University College London

Development of Novel Metabolically Stable Analogues of PAR-1 Antagonist RWJ-58259

By

Eifion Robinson

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Declaration

I, Eifion Robinson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Eifion Robinson

September 2014

Abstract

This thesis describes the investigation into the development of novel antagonists of the major thrombin receptor, protease-activated receptor-1, in the aim of generating analogues of the project lead compound with improved metabolic stability. Chapter 1 provides an introduction to the research project and the biological rationale for targeting protease-activated receptor-1 for inhibition. Chapter 2 describes modifications made to the synthetic route to produce RWJ-58259, the project lead, and subsequent biological evaluation carried out in order to confirm the compound's identity and verify batch quality. Chapter 3 describes the investigation into the metabolic profile of RWJ-58259 to identify major sites of metabolism. Presented in this chapter are details of the syntheses of analogues of RWJ-58259 incorporating substituents designed to block possible sites of metabolism in the aim of improving *in vivo* stability. The development of an *in vitro* assay suited to screen analogues efficiently is described and the results from the biological evaluation of compounds synthesised are presented with discussion of the data generated.

Contents

Acknowledgments

I would like to thank my supervisor Professor Stephen Caddick for giving me the opportunity to carry out this research and for his support and guidance throughout my PhD studies. I would also like to thank Professor Rachel Chambers and Graham Inglis for their collaboration on this project, providing valuable advice and enthusiasm during the research.

Thanks to all the members of the KLB 230 and 237 labs for making my time there so enjoyable and especially to the Caddick group members, past and present, for their friendship and support. Special thanks to Vijay, Rachel and Ramiz for proof reading my thesis.

I would also like to thank Dr Natalia Smoktunowicz and Dr Ricardo José for their help with the biological work. Thank you to Abil Aliev for assistance with NMR, Lisa Haigh and Kersti Karu with mass spectrometry and Sam Ranasinghe for support with the FLIPR instrument.

Also, thanks to GSK for the opportunity to work at the Stevenage site for six months during my research and for making their facilities available to me.

Finally, thank you to my family for their unwavering support and especially Linda and Tristan for your sacrifices, patience and belief in me.

Abbreviations

$[\alpha]_D$	-	Optical rotation
AcOH	-	Acetic acid
AECs	-	Alveolar epithelial cells
Ala	-	Alanine
APC	-	Activated protein C
API	-	Active pharmaceutical ingredient
Arg	-	Arginine
AT2	-	Alveolar epithelial Type-II
ATP	-	Adenosine 5'-triphosphate
BAL	-	Bronchoalveolar lavage
Boc	-	*tert*-Butoxycarbonyl
Bn	-	Benzyl
b.p.	-	Boiling point
CCL	-	Chemokine (C-C motif) ligand
CNS	-	Central nervous system
CPFE	-	Combined pulmonary fibrosis and emphysema
CTGF	-	Connective tissue growth factor
CYP	-	Cytochrome P450
DAG	-	Diacyl glycerol
Dbu	-	α,γ-Diaminobutyric acid

DCC	-	*N,N'*-Dicyclohexylcarbodiimide
DIC	-	*N,N'*-Diisopropanylcarbodiimide
DIP	-	Desquamative interstitial pneumonia
DIPEA	-	*N,N*-Diisopropylethylamine
DIU	-	*N,N'*-Diisopropyl urea
DMAP	-	4-(Dimethylamino)pyridine
DMF	-	*N,N*-Dimethylformamide
DM/PK	-	Drug metabolism/Pharmacokinetics
DMSO	-	Dimethylsulfoxide
DNA	-	Deoxyribonucleic acid
DPLDs	-	Diffuse parenchymal lung diseases
EBV	-	Epstein-Barr virus
ECM	-	Extracellular matrix
ECs	-	Extracellular loops
EDC	-	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
ELISA	-	Enzyme-linked immunosorbent assay
EMT	-	Epithelial-mesenchymal transition
ER	-	Endoplasmic reticulum
ERK	-	Extracellular signal-regulated kinases
Et	-	Ethyl
FDA	-	Food and Drug Administration

FII-XIII(a)	-	Coagulation factors II to XIII (activated)
FLIPR	-	Fluorescence Imaging Plate Reader
Fmoc	-	Fluorenylmethyloxycarbonyl
FT-IR	-	Fourier transform infrared
GI	-	Gastrointestinal
GOR	-	Gastro-oesophageal reflux
GPCR	-	Guanosine nucleotide-binding protein coupled receptor
GRKs	-	GPCR kinases
GTPase	-	Guanosine triphosphate hydrolase
HBTU	-	O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate
HCV	-	Hepatitis C virus
HEK	-	Human embryonic kidney
HIV	-	Human immunodeficiency virus
HLF	-	Human lung fibroblast
HOBt	-	Hydroxybenzotriazole
HRMS	-	High resolution mass spectrometry
HTS	-	High throughput screening
IC_{50}	-	Half maximal inhibitory concentration
IIPs	-	Idiopathic interstitial pneumonias
IL	-	Interleukin
ILDs	-	Interstitial lung diseases

IP$_3$	-	Inositol 1,4,5-triphosphate
IPF	-	Idiopathic pulmonary fibrosis
^{i}Pr	-	Isopropyl
LAP	-	Latency-associated peptide
LC/MS	-	Liquid chromatography/mass spectrometry
Leu	-	Leucine
LHS	-	Left-hand side
LPS	-	Lipopolysaccharide
Lys	-	Lysine
MAPK	-	Mitogen-activated protein kinases
Me	-	Methyl
MeCN	-	Acetonitrile
MeOH	-	Methanol
MI	-	Myocardial infarction
MLF	-	Murine lung fibroblast
MMP	-	Matrix metalloprotease
mRNA	-	Messenger ribonucleic acid
Ms	-	Mesylate
NAM	-	Negative allosteric modulator
NIH	-	National Institutes of Health
NMR	-	Nuclear magnetic resonance

Orn	-	Ornithine
P1-5	-	Position 1 to 5
PAR	-	Protease-activated receptor
Ph	-	Phenyl
PH	-	Pulmonary hypertension
Phe	-	Phenylalanine
PIP$_2$	-	Phosphatidylinositol 4,5-biphosphate
PKC	-	Protein kinase C
PLC	-	Phospholipase C
PPIs	-	Proton pump inhibitors
RHS	-	Right-hand side
ROS	-	Reactive oxygen species
SAR	-	Structure-activity relationship
Ser	-	Serine
α-SMA	-	α-Smooth muscle actin
SMC	-	Smooth muscle cell
SP	-	Surfactant protein
$t_{1/2}$	-	Half-life
TF	-	Tissue factor
TFA	-	Trifluoroacetic acid
TFPI	-	Tissue factor pathway inhibitor

TGF-β	-	Transforming-growth factor-β
THF	-	Tetrahydrofuran
TIA	-	Transient ischemic attack
TLC	-	Thin layer chromatography
7TM	-	Seven-transmembrane
TNF	-	Tumour necrosis factor
TRAP	-	Thrombin receptor activating peptide
Tyr	-	Tyrosine
UPR	-	Unfolded protein response

Chapter 1 **Introduction**

1.1 Fibrosis

Accounting for almost 45% of deaths in the developed world, chronic fibroproliferative diseases represent one of the biggest challenges for modern medicine.[1] The economic cost of the burden placed on health care services, by treating diseases resulting in fibrosis, is estimated to be in the tens of billions of dollars per year for the USA alone.[2] As such, development of therapeutics to treat fibrosis would not only improve human health considerably but could also plug a significant economic drain.

Fibrosis is defined as the formation of scar tissue due to excessive deposition and accumulation of extracellular matrix (ECM) components. The ECM is a dynamic network of various fibrous proteins and glycosaminoglycans, which provides structural support which is unique in its arrangement for specific anatomical locations.[3] The biochemical make-up of the ECM allows for the sequestration of growth factors and cytokines, regulating access to which, provides a means for intercellular communication and the transmission of mechanical cues to the surrounding cells. ECM components are synthesised within and excreted from cells for which the ECM in turn provides structural support. Fibrosis is believed to stem from a combination of persistent local injury and dysregulation of the natural wound healing process, and can occur in all tissues and organ systems.[4] Examples of major contributing diseases (and the organs affected) include: pulmonary fibrosis (lungs), cirrhosis (liver), chronic kidney disease (kidneys), endomyocardial fibrosis (heart), scleroderma/systemic sclerosis (skin/various organs) and Crohn's disease (intestines).

The origin of the initial injury in various fibrotic diseases may vary greatly for different organs, and in some cases the origin is unknown. For example, major sources of initial injury to hepatocytes are chronic viral hepatitis B and C infection, or aberrant lipid metabolism caused by alcohol-related disorders or high fat diets.[5] Likewise, sources of injury to kidney cells can range from diabetes or autoimmune disease to xenobiotic toxicity.[4] Regardless of the source of initial insult, if persistent,

chronic tissue injury results and may give rise to fibrosis. Gross remodelling then distorts the normal tissue architecture of the organ system leading to progressive loss of function. The consequences are dire in diseases such as idiopathic pulmonary fibrosis (IPF), liver cirrhosis, cardiovascular fibrosis, systemic sclerosis and nephritis, where organ failure ensues and ultimately leads to death.[1]

Despite the contribution of fibroproliferative diseases towards mortality in the developed world, research in this field is relatively new. Mechanisms involved in disease pathogenesis are poorly understood and as a result there are currently no clinically proven therapeutic options to combat fibrosis. Progress in the field over the last quarter of a century, however, has led to the recognition of common features in different fibrotic diseases.[6] This has begun to generate a clearer picture of the possible processes involved in disease progression, making therapeutic intervention a realistic goal for the near future.

1.2 Wound Repair and Regeneration

The ability to detect injury and respond quickly is a process that is well-established in most multicellular organisms. Delivery of essential components to fight infection and repair or replace injured tissues reduces the damage sustained and increases the chance of survival.

Within a healthy organism the physiological environment is described as being in an anticoagulative-profibrinolytic state, meaning conditions are held so as to deter coagulation and favour fibrinolysis. Upon injury to cells, the wound healing process (Figure 1) is triggered involving a complex network of responses, which temporarily alter the normal state of the body towards a procoagulative-antifibrinolytic state. The balance between these counterpart positions is maintained until the healing process is complete and the anticoagulative-profibrinolytic condition resumes. The process of repair is generally recognised as being comprised of three distinct phases: inflammation, fibroplasia and maturation, each of which will be discussed in detail below (see sections 1.2.1-1.2.3).[7] In the normal state, following initial injury, the healing process culminates in the restoration or replacement of damaged tissue to its normal functioning state and concludes leaving little or no scarring. However, in a diseased state, sustained injury leads to unbalanced responses and abnormalities in

the healing process, resulting in excessive deposition of ECM components, scarring and distortion of the normal tissue architecture.
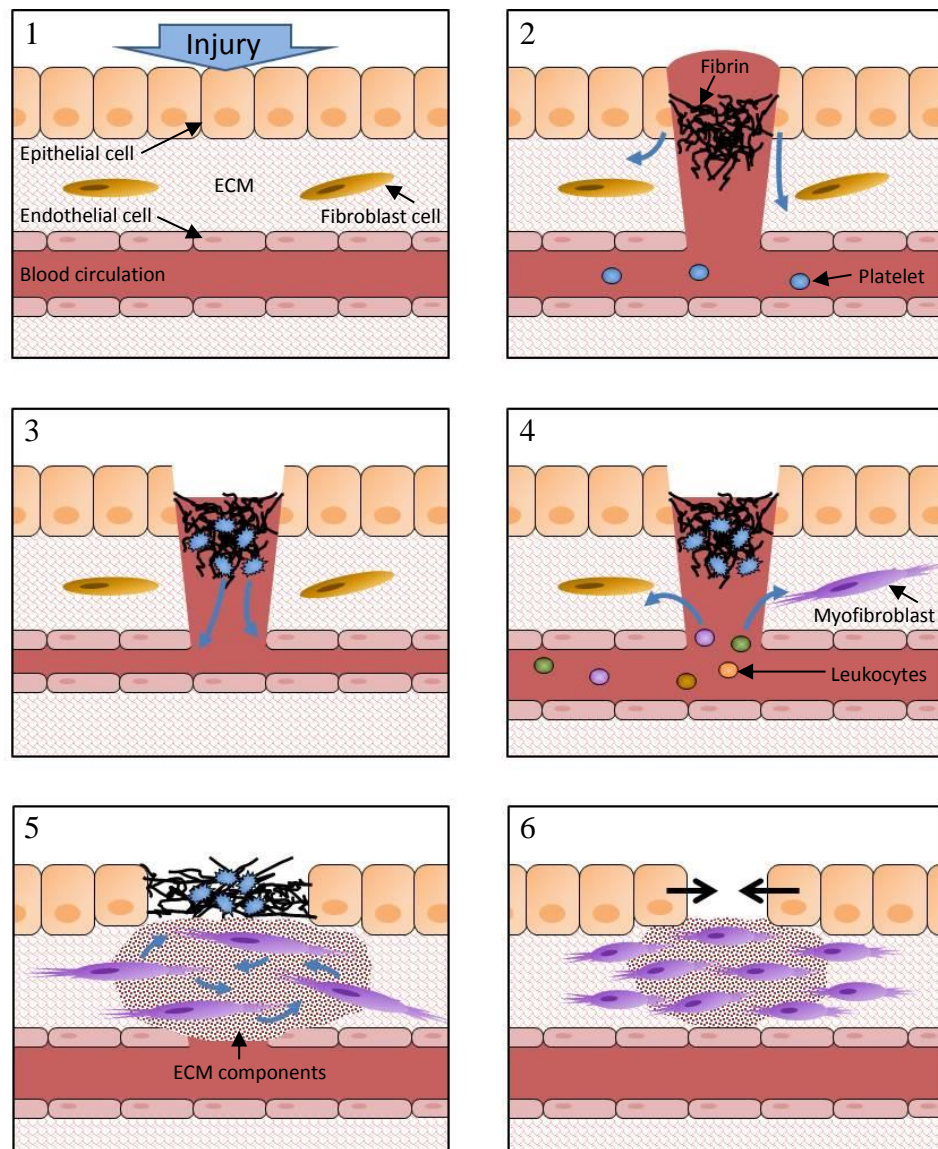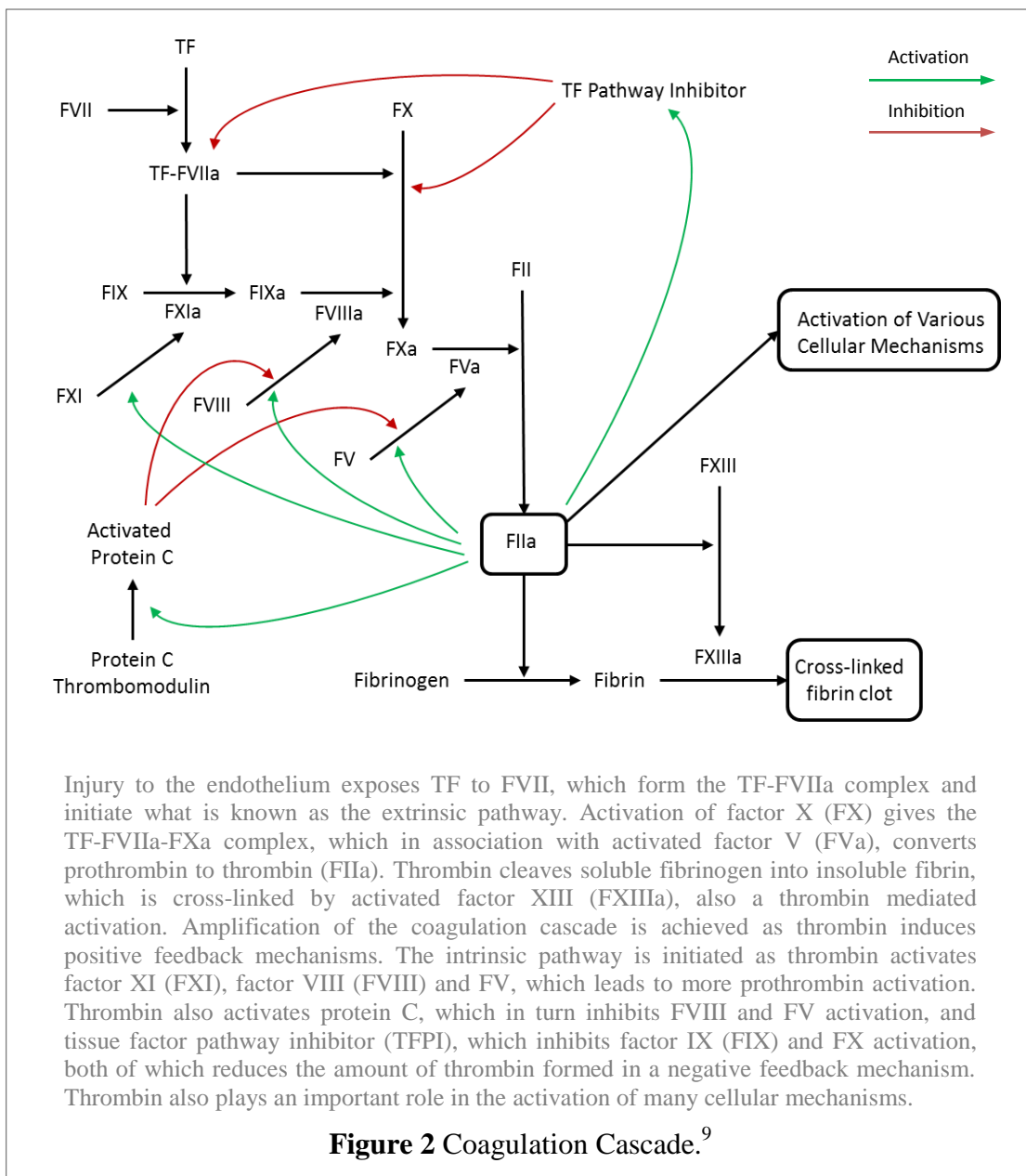

1 – Injury occurs to healthy tissue; 2 – following injury growth factors and chemokines released from damaged epithelial cells trigger various responses in the local environment including vasoconstriction, stress responses in neighbouring cells and transforming growth factor (TGF)-β activation. Activation of the coagulation cascade leads to the formation of a fibrin clot; 3 – activated platelets adhere to the fibrin clot and degranulation leads to the release of numerous growth factors and vasoactive substances resulting in vasodilation and increased blood vessel permeability; 4 - leukocytes recruited from circulation release various cytokines and growth factors, which direct the healing process. Disruption of the ECM and TGF-β activation leads to development of the fibroblast cell into a myofibroblast; 5 - myofibroblasts from various sources migrate to the injury site and deposit ECM components; 6 – myofibroblasts contract and close the wound, allowing re-epithelialisation of the surface. Myofibroblasts undergo apoptosis and matrix metalloproteases (MMPs) degrade excess ECM to restore the tissue to its normal state.

