727
F132_L0	-11.631	-74.575	1.509	-10.755	-48.762	1.095	0.725
F54_L0	-11.291	-86.771	1.756	-10.147	-56.685	1.273	0.725
F27_L0	-11.749	-74.357	1.505	-10.427	-48.567	1.091	0.725
F71_L2	-11.446	-82.439	1.668	-10.449	-53.737	1.207	0.723
F37_L1	-13.197	-80.098	1.621	-11.119	-52.199	1.172	0.723
F75_L0	-9.657	-82.973	1.679	-10.197	-54.028	1.213	0.722

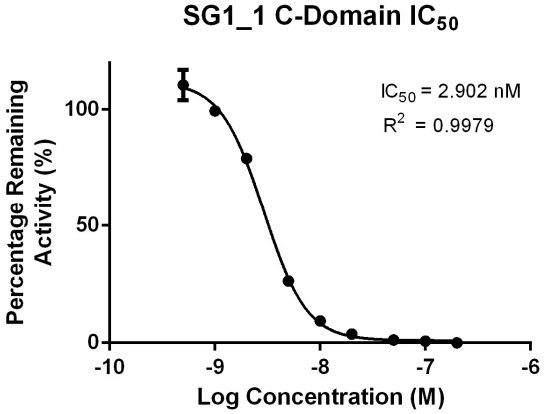
F43_L2	-11.674	-68.262	1.381	-11.410	-44.443	0.998	0.722
F25_L0	-11.896	-82.278	1.665	-10.958	-53.564	1.203	0.722
F33_L0	-11.457	-62.310	1.261	-10.302	-40.564	0.911	0.722
F71_L1	-11.588	-68.101	1.378	-9.875	-44.278	0.994	0.721
F65-1_L0	-11.381	-61.693	1.249	-10.090	-40.040	0.899	0.720
F41_L0	-9.648	-67.060	1.357	-9.925	-43.491	0.977	0.720
F193_L1	-12.052	-60.884	1.232	-10.837	-39.479	0.886	0.719
F180_L3	-12.609	-75.288	1.524	-10.658	-48.630	1.092	0.717
F72_L1	-11.298	-76.697	1.552	-10.032	-49.453	1.110	0.715
F96_L0	-12.465	-76.540	1.549	-10.851	-49.268	1.106	0.714
F51_L1	-11.341	-74.564	1.509	-10.450	-47.950	1.077	0.714
F74_L0	-11.396	-64.275	1.301	-10.104	-41.195	0.925	0.711
F159_L0	-12.762	-71.059	1.438	-10.670	-45.526	1.022	0.711
F51_L0	-7.788	-60.468	1.224	-10.550	-38.715	0.869	0.710
F77_L0	-11.164	-61.115	1.237	-11.812	-39.128	0.879	0.710
F91_L0	-11.216	-68.043	1.377	-9.855	-43.514	0.977	0.710
F49_L2	-11.607	-77.828	1.575	-10.500	-49.761	1.117	0.709
F97_L0	-13.308	-86.298	1.746	-10.023	-55.175	1.239	0.709
F176_L0	-11.003	-85.102	1.722	-11.759	-54.399	1.221	0.709
F190_L0	-11.718	-55.487	1.123	-9.775	-35.293	0.792	0.706
F33_L3	-12.482	-66.997	1.356	-12.354	-42.515	0.955	0.704
F66_L1	-11.348	-84.950	1.719	-11.772	-53.892	1.210	0.704
F147_L0	-12.862	-90.802	1.838	-11.667	-57.564	1.293	0.703
F34_L4	-12.810	-73.484	1.487	-12.079	-46.568	1.046	0.703
F3_L0	-12.292	-48.903	0.990	-10.334	-30.957	0.695	0.702
F2_L0	-12.197	-57.404	1.162	-13.098	-36.332	0.816	0.702
F179_L3	-12.785	-83.859	1.697	-11.530	-52.949	1.189	0.701
F81_L1	-11.459	-86.042	1.741	-10.693	-54.318	1.220	0.700
F137_L0	-11.125	-96.786	1.959	-11.082	-60.855	1.366	0.698
F184_L0	-11.519	-66.903	1.354	-10.396	-41.980	0.943	0.696
F1_L4	-11.777	-62.080	1.256	-10.750	-38.946	0.874	0.696
F88-1_L0	-11.683	-58.868	1.191	-10.546	-36.806	0.826	0.694
F57_L2	-11.642	-78.733	1.593	-10.601	-49.099	1.102	0.692
F26_L1	-12.245	-55.746	1.128	-10.899	-34.761	0.781	0.692
F197_L0	-11.345	-66.489	1.346	-10.507	-41.373	0.929	0.690
F95_L2	-12.433	-83.970	1.699	-11.794	-52.237	1.173	0.690
F29_L1	-11.756	-77.792	1.574	-11.149	-48.303	1.085	0.689
F105_L0	-11.654	-89.163	1.804	-10.477	-55.076	1.237	0.685
F173_L0	-12.432	-89.294	1.807	-10.629	-55.061	1.236	0.684
F95_L0	-11.353	-84.572	1.712	-10.477	-52.081	1.169	0.683
F69_L3	-11.649	-65.039	1.316	-9.999	-39.921	0.896	0.681
F7_L1	-11.703	-65.410	1.324	-11.418	-40.138	0.901	0.681
F9_L0	-11.300	-62.136	1.257	-10.151	-38.122	0.856	0.681
F178_L2	-13.537	-70.408	1.425	-11.564	-43.165	0.969	0.680

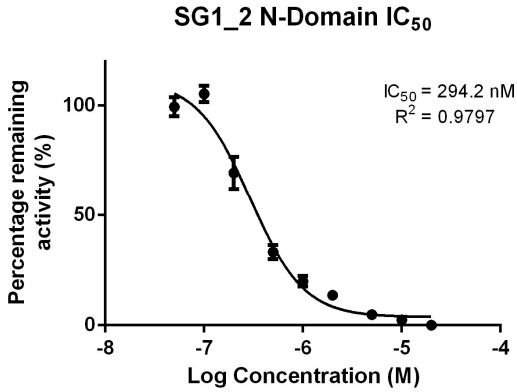
F61_L2	-13.056	-71.380	1.445	-10.412	-43.732	0.982	0.680
F133_L2	-12.421	-79.276	1.604	-10.754	-48.125	1.081	0.674
F113_L0	-11.527	-69.859	1.414	-10.313	-42.394	0.952	0.673
F197_L1	-12.040	-54.996	1.113	-10.998	-33.365	0.749	0.673
F165_L0	-11.714	-70.238	1.421	-10.822	-42.496	0.954	0.671
F14_L0	-11.335	-70.026	1.417	-10.183	-42.125	0.946	0.667
F39_L0	-11.404	-71.950	1.456	-10.016	-43.075	0.967	0.664
F6_L0	-11.369	-51.692	1.046	-10.956	-30.916	0.694	0.664
F156_L0	-12.746	-75.621	1.530	-11.563	-45.153	1.014	0.662
F186_L0	-11.373	-68.059	1.377	-12.199	-40.454	0.908	0.660
F135_L0	-11.384	-75.524	1.528	-10.691	-44.498	0.999	0.654
F177_L3	-11.750	-85.933	1.739	-12.023	-50.380	1.131	0.650
F73_L1	-11.698	-85.848	1.737	-9.886	-50.195	1.127	0.649
F29_L0	-11.696	-69.100	1.398	-11.162	-40.191	0.902	0.645
F187_L3	-12.193	-54.061	1.094	-10.735	-31.417	0.705	0.645
F182_L3	-11.897	-66.970	1.355	-11.386	-38.812	0.871	0.643
F45_L1	-12.197	-66.992	1.356	-11.719	-38.755	0.870	0.642
F184_L3	-12.790	-62.135	1.257	-10.800	-35.797	0.804	0.639
F103_L0	-11.443	-85.989	1.740	-11.090	-49.461	1.111	0.638
F135_L1	-11.631	-76.521	1.549	-10.763	-43.925	0.986	0.637
F49_L0	-11.204	-73.373	1.485	-10.213	-41.873	0.940	0.633
F35_L4	-12.355	-85.025	1.721	-13.304	-48.056	1.079	0.627
F41_L1	-11.499	-87.129	1.763	-10.454	-49.211	1.105	0.627
F52_L1	-12.093	-80.820	1.636	-11.216	-45.036	1.011	0.618
F47_L0	-11.392	-67.876	1.374	-10.532	-37.816	0.849	0.618
F48_L2	-11.816	-88.539	1.792	-10.486	-49.029	1.101	0.614
F61_L1	-12.605	-65.186	1.319	-10.817	-35.958	0.807	0.612
F70_L1	-11.401	-85.499	1.730	-10.374	-47.136	1.058	0.612
F82_L1	-11.802	-89.059	1.802	-10.550	-49.060	1.102	0.611
F27_L1	-12.755	-61.783	1.250	-10.676	-33.989	0.763	0.610
F74_L2	-12.513	-92.595	1.874	-10.735	-50.862	1.142	0.609
F36_L0	-12.173	-85.337	1.727	-11.267	-46.746	1.050	0.608
F1_L0	-11.311	-52.840	1.069	-9.791	-28.927	0.650	0.607
F39_L1	-11.628	-70.030	1.417	-10.945	-38.204	0.858	0.605
F168_L0	-12.490	-103.56	2.096	-12.429	-56.322	1.265	0.603
F133_L0	-11.074	-84.576	1.712	-11.889	-45.990	1.033	0.603
F64_L2	-11.730	-79.199	1.603	-11.145	-42.966	0.965	0.602
F131_L0	-11.450	-87.652	1.774	-10.613	-47.538	1.067	0.602
F144_L0	-10.678	-104.25	2.110	-11.055	-56.146	1.261	0.598
F6_L1	-11.623	-65.134	1.318	-10.249	-35.075	0.788	0.598
F20_L2	-11.635	-95.018	1.923	-10.695	-51.151	1.149	0.597
F122_L0	-9.678	-71.580	1.449	-9.974	-38.496	0.864	0.597
F1_L3	-11.562	-67.901	1.374	-10.495	-36.415	0.818	0.595
F64_L3	-11.855	-78.093	1.580	-12.192	-41.839	0.939	0.594