Figure 1 Wound Healing Process.[8]

1.2.1 Inflammation Phase

The initial response following injury to the epithelium is a period of intense vasoconstriction triggered by a combination of factors, including the release of components from damaged epithelial cells. Activation of the coagulation cascade is one of the earliest responses to tissue injury and is triggered as circulating factor VII (FVII) is exposed to tissue factor (TF), which is expressed in tissues that are normally concealed from plasma. The coagulation cascade (Figure 2) involves the sequential activation of zymogens in circulation to their functional forms, which ultimately results in the formation of a blood clot through the conversion of soluble fibrinogen to insoluble fibrin by activated thrombin (FIIa).

Injury to the endothelium exposes TF to FVII, which form the TF-FVIIa complex and initiate what is known as the extrinsic pathway. Activation of factor X (FX) gives the TF-FVIIa-FXa complex, which in association with activated factor V (FVa), converts prothrombin to thrombin (FIIa). Thrombin cleaves soluble fibrinogen into insoluble fibrin, which is cross-linked by activated factor XIII (FXIIIa), also a thrombin mediated activation. Amplification of the coagulation cascade is achieved as thrombin induces positive feedback mechanisms. The intrinsic pathway is initiated as thrombin activates factor XI (FXI), factor VIII (FVIII) and FV, which leads to more prothrombin activation. Thrombin also activates protein C, which in turn inhibits FVIII and FV activation, and tissue factor pathway inhibitor (TFPI), which inhibits factor IX (FIX) and FX activation, both of which reduces the amount of thrombin formed in a negative feedback mechanism. Thrombin also plays an important role in the activation of many cellular mechanisms.

Figure 2 Coagulation Cascade.[9]

Sub-endothelial ECM components and proteases from the coagulation cascade trigger platelet aggregation, contributing to the fibrin clot and together they form the provisional matrix. Platelets are storehouses of numerous growth factors and vasoactive substances. Upon activation, degranulation occurs, promoting vasodilation and increased blood vessel permeability, which assists in the recruitment of pro-inflammatory cells.[8, 10]

Epithelial cells respond to injury in various ways, which initiate pathways to assist the wound healing process. If injury is severe, cell death occurs, releasing cell components, which stimulate stress responses in neighbouring cells.[6, 11] In episodes of less severity, the cell survives initial injury but a variety of stress responses are triggered as a result. Changes in the cell's local environment as a result of tissue damage can lead to complications in the folding of proteins, known as endoplasmic reticulum (ER) stress, as gene expression is increased in an attempt to aid survival. [12] The unfolded protein response (UPR) is initiated as a consequence to help clear the accumulation of misfolded proteins or, for cases in which the situation cannot be resolved, initiate apoptosis. Sources and mechanisms by which energy for metabolism is generated within cells, in the form of adenosine 5'-triphosphate (ATP), are also affected under stress leading to the formation of damaging reactive oxygen species (ROS) as by-products, which stimulate pro-fibrotic pathways.[13]

High levels of integrin αvβ6 are expressed in epithelial cells in response to stress. This trans-membrane protein binds to the latency-associated peptide (LAP) portion of latent transforming-growth factor (TGF)-β (Figure 1, page 3) stored in the ECM. Activation of cell surface receptors by mediators released or activated as part of the stress response, such as the coagulation protease thrombin, initiates epithelial cell contraction. Transmission of this force to the LAPs *via* integrin αvβ6 results in TGF-β activation.[14] TGF-β is a major component in the wound healing process affecting various responses from many different cell types. These include: i) connective tissue growth factor (CTGF) expression, which stimulates fibroblast proliferation and ECM synthesis; ii) αvβ6 integrin expression, which creates a positive feedback mechanism by activation of more latent TGF-β; and iii) alteration in gene expressions in epithelial cells so as to trigger epithelial-mesenchymal transition (see section 1.4.3.1, page 14).[6, 15, 16]

Recruitment of leukocytes, following the secretion of growth factors and chemokines by epithelial and endothelial cells, is assisted through disruption of the basement membrane by matrix metalloproteases (MMPs) produced by epithelial and/or endothelial cells and myofibroblasts.[8] The earliest to respond and migrate are neutrophils, which eliminate any invading organisms by phagocytosis and the formation of phagolysosomes. However, they often release free radicals as a result, destroying many healthy host cells in the local environment.[17] Macrophages arrive later and clear the wound of tissue debris and expended neutrophils. Also recruited, but later still, are mast cells. Together, these different leukocytes secrete a cocktail of cytokines and growth factors, which direct the wound healing response and amplify the inflammatory signal drawing more leukocytes to the site of injury. Recruitment and activation of T-cells promotes the secretion of pro-fibrotic cytokines such as interleukin-13 (IL-13) and TGF-β.

1.2.2 Fibroplasia Phase

Fibroblast cells present in connective tissues become activated as a result of disruption to their protective cross-linked ECM micro-environment during injury. Normally a relatively dormant cell, upon activation the fibroblast develops into a migratory phenotype, termed "proto-myofibroblast", and travels towards the wound.[18] Further activation by TGF-β completes the transformation to a myofibroblast (a collagen-secreting α-smooth muscle actin-positive (α-SMA$^+$) fibroblast). The myofibroblast is the major cell type involved in the production and deposition of granulation tissue, comprised mostly of ECM components collagen and fibronectin, which replace the provisional matrix.[19] Myofibroblasts differ from fibroblasts by their expression of α-SMA and the presence of contractile apparatus. Bundles of cytoplasmic α-SMA incorporated actin microfilaments align and converge at sites on the cell membrane. These are known as focal adhesions in most cell types, or fibronexus in myofibroblasts, and are sites where transmembrane integrins link the intracellular actin with extracellular fibronectin fibrils.[18, 20-22] The force generated by the actin stress fibres is transmitted to the surrounding ECM, leading to contraction of the tissue and closing of the wound. Regeneration of the damaged tissue is achieved as epithelial and endothelial cells divide and migrate across the basal layer.

1.2.3 Maturation Phase

Finally, myofibroblasts undergo apoptosis, reducing the rate of ECM deposition. A balance is reached between synthesis and MMP degradation of ECM components. Reorganisation of the collagen fibres occurs over a period of up to 2 years, progressively strengthening the tissue.[7, 8, 23]

The wound healing process as a whole is a complex mix of balanced responses to various stimuli, which ordinarily play out without issue and result in the regeneration of damaged tissues. However, complications arise with serious consequences when responses are uncontrolled, as discussed in the following sections.

1.3 Pulmonary Fibrosis

Pulmonary fibrosis is a devastating progressive condition, which results from scarring of the lung tissue leaving it thickened and stiff. The walls of the alveoli, through which gaseous exchange occurs with blood in the pulmonary capillaries, are left unable to function as a suitable membrane and the body suffers oxygen deficiency and carbon dioxide build-up as a result. Patients diagnosed with the condition face a devastating prognosis - in the absence of therapeutic options, organ transplantation gives the only realistic hope of survival. However, there is a shortage of donors and for many the waiting list is too long.

Rather than a single disease, pulmonary fibrosis represents the common end stage of many diffuse parenchymal lung diseases (DPLDs). DPLDs, also known as interstitial lung diseases (ILDs), are a group of disorders which affect the alveolar space and/or the pulmonary interstitium – the space between the alveolar epithelial and capillary endothelial basement membranes. There are over 200 distinct entities of DPLDs, each characterised by varying levels of inflammation and fibrosis, and are classified by their differing etiology, clinical course, radiologic presentation and histopathology.[24] Not all DPLDs progress towards pulmonary fibrosis and in some cases early and accurate classification of the specific disease can allow the cause of injury to be treated, avoiding disease progression.[25] Classification, however, is a difficult process due to the complexity of mechanisms involved in progressing DPLDs towards the pulmonary fibrosis stage. Continued research in the area means

that previously agreed classes of disease are revised regularly as experience with and understanding of the various diseases grows.

1.3.1 DPLD Classification

The classification of DPLDs has challenged clinicians and investigators for over a century. Sir William Osler, in 1892, used the term "cirrhosis of the lung" and described patients as having "a fibrinoid change, which may have its starting point in the tissue about the bronchi and blood vessels, the interlobular septa, the alveolar walls of the pleura".[26] He also stated that "so diverse are the different forms and so varied the conditions under which this change occurs that proper classification is extremely difficult" – as has been the case in the time since.[26]

Cases of the different diseases are relatively uncommon and therefore very few clinicians have extensive experience; and yet collectively DPLDs account for approximately 15% of diseases seen in pulmonary medicine, which has led to discrepancies in diagnoses.[27] Accurate classification of the disease at hand is imperative since treatment options and prognosis can vary significantly for different DPLDs. Sarcoidosis, for example, in most cases does not require systemic therapy and resolves itself over a period of 1-2 years. However, chronic forms, in which the disorder persists for more than 2-5 years, may require long-term therapy involving corticosteroids for more than 5 years.[28] Prognosis for collagen vascular diseases, such as systemic sclerosis, can vary greatly depending upon the extent and severity of the disease, with reports suggesting 10- and 20-year survival rates of 70% and 45%, respectively, and initial treatment with cyclophosphamide to be efficacious.[29, 30]

The work of Hamman and Rich in the 1930's, who are generally considered as the first to describe IPF as a new clinical and pathological entity, enhanced awareness of different classes of fibrotic lung diseases. Over the following decades associations were made for certain symptoms with known causes such as systemic disease (collagen vascular diseases), drugs, inherited conditions and occupational or environmental exposures (organic/inorganic dust).[31] However, a substantial number of cases remained unclassified with regards to causative agents and formed a category of diseases now generally referred to as idiopathic interstitial pneumonias (IIPs).[32, 33]

In 2002, the American Thoracic Society and the European Respiratory Society published a joint international consensus statement to aid in the classification of IIPs.[34] The objective was to provide an integrated multidisciplinary approach to IIP classification to avoid confusion and misclassification of patients. The statement was updated in 2013, reflecting the continued progress made with regards to disease classification.[35] However, although accurate classification of IIPs may provide an improved prognosis for some patients, the fact remains that in most cases therapy is based on relieving symptoms rather than treating the disease. IPF is the major contributor to cases of IIP and as such has been the driving force behind this research project.

1.4 Idiopathic Pulmonary Fibrosis

IPF represents the most common form of diseases culminating in pulmonary fibrosis, with an estimated 5000 new cases diagnosed each year in the UK alone, and with a total of 500,000 people currently affected in the USA and Europe.[36, 37] There are currently no clinically proven therapeutic options for patients diagnosed with IPF, as such it represents a major unmet medical need.

1.4.1 Diagnosis and Prognosis of Patients with IPF

Of all DPLDs IPF commands the most devastating prognosis, with a mean survival time from diagnosis of only 3 years and 80% mortality within 5 years. Defined as a specific form of chronic, progressively fibrosing IIP limited to the lungs, the disease occurs predominantly in middle-aged and elderly adults with a mean age of 66 years at diagnosis and prevalence is generally higher in men than in women.[38, 39]

IPF has a long asymptotic period (months to years) when it is believed micro-injuries occur in the lungs, progressing towards a threshold beyond which the damage is sufficient to cause the onset of symptoms. At the time medical attention is sought symptoms observed include chronic and progressive dyspnoea, non-productive cough with exertion, bibasilar inspiratory crackles and, frequently, clubbing of the fingers. Most patients have these symptoms for an average of 24 months before diagnosis.[40, 41] The clinical course following diagnosis of IPF is recognised as being heterogeneous and distinct patterns of comorbidities are becoming associated with

different clinical phenotypes.[40] Factors correlating with a poorer prognosis include old age (>70 years), a history of smoking, a low body-mass index, digital clubbing, severe physiological impairment, large radiological extent of disease, and pulmonary hypertension.[42]

The typical clinical course in most cases of IPF follows a stable or slowly progressive deterioration in health and the patient's symptoms are carefully managed over a period of months to years. In an estimated 10-20% of cases, however, patients experience periods of rapid deterioration with no discernible cause. Referred to as acute exacerbations, episodes of abrupt declines in clinical, physiological and radiological findings precede, and possibly initiate, the terminal phase of the disease. Less than 40% of hospital admissions with acute exacerbations survive, and there is >90% mortality among survivors within 6 months.[43, 44]

An accelerated variant of IPF is seen in a subgroup of patients and is known as accelerated IPF. This generally affects male cigarette smokers, and this clinical phenotype exhibits a rapidly progressive course with significant reduction in survival time. Despite presenting no divergent diagnostic features at the time of diagnosis from the standard slowly progressive form, the transcriptional profile of this rapidly progressing form indicates up-regulation of a number of functional pathways that operate mainly in alveolar epithelial and mesenchymal cells. Cell motility, myofibroblast differentiation, oxidative stress, coagulation and development may be affected as a result of the differing pattern of gene expression.[45]

The diagnosis of IPF in combination with other disorders of the lung is seen to affect prognosis. A number of patients, mainly men who are heavy cigarette smokers, diagnosed with IPF present lesions associated with emphysema, which has been reported to have a negative effect on prognosis compared with IPF patients without emphysema.[46] It is unclear whether combined pulmonary fibrosis and emphysema (CPFE) is a distinct condition or a subgroup of IPF patients with a different prognosis.[47]

Pulmonary hypertension (PH) has been reported to manifest in almost a third of IPF cases and could possibly indicate the beginning of the terminal phase.[48-50] PH

correlates with a higher mortality not only in IPF but also in other DPLDs and patients with CPFE are seen to promptly develop severe PH.[47, 48, 51]

High incidence of lung cancer has also been associated with IPF (10-38% of cases).[52] Although the mechanisms by which fibrosis leads to the development of lung carcinoma are obscure, regions of honeycombing lesions distinctive of IPF seem to be highly susceptible, suggesting a contributory relationship.[53]

1.4.2 Etiology of IPF

By definition, IPF is a disease of unknown origin. Potential risk factors have been proposed, such as cigarette smoking, viral infections, environmental exposures, gastro-oesophageal reflux (GOR) and surfactant protein polymorphism, however specific etiology remains poorly understood.[54, 55] A combination of genetic predisposition and exposure to injurious environmental agents is currently the generally accepted cause of persistent injury.

1.4.2.1 Genetic factors

Although the occurrence of IIPs in two or more members of the same primary biological family is reportedly seen in only 0.5-2% of patients, sufficient clustering of cases occurs to suggest the involvement of genetic factors in the development of disease.[56]

Mutations in surfactant protein (SP)-C have long been associated with familial cases of interstitial pneumonias, with a recent study reporting 25% of unrelated cases of familial pulmonary fibrosis showing SP-C mutations that were absent in sporadic cases.[57] Expression of SP-A & SP-C mutations leads to dysfunction in alveolar epithelial Type-II (AT2) cells and, together with mutations associated with cell stability such as telomerase mutations (hTERT and hTER), can compromise normal healing of the epithelium.[58]

Reports linking IPF to the ELMOD2 gene implicate genetic alterations in disease development for both familial and sporadic cases. mRNA expression of the ELMOD2 gene, shown to be involved in the regulation of anti-viral responses, is significantly decreased in diseased lungs by comparison to healthy individuals.[59]

This observation gives strength to the view that a combination of genetic and environmental factors may be responsible.

Increased existence in polymorphisms of genes encoding for various cytokines, enzymes, pro-fibrotic molecules, and other mediators of tissue repair and regeneration have been reported in cases of IPF, but none have been validated to date.[54]

1.4.2.2 Injurious agents

Various harmful environmental factors have been proposed as possible causes of injury, prominent among these are viral infections and GOR.

The main focus of research into viruses linked with IPF has centred on Epstein-Barr virus (EBV) and hepatitis C. A number of studies report the detection of both EBV protein and DNA expression in cases of IPF.[60, 61] Tang *et al.* reported that almost two-thirds of both familial and sporadic cases of IPF showed the presence of EBV in alveolar epithelial cells and their data suggested a decrease in the rate of decline in pulmonary function following anti-viral therapy.[62] Other herpes viruses were also tested for and it was found that one or more viruses were present in 96% of cases with IPF compared with 36% of the control population. Studies looking into the possible causative relationship between hepatitis C virus (HCV) and IPF have returned varied results. Significant HCV prevalence is recognised in cases of the disease; however, anti-HCV antibody levels did not differ from other lung diseases with known etiology.[54, 63]

A relationship between IPF and GOR has been suggested since the 1970's and current research estimates the prevalence of GOR in cases of IPF to be as much as 90%.[64, 65] Abnormal acid GOR was shown to be significantly higher in IPF patients than in patients with intractable asthma, a population shown to have high prevalence of GOR. However, 63% of IPF patients treated with the standard dose of proton pump inhibitors (PPIs), the primary medication for GOR, continued to display abnormal levels of acid GOR.[65] This may signify a reversal in the suspected roles in the relationship, with GOR possibly a symptom in most cases of IPF. Another interesting study investigated cases of asymmetrical IPF and found a significantly increased prevalence of GOR in patients with asymmetrical (63%) compared with

symmetrical IPF (31%).[66] Acute exacerbations were also seen to occur more often in cases of asymmetrical IPF, possibly suggesting the role of specific regional conditions – affected by GOR, for example – in the development and/or progression of fibrosis.

1.4.3 Disease Progression in IPF

Inflammation is accepted as having an important role in many DPLDs and, in chronic forms, can progress towards fibrosis. However, hypotheses suggesting inflammation leads to fibrosis in IPF have been put in doubt as understanding of the disease has grown and various studies have failed to confirm this relationship.[41] Locations of extensive fibrosis known as fibroblastic foci, which are characteristic of IPF, were previously thought to develop from lesions seen in desquamative interstitial pneumonia (DIP) – an IIP featuring elevated accumulation of macrophages in the alveolar interstitium, *i.e.* inflammation, resulting from insults caused by cigarette smoking. However, these fibroblastic foci are now understood to occur *via* different mechanisms since elevated levels of macrophages are not seen in these locations. Some evidence suggests inflammation is not required for fibrosis to occur and that, in the absence of certain biological factors, exaggerated inflammation can be induced without leading to fibrosis. Also, the extent of inflammation fails to correlate with staging of fibrosis and the absence of a response in IPF patients to anti-inflammatory therapy would suggest that development of disease is driven by an alternative mechanism.[41]

Current understanding of mechanisms involved in the progression of IPF towards end-stage pulmonary fibrosis suggests an epithelial-driven hypothesis. Persistent injury, possibly resulting from a combination of genetic predisposition and exposure to a harmful environmental agent, affects the ability of alveolar epithelial cells (AECs) to regenerate and replenish the alveolar surface (Figure 3). Subsequent destruction of the basement membrane allows pro-fibrotic cells and mediators to infiltrate. A malfunctioning response to tissue injury results in uncontrolled deposition of ECM components whilst processes such as apoptosis and phagocytosis are retarded. Normal lung architecture is distorted by fibroblastic foci leading to loss of function and ultimately death.
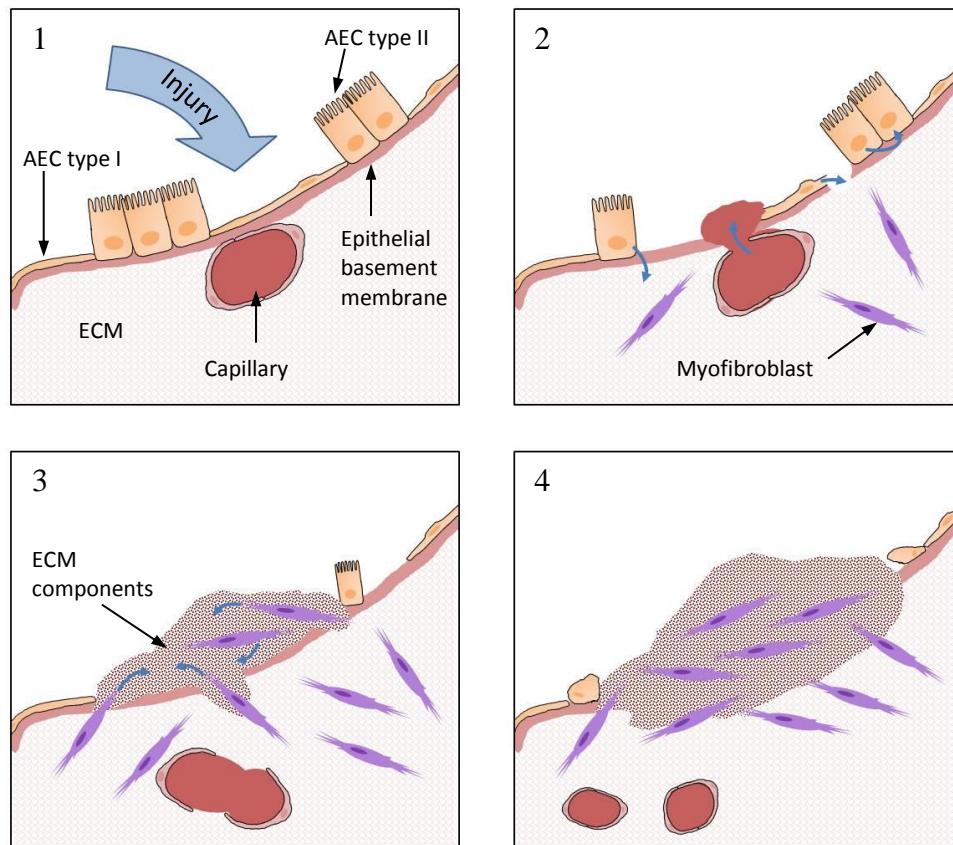
1 - Persistent injury leads to increase levels of alveolar epithelial cell (AEC) death; 2 - release of growth factors and chemokines from damaged AECs trigger stress responses in neighbouring cells and surrounding tissue. Increased basement membrane permeability allows leakage of plasma, containing activated coagulation proteases, into the site of injury triggering a fibrotic response; 3 – further disruption of the basement membrane allows myofibroblasts to infiltrate and deposit ECM components preventing re-epithelialisation; 4 – fibroblastic foci develop as dysregulation of the wound healing process leads to fibrosis.

Figure 3 Pulmonary Fibrosis[41]

A number of key aspects of IPF are implicated in contributing towards disease progression, namely increased myofibroblast cell population, increased levels of pro-inflammatory cytokines and chemokines, elevated numbers of cells initiating stress responses, and increased levels of proteases as a result of activation of the coagulation cascade. These are outlined in more detail in the following sections.

1.4.3.1 Increased Myofibroblast Population

Injury to AECs is understood to be a key process in the pathogenesis of IPF; hyperplasia of alveolar epithelial type 2 (AT2) cells within fibroblastic foci has been

reported together with increased apoptosis and denudation of the epithelial basement membrane in cases of IPF. As mentioned earlier, expression of gene mutations in AECs lead to cell instability and affects their ability to regenerate normal epithelium. As much as 70-80% of AT2 cells have been shown to be in stages of apoptosis in patients with IPF, which in mouse models has been linked to the development of lung fibrosis.[67, 68] Damage to AT2 cells induces fibroblasts resident in the interstitium to differentiate into myofibroblasts, the other key cell type in the progression of IPF.

Myofibroblasts are the major cell type responsible for the production and secretion of ECM in the interstitium. Currently, it is understood that there are three possible sources of myofibroblast cell replenishment at the site of injury, though the relative contribution of each to the cell population in IPF is unclear. The first, already mentioned, is the activation of resident lung fibroblasts by the potent pro-fibrotic mediator, TGF-β (section 1.2.2, page 6). Increased numbers of cells experiencing stress and undergoing apoptosis due to chronic injury in IPF, results in elevated levels of TGF-β, promoting more fibroblast differentiation. The second possible source is the transformation of epithelial cells, following stress/injury, to mesenchymal cells, which can then differentiate into myofibroblasts.[16] It has been demonstrated *in vitro* that TGF-β is able to induce this process, known as epithelial-mesenchymal transition (EMT), whereby epithelial cells lose their characteristic markers, undergo cytoskeletal reorganisation to adopt a spindle-shaped morphology and acquire typical mesenchymal markers.[69] Markers for both cell types have been shown to co-localise on hyperplastic AT2 cells in IPF lung tissue suggesting the occurrence of EMT to provide an additional source of myofibroblasts in lung fibrosis.[16] The hypothesis behind the third source of myofibroblasts is that they may develop from circulating fibrocytes or other bone marrow-derived progenitor cells, the numbers of which have been shown to correlate with disease progression in IPF.[55, 70]

1.4.3.2 Pro-inflammatory Mediators

Many of the various cytokines and chemokines involved in the complex interactions in response to tissue injury, when unregulated, have been implicated in promoting the generation of a pro-fibrotic environment contributing to the progression of the

disease. Tumour necrosis factor (TNF)-α plays a central role in many processes involved in an inflammatory response and mice deficient in this cytokine have been shown not to develop fibrosis following treatment with bleomycin – a glycoprotein antibiotic known to induce pulmonary fibrosis.[71] However, no beneficial effect was seen in IPF patients treated with TNF inhibitor Etanercept for 48 weeks.[72]

TGF-β is a pleiotropic cytokine produced by various cell types and is currently recognised as the most potent mediator of fibrosis in the lung. TGF-β mRNA expression has been shown to be up-regulated in the lungs of mice treated with bleomycin.[73] Fibrosis is promoted through increased recruitment and activation of monocytes and fibroblasts by elevated levels of TGF-β, which also increase ECM gene transcription and suppression of proteins involved in collagen degradation, such as matrix metalloproteinases, plasminogen activators and elastases.[74]

Despite the theory of inflammation as a driving force being discounted in IPF, a number of pro-inflammatory chemokines, including CCL2, CCL3, CCL17 and CCL22, have been shown to be up-regulated in the fibrotic lung, with benefits seen in their absence or when neutralised in animal models of fibrosis.[74-78] These chemokines are mainly responsible for the recruitment of monocytes and lymphocytes, activation of which can in turn release pro-fibrotic cytokines.[79]

1.4.3.3 Cell Stress Response

Chronic injury leads to sustained levels of cell stress, likely to result in fewer cells surviving the UPR and resulting in increased levels of apoptosis.[80] Oxidative stress may play a role in IPF progression as increased numbers of cells, such as neutrophils, generating ROSs outweigh the capacity to neutralise them. Lung myofibroblasts have also been shown to generate hydrogen peroxide, possibly leading to epithelial apoptosis.[81]

1.4.3.4 Activation of the Coagulation Cascade

Although its primary function is the formation of a blood clot to stem the loss of blood and prevent infection, the coagulation cascade is increasingly understood to have an influence in wound healing beyond haemostasis. Understanding of the roles that coagulation proteases, such as thrombin and FXa, have in the initiation of

cellular mechanisms has shed light on the wider influence the coagulation cascade has on the inflammation and tissue repair processes.

Thrombin's cellular responses are mediated by a specific family of cell-surface receptors known as protease-activated receptors (PARs), of which four members have been identified to date, PAR-1 to PAR-4. PAR-1 demonstrates high affinity for thrombin and has been shown to be expressed in numerous cell types present in the lung, including pulmonary epithelial cells, fibroblasts and macrophages. Activation of PAR-1 in these cell types leads to secretion of potent pro-inflammatory and pro-fibrotic mediators, and in fibroblasts promotes proliferation and differentiation into myofibroblasts.[82-87] Coagulation factor X has recently been reported to be up-regulated in patients with IPF and that bronchial and alveolar epithelial cells are major sources of this zymogen.[88] The fact that sources of coagulation proteases are found outside of the circulatory system strengthens the recognition of their importance in processes other than haemostasis.

Investigations into the participation of the coagulation cascade and its mediators in the progression of fibrosis have been carried out in mice using various strategies to block either the cascade itself or the activated proteases responsible for initiating cellular actions. Results have shown a reduction in lung collagen accumulation and fewer fibrotic lesions developing.[89] Mice deficient in the major thrombin receptor, PAR-1, also showed a large reduction in fibrosis, inflammatory cell recruitment and protection from experimental lung oedema.[82, 90]

Given these results, obtained by our collaborators in Professor Chambers' research group, there is a clear interest in further investigation into the possible role of PAR-1 in the progression of fibrotic lung diseases.

1.5 Protease-Activated Receptors

Since the 1970's it has been recognised that the role of thrombin extends beyond that of a key component central to the mechanisms of the coagulation cascade.[91] Proteolytic cleavage of various zymogens in the coagulation cascade by the serine protease thrombin not only converts fibrinogen to fibrin but also generates both positive and negative feedback mechanisms, resulting in regulation of the cascade

(Figure 2, page 4). Further to this role, thrombin's hormone-like behaviour in stimulating cellular actions on various cell types led to the belief in the existence of a thrombin receptor.

1.5.1 Discovery of PARs

It was not until the early 1990's that Vu and co-workers isolated cDNA encoding a functional human thrombin receptor, or PAR-1 as it is presently known.[92] Their work in determining the amino acid sequence of the receptor revealed it to be a new member of the guanosine nucleotide-binding protein coupled receptor (GPCR) superfamily, with an unusually extended amino-terminal extracellular chain.
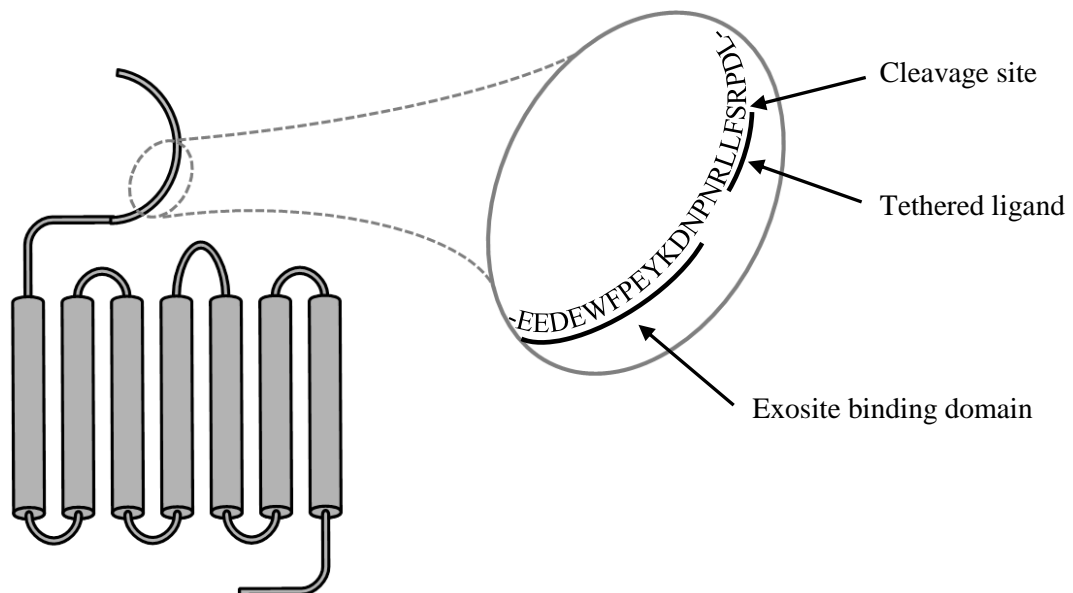

Figure 4 Protease-Activated Receptor-1.