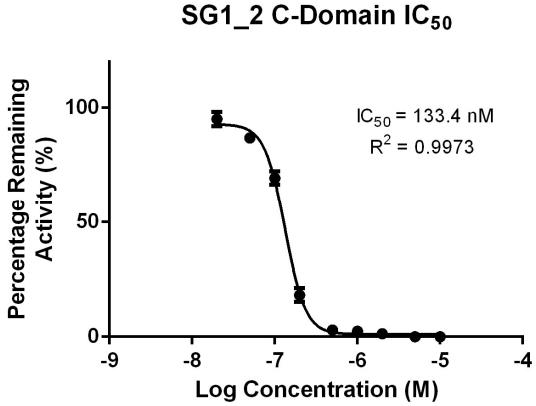
F57_L0	-11.449	-77.057	1.559	-10.832	-41.163	0.924	0.593
F68_L0	-11.498	-52.088	1.054	-12.067	-27.730	0.623	0.591
F12_L0	-11.324	-62.542	1.266	-10.061	-33.241	0.746	0.590
F44_L2	-13.684	-69.527	1.407	-11.154	-36.787	0.826	0.587
F134_L0	-12.797	-78.840	1.596	-11.060	-41.532	0.933	0.584
F158_L0	-12.658	-87.836	1.778	-10.772	-46.032	1.034	0.581
F129_L0	-10.403	-61.485	1.244	-10.565	-32.085	0.720	0.579
F121_L0	-11.851	-88.787	1.797	-11.135	-45.985	1.033	0.575
F73_L0	-11.767	-88.030	1.782	-10.236	-45.211	1.015	0.570
F1_L2	-11.477	-52.024	1.053	-10.077	-26.658	0.599	0.569
F60_L2	-11.682	-88.334	1.788	-11.558	-44.899	1.008	0.564
F63_L4	-11.781	-71.910	1.455	-11.972	-36.221	0.813	0.559
F5_L0	-11.166	-61.519	1.245	-9.843	-30.828	0.692	0.556
F52_L0	-11.892	-69.193	1.400	-9.916	-34.400	0.772	0.552
F166_L0	-12.715	-87.956	1.780	-12.013	-43.064	0.967	0.543
F10_L0	-11.143	-83.032	1.680	-10.167	-40.117	0.901	0.536
F67_L0	-11.843	-63.411	1.283	-12.143	-29.844	0.670	0.522
F78_L0	-11.735	-83.300	1.686	-10.635	-39.094	0.878	0.521
F36_L1	-12.641	-83.442	1.689	-11.404	-38.970	0.875	0.518
F187_L2	-11.891	-70.177	1.420	-11.146	-32.641	0.733	0.516
F11_L1	-11.942	-65.624	1.328	-10.332	-30.483	0.684	0.515
F87_L1	-11.790	-64.766	1.311	-10.827	-29.587	0.664	0.507
F7_L0	-11.480	-65.231	1.320	-10.344	-27.839	0.625	0.474
F63_L2	-11.596	-74.852	1.515	-10.276	-31.617	0.710	0.469

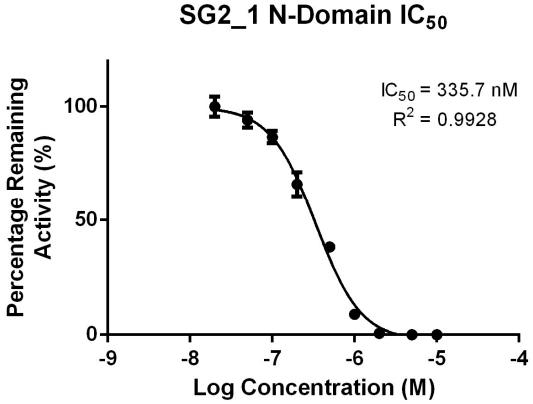
Appendix 6.1 – Diprolyl Series Dose Response Curves