Two observations were made regarding the sequence of amino acids in the extracellular N-terminus: i) the residue sequence leading up to the arginine at position 41 (LDPR^{41}S^{42}) mimicked that of the activation cleavage site in the zymogen protein C (LDPR/I, where "/" denotes cleavage site); ii) nearby to this possible cleavage site on the carboxyl side, a succession of acidic residues (D^{50}KYEPFWEDEE^{60}) resembled those found in hirudin - a potent anti-thrombin protein found in the saliva of leaches, which binds to the thrombin's anion-binding exosite, inhibiting protease activity and blood coagulation (Figure 4).[93, 94] A novel mode of activation was proposed whereby cleavage of the extracellular N-terminus

chain between residues R^{41} and S^{42} uncovers a sequence of amino acids, termed a "tethered ligand", which in turn binds to the receptor activation site. A synthesised peptide, representing the 14 amino acid chain uncovered following cleavage, was shown to be able to activate the receptor in the absence of thrombin. It was later demonstrated that the minimum length required for efficient activation was five residues (SFLLR-NH$_2$) and synthetic peptides containing this motif became known as thrombin receptor activating peptides (TRAPs).[95] Site-directed mutagenesis of the receptor, replacing arginine with alanine at the proposed cleavage site, rendered the receptor unresponsive to thrombin even at high concentrations (10 nM), yet activation was achieved when treated with TRAPs.[92] Thus, the novel mode of activation was accepted.

Since the discovery of the first thrombin sensitive receptor a further three receptors bearing a cleavage site in their extended extracellular N-terminus chain have been isolated. In order of their discovery members of this sub-group of GPCRs were named PAR-1 to -4.[96-99] Some interesting differences and associations between the PARs have been documented over the years since their uncovering. All PARs reveal a unique sequence of amino acids following proteolytic cleavage: PAR-1 - S^{42}FLLRN; PAR-2 – S^{35}LIGKV; PAR-3 – T^{39}FRGAP; and PAR-4 – G^{48}YPGQV. Thrombin is the major, though not exclusive, activating protease for PAR-1, -3 & -4, whereas PAR-2 is activated by trypsin in most cases, a serine protease generated in the pancreas, which is itself activated by enteropeptidase. Interestingly, although PAR-3 demonstrates high affinity for thrombin, due to the presence of a hirudin-like sequence on the extracellular N-terminus chain, cleavage fails to initiate cell signalling pathways. Yet PAR-3 has been shown to play a vital role in the activation of mouse platelets, which, unlike human platelets, do not express PAR-1.[100] Nakanishi-Matsui and co-workers revealed that in the absence of PAR-1, a PAR-3/PAR-4 affiliation operates whereby PAR-3 acts as a cofactor for PAR-4 activation, which lacks the hirudin-like sequence in its extracellular N-terminus chain. They propose that following cleavage of the PAR-3 extracellular chain, the thrombin active site is made available for cleavage of the neighbouring PAR-4 extracellular chain, which uncovers the tethered ligand leading to activation of the receptor and initiation of cell signalling pathways. PAR-3 has also been

suggested to play a regulatory role with PAR-1 by forming heterodimers and altering the extent to which cell signalling pathways are initiated.[101]

1.5.2 PAR-1 Signalling

In recent years, the mechanisms involved in PAR-1 activation and subsequent signal transduction have been revealed to be of much greater complexity than initially anticipated. As with other GPCRs, PAR-1 can initiate a variety of cell signalling pathways by affiliation with different G-proteins, as well as the G-protein-independent β-arrestin-mediated pathways - this is known as functional selectivity or biased agonism (Figure 5).[102-104] For PAR-1, the pathway is dependent upon the nature of the activating protease and, in some cases, the locality of the receptor on the cell surface. Although thrombin is the major activating protease associated with PAR-1, other proteases are also capable of cleaving the extracellular N-terminus chain. Cleavage can occur at various sites, some of which result in the removal of the SFLLR sequence, effectively deactivating the receptor in relation to thrombin activation.[105] Alternatively, unmasking a different sequence of amino acids, *i.e.* a different ligand, can lead to activation and preferential signalling *via* alternative pathways.[102]

Conventional activation of PAR-1 on endothelial cells by thrombin leads to signalling mediated by the small GTPase RhoA, which induces pro-inflammatory responses and increases endothelial barrier permeability. However, compartmentalisation of PAR-1 into micro-domains, called caveolae, on the epithelial cell surface can lead to activated protein C (APC) triggered activation, resulting in Rac1 mediated signalling.[106] Different cellular actions result, though mediated by the same receptor, as APC activation is anti-inflammatory and stabilises the endothelial barrier.[107]
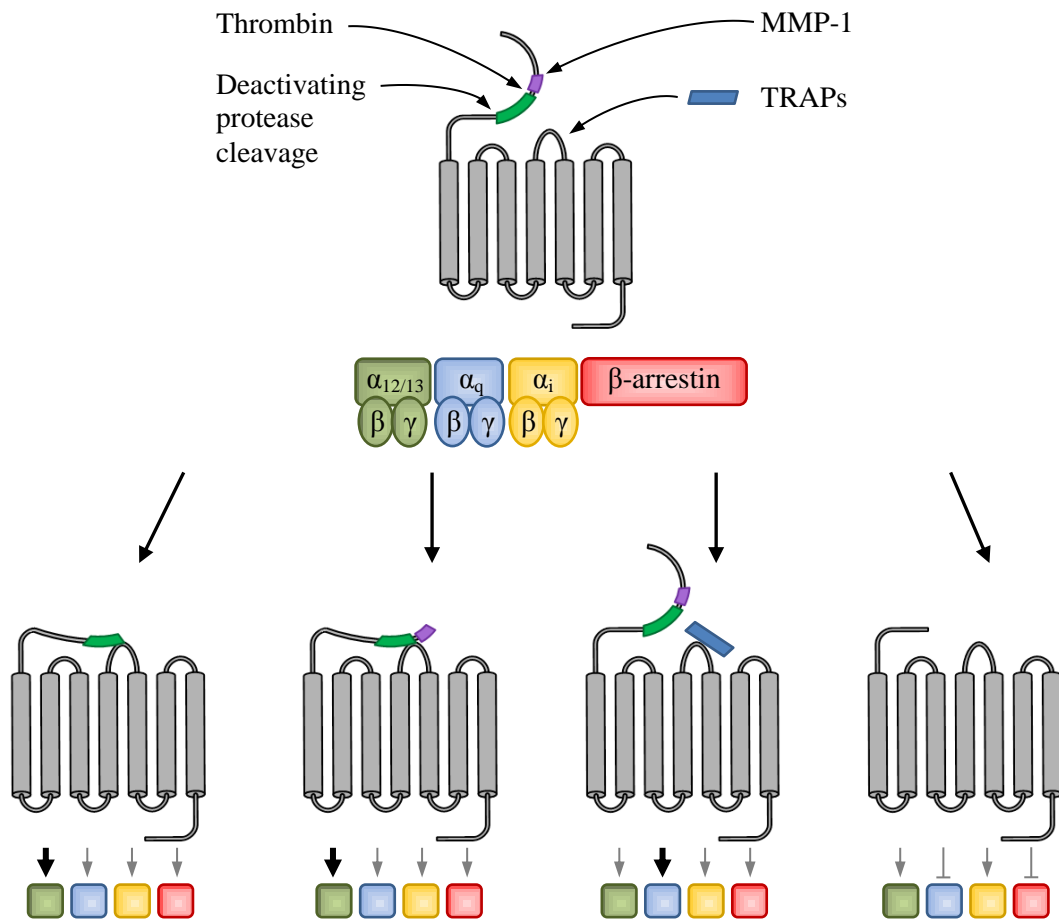
Cleavage by thrombin or MMP-1 favours signalling *via* $G\alpha_{12/13}$ whereas activation by TRAPs preferentially leads $G\alpha_q$ signalling. Cleavage by deactivating proteases may arrest conventional signalling pathways whilst allowing alternative pathways to be triggered by different ligands.

Figure 5 Functional Selectivity/Biased Agonism in PAR-1.[102]

Studies investigating the roles of FXa (Figure 2, page 4) have demonstrated that, as part of the TF-FVIIa-FXa complex, this protease could affect cell signalling in endothelial cells *via* cleavage and activation of PAR-1.[108] Similarly, matrix metalloprotease (MMP)-1 has been shown to activate PAR-1 and, as with thrombin, induce signalling *via* the RhoA pathway; the cleavage site was deduced as being two residues towards the N-terminus from the thrombin cleavage site, to afford P^{40}RSFLLR as the activating sequence.[109] Human tissue kallikreins have also been shown to cleave the N-terminus of PAR-1 leading to activation and Ca^{2+} mobilisation in human embryonic kidney (HEK) cells.[110] McLaughlin and co-workers demonstrated that PAR-1 activation by TRAPs led to preferential signalling *via* $G\alpha_q$, leading to Ca^{2+} mobilisation, rather than $G\alpha_{12/13}$ as favoured following thrombin activation.[111] This bias in signalling pathway is significant since both thrombin and TRAPs are used in biological assays investigating PAR-1 activity.

Conventionally, GPCRs are activated by the binding of an agonist released from a source in response to a stimulus. The docking of the ligand is reversible and the receptor alternates between an activated and deactivated conformation, with the degree of activation dependent upon the concentration of the agonist. With higher concentrations of agonist, *e.g.* resulting from persistent stimulus, receptor occupancy is high and desensitisation occurs. To ensure responsiveness to new stimuli, intracellular events lead to endocytosis of the receptor, which is then recycled to the cell membrane ready to function again, or directed towards lysosomes for degradation.

A consequence of the novel mechanism of activation observed in PARs is that the activating ligand is tethered to the receptor and once exposed affects what is essentially an irreversible interaction with the activation site. To overcome the issue of desensitisation, following activation PAR-1 is rapidly phosphorylated by several GPCR kinases (GRKs). Deubiquitylation of the activated receptor then leads to clathrin- and dynamin-dependent endocytosis and targeting to the relevant lysosomes for degradation.[112] Re-sensitisation is achieved through mobilisation of intracellular stores of PARs to the cell membrane, while up-regulation of PAR-1 expression replenishes these stores.[105]

1.5.3 PARs in other diseases

PARs have been shown to be expressed by a variety of cell types and are found in almost all tissues throughout the body. It is therefore understandable that investigations have uncovered many roles for PARs in normal and various diseased states of the body.

1.5.3.1 Cardiovascular Disease

Based on the importance of thrombin in haemostasis, a major role in the cardiovascular system for a thrombin receptor was predicted decades before the discovery and isolation of PAR-1. Subsequent findings of elevated levels of expression of this receptor in platelets and endothelial cells supported the prediction. All PARs are now recognised to be expressed throughout the vasculature and have

roles in the regulation of platelet functions, endothelial permeability, leukocyte adhesion, nitric oxide release and the formation of atherosclerotic plaques.[102]

Activation of PAR-1 and PAR-2 on the endothelium triggers nitric oxide-mediated vasodilation in undamaged endothelium.[113] However, when the endothelium is disrupted, PAR-1 and PAR-2 activation leads to vasoconstriction.[114] Up-regulation of PAR-1 expression has been implicated in vascular vasospasm following subarachnoid haemorrhage, which is a major contributor to post-stroke morbidity and mortality.[115] PAR-1 inhibition has been shown to attenuate *in vitro* human platelet aggregation and thrombosis in animal models, suggesting drugs targeting this receptor in the treatment of thrombosis may block platelet plug formation without causing bleeding diathesis.[116] Although discovery of the ability of PAR-4 to stimulate human platelet aggregation at higher concentrations of thrombin suggests antagonising both PAR-1 and PAR-4 may be required to be an effective treatment for thrombosis.

1.5.3.2 Musculoskeletal Disorders

Vergnolle and co-workers established a role for PAR-1 in peripheral inflammation when in a rat paw oedema model the receptor was shown to mediate a pro-inflammatory response induced by thrombin activation.[117] Evidence since this has indicated a significant role for PARs in inflammation, with both pro- and anti-inflammatory, as well as nociception, responses to activation by different proteases being observed. Moreover, diminished inflammation was observed in PAR-deficient animals.[105, 118] PAR-2 is implicated in joint inflammation and pain, with symptoms nullified in a rodent model of acute joint inflammation using PAR-2 deficient mice.[119]

1.5.3.3 Gastrointestinal Disease

All PARs are expressed on various cell types throughout the gastrointestinal (GI) tract and regulate ion transport, mucosal permeability, epithelial cell motility, immune cell recruitment and nociception.[102] Activation of epithelial PAR-1 and PAR-2 can stimulate chloride secretion, which leads to diarrhoea, a common symptom in conditions such as inflammatory bowel disease. Epithelial cell apoptosis

triggered by PAR-1 activation increases intestinal permeability, as does PAR-2 through disruption of epithelial cell organisation at tight junctions. Blocking signal transduction mediated by these receptors could reduce infiltration by pro-inflammatory mediators and combat GI inflammation in conditions such as colitis.[120, 121] However, some reports suggest PAR-1 also exhibits an anti-inflammatory property and may protect against infection in the gut.[122] PAR-1 and PAR-2 have roles to play in both contraction and relaxation of the muscle tissue in the GI tract, which is responsible for motility, disruption of which leads to increased GI transit time.[123, 124]

1.5.3.4 Respiratory Disease

As well as the aforementioned role in fibrosis, which is the main concern of this research, PARs are also believed to have an influence on allergen-induced asthma. Neutralising the PAR-2 activating ability of an allergen led to a reduction in eosinophil recruitment in a murine model of asthma.[125] However, there is some evidence that suggests PAR-2 activation also has an anti-inflammatory role in the airways.[126] This suggests that targeting this receptor with regards to the treatment of asthma may not be as straight forward as it might seem.

1.5.3.5 Central Nervous System (CNS) Disorders

PARs and their activating proteases are expressed on neuronal and non-neuronal cells of the CNS and research suggests a variety of roles in both normal and diseased states. PAR-1 demonstrates bimodal effects dependent upon the concentration of activating agents, with basal activation providing protection against damaging stimuli, whereas high levels of activation can have deleterious effects on neurons.[127, 128] In low thrombin concentration, PAR-1 activation has been shown to reduce ceramide-induced astrocyte death, and activation by APC rather than thrombin may promote neurogenesis and neovascularisation in post-ischaemic brain.[129, 130] However, studies suggest PAR-1 predominantly has a neurodegenerative and/or pro-inflammatory role in response to trauma to the brain. PAR-1 deficiency in mice or treatment with a PAR-1 antagonist shows a beneficial effect in models for stroke, brain haemorrhage and neurotrauma, and increased PAR-1 expression in astrocytes has been discovered in patients with HIV-associated dementia.[131, 132] On the other

hand, PAR-2 activation has been linked to neuronal excitability, synaptic transmission and plasticity, and demonstrates equally both protective and damaging roles.[102]

1.5.3.6 Cancer

A number of proteases capable of activating PAR-1, including APC, the FVIIa-Xa complex, MMP1 and kallikreins, have been detected in tumours. [133-135] Together with the known PAR-mediated mitogenic actions of thrombin and trypsin, a possible role for this receptor in tumour growth has been proposed. Tumour tissues from prostate and colon cancer, as well as malignant melanoma, have revealed upregulation in the expression of PARs and possible activating proteases.[136-138] PAR-1 has been linked with promoting tumour cell invasion and metastasis, with the level of receptor expression correlating directly with metastatic potential.[139, 140]
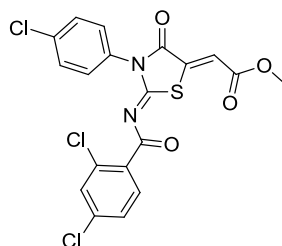
1.6 PAR-1 Antagonist Development

Given the broad spectrum of diseases in which PAR-1 has been implicated as having a role in pathogenesis, development of therapeutics that work by interacting with this receptor is an attractive, if somewhat challenging, target. As a consequence of its unusual mode of activation many different approaches to the inhibition of PAR-1 signalling have been employed. Attention has been mainly focussed on: i) interfering between the activating protease and the extracellular N-terminal domain with antibodies targeting the cleavage site; ii) cell penetrating molecules targeting and disrupting the interaction of the receptor with signal transducing proteins; and iii) small molecule attachment to the binding site blocking its interaction with the tethered ligand. The latter of these approaches has yielded the most success to date and will be discussed in the following sections. Work in this area has adopted two distinct approaches towards the development of active compounds: i) optimisation of non-peptidic compounds discovered through target based high throughput screening (HTS); and ii) rationally designed compounds derived from SAR data obtained for PAR-1 activating peptides.