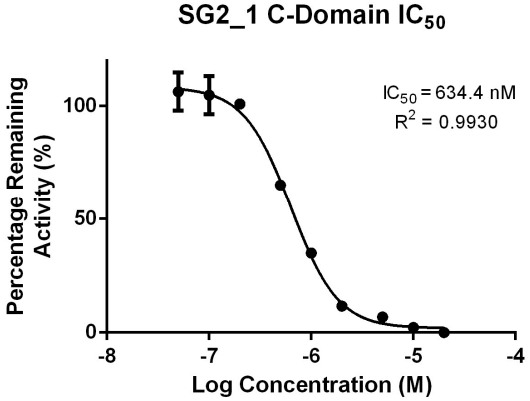


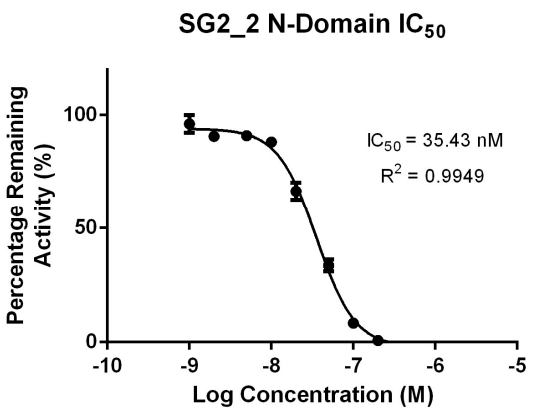


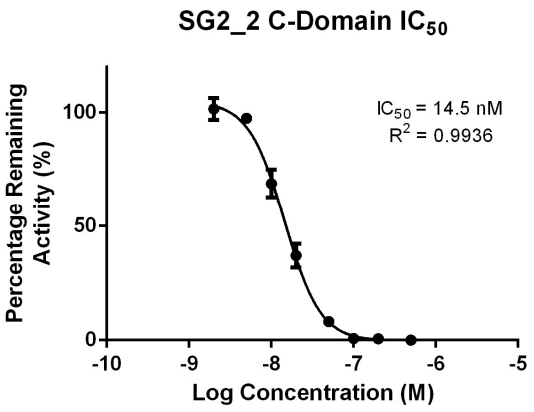


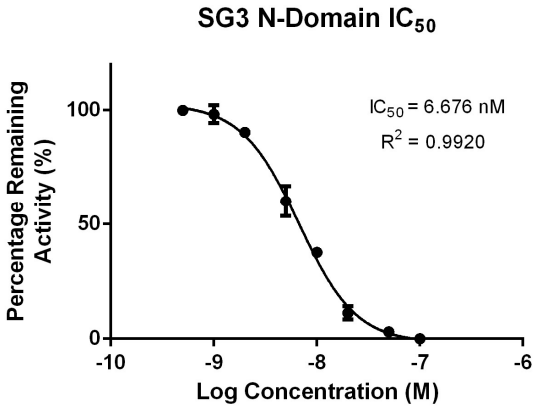


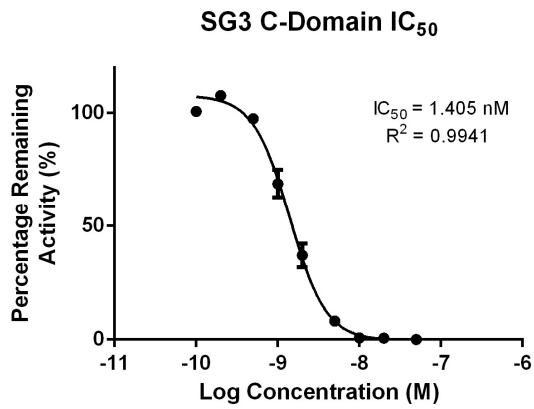


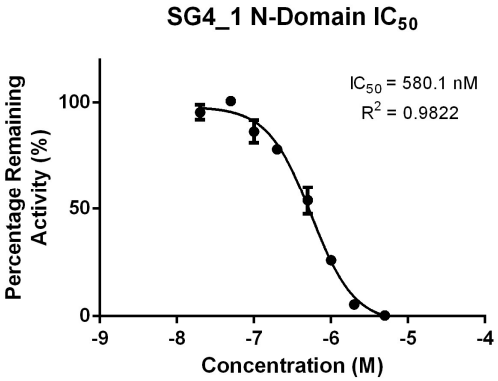


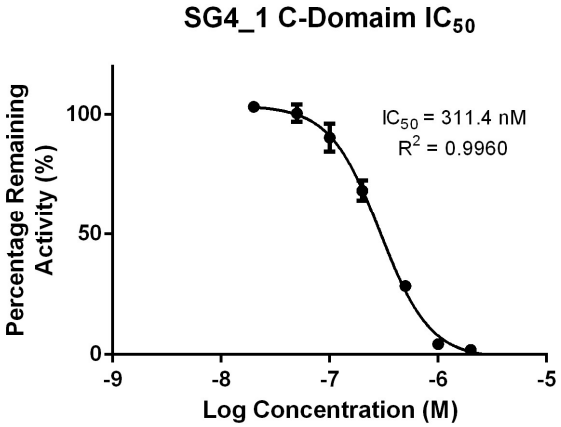

SG2_2 N-Domain IC50
Percentage Remaining Activity (%)
IC50 = 35.43 nM
R2 = 0.9949
Log Concentration (M)

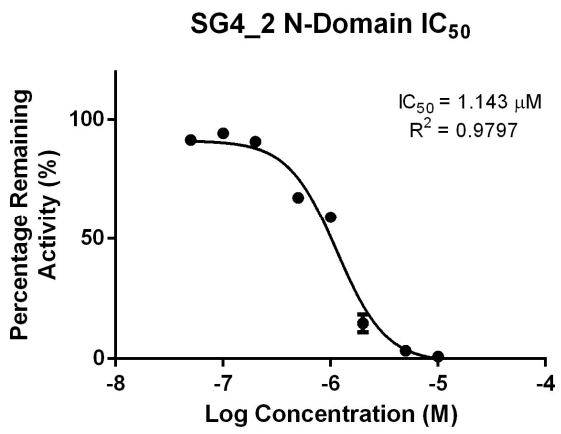

SG2_2 C-Domain IC50
Percentage Remaining Activity (%)
IC50 = 14.5 nM
R2 = 0.9936
Log Concentration (M)

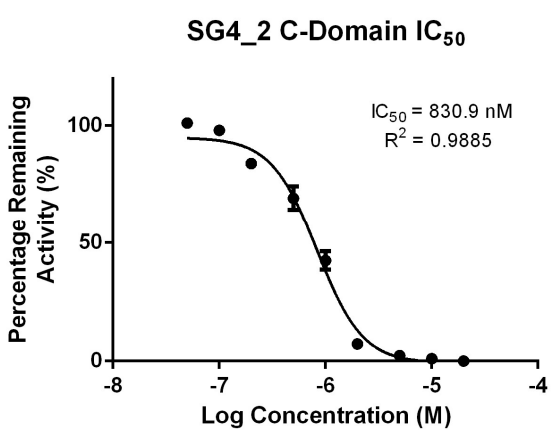

SG3 N-Domain IC50
Percentage Remaining Activity (%)
IC50 = 6.676 nM
R2 = 0.9920
Log Concentration (M)

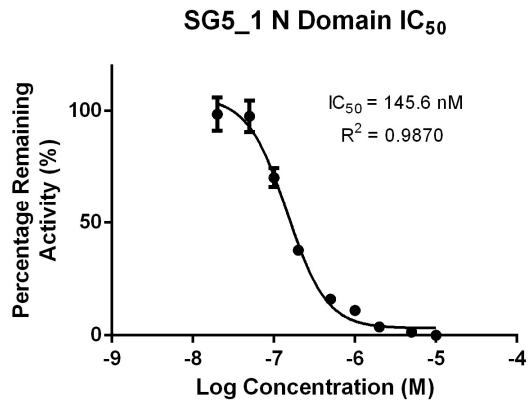

SG3 C-Domain IC50
Percentage Remaining Activity (%)
IC50 = 1.405 nM
R2 = 0.9941
Log Concentration (M)

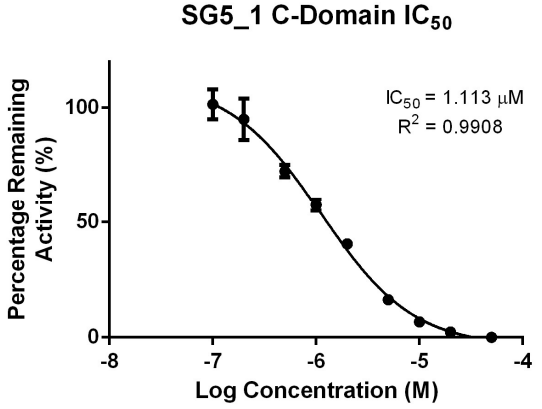

SG4_1 N-Domain IC50
Percentage Remaining Activity (%)
IC50 = 580.1 nM
R2 = 0.9822
Concentration (M)

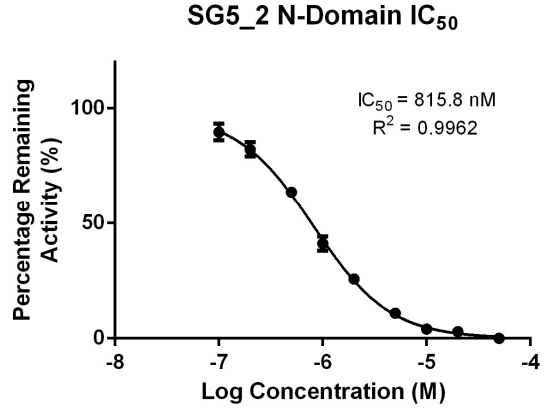

SG4_1 C-Domaim IC50
Percentage Remaining Activity (%)
IC50 = 311.4 nM
R2 = 0.9960
Log Concentration (M)

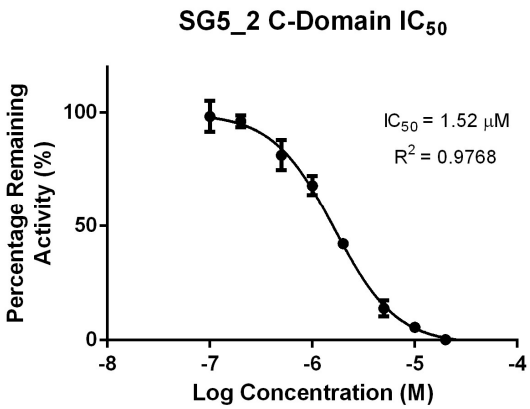

SG4_2 N-Domain IC50
IC50 = 1.143 μM
R2 = 0.9797
Percentage Remaining Activity (%)
Log Concentration (M)

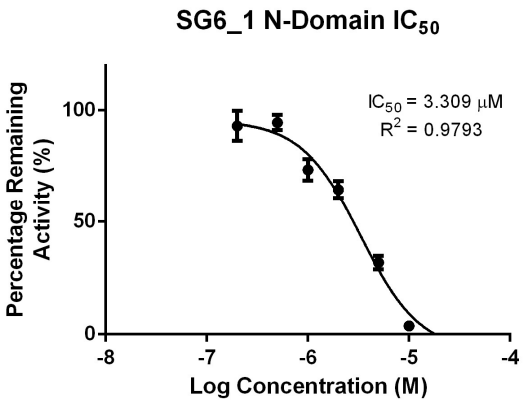

SG4_2 C-Domain IC50
IC50 = 830.9 nM
R2 = 0.9885
Percentage Remaining Activity (%)
Log Concentration (M)

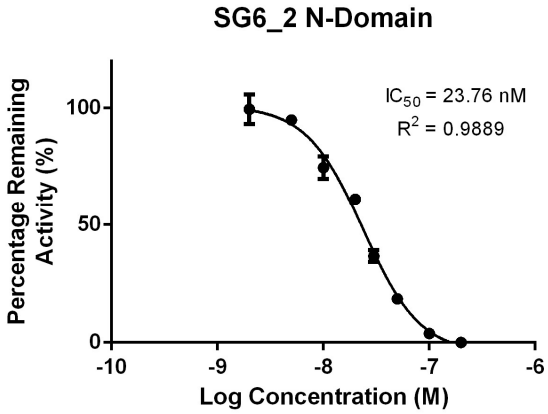

SG5_1 N Domain IC50
IC50 = 145.6 nM
R2 = 0.9870
Percentage Remaining Activity (%)
Log Concentration (M)

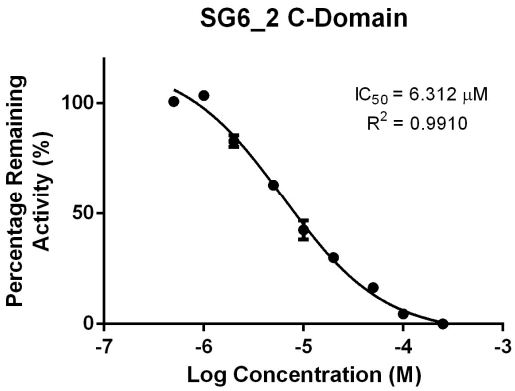

SG5_1 C-Domain IC50
IC50 = 1.113 μM
R2 = 0.9908
Percentage Remaining Activity (%)
Log Concentration (M)

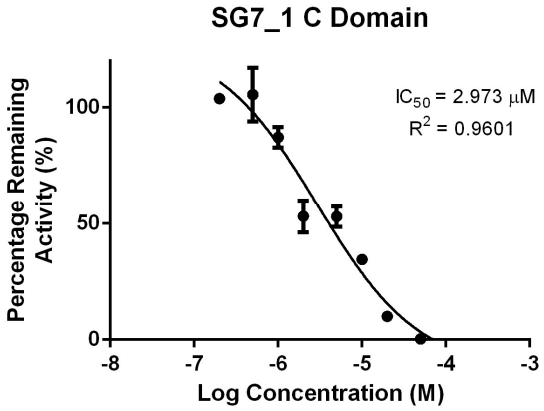

SG5_2 N-Domain IC50
IC50 = 815.8 nM
R2 = 0.9962
Percentage Remaining Activity (%)
Log Concentration (M)

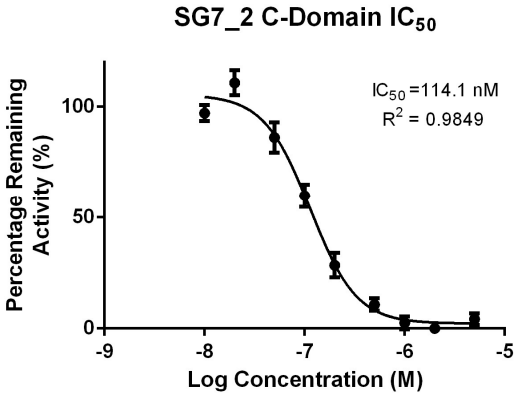

SG5_2 C-Domain IC50
IC50 = 1.52 μM
R2 = 0.9768
Percentage Remaining Activity (%)
Log Concentration (M)

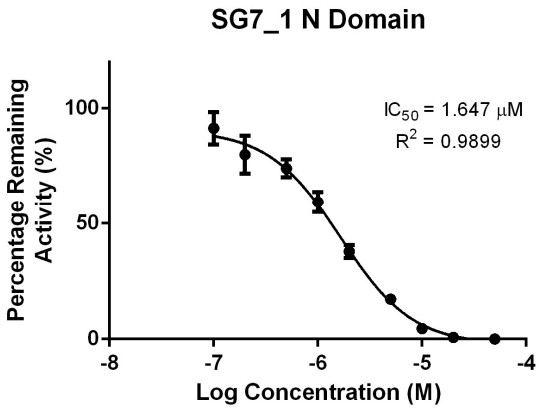

SG6_1 N-Domain IC50
IC50 = 3.309 μM
R2 = 0.9793
Percentage Remaining Activity (%)
Log Concentration (M)

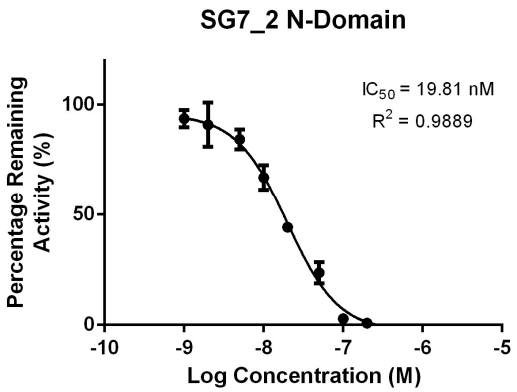

SG6_2 N-Domain
IC50 = 23.76 nM
R2 = 0.9889
Percentage Remaining Activity (%)
Log Concentration (M)


SG6_2 C-Domain
IC50 = 6.312 μM
R2 = 0.9910
Percentage Remaining Activity (%)
Log Concentration (M)


SG7_1 C Domain
IC50 = 2.973 μM
R2 = 0.9601
Percentage Remaining Activity (%)
Log Concentration (M)


SG7_2 C-Domain IC50
IC50 =114.1 nM
R2 = 0.9849
Percentage Remaining Activity (%)
Log Concentration (M)


SG7_1 N Domain
Percentage Remaining Activity (%)
100
50
0
-8
-7
-6
-5
-4
Log Concentration (M)
IC50 = 1.647 μM
R2 = 0.9899


SG7_2 N-Domain
Percentage Remaining Activity (%)
100
50
0
-10
-9
-8
-7
-6
-5
Log Concentration (M)
IC50 = 19.81 nM
R2 = 0.9889