Due to the functional selectivity demonstrated by PAR-1, assays investigating the activity of new compounds on the receptor are routinely tested against both thrombin

and TRAPs containing the peptide sequence TFLLR, which has been shown to selectively activate PAR-1 over other PARs.[141] Doing so ensures identification of compounds active against PAR-1 signalling *via* $G\alpha_q$, $G\alpha_{12/13}$ or both.

1.6.1 Nonpeptidic PAR-1 Antagonists

1.6.1.1 FR171113

1 $IC_{50} = 0.29$ μM

Figure 6 PAR-1 antagonist FR17113

The first non-peptidic PAR-1 antagonist was developed at Fujisawa Pharmaceutical through optimisation of an undisclosed compound identified by HTS. *In vitro* testing of the optimised compound, FR171113 **1** (Figure 6), in a human washed platelet aggregation assay demonstrated IC_{50} values of 0.15 μM and 0.29 μM versus TRAP-6 (1 μM) and thrombin (0.1 NIH unit/mL or 0.9 nM), respectively.[142] FR171113 was also shown to inhibit $FeCl_3$-induced carotid artery thrombosis in a dose-dependent manner in *in vivo* studies of arterial thrombosis in guinea pigs.[143] The compound is currently commercially available as a biological tool for investigating the role of PAR-1 in various diseases.

1.6.1.2 Tri-substituted Ureas

Barrow and co-workers at Merck research laboratories reported the identification of a potent, non-peptidic PAR-1 antagonist **2** (Figure 7) following the screening of combinatorial libraries of compounds.[144] This tri-substituted urea demonstrated sub-micromolar activity in a human washed platelet activation assay in competition with TRAP-5 (3 μM) as agonist.
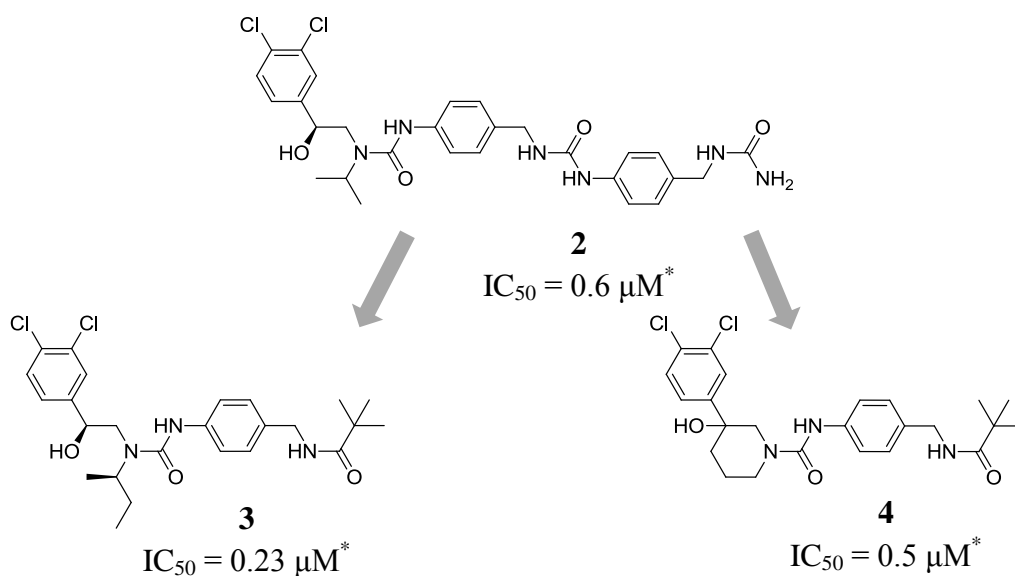
Figure 7 Tri-substituted ureas

Optimisation of the initial hit included replacement of the central urea in compound **2** with a pivaloyl protected amine. The shortened analogues gave compounds of similar potency but improved solubility and reduced molecular weight. However, loss of activity occurred in the absence of the amide. Other key residues, leading to significant reductions in potency in their absence, were identified as the 3,4-dichlorophenyl ring, the secondary OH and the urea NH. A slight increase in bulk of the urea *N*-alkyl substituent led to the development of analogues with reduced IC$_{50}$ values, the *N-sec*-butyl analogue **3** (Figure 7) proving to be the most potent with an IC$_{50}$ = 0.23 μM[*]. However, room for development here proved to be limited since the *N*-cyclopentyl analogue turned out to be inactive. The introduction of rigidity into a bioactive molecule is often favourable in order to hold pharmacophoric moieties in optimal spatial arrangements. In this case, incorporating the tertiary amine of the urea into a six-membered ring also gave a potent compound, **4** (IC$_{50}$ = 0.5 μM[*]) (Figure 7).

1.6.1.3 Isoxazoles

Also developed within the Merck laboratories were a series of PAR-1 active isoxazoles, which were discovered following directed screening based on tri-substituted urea **5** (Figure 8).

---

[*] Platelet aggregation assay *versus* TRAP-5 (3 μM)
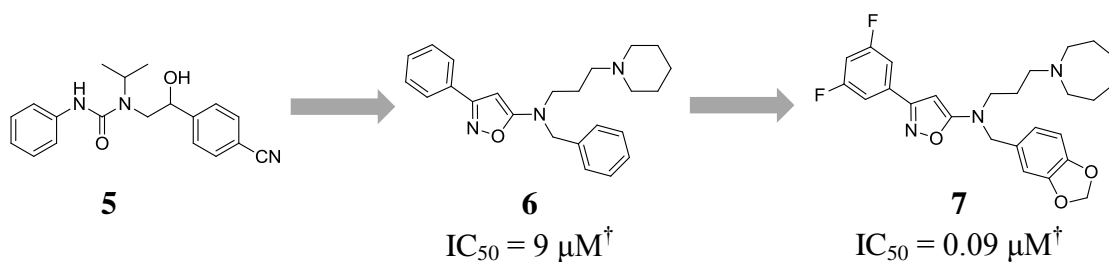
Figure 8 Development of isoxazole analogues

Isoxazole **6** (Figure 8), represented a structurally novel compound with moderate potency (IC_{50} = 9 μM[†]), which improved significantly following optimisation.[145] Modifications involved substitution on the *N*-benzyl substituent, expansion of the piperidine ring and substitution on the 3-aryl substituent. Optimisation resulted in compound **7** (Figure 8) which had IC_{50} values of 0.09 μM and 0.51 μM in a human washed platelet activation assay versus TRAP-5 (3 μM) and thrombin (1 nM) as agonists, respectively.
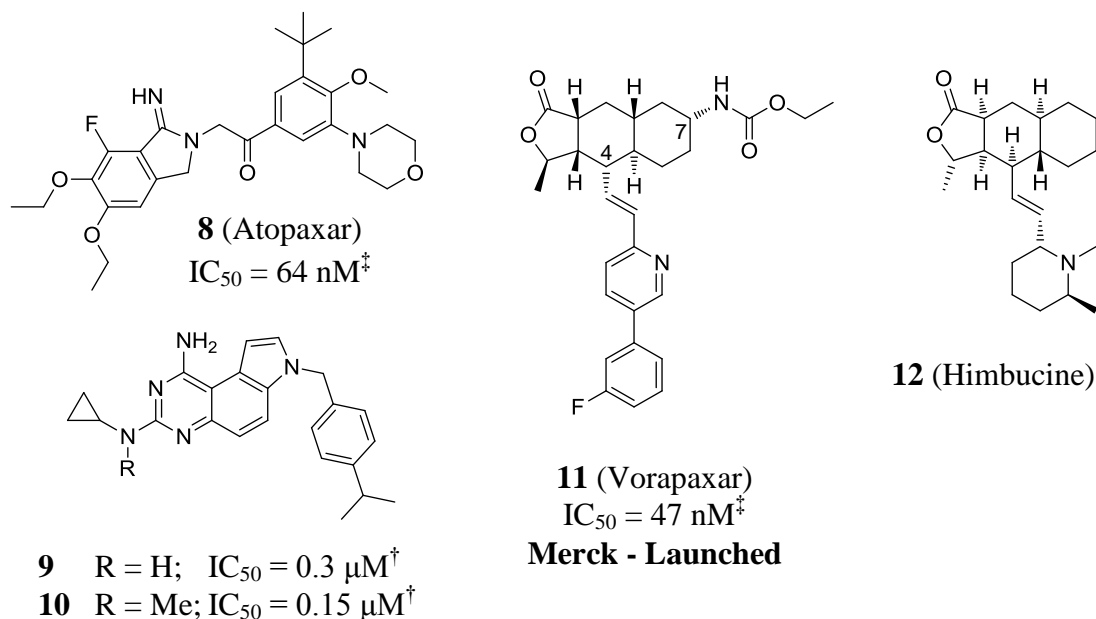
1.6.1.4 E5555 (Atopaxar)

Figure 9 Small molecule PAR-1 antagonists

Atopaxar **8** (Figure 9) is a potent small molecule PAR-1 antagonist developed by Eisai, which demonstrates a high degree of affinity for the receptor. In a thrombosis model in guinea pigs, Atopaxar exhibited an anti-thrombotic effect and was also able

---

[†] Platelet aggregation assay *versus* TRAP-5 (3 μM)
[‡] Platelet aggregation assay *versus* thrombin

to inhibit neo-intimal hyperplasia in a rat balloon injury model.[146] Following very encouraging results in Phase I studies, the compound was entered into two Phase II clinical trials where it achieved a significant reduction in heart problems compared with the control group.[147] The clinical trials showed that although increased bleeding was seen as the dose of Atopaxar was increased, the risk of serious bleeding was not significantly higher. However, based on the number of test subjects experiencing dose-dependent hepatic dysfunction and QT prolongation, these trials had to be terminated on safety grounds; Eisai decided to discontinue the development of Atopaxar in 2012.

1.6.1.5 SCH79797

A series of pyrroloquinazolines, reported by Schering-Plough to be active small non-peptidic PAR-1 antagonists, were discovered by HTS of the company's chemical compound library.[148] The initial hit required minor optimisation to yield two potent compounds, SCH79797 **9** (Figure 9) and SCH203099 **10** (Figure 9) with IC_{50} values of 0.3 μM and 0.15 μM, respectively, in a TRAP-induced platelet aggregation assay.

## 1.6.1.6 SCH530348 (Vorapaxar)

Also developed by Schering-Plough scientists initially, and later by Merck, is Vorapaxar **11** (Figure 9), which is currently the most potent PAR-1 antagonist known. The original hit in the developmental series for this compound was the natural product himbacine **12** (Figure 9), which was actually of interest as part of a research program into Alzheimer's disease, due to its activity at muscarinic receptors.[149] One of the analogues in the developmental process was identified as having PAR-1 antagonist activity and was progressed further in this alternative direction. Early in the optimisation process the enantiomer of himbacine was found to be greater than 10-fold more potent than the naturally occurring isomer.[150] Further developments involved variation in substitution at the 4-position of the tricycle, before later incorporation of the ethyl carbamate at the 7-position to tackle clearance issues with the methylene precursor.[151] Vorapaxar demonstrated exceptional potency in *ex vivo* human platelet aggregation assays induced by TRAP (15 μM) and thrombin (10 nM) with IC_{50} values of 25 nM and 47 nM, respectively.

In 2011, Vorapaxar was assessed in two separate Phase III clinical trials for the treatment of myocardial infarction (MI) in patients with atherosclerosis (trial name: TRA2°P-TIMI 50) and patients with acute coronary syndrome (trial name: TRA-CER). Results from the TRA2°P-TIMI 50 trial found that Vorapaxar did not reduce the risk of cardiovascular death, myocardial infarction, or stroke in patients with peripheral artery disease.[152] However, the compound was found to significantly reduce acute limb ischemia and peripheral revascularisation but an increased risk of bleeding accompanied these beneficial effects, which resulted in the termination of the trial. Initial results from TRA-CER indicated that Vorapaxar was associated with a reduction of MI, however, the trial was halted as a result of the findings in the TRA2°P-TIMI 50 study.[153]

Vorapaxar underwent review by the U.S. Food and Drug Administration for approval to be used for the secondary prevention of cardiovascular events in a sub-population of patients with a history of myocardial infarction or peripheral arterial disease but no history of stroke or transient ischemic attack (TIA). In May 2014, Vorapaxar was approved as the first-in-class PAR-1 antagonist for the treatment of patients with a history of health corresponding to that described above and is the API in Zontivity[TM].

1.6.2 Peptidomimetic PAR-1 Antagonists

Selective partial agonists and antagonists of the thrombin receptor, with low potency, had been reported even before the isolation and cloning of PAR-1.[154, 155] However, following elucidation of the amino acid sequence of the receptor, and in particular the discovery of the tethered ligand sequence SFLLR-NH$_2$ **13** (Figure 10), sharp progress in the development of antagonists ensued.[92] The later discovery that all PARs had unique tethered ligand sequences gave reassurance that this approach may lead to selective PAR-1 active compounds.
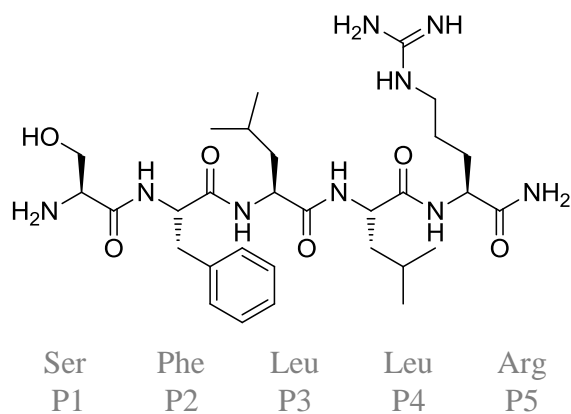
13 $EC_{50} = 6.0\ \mu M$

Figure 10 PAR-1 activating pentapeptide sequence (TRAP-5)

A number of SAR studies of the TRAPs determined the minimal structural requirements for PAR-1 agonism to be attained. These were reported to be an N-terminal pentapeptide with an aromatic residue at position 2 (P2), a basic residue at P5 and a large hydrophobic residue at P4.[95, 156] A small side chain and free amine at P1 are also important, but substitution at P3 is widely tolerated. Investigations into the backbone conformation and *N*-methylation of TRAPs were carried out by Ceruso and co-workers suggesting that an extended conformation is required for receptor recognition, as well as the amide nitrogen between residues 1 and 2.[157] Carbonyl groups along the backbone were suggested to be involved in hydrogen bonding with the receptor.

The proposition that PAR-1 was activated by the docking of the tethered ligand on the extracellular loops of the receptor was generally accepted; however, information regarding the actual binding site and specific interactions with the ligand was scarce. GPCRs traverse the cell membrane and isolation of this type of receptor in order to obtain crystal structures is notoriously difficult due to their instability when not supported by the cell membrane. The only example of a resolved structure for a 7-transmembrane (7TM) receptor at the time of development of early peptidomimetic PAR-1 active compounds was bacteriorhodopsin. Homology modelling of GPCRs based on the structure of bacteriorhodopsin was generally unreliable due to low sequence identity (*i.e.* similarities in amino acid sequence). Beavers and co-workers, however, reported a novel approach to building a model of PAR-1, which also incorporated a mutagenesis study of the receptor. The results of

this study highlighted important residues required for the effective activation of PAR-1.[158] The extracellular loops (ECs) were noted as containing many acidic residues and individual mutagenesis of these with alanine identified D^{256} on EC2 and E^{347} on EC3 as having a major influence on ligand binding and receptor activation. Replacement of E^{347} with serine also led to loss of function, implying the requirement of an acidic residue rather than simply a hydrogen bond donor, and thus a likely electrostatic interaction with the ligand at this position. Loss of the di-sulfide bridge between cysteine residues C^{175} and C^{254} resulted in moderate loss of activity suggesting a role in receptor conformation.