References

1. Skeggs, L. T.; Kahn, J. R.; Shumway, N. P., The preparation and function of the hypertensin-converting enzyme. *J. Exp. Med.* **1956,** *103* (3), 295-299.
2. Soubrier, F.; Alhenc-Gelas, F.; Hubert, C.; Allegrini, J.; John, M.; Tregear, G.; Corvol, P., Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proc. Nat. Acad. Sci.* **1988,** *85* (24), 9386-9390.
3. (a) Wei, L.; Alhenc-Gelas, F.; Corvol, P.; Clauser, E., The two homologous domains of human angiotensin I-converting enzyme are both catalytically active. *J. Biol. Chem.* **1991,** *266* (14), 9002-9008; (b) Wei, L.; Clauser, E.; Alhenc-Gelas, F.; Corvol, P., The two homologous domains of human angiotensin I-converting enzyme interact differently with competitive inhibitors. *J. Biol. Chem.* **1992,** *267* (19), 13398-13405.
4. Fuchs, S.; Xiao, H. D.; Hubert, C.; Michaud, A.; Campbell, D. J.; Adams, J. W.; Capecchi, M. R.; Corvol, P.; Bernstein, K. E., Angiotensin-converting enzyme C-terminal catalytic domain is the main site of angiotensin I cleavage in vivo. *Hypertens.* **2008,** *51* (2), 267-274.
5. Fuchs, S.; Xiao, H. D.; Cole, J. M.; Adams, J. W.; Frenzel, K.; Michaud, A.; Zhao, H.; Keshelava, G.; Capecchi, M. R.; Corvol, P., Role of the N-terminal catalytic domain of angiotensin-converting enzyme investigated by targeted inactivation in mice. *J. Biol. Chem.* **2004,** *279* (16), 15946-15953.
6. Skeggs, L. T.; Lentz, K. E.; Kahn, J. R.; Shumway, N. P.; Woods, K. R., The amino acid sequence of hypertensin II. *J. Exp. Med.* **1956,** *104* (2), 193-197.
7. Skeggs, L. T.; Marsh, W. H.; Kahn, J. R.; Shumway, N. P., The existence of two forms of hypertensin. *J. Exp. Med.* **1954,** *99* (3), 275-282.
8. Helmer, O. In *A factor in plasma that enhances contraction produced by angiotonin on rabbit aortic strips*, Fed. Proc, 1955; p 728.
9. Goldblatt, H., *The renal origin of hypertension*. Am Physiological Soc: 1948.
10. Romero, C. A.; Orias, M.; Weir, M. R., Novel RAAS agonists and antagonists: clinical applications and controversies. *Nat. Rev. Endocrin.* **2015**.
11. e Silva, M. R.; Beraldo, W. T.; Rosenfeld, G., Bradykinin, a hypotensive and smooth muscle stimulating factor released from plasma globulin by snake venoms and by trypsin. *Am. J. Physiol.* **1949,** *156* (2), 261-273.
12. Ferreira, S.; e Silva, M. R., Potentiation of bradykinin and eledoisin by BPF (bradykinin potentiating factor) fromBothrops jararaca venom. *Experientia* **1965,** *21* (6), 347-349.
13. Bakhle, Y., Conversion of angiotensin I to angiotensin II by cell-free extracts of dog lung. **1968**.
14. Ondetti, M. A.; Williams, N. J.; Sabo, E.; Pluscec, J.; Weaver, E. R.; Kocy, O., Angiotensin-converting enzyme inhibitors from the venom of Bothrops jararaca. Isolation, elucidation of structure, and synthesis. *Biochem* **1971,** *10* (22), 4033-4039.
15. Collier, J.; Robinson, B.; Vane, J., Reduction of pressor effects of angiotensin I in man by synthetic nonapeptide (BPP 9a or SQ 20,881) which inhibits converting enzyme. *Lanc.* **1973,** *301* (7794), 72-74.
16. DORER, F. E.; KAHN, J. R.; LENTZ, K. E.; LEVINE, M.; SKEGGS, L. T., Hydrolysis of bradykinin by angiotensin-converting enzyme. *Circ. Res.* **1974,** *34* (6), 824-827.
17. Cushman, D. W.; Cheung, H.; Sabo, E.; Ondetti, M., Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids. *Biochem* **1977,** *16* (25), 5484-5491.
18. Schechter, I.; Berger, A., Protease subsite nomenclature. *Biochem. Biophys. Res. Commun* **1967,** *27*, 157-162.
19. Byers, L. D.; Wolfenden, R., Binding of the by-product analog benzylsuccinic acid by carboxypeptidase A. *Biochem* **1973,** *12* (11), 2070-2078.
20. Patchett, A.; Harris, E.; Tristram, E.; Wyvratt, M.; Wu, M.; Taub, D.; Peterson, E.; Ikeler, T.; Ten Broeke, J.; Payne, L., A new class of angiotensin-converting enzyme inhibitors. **1980**.

21. Subissi, A.; Evangelista, S.; Giachetti, A., Preclinical profile of zofenopril: an angiotensin converting enzyme inhibitor with peculiar cardioprotective properties. *Cardiovasc. Drug Rev.* **1999,** *17* (2), 115-133.
22. Takeyama, K.; Minato, H.; Fukuya, F.; Kawahara, S.; Hosoki, K.; Kadokawa, T., Antihypertensive activity of alacepril, an orally active angiotensin converting enzyme inhibitor, in renal hypertensive rats and dogs. *Arzneimittel-Forschung* **1985,** *35* (10), 1502.
23. Yamauchi, H.; Nishimura, K.; Nakata, K.; Suda, H.; Iso, T.; Shimizu, M.; Hiramatsu, Y., General pharmacological properties of the potent angiotensin converting enzyme inhibitor rentiapril. *Arzneimittel-Forschung* **1987,** *37* (2), 157-164.
24. Frank, G. J.; Knapp, L. E.; Olson, S. C.; Phelps, M. C.; Quade, M. M.; Rieger, M. M.; Sedman, A. J., Overview of quinapril, a new ACE inhibitor. *J. Cardiovasc. Pharmacol.* **1990,** *15*, S14-S23.
25. Attwood, M. R.; Hassall, C. H.; Kröhn, A.; Lawton, G.; Redshaw, S., The design and synthesis of the angiotensin converting enzyme inhibitor cilazapril and related bicyclic compounds. *J. Chem. Soc.* **1986**, 1011-1019.
26. Webb, R. L.; Miller, D.; Traina, V.; Gomez, H. J., Benazepril. *Cardiovasc. Drug Rev.* **1990,** *8* (2), 89-104.
27. Smith, E. M.; Swiss, G. F.; Neustadt, B. R.; McNamara, P.; Gold, E. H.; Sybertz, E. J.; Baum, T., Angiotensin converting enzyme inhibitors: spirapril and related compounds. *J. Med. Chem.* **1989,** *32* (7), 1600-1606.
28. Grass, G. M.; Morehead, W. T., Evidence for site-specific absorption of a novel ACE inhibitor. *Pharmaceut Res.* **1989,** *6* (9), 759-765.
29. Brown, N. L.; Badel, M.-Y.; Benzoni, F.; Zanirato, J.; Vincent, J.-C.; Fichelle, J.; Worcel, M., Angiotensin-converting enzyme inhibition, anti-hypertensive activity and hemodynamic profile of trandolapril (RU 44570). *Eur J. Pharmacol.* **1988,** *148* (1), 79-91.
30. Todd, P. A.; Benfield, P., Ramipril. *Drug* **1990,** *39* (1), 110-135.
31. Unger, T.; Moursi, M.; Ganten, D.; Hermann, K.; Lang, R. E., Antihypertensive action of the converting enzyme inhibitor perindopril (S9490-3) in spontaneously hypertensive rats: comparison with enalapril (MK421) and ramipril (Hoe498). *J. Cardiovasc. Pharmacol.* **1986,** *8* (2), 276-285.
32. Shionoiri, H.; Yasuda, G.; Ikeda, A.; Ohta, T.; Miyajima, E.; Kaneko, Y., Pharmacokinetics and depressor effect of delapril in patients with essential hypertension. *Clin. Pharm. Ther.* **1987,** *41* (1), 74-79.
33. Nawano, M.; Anai, M.; Funaki, M.; Kobayashi, H.; Kanda, A.; Fukushima, Y.; Inukai, K.; Ogihara, T.; Sakoda, H.; Onishi, Y., Imidapril, an angiotensin-converting enzyme inhibitor, improves insulin sensitivity by enhancing signal transduction via insulin receptor substrate proteins and improving vascular resistance in the Zucker fatty rat. *Metabol.* **1999,** *48* (10), 1248-1255.
34. Nakashima, M.; Yamamoto, J.; Shihata, M.; Uematsu, T.; Shinjo, H.; Akahori, T.; Shioya, H.; Sugiyama, K.; Kawahara, Y., Pharmacokinetics of temocapril hydrochloride, a novel angiotensin converting enzyme inhibitor, in renal insufficiency. *Eur. J. Clin. Pharmacol.* **1992,** *43* (6), 657-659.
35. Pool, J., Antihypertensive effect of fosinopril, a new angiotensin converting enzyme inhibitor: findings of the Fosinopril Study Group II. *Clin. Ther.* **1989,** *12* (6), 520-533.
36. DeForrest, J. M.; Waldron, T. L.; Harvey, C.; Scalese, R.; Hammerstone, S.; Powell, J. R.; Karanewsky, D., Ceranapril (SQ 29,852), an orally active inhibitor of angiotensin converting enzyme (ACE). *J. Cardiovasc. Pharmacol.* **1990,** *16* (1), 121-127.
37. Das, M.; Soffer, R. L., Pulmonary angiotensin-converting enzyme. Structural and catalytic properties. *J. Biol. Chem.* **1975,** *250* (17), 6762-6768.
38. El-Dorry, H.; Bull, H. G.; Iwata, K.; Thornberry, N. A.; Cordes, E. H.; Soffer, R., Molecular and catalytic properties of rabbit testicular dipeptidyl carboxypeptidase. *J. Biol. Chem.* **1982,** *257* (23), 14128-14133.
39. Jaspard, E.; Wei, L.; Alhenc-Gelas, F., Differences in the properties and enzymatic specificities of the two active sites of angiotensin I-converting enzyme (kininase II). Studies with bradykinin and other natural peptides. *J. Biol. Chem.* **1993,** *268* (13), 9496-9503.