During the current research the stabilised crystal structure of PAR-1 bound with the potent antagonist Vorapaxar was reported, which may assist with future developments towards PAR-1 antagonists.[159]

1.6.2.1 BMS-200261

14
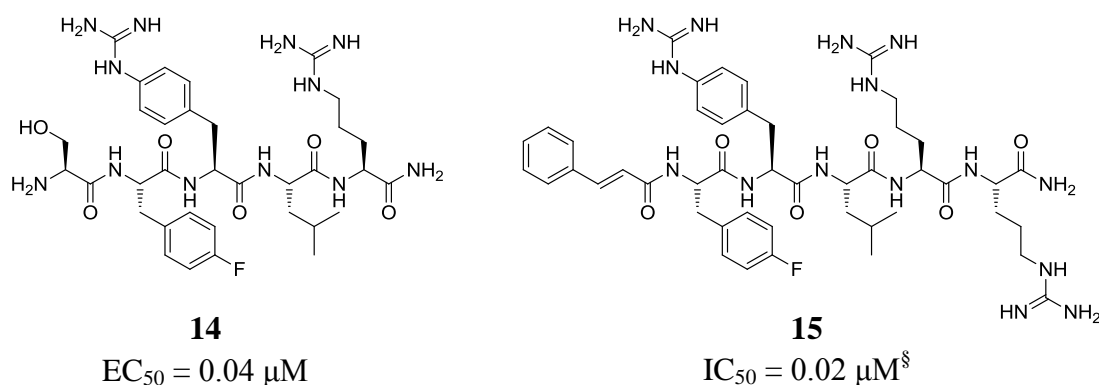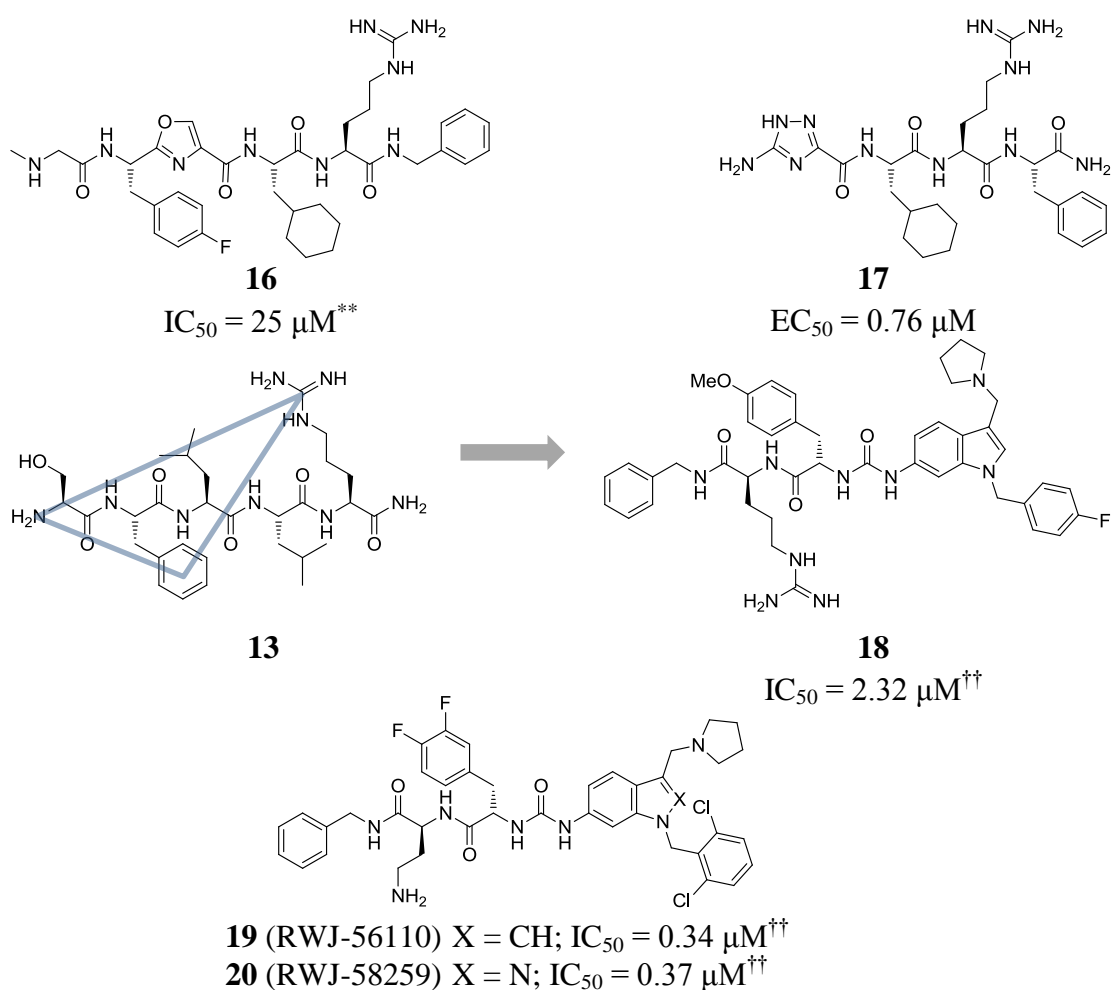$EC_{50} = 0.04\ \mu M$

15
$IC_{50} = 0.02\ \mu M^{\S}$

Figure 11 PAR-1 active synthetic peptides

Systematic substitution, by Bernatowicz and co-workers at Bristol-Myers Squibb, of the second and third residues in a pentapeptide analogue of TRAP-5 led to agonists with enhanced potency – S(*p*-fluoro)F(*p*-guanidino)FLR-NH$_2$ **14** (Figure 11) proved to be ~10-fold more potent than TRAP-5.[160] This was followed by substitution of the serine residue leading to the development of partial agonists initially, and eventually producing full antagonists. BMS-200261 **15** (Figure 11) is still one of the most potent PAR-1 antagonists for the inhibition of TRAP-7-induced platelet aggregation, with an IC_{50} ~20 nM. However, this compound was unable to inhibit platelet aggregation induced by thrombin.

---

§ Platelet aggregation assay *versus* TRAP-7 (3 μM)

1.6.2.2 RWJ-56110

Expanding on the SAR work done previously on TRAPs, Hoekstra and co-workers looked to introduce a central constraint on the compounds by incorporating a heterocycle at the 2/3 linker position.[161] Results showed that introducing an oxazole ring to yield compound **16** (Figure 12) was well tolerated but fell short of enhancing antagonist potency. McComsey and co-workers also looked at introducing rigidity to the agonist backbone, by replacing serine with heterocycles to make heterocycle-peptide hybrids such as compound **17** (Figure 12), resulting in potent PAR-1 agonists and partial agonists.[162]

16
$IC_{50} = 25 \ \mu M^{**}$

17
$EC_{50} = 0.76 \ \mu M$

13

18
$IC_{50} = 2.32 \ \mu M^{\dagger\dagger}$

19 (RWJ-56110) X = CH; $IC_{50} = 0.34 \ \mu M^{\dagger\dagger}$
20 (RWJ-58259) X = N; $IC_{50} = 0.37 \ \mu M^{\dagger\dagger}$

The triangle over compound **13** illustrates the three-point model devised by Andrade-Gordon and co-workers in their development of PAR-1 antagonists.

Figure 12 Peptidomimetic PAR-1 active compounds

[**] Platelet aggregation assay *versus* thrombin (0.05-0.5 nM)
[††] Platelet aggregation assay *versus* thrombin (0.15 nM)

A *de novo* approach was employed by Andrade-Gordon and co-workers, which considered the SAR data for PAR-1 agonist peptide epitopes developed to date, the spatial constraints in the TRAP-5 ligand and the desirable integration of a rigid heterocyclic template in order to display key functional groups in the correct spatial positions.[163] A three-point model was proposed giving ideal distance ranges between the amine group (P1), the centre of the aromatic ring (P2) and the central carbon atom of the arginine residue (P5) (**13**, Figure 12).

Various aromatic rings and heterocycles were investigated before incorporation of an indole framework led to an early lead compound with RWJ-53052 **18** (Figure 12). An SAR study of this compound resulted in the development of potent PAR-1 antagonist RWJ-56110 **19** (Figure 12), demonstrating IC_{50} values of 0.16 μM and 0.34 μM in a TRAP-6 (2 μM) and thrombin (0.15 nM) induced platelet aggregation assay, respectively.

However, in a guinea pig model of *ex vivo* platelet aggregation to determine the ability of RWJ-56110 to inhibit the actions of thrombin on platelets, i.v. administration of the compound was discovered to cause severe hypertension.[164]

1.7 RWJ-58259

Investigations into the hypertension witnessed in the *ex vivo* study involving RWJ-56110 and analogues suggested that the size of the substituent at the 2-position of the indole may contribute towards the unwanted cardiovascular side-effects. Exchanging the indole framework for indazole meant that the 2-position was occupied by a lone pair of electrons only, which led to the development of the potent and selective PAR-1 antagonist RWJ-58259 **20** (Figure **12**). Not only did this compound avoid the hypertensive issues seen with previous analogues, but it also demonstrated an increased binding affinity for PAR-1 whilst maintaining potency against TRAP-6 (2 μM) and thrombin (0.15 nM) in the platelet aggregation assay with IC_{50} = 0.11 μM and 0.37 μM, respectively.[164] Further RWJ-58259 *in vitro* studies have also shown its ability to inhibit thrombin-induced Ca^{2+} mobilisation in human and rat aortic smooth muscle cells (SMCs) (IC_{50} = 0.14 μM and 0.07 μM, respectively), cell proliferation in rat vascular SMCs (IC_{50} = 2.3 μM) and interleukin (IL)-6 release in human vascular SMCs (IC_{50} = 3.6 μM).[165]

Ex vivo studies have demonstrated that RWJ-58259 inhibits low concentration thrombin-induced guinea pig platelet aggregation in a dose-dependent manner following i.v. administration at 0.3, 1.0 and 3.0 mg/kg. At high thrombin concentration, RWJ-58259 failed to inhibit platelet aggregation due to the ability of PAR-4, also expressed on guinea pig platelets, to activate this cell type when exposed to higher levels of thrombin. RWJ-58259 also failed to reduce the weight of thrombus and time to occlusion in two different *in vivo* models in guinea pigs, which was reasoned to be due to sufficient presence of thrombin *in vivo* to enable activation of PAR-4 *via* the PAR-3/PAR-4 heterodimer. However, in cynamolgus monkeys where platelets do not express PAR-3, as in humans, time to occlusion was significantly increased in an electrolytic injury-induced carotid thrombosis model.[166] Carotid arteries from three out of five RWJ-58259-treated animals failed to occlude within the observation period (60 min) and the thrombi that did form were found to have a reduced platelet count. Also, RWJ-58259 was found not to affect the numbers of white/red blood cells and platelets in circulation, nor did it affect coagulation, with respect to activated clotting time. These findings are all consistent with the selectivity of RWJ-58259 for PAR-1 mediated cellular responses to thrombin.

A rat model of balloon-injury induced carotid stenosis investigated the use of RWJ-58259 to inhibit PAR-1 mediated effects of thrombin on vascular SMCs during injury.[165] Normal responses include cell proliferation, ECM synthesis, inflammatory cytokine release and changes in vascular tone. Perivascular administration of RWJ-58259 resulted in a reduction in intimal thickness and area; these results can be deemed to be independent of platelet action since rat platelets do not express PAR-1.

1.8 Research Objectives

As a result of the proven *in vivo* selectivity and potency of RWJ-58259 for PAR-1 antagonism, this compound was selected by our collaborators in the Chambers research group as a tool compound to investigate the role of PAR-1 in the pathogenesis of IPF. However, some concerns were raised over the bioavailability of the compound following the report of studies in rats having been i.v. administered a dose of RWJ-58259 measuring a half-life ($t_{1/2}$) of 0.32 h.[166]

A short $t_\frac{1}{2}$ would not be a major issue for a drug in the treatment of pulmonary fibrosis since administration as an inhaled formulation would allow direct access to the site where treatment would be required. This would avoid the need for the drug to survive first-pass metabolism. In fact, given the broad distribution of PAR-1 throughout the body and its possible roles in various diseases, a short $t_\frac{1}{2}$ may be beneficial for a drug acting as a PAR-1 antagonist taken as an inhaled formulation. Side effects associated with off target PAR-1 inhibition may be minimised should the drug be rapidly metabolised once in circulation.

Concerns do arise, however, when considering the use of RWJ-58259 as a research tool compound in animal models of pulmonary fibrosis. Although administering the drug as an aspirated formulation to allow inhalation by the animal is possible, inconsistencies in the amount of drug actually taken in by each animal has led to a preference for intraperitoneal administration. As such, systemic distribution is required in order to deliver the compound to the site of interest, which means such a short $t_\frac{1}{2}$ would be an unfavourable characteristic.

The main objective of this research was to develop analogues of the project lead compound RWJ-58259, which maintained its potency as a PAR-1 antagonist whilst improving its metabolic stability. This would involve expanding the currently limited drug metabolism/pharmacokinetic (DM/PK) profile of RWJ-58259 by investigating the major pathways responsible for transforming the parent compound into its metabolites. Identifying properties of the metabolites formed would allow for a better understanding of this compound's activity in a biological setting, which may assist in deciphering mechanisms involved in the progression of fibrosis. Guided by knowledge of the metabolic pathways involved in the transformation of RWJ-58259, analogues would be synthesised aiming to improve metabolic stability. These compounds would then be used as tool compounds in future investigations into the pathogenesis of pulmonary fibrosis carried out by our collaborators in Prof Chambers' research group.

1.9 Review of RWJ-series SAR data

During the optimisation process of the initial hit RWJ-53052 **18** (Figure 12, page 33) a large amount of SAR data was generated by the investigators at Johnson & Johnson

for compounds synthesised and tested in a platelet aggregation assay. Following the discovery of RWJ-58259, this information was made available through publication of patents and literature articles covering their discoveries. With the intention of synthesising analogues of RWJ-58259 in mind, a review of this SAR data was important in order to guide any decisions made on proposed synthetic targets in the aim of maintaining PAR-1 antagonist activity.

Biological activity data for 256 different compounds in the platelet aggregation assay was published allowing for important deductions to be made regarding the effect of structural variations in different regions of the compound on PAR-1 activity. Figure 13 illustrates the balance of variations and the following sections will summarise the main findings for structural variation in the different regions of the molecule.
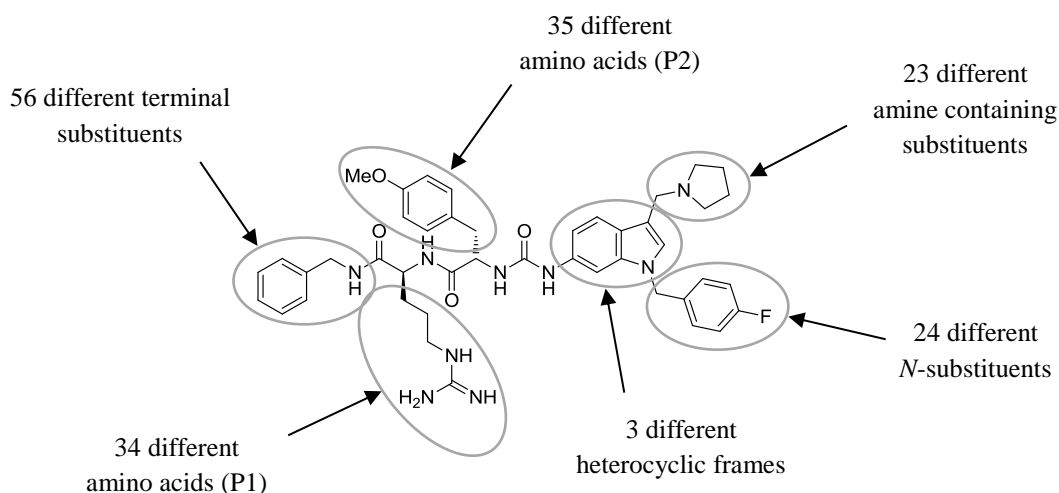

Figure 13 Balance of structural variation seen from published SAR data

1.9.1 Terminal substituent variation

Investigation of the left-hand side (as drawn) terminal substituent saw the greatest variation of all the positions highlighted, with 56 structurally different substituents being tested. Though no significant improvement in IC_{50} values was demonstrated by analogues over the original benzylamine, results suggested that slight enhancement in binding affinity was achieved for compounds with a basic moiety at the *para* position or at the end of an alkyl chain of similar distance. Interestingly, it was also recognised that such compounds with a basic moiety at the terminal position showed

a greater tolerance for variation at the neighbouring amino acid position (P1, Figure **13**).

Substitution at the benzylic position was investigated with incorporation of a (R)-CH$_3$, (S)-CO$_2$H or (S)-CO$_2$CH$_3$ group and were shown to be tolerated quite well with only a slight reduction in potency.

1.9.2 Amino acid variation at P1

In the absence of a basic moiety on the terminal substituent it was found that a basic residue on the side chain of this first amino acid position was essential. Conflicting results were given, however, with regards to the alkyl chain length to the amine group. Binding affinity was shown to increase with chain length going from Dbu to Orn, and again to Lys; though IC$_{50}$ values in the platelet aggregation assay improved in the opposite direction with the best result seen for Dbu. Binding affinity was seen to increase with no loss of potency when the amine of Dbu and Orn residues was converted to the bidentate acetimidoamide group. The tightest binding analogue comprised a combination of methionine at the first amino acid position with ethane-1,2-diamine as the terminal substituent - however, this compound proved to be cytotoxic to platelets at higher concentrations.[163]

1.9.3 Amino acid variation at P2

Biological data suggests the second amino acid position tolerates a wide range of variation when in combination with an Arg residue at P1. Tolerance of large lipophilic substituents such as Tyr(Bn) and di-Ph-Ala, increasing the IC$_{50}$ value by less than an order of magnitude, suggests the presence of a significant lipophilic pocket close to this position. Variation became more restricted, however, when Arg was replaced with Dbu at the first amino acid position; though the best performing compounds involved the Dbu (P1) with 3,4-diF-Phe (P2) combination.