40. Lenfant, M.; Wdzieczak-Bakala, J.; Guittet, E.; Prome, J.-C.; Sotty, D.; Frindel, E., Inhibitor of hematopoietic pluripotent stem cell proliferation: purification and determination of its structure. *Proc. Nat. Acad. Sci.* **1989,** *86* (3), 779-782.
41. Robinson, S.; Lenfant, M.; Wdzieczak-Bakala, J.; Melville, J.; Riches, A., The mechanism of action of the tetrapeptide acetyl-N-Ser-Asp-Lys-Pro (AcSDKP) in the control of haematopoietic stem cell proliferation. *Cell Prolif.* **1992,** *25* (6), 623-632.
42. Rieger, K.; Saez-Servent, N.; Papet, M.; Wdzieczak-Bakala, J.; Morgat, J.; Thierry, J.; Voelter, W.; Lenfant, M., Involvement of human plasma angiotensin I-converting enzyme in the degradation of the haemoregulatory peptide N-acetyl-seryl-aspartyl-lysyl-proline. *Biochem. J* **1993,** *296*, 373-378.
43. Rousseau, A.; Michaud, A.; Chauvet, M.-T.; Lenfant, M.; Corvol, P., The hemoregulatory peptide N-acetyl-Ser-Asp-Lys-Pro is a natural and specific substrate of the N-terminal active site of human angiotensin-converting enzyme. *J. Biol. Chem.* **1995,** *270* (8), 3656-3661.
44. (a) Azizi, M.; Rousseau, A.; Ezan, E.; Guyene, T.-T.; Michelet, S.; Grognet, J.-M.; Lenfant, M.; Corvol, P.; Ménard, J., Acute angiotensin-converting enzyme inhibition increases the plasma level of the natural stem cell regulator N-acetyl-seryl-aspartyl-lysyl-proline. *J. Clin. Inv.* **1996,** *97* (3), 839; (b) Rousseau-Plasse, A.; Wdzieczak-Bakala, J.; Lenfant, M.; Ezan, E.; Genet, R.; Robinson, S.; Briscoe, T.; Melville, J.; Riches, A., Lisinopril, an angiotensin I-converting enzyme inhibitor, prevents entry of murine hematopoietic stem cells into the cell cycle after irradiation in vivo. *Exp. Hemat.* **1998,** *26* (11), 1074-1079.
45. Sun, Y.; Cleutjens, J. P.; Diaz-Arias, A. A.; Weber, K. T., Cardiac angiotensin converting enzyme and myocardial fibrosis in the rat. *Cardiovascular research* **1994,** *28* (9), 1423-1432.
46. Brooks, W. W.; Bing, O. H.; Robinson, K. G.; Slawsky, M. T.; Chaletsky, D. M.; Conrad, C. H., Effect of angiotensin-converting enzyme inhibition on myocardial fibrosis and function in hypertrophied and failing myocardium from the spontaneously hypertensive rat. *Circ.* **1997,** *96* (11), 4002-4010.
47. (a) Peng, H.; Carretero, O. A.; Raij, L.; Yang, F.; Kapke, A.; Rhaleb, N.-E., Antifibrotic effects of N-acetyl-seryl-aspartyl-lysyl-proline on the heart and kidney in aldosterone-salt hypertensive rats. *Hypertens.* **2001,** *37* (2), 794-800; (b) Rhaleb, N.-E.; Peng, H.; Yang, X.-P.; Liu, Y.-H.; Mehta, D.; Ezan, E.; Carretero, O. A., Long-term effect of N-acetyl-seryl-aspartyl-lysyl-proline on left ventricular collagen deposition in rats with 2-kidney, 1-clip hypertension. *Circ.* **2001,** *103* (25), 3136-3141.
48. Peng, H.; Carretero, O. A.; Peterson, E. L.; Rhaleb, N.-E., Ac-SDKP inhibits transforming growth factor-β1-induced differentiation of human cardiac fibroblasts into myofibroblasts. *Am. J. Phys-Heart Circ. Phys.* **2010,** *298* (5), H1357-H1364.
49. Gavras, H.; Gavras, I., Angiotensin converting enzyme inhibitors. Properties and side effects. *Hypertens.* **1988,** *11* (3 Pt 2), II37.
50. Agostoni, A.; Cicardi, M.; Cugno, M.; Zingale, L. C.; Gioffré, D.; Nussberger, J., Angioedema due to angiotensin-converting enzyme inhibitors. *Immunopharmacol.* **1999,** *44* (1), 21-25.
51. Michaud, A.; Williams, T. A.; Chauvet, M.-T.; Corvol, P., Substrate Dependence of Angiotensin I-Converting Enzyme Inhibition: Captopril Displays a Partial Selectivity for Inhibition ofN-Acetyl-Seryl-Aspartyl-Lysyl-Proline Hydrolysis Compared with That of Angiotensin I·. *Mol. Pharmacol.* **1997,** *51* (6), 1070-1076.
52. Deddish, P. A.; Marcic, B.; Jackman, H. L.; Wang, H.-Z.; Skidgel, R. A.; Erdös, E. G., N-Domain–Specific Substrate and C-Domain Inhibitors of Angiotensin-Converting Enzyme Angiotensin-(1–7) and Keto-ACE. *Hypertens.* **1998,** *31* (4), 912-917.
53. Thomas, D. A.; Grant, P. J.; Turner, A. J.; Hooper, N. M., Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism. *Biochem. J.* **2004,** *383* (1), 45-51.
54. Weare, J. A.; Stewart, T. A.; Gafford, J. T.; Erdös, E., Inhibition of human converting enzyme in vitro by a novel tripeptide analog. *Hypertens.* **1981,** *3* (3 Pt 2), I50.
55. (a) Jirácek, J.; Yiotakis, A.; Vincent, B.; Lecoq, A.; Nicolaou, A.; Checler, F.; Dive, V., Development of highly potent and selective phosphinic peptide inhibitors of zinc endopeptidase 24-

15 using combinatorial chemistry. *J. Biol. Chem.* **1995,** *270* (37), 21701-21706; (b) Jiráček, J.; Yiotakis, A.; Vincent, B.; Checler, F.; Dive, V., Development of the first potent and selective inhibitor of the zinc endopeptidase neurolysin using a systematic approach based on combinatorial chemistry of phosphinic peptides. *J. Biol. Chem.* **1996,** *271* (32), 19606-19611.

56. Dive, V.; Cotton, J.; Yiotakis, A.; Michaud, A.; Vassiliou, S.; Jiracek, J.; Vazeux, G.; Chauvet, M.-T.; Cuniasse, P.; Corvol, P., RXP 407, a phosphinic peptide, is a potent inhibitor of angiotensin I converting enzyme able to differentiate between its two active sites. *Proc. Nat. Acad. Sci.* **1999,** *96* (8), 4330-4335.

57. Georgiadis, D.; Beau, F.; Czarny, B.; Cotton, J.; Yiotakis, A.; Dive, V., Roles of the Two Active Sites of Somatic Angiotensin-Converting Enzyme in the Cleavage of Angiotensin I and Bradykinin Insights From Selective Inhibitors. *Circ. Res.* **2003,** *93* (2), 148-154.

58. Georgiadis, D.; Cuniasse, P.; Cotton, J.; Yiotakis, A.; Dive, V., Structural determinants of RXPA380, a potent and highly selective inhibitor of the angiotensin-converting enzyme C-domain. *Biochem* **2004,** *43* (25), 8048-8054.

59. Natesh, R.; Schwager, S. L.; Sturrock, E. D.; Acharya, K. R., Crystal structure of the human angiotensin-converting enzyme–lisinopril complex. *Nature* **2003,** *421* (6922), 551-554.

60. Corradi, H. R.; Schwager, S. L.; Nchinda, A. T.; Sturrock, E. D.; Acharya, K. R., Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a structural basis for domain-specific inhibitor design. *J. Mol. Biol.* **2006,** *357* (3), 964-974.

61. Anthony, C. S.; Corradi, H. R.; Schwager, S. L.; Redelinghuys, P.; Georgiadis, D.; Dive, V.; Acharya, K. R.; Sturrock, E. D., The N Domain of Human Angiotensin-I-converting Enzyme. *J. Biol. Chem.* **2010,** *285* (46), 35685-35693.