A handful of analogues were synthesised where the urea linker was eliminated by reversing the order of the amino acid and coupling the indazole amine directly to the acid of the amino acid residue at P2. This resulted in loss of activity in most cases and a significant reduction in others and so the urea linker was retained.

1.9.4 Heterocyclic frame variation

Three different heterocyclic frames were investigated during the optimisation process and the transition from indole to indazole has already been discussed (see section 1.7, page 34). Also investigated was the benzimidazolone heterocycle, which despite only slightly reducing potency against low thrombin concentration (0.15 nM), showed a significant drop in binding affinity suggesting the effectiveness of such compounds may suffer against higher thrombin concentrations.

The substitution position on the benzene ring of the heterocycle was also investigated but deviation from the 6-position for linkage to the dipeptide fragment failed to produce any compounds of interest as PAR-1 antagonists.

1.9.5 Indole *N*-substituent variation

Substitution at the 1-position nitrogen of the heterocyclic ring system seems to be driven by sterics rather than electronics as both electron-donating and electron-withdrawing groups are tolerated at the *para* position of a benzyl group - any substitution at the *meta* position is less tolerated. Slight improvements in IC_{50} values were demonstrated for the 2,6-dichloro and the 2-methylbenzyl derivatives. Extending the ring system to a 2-naphthyl, or the alkyl chain to two CH_2 units, resulted in a reduction in activity suggesting the presence of quite a tight fitting pocket into which this substituent lies.

No information on biological data or the synthesis of compounds involving substitution at this benzylic position was reported.

1.9.6 Indole 3-position variation

From the data available it would seem that the heterocycle 3-position substitution is the least tolerant of all regions of the molecule. Any variation away from the original pyrrolidin-1-ylmethyl substituent results in a significant decrease in potency. The nearest performing analogue features an azetidine ring rather than a pyrrolidine ring, however this compound decreases potency by more than an order of magnitude.

1.9.7 Summary

Collectively, this SAR data provides valuable information that will assist in decisions made, following the determination of major metabolites from RWJ-58259, on the synthesis of analogues aimed at improving metabolic stability while maintaining PAR-1 antagonist activity. Encouragingly, it would seem that there is a level of tolerance in most regions of the molecule for small changes without impacting its activity on PAR-1. However, the drop in activity of any variation of the pyrrolidine is disappointing as options for analogue synthesis seem to be limited should this region prove to be the site of major metabolism.

Chapter 2 Lead Compound Validation

2.1 Introduction

It is imperative in every medicinal chemistry project that the identity and quality of the lead compound, against which all analogues are be compared, be verified in order to avoid the generation of data based on a misconception regarding the active compound. Typically, this simply requires analysis of the sample by analytical techniques such as high resolution mass spectrometry (HRMS) and nuclear magnetic resonance (NMR) spectroscopy. The sample of RWJ-58259, **20** (Figure 14), previously used within the Chambers research group was donated by Dr Claudia Derian, Johnson & Johnson. Unfortunately, very little remained at the beginning of this research project and was deemed insufficient for the planned work.
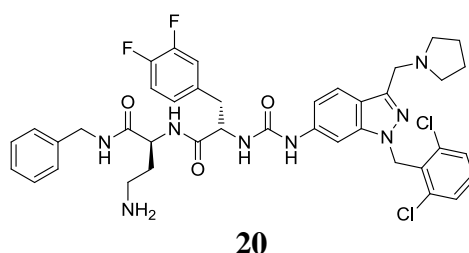

20

Figure 14 RWJ-58259

RWJ-58259 is not commercially available and thus it was necessary to synthesise a new batch of the lead compound. The few aliquots that remained of the original batch were insufficient for analysis by NMR spectroscopy and were to be used as a standard to validate the new batch of synthesised RWJ-58259.

2.2 Synthesis of RWJ-58259

During the development of analogues of the initial hit in the RWJ-series of PAR-1 antagonists, compound RWJ-53052 **18** (Figure 12, page 33), the investigators utilised a solid-phase synthesis approach. Three strategies were employed for attachment of the anchoring substituent to the resin (Figure 15): i) following coupling of Fmoc-Dbu(Boc)-OH with benzylamine and Boc deprotection, Dbu containing analogues were loaded onto a 2-chloro-trityl resin *via* the primary γ-N; ii) analogues capped with ethane-1,2-diamine were attached *via* the terminal primary amine, allowing for

variation of the first amino acid position; and iii) an amide-type resin, such as Tentagel S AM, was used to attach *via* a tertiary amide to the first amino acid position (Figure 15).[163, 167]
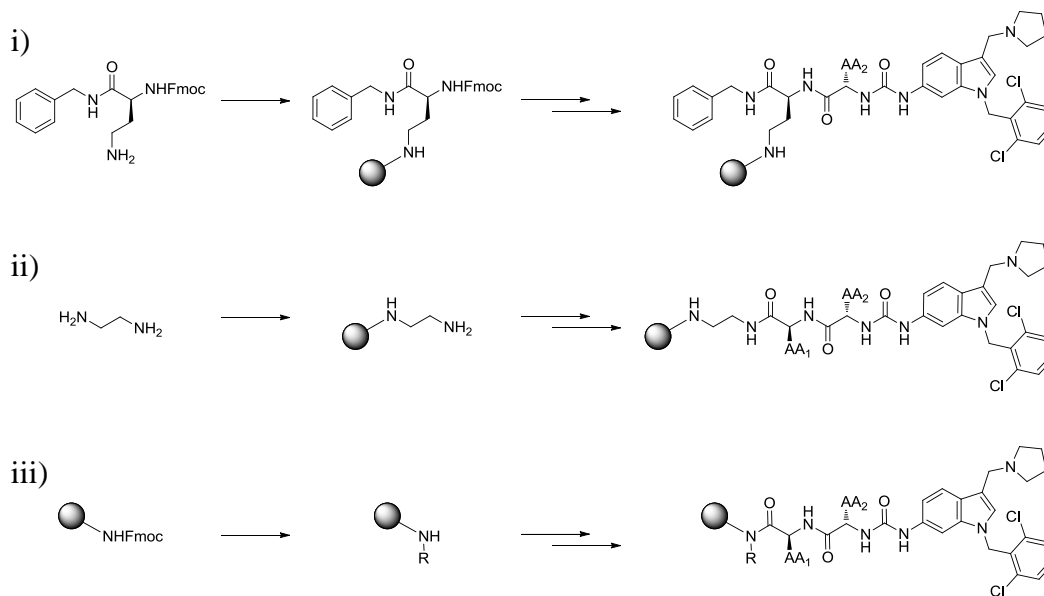
i)

ii)

iii)

**Figure 15** Solid-phase synthetic strategies employed in RWJ-series

Later, a convergent solution-phase method was described for the synthesis of analogues and it was this approach that Zhang and co-workers reportedly utilised for their synthesis of RWJ-58259.[164] From 6-nitroindole, the right-hand side (RHS) fragment **25** (Scheme 1) was reportedly achieved in 4-steps in an overall yield of 24%.
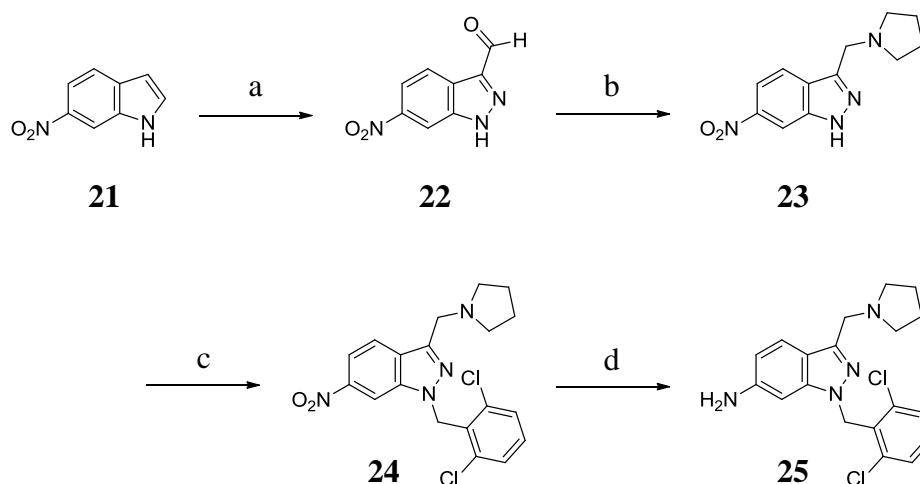
Reagents and conditions: a) NaNO$_2$, 6 N HCl, 5 h; b) pyrrolidine, NaB(OAc)$_3$H, ClCH$_2$CH$_2$Cl/DMF/AcOH (90:9:1), 2.5 h; c) 2,6-dichlorobenzyl bromide, KOH, THF, 1 h; d) FeCl$_3$.6H$_2$O, activated charcoal, Me$_2$NNH$_2$, MeOH, reflux, 2 h.

Scheme 1 RHS fragment synthesis

The first step involves a rearrangement of the indole starting material, as described by Buchi and co-workers, to form aldehyde **22**.[168] Nitrosation of 6-nitroindole **21** (Scheme 2) with sodium nitrite under aqueous acidic conditions leads to the formation of oxime **26**. Nucleophilic attack by water at the 2-position and a second nitrosation of the molecule at the 1-position nitrogen gives an α-hydroxy nitrosamine, which in the presence of acid collapses, cleaving the C-N bond, to form the aldehyde and the hydroxydiazene moieties of compound **27**. Loss of water drives the closure of the ring as the oxime at the pseudo 3-position reverts to the nitroso formation leading to nucleophilic attack on the pseudo 2-position nitrogen. Loss of nitric oxide restores aromaticity as 6-nitro-1*H*-indazole-3-carbaldehyde **22** (Scheme 1) is formed.
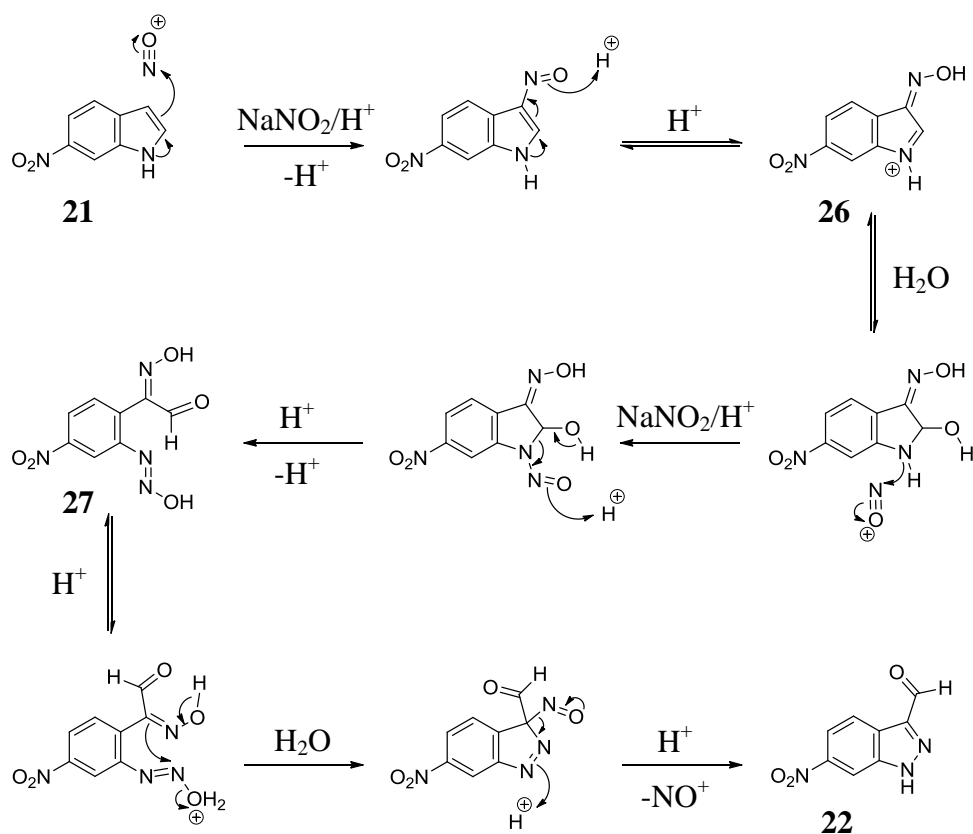

Scheme 2 Indazole Rearrangement Mechanism

Reductive amination of aldehyde **22** (Scheme 1) with pyrrolidine and sodium triacetoxyborohydride, followed by *N*-alkylation gives the 1,3-disubtituted 6-nitroindazole compound **24**. Reduction of the nitro group is achieved using the mild conditions first described by Boothroyd and Kerr.[169] Rather than using

hydrazine, unconventional *N,N*-dimethylhydrazine is used as the reducing agent with a catalytic amount of ferric chloride hexahydrate to give amine **25**.

Left-hand side (LHS) fragment **33** (Scheme 3) was reported as being prepared from Fmoc-Dbu(Boc)-OH **28** in 65% yield over 4 steps. The first amide coupling installs the benzylamine cap and reaction of deprotected amine **30** with Fmoc-3,4-difluoroPhe-OH **31**, followed by a second Fmoc deprotection, yields the dipeptide LHS fragment **33**.
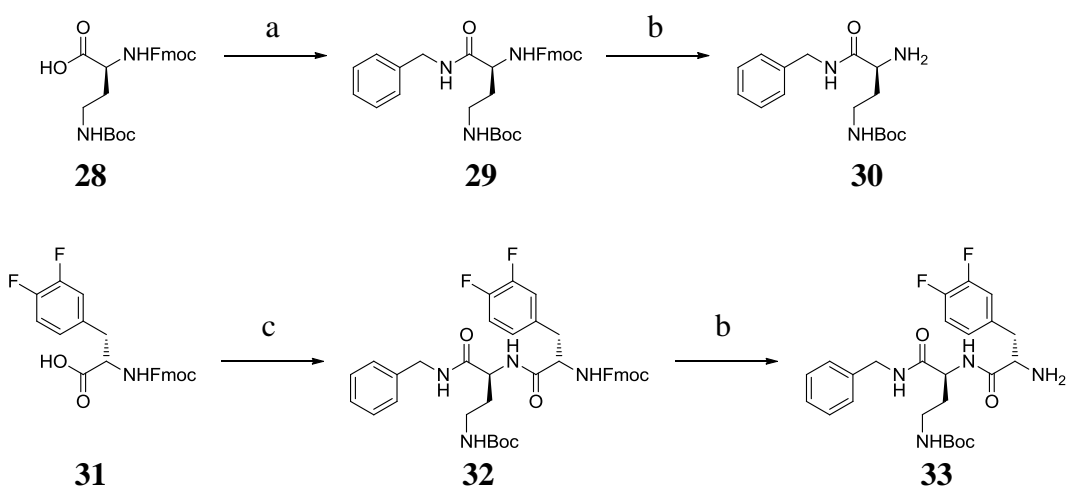
Reagents and conditions: a) BnNH$_2$, DCC, HOBt, MeCN, 3 h; b) Et$_2$NH, MeCN, 2 h; c) DIC, HOBt, **30**, MeCN, 16 h

Scheme 3 LHS fragment synthesis

Converging of the syntheses of RHS and LHS fragments **25** and **33** involves formation of a urea linker. Zhang and co-workers reported the use of 4-nitrophenyl chloroformate as the coupling reagent for this step and, following the subsequent Boc deprotection, was accomplished in 26% yield to give RWJ-58259 **20** (Scheme **4**) as its TFA salt.
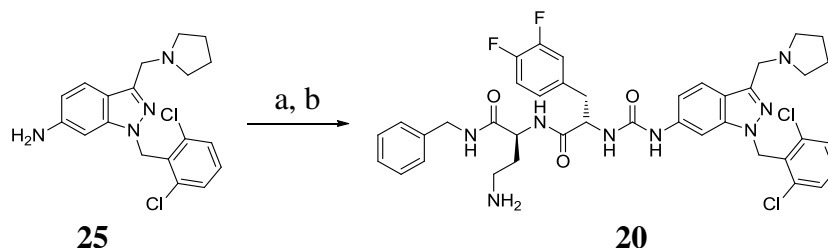
Reagents and conditions: a) 4-nitrophenyl chloroformate, ^{i}Pr$_2$NEt, CH$_2$Cl$_2$, **33**, -20 to 23 °C, 6 h; b) CF$_3$CO$_2$H, CH$_2$Cl$_2$, 2 h.