62. Natesh, R.; Schwager, S. L.; Evans, H. R.; Sturrock, E. D.; Acharya, K. R., Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. *Biochem* **2004,** *43* (27), 8718-8724.

63. Watermeyer, J. M.; Sewell, B. T.; Schwager, S. L.; Natesh, R.; Corradi, H. R.; Acharya, K. R.; Sturrock, E. D., Structure of testis ACE glycosylation mutants and evidence for conserved domain movement. *Biochem* **2006,** *45* (42), 12654-12663.

64. Mohd, A.; Sylva, L. S.; Colin, S. A.; Bertrand, C.; Fabrice, B.; Vincent, D.; Edward, D. S., Novel mechanism of inhibition of human angiotensin-I-converting enzyme (ACE) by a highly specific phosphinic tripeptide. *Biochem. J.* **2011,** *436* (1), 53-59.

65. Akif, M.; Masuyer, G.; Schwager, S. L.; Bhuyan, B. J.; Mugesh, G.; Isaac, R. E.; Sturrock, E. D.; Acharya, K. R., Structural characterization of angiotensin I-converting enzyme in complex with a selenium analogue of captopril. *FEBS J.* **2011,** *278* (19), 3644-3650.

66. Watermeyer, J. M.; Kröger, W. L.; O'Neill, H. G.; Sewell, B. T.; Sturrock, E. D., Probing the Basis of Domain-Dependent Inhibition Using Novel Ketone Inhibitors of Angiotensin-Converting Enzyme†‡. *Biochem* **2008,** *47* (22), 5942-5950.

67. Watermeyer, J.; Kroger, W.; O'Neill, H.; Sewell, B.; Sturrock, E., Characterization of domain-selective inhibitor binding in angiotensin-converting enzyme using a novel derivative of lisinopril. *Biochem. J* **2010,** *428*, 67-74.

68. Masuyer, G.; Schwager, S. L.; Sturrock, E. D.; Isaac, R. E.; Acharya, K. R., Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural inhibitory peptides. *Scientific reports* **2012,** *2*.

69. Masuyer, G.; Akif, M.; Czarny, B.; Beau, F.; Schwager, S. L.; Sturrock, E. D.; Isaac, R. E.; Dive, V.; Acharya, K. R., Crystal structures of highly specific phosphinic tripeptide enantiomers in complex with the angiotensin-I converting enzyme. *FEBS J.* **2014,** *281* (3), 943-956.

70. Yates, C. J.; Masuyer, G.; Schwager, S. L.; Akif, M.; Sturrock, E. D.; Acharya, K. R., Molecular and thermodynamic mechanisms of the chloride-dependent human angiotensin-I-converting enzyme (ACE). *J. Biol. Chem.* **2014,** *289* (3), 1798-1814.

71. Kramer, G. J.; Mohd, A.; Schwager, S. L.; Masuyer, G.; Acharya, K. R.; Sturrock, E. D.; Bachmann, B. O., Interkingdom pharmacology of angiotensin-I converting enzyme inhibitor phosphonates produced by actinomycetes. *ACS Med. Chem. Lett* **2014,** *5* (4), 346-351.
72. Douglas, R. G.; Sharma, R. K.; Masuyer, G.; Lubbe, L.; Zamora, I.; Acharya, K. R.; Chibale, K.; Sturrock, E. D., Fragment-based design for the development of N-domain-selective angiotensin-1-converting enzyme inhibitors. *Clin. Sci.* **2014,** *126* (4), 305-313.
73. Masuyer, G.; Douglas, R. G.; Sturrock, E. D.; Acharya, K. R., Structural basis of Ac-SDKP hydrolysis by Angiotensin-I converting enzyme. *Scientific reports* **2015,** *5*.
74. Corradi, H. R.; Chitapi, I.; Sewell, B. T.; Georgiadis, D.; Dive, V.; Sturrock, E. D.; Acharya, K. R., The structure of testis angiotensin-converting enzyme in complex with the C domain-specific inhibitor RXPA380. *Biochem* **2007,** *46* (18), 5473-5478.
75. Burger, D.; Reudelhuber, T. L.; Mahajan, A.; Chibale, K.; Sturrock, E. D.; Touyz, R. M., Effects of a domain-selective ACE inhibitor in a mouse model of chronic angiotensin II-dependent hypertension. *Clin. Sci.* **2014,** *127* (1), 57-63.
76. Denti, P.; Sharp, S.-K.; Kröger, W. L.; Schwager, S. L.; Mahajan, A.; Njoroge, M.; Gibhard, L.; Smit, I.; Chibale, K.; Wiesner, L., Pharmacokinetic evaluation of lisinopril-tryptophan, a novel C-domain ACE inhibitor. *Eur. J. Pharmaceut. Sci.* **2014,** *56*, 113-119.
77. Kröger, W. L.; Douglas, R. G.; O'Neill, H. G.; Dive, V.; Sturrock, E. D., Investigating the domain specificity of phosphinic inhibitors RXPA380 and RXP407 in angiotensin-converting enzyme. *Biochem* **2009,** *48* (35), 8405-8412.
78. Bergmann, R.; Linusson, A.; Zamora, I., SHOP: Scaffold HOPping by GRID-based similarity searches. *J. Med. Chem.* **2007,** *50* (11), 2708-2717.
79. Lyne, P. D., Structure-based virtual screening: an overview. *Drug Disc. Tod.* **2002,** *7* (20), 1047-1055.
80. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Disc.* **2004,** *3* (11), 935-949.
81. Bondi, A., van der Waals volumes and radii. *J. Phys. Chem.* **1964,** *68* (3), 441-451.
82. Goodsell, D. S.; Olson, A. J., Automated docking of substrates to proteins by simulated annealing. *Proteins: Struct. Funct. Bioinf.* **1990,** *8* (3), 195-202.
83. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997,** *267* (3), 727-748.
84. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **1983,** *4* (2), 187-217.
85. Weiner, P. K.; Kollman, P. A., AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J. Comp. Chem.* **1981,** *2* (3), 287-303.
86. Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988,** *110* (6), 1657-1666.
87. Donini, O. A.; Kollman, P. A., Calculation and prediction of binding free energies for the matrix metalloproteinases. *J. Med. Chem.* **2000,** *43* (22), 4180-4188.
88. Vorobjev, Y. N.; Almagro, J. C.; Hermans, J., Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model. *Proteins: Struct. Funct. Bioinf.* **1998,** *32* (4), 399-413.
89. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P., Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997,** *11* (5), 425-445.
90. Leach, A. R., Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.* **1994,** *235* (1), 345-356.

91. Wei, B. Q.; Weaver, L. H.; Ferrari, A. M.; Matthews, B. W.; Shoichet, B. K., Testing a flexible-receptor docking algorithm in a model binding site. *J. Mol. Biol.* **2004,** *337* (5), 1161-1182.
92. Whalen, K. L.; Chang, K. M.; Spies, M. A., Hybrid Steered Molecular Dynamics-Docking: An Efficient Solution to the Problem of Ranking Inhibitor Affinities Against a Flexible Drug Target. *Mol. Inform.* **2011,** *30* (5), 459-471.
93. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004,** *47* (7), 1739-1749.
94. Jones, G.; Willett, P.; Glen, R. C., Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995,** *245* (1), 43-53.
95. Abagyan, R.; Totrov, M.; Kuznetsov, D., ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *Journal of computational chemistry* **1994,** *15* (5), 488-506.
96. Goodsell, D. S.; Morris, G. M.; Olson, A. J., Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recogn.* **1996,** *9* (1), 1-5.
97. Vilar, S.; Cozza, G.; Moro, S., Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery. *Curr. Top. Med. Chem.* **2008,** *8* (18), 1555-1572.
98. Rarey, M.; Kramer, B.; Lengauer, T. In *Time-efficient docking of flexible ligands into active sites of proteins*, Ismb, 1995; pp 300-308.
99. (a) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S., A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006,** *49* (20), 5912-5931; (b) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C., Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *J. Chem. Inf. Mod.* **2009,** *49* (6), 1455-1474.
100. Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M., LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graph. Model.* **2003,** *21* (4), 289-307.
101. Leeson, P. D.; Springthorpe, B., The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Disc.* **2007,** *6* (11), 881-890.
102. Irwin, J. J.; Shoichet, B. K., ZINC-a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Mod.* **2005,** *45* (1), 177-182.
103. Weininger, D., SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comp. Sci.* **1988,** *28* (1), 31-36.
104. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2012,** *64*, 4-17.
105. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'rule of three'for fragment-based lead discovery? *Drug Disc. Tod.* **2003,** *8* (19), 876-877.
106. Agrawal, A.; Johnson, S. L.; Jacobsen, J. A.; Miller, M. T.; Chen, L. H.; Pellecchia, M.; Cohen, S. M., Chelator fragment libraries for targeting metalloproteinases. *Chem. Med. Chem.* **2010,** *5* (2), 195-199.
107. Downs, G. M.; Barnard, J. M., Clustering methods and their uses in computational chemistry. *Rev. Comp. Chem.* **2002,** *18*, 1-40.
108. Hartenfeller, M.; Schneider, G., De novo drug design. *Chemoinf. Comp. Chem. Biol.* **2011**, 299-323.
109. (a) Pierce, A. C.; Rao, G.; Bemis, G. W., BREED: Generating novel inhibitors through hybridization of known ligands. Application to CDK2, p38, and HIV protease. *J. Med. Chem.* **2004,** *47* (11), 2768-2775; (b) Nisius, B.; Rester, U., Fragment shuffling: an automated workflow for three-dimensional fragment-based ligand design. *J. Chem. Inf. Mod.* **2009,** *49* (5), 1211-1222.

110. Böhm, H.-J., Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput. Aided Mol. Des.* **1998,** *12* (4), 309-309.
111. Kutchukian, P. S.; Lou, D.; Shakhnovich, E. I., FOG: Fragment Optimized Growth algorithm for the de novo generation of molecules occupying druglike chemical space. *J. Chem. Inf. Mod.* **2009,** *49* (7), 1630-1642.
112. Lewell, X. Q.; Judd, D. B.; Watson, S. P.; Hann, M. M., Recap retrosynthetic combinatorial analysis procedure: a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. *J. Chem. Inf. Comp. Sci.* **1998,** *38* (3), 511-522.
113. Patchett, A.; Harris, E.; Tristram, E.; Wyvratt, M.; Wu, M.; Taub, D.; Peterson, E.; Ikeler, T.; Ten Broeke, J.; Payne, L., A new class of angiotensin-converting enzyme inhibitors. *Nature* **1980,** *288* (5788), 280-283.
114. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W., Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000,** *33* (12), 889-897.
115. Lyne, P. D.; Lamb, M. L.; Saeh, J. C., Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **2006,** *49* (16), 4805-4808.
116. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Mod.* **2010,** *51* (1), 69-82.
117. Ballatore, C.; Huryn, D. M.; Smith, A. B., Carboxylic acid (bio) isosteres in drug design. *Chem. Med. Chem.* **2013,** *8* (3), 385-395.
118. Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R. A., The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins: Struct. Funct. Bioinf.* **2011,** *79* (10), 2794-2812.
119. (a) Haj-Yehia, A.; Khan, M.; Qadri, B. Ace-inhibitors having antioxidant and no-donor activity. US20060166894 A1, 2003; (b) Hardy, G. W.; Bull, D.; Jones, H. T.; Mills, G.; Allan, G., Synthesis of A575C, a combined angiotensin converting enzyme inhibitor-beta adrenoceptor antagonist. *Tetrahedron Lett.* **1988,** *29* (7), 799-802; (c) Sattigeri, J. A.; Ahmed, S.; Sethi, S.; Gadhave, A. G.; Kaur, K.; Sharma, L.; Jadhav, B. G.; Datta, D.; Chilla, S. M.; Bhatnager, P. K. Proline derivatives for use in the treatment of diabetes. WO2011092671 A1, 2011.
120. Cotton, J.; Hayashi, M. A.; Cuniasse, P.; Vazeux, G.; Ianzer, D.; De Camargo, A. C.; Dive, V., Selective inhibition of the C-domain of angiotensin I converting enzyme by bradykinin potentiating peptides. *Biochem* **2002,** *41* (19), 6065-6071.
121. Carpino, L. A.; El-Faham, A.; Albericio, F., Efficiency in peptide coupling: 1-hydroxy-7-azabenzotriazole vs 3, 4-dihydro-3-hydroxy-4-oxo-1, 2, 3-benzotriazine. *J. Org. Chem.* **1995,** *60* (11), 3561-3564.
122. Harfenist, M.; Hoerr, D. C.; Crouch, R., Enantiospecific synthesis of the trans-9-[3-(3, 5-dimethyl-1-piperazinyl) propyl] carbazoles. *J. Org. Chem.* **1985,** *50* (9), 1356-1359.
123. Greenlee, W. J.; Allibone, P. L.; Perlow, D. S.; Patchett, A. A.; Ulm, E. H.; Vassil, T. C., Angiotensin-converting enzyme inhibitors: synthesis and biological activity of acyltripeptide analogs of enalapril. *J. Med. Chem.* **1985,** *28* (4), 434-442.
124. McDermott, J. R.; Benoiton, N. L., N-Methylamino acids in peptide synthesis. III. Racemization during deprotection by saponification and acidolysis. *Can. J. Chem.* **1973,** *51* (15), 2555-2561.
125. Fischer, J. D.; Fodor, T. D.; Dobay, L.; Stefko, B. D. Verfahren zur herstellung von 1-(3'-(mercapto)-(2's)-(methyl)-propionyl)-pyrrolidin-(2s)-carbonsaeure oder ihren salzen. DE3721430 A1, 1988.
126. Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B., ChEMBL: a large-scale bioactivity database for drug discovery. *Nucl. Acids Res.* **2012,** *40* (D1), D1100-D1107.

127. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K., BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucl. Acids Res.* **2007,** *35* (suppl 1), D198-D201.
128. Cushman, D.; Cheung, H., Spectrophotometric assay and properties of the angiotensin-converting enzyme of rabbit lung. *Biochem. Pharm.* **1971,** *20* (7), 1637-1648.
129. Mottley, C.; Toy, K.; Mason, R., Oxidation of thiol drugs and biochemicals by the lactoperoxidase/hydrogen peroxide system. *Mol. Pharmacol.* **1987,** *31* (4), 417-421.
130. (a) Stanton, J. L.; Watthey, J. W.; Desai, M. N.; Finn, B. M.; Babiarz, J. E.; Tomaselli, H. C., Angiotensin converting enzyme inhibitors: structure-activity profile of 1-benzazepin-2-one derivatives. *J. Med. Chem.* **1985,** *28* (11), 1603-1606; (b) Ksander, G. M.; Erion, M.; Yuan, A. M.; Diefenbacher, C. G.; El-Chehabi, L.; Cote, D.; Levens, N., Dual angiotensin converting enzyme/thromboxane synthase inhibitors. *J. Med. Chem.* **1994,** *37* (12), 1823-1832.
131. (a) Gruenfeld, N.; Stanton, J. L.; Yuan, A. M.; Ebetino, F. H.; Browne, L. J.; Gude, C.; Huebner, C. F., Angiotensin converting enzyme inhibitors: 1-glutarylindoline-2-carboxylic acid derivatives. *J. Med. Chem.* **1983,** *26* (9), 1277-1282; (b) Ksander, G. M.; Yuan, A. M.; Diefenbacher, C. G.; Stanton, J. L., Angiotensin converting enzyme inhibitors: N-substituted D-glutamic acid. gamma.-dipeptides. *J. Med. Chem.* **1985,** *28* (11), 1606-1611.
132. Aoyagi, Y., An angiotensin-I converting enzyme inhibitor from buckwheat (Fagopyrum esculentum Moench) flour. *Phytochem.* **2006,** *67* (6), 618-621.
133. Piquilloud, Y.; Reinharz, A.; Roth, M., Studies on the angiotensin converting enzyme with different substrates. *Biochim. Biophys. Acta - Enzym.* **1970,** *206* (1), 136-142.
134. Danilov, S. M.; Balyasnikova, I. V.; Albrecht, R. F.; Kost, O. A., Simultaneous determination of ACE activity with 2 substrates provides information on the status of somatic ACE and allows detection of inhibitors in human blood. *J. Cardiovasc. Pharmacol.* **2008,** *52* (1), 90-103.
135. Araujo, M. C.; Melo, R. L.; Cesari, M. H.; Juliano, M. A.; Juliano, L.; Carmona, A. K., Peptidase specificity characterization of C-and N-terminal catalytic sites of angiotensin I-converting enzyme. *Biochem* **2000,** *39* (29), 8519-8525.