Scheme 4 Urea coupling

More recently, Valdivielso and co-workers reported an optimisation of the synthesis of RWJ-58259, having experienced difficulties when utilising the original conditions for two of the synthetic steps.[170] They found that low yields were obtained for the *N*-alkylation of indazole **23** (Scheme 5), and decided to investigate the use of alternative bases to promote the reaction. They reported that the use of either sodium hydride or triethylamine resulted in the formation of the undesired bromide of the di-alkylated compound as the major product − the second alkylation being on the pyrrolidine nitrogen (**34**, Scheme 5). However, an improved yield was achieved with the use of caesium carbonate, although a longer reaction time was required, giving the desired mono-alkylated product in 66% isolated yield.
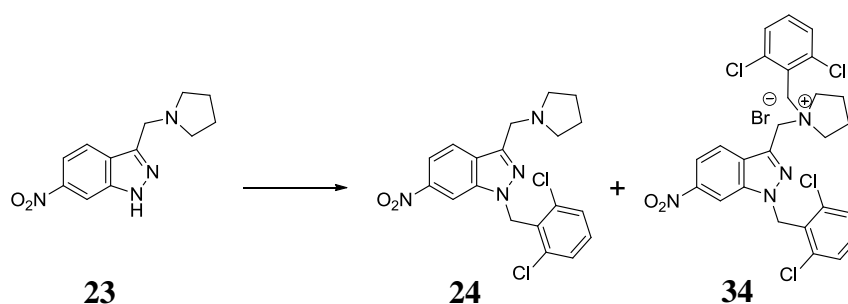
Reagents and conditions: 2,6-dichlorobenzyl bromide, Cs$_2$CO$_3$, THF (dry), argon, 24 h.

**Scheme 5** *N*-alkylation with modified conditions

Valdivielso and co-workers also reported difficulties in reproducing the urea formation product in a reasonable yield and looked to employ an alternative approach.[170] This was achieved through *in situ* formation of the isocyanate of the indazole amine by reacting with bis(trichloromethyl)carbonate in the presence of propylene oxide as an HCl scavenger. They reported that, following addition of the dipeptide amine, the desired urea product was obtained in a 91% yield.

2.2.1 Synthesis of RHS fragment

Encouraged by the methodological advances reported by Valdivielso and co-workers, synthesis of RWJ-58259 was initiated following their improved methodology. As described earlier, treatment of 6-nitroindole with sodium nitrite under acidic conditions led to nitrosation and subsequent rearrangement to 6-nitroindazole-3-carbaldehyde **22** (Scheme 1, page 42). Purification of this compound by flash column chromatography proved to be problematic due to the

instability of the aldehyde on silica. Thus the crude material was carried through to the next step without further purification. Gratifyingly, reductive amination of the aldehyde with pyrrolidine installed the tertiary amine moiety of **23** in 62% yield over 2 steps. *N*-alkylation of **23** with 2,6-dichlorobenzyl bromide was achieved in good yield following the improved methodology developed by Valdivielso and co-workers, which employed caesium carbonate as the base (Scheme 5).[170]

Subsequent nitro reduction under the mild reaction conditions stated above (Scheme **6**) afforded RHS fragment **25** in good yield giving an overall yield of 28% over 4 steps.
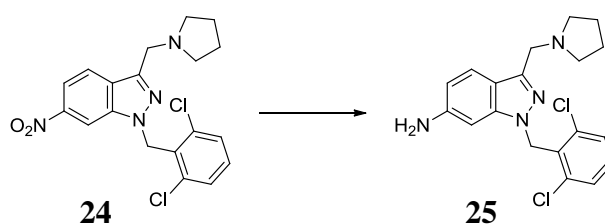

Reagents and conditions: Me$_2$NNH$_2$, FeCl$_3$.6H$_2$O, act. charcoal, MeOH, reflux, 2 h.

Scheme 6 Nitro reduction

2.2.2 Synthesis of LHS fragment

Synthesis of LHS dipeptide fragment **33** (Scheme 3, page 44) was initially attempted following the methodology described by Zhang and co-workers, using DCC with HOBt as coupling reagents for the first amide bond formation and DIC with HOBt for the second coupling. Purification of the dipeptide product following Fmoc deprotection proved difficult as the diisopropyl urea (DIU) by-product was found to co-elute with the desired product. Although dipeptide **33** was obtained in a calculated yield of 25%, this was contaminated with an equal amount of DIU by-product.

It was envisaged that the use of EDC might result in better separation of by-products, since its water solubility would improve its removal during an acid work up. As such, the second amide coupling step was repeated using EDC, in conjunction with HOBt, to activate the acid (Scheme 7). Although purification of the desired product was achieved, it was found that the reaction yield dropped significantly to a disappointing 7%.
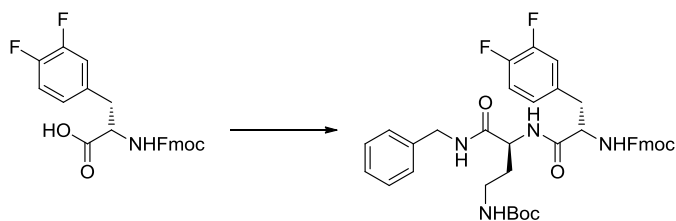
Reagents and conditions: EDC, HOBt, **30**, MeCN, 16 h.

Scheme 7 Second amide coupling step with EDC

As a result of previous success within our laboratories with the use of HBTU in amide couplings, synthesis of LHS fragment **33** was attempted utilising this reagent with HOBt and stirring with DIPEA to ensure pre-activation of the acid before addition of the benzylamine. To our satisfaction, purification of the desired product proved less problematic and using HBTU/HOBt for both amide coupling steps resulted in the isolation of LHS fragment **33** in 57% yield over 4 steps (Scheme 8).
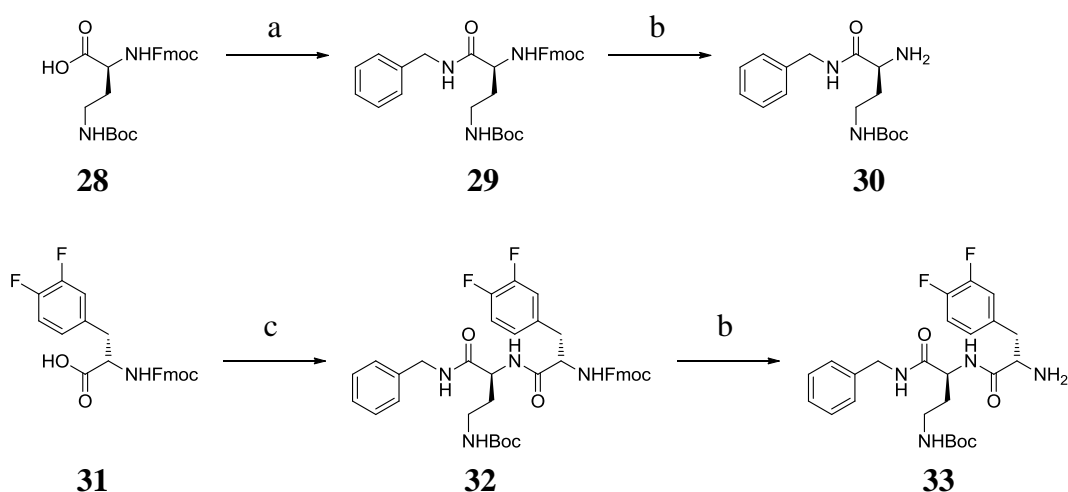
Reagents and conditions: a) HBTU, HOBt, DIPEA, BnNH₂, DMF, 5 h; b) Et₂NH, DMF, 5 h; c) HBTU, HOBt, DIPEA, **30**, DMF, 16 h

Scheme 8 LHS fragment synthesis

2.2.3 Urea coupling step

Unfortunately, despite Valdivielso *et al.* reporting formation of the desired product in a 91% yield, repeated attempts at the urea coupling reaction using this methodology failed to give more than a 28% yield for the target compound, RWJ-58259 **20** (Scheme 9). As this yield was too low to justify carrying large quantities of relatively

expensive reagents through the preceding steps, alternative coupling conditions were investigated.
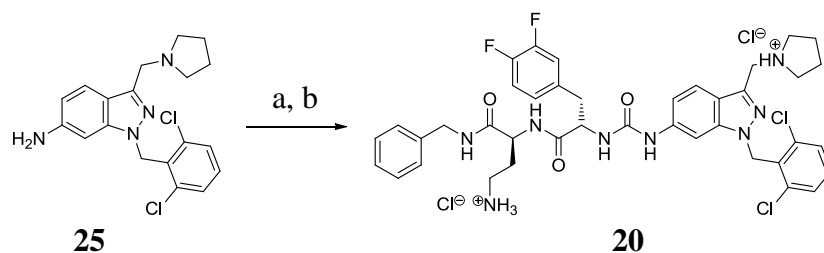
Reagents and conditions: a) triphosgene, propylene oxide, THF (dry), **33**, 0 to 20 °C, 1 h; b) HCl (4 N) in MeOH, 2 h.

Scheme 9 Urea coupling and deprotection

It was anticipated that use of a nucleophilic/base catalyst such as DMAP, which could act as both catalyst and HCl scavenger in the reaction, would promote the *in situ* formation of the isocyanate. To investigate this prediction in a model system, H-Phe-OMe (**36**) was used in place of LHS fragment **33** to react with RHS fragment **25** on a test scale (20.0 mg) to give di-substituted urea **35** (Scheme 10). NMR analysis of the crude reaction was used in order to determine the success of a number of reaction conditions based on the degree of conversion of starting materials to products.
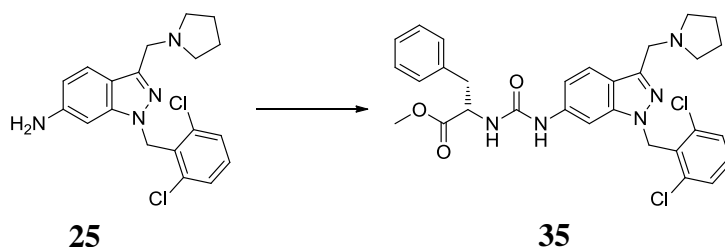
Reagents and conditions: triphosgene, DMAP or propylene oxide, THF (dry), H-Phe-OMe, 0 to 20 °C, 1 h.

Scheme 10 Urea coupling investigation model system

The homo-coupled (2*S*,2'*S*)-dimethyl 2,2'-[carbonylbis(azanediyl)]bis(3-phenylpropanoate) **37**, is a potential by-product of this reaction. In order to identify and exclude any signals in the crude NMR relating to **37**, the compound was synthesised, purified and analysed by NMR.

Compound **37,** was prepared by stirring amine **36** with bis(trichloromethyl) carbonate and Et₃N (Scheme 11) at room temperature for 1 h. The ^{1}H NMR spectrum

of the product obtained allowed for the identification of chemical shift changes for specific protons of interest upon urea coupling, which are indicated in Scheme 11.

36 → **37**
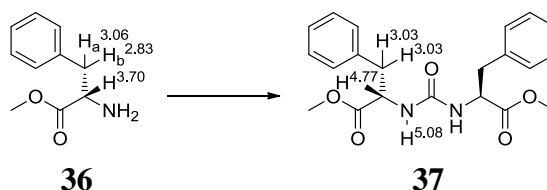
Reagents and conditions: triphosgene, Et₃N, THF (dry), 20 °C, 1 h.

Scheme 11 Di-Phe-OMe urea synthesis

In order to compare the use of DMAP with that of propylene oxide, as used by Valdivielso and co-workers, the reaction conditions reported for the modified RWJ-58259 synthesis were employed in the model system (Scheme 10). RHS fragment **25** was stirred with bis(trichloromethyl) carbonate and propylene oxide at 0 °C for 5 min before addition of aliphatic amine **36** at room temperature and stirring for 1 h. ^{1}H NMR spectroscopy showed no discernible peaks relating to desired product **35** nor for the possible homo-coupled products - di-Phe-OMe **37** and di-RHS fragment **38**.

The reaction was repeated with the use of DMAP in place of propylene oxide and, following a 30 min period of stirring at ambient temperature, aliphatic amine **36** was added at 0 °C. This time new peaks were observed in the ^{1}H NMR spectrum of the crude material. Peaks recognised as those representing starting materials **25** and **36** were seen – the α-*H* doublet of doublets (dd) at δ 3.70 and neighbouring diastereomeric *H*s doublet of dd (ddd) at δ 2.83 and 3.06, as well as the aromatic *H*s of the indazole positions 4, 5 and 7 at δ 7.45 (d), 6.53 (dd) and 6.49 (d), respectively. Also present were new peaks of interest which did not match those seen for the starting materials nor the homo-coupled by-product **37**, indicating successful urea coupling to give desired product **35**. Assisted by H-H correlation spectroscopy, peaks corresponding to the protons mentioned above were identified in a new environment. The α-*H* at δ 4.82 coupled to diastereomeric *H*s at δ 3.06 and 3.14, and to the N*H* at δ 5.86; the integration of these peaks matched precisely with those for aromatic *H*s at δ 7.49, 6.70 and 7.70, which correlated to each other as seen in indazole **25** for positions 4, 5 and 7. Analysis of the sample by mass spectrometry also showed the presence of a mass ion corresponding to [compound **35**+H]$^+$.

Confident that the desired urea coupling product **35** could be identified from the crude reaction mixture by ^{1}H NMR spectroscopy, variations in the reaction conditions were investigated for an improvement in conversion to the desired product.
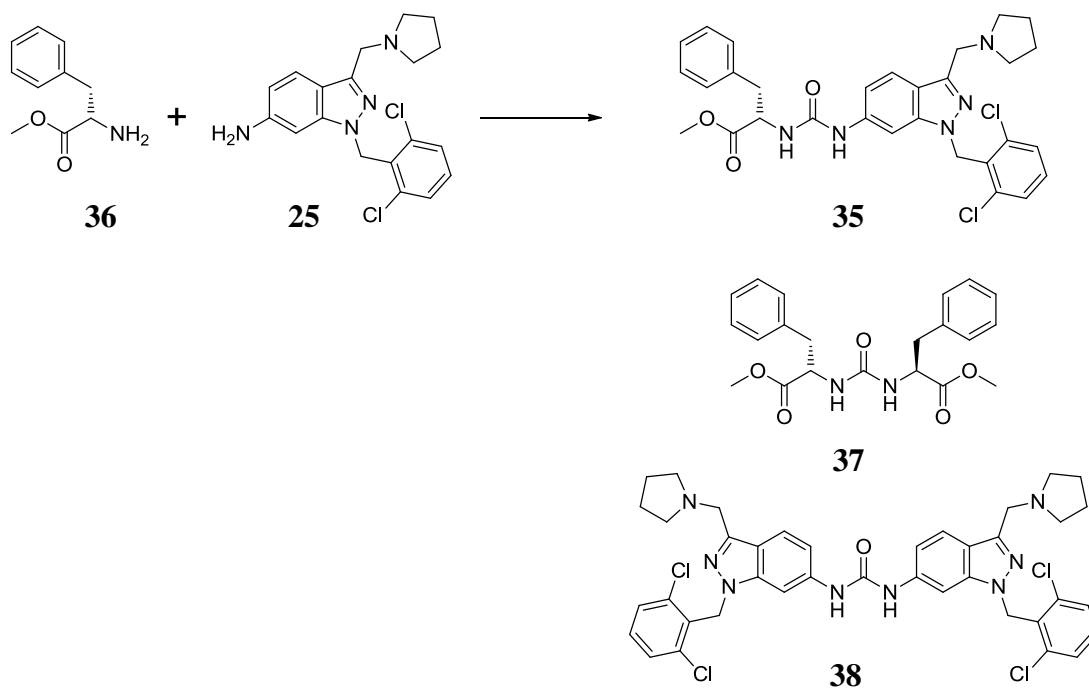
Reagents and conditions: triphosgene, DMAP, 0 to 20 °C, 1 h, Method A-E.

Scheme 12 Reaction scheme for urea coupling investigation

Table 1 Showing products identified from urea coupling reaction (Scheme 12)

Method	Solvent	Starting Materials 36	Starting Materials 25	Products 35	Products 37	Ratio of 25 in products 35 : 38 : 25 : other[d]
A	THF	1 eq	2 eq	64%[c]	0%	0.32 : 0.13 : 0.13 : 0.42
B	CH$_2$Cl$_2$	1 eq	2 eq	45%[c]	0%	0.23 : 0.52 : 0.02 : 0.24
C[a]	THF	1 eq	2 eq	77%[c]	0%	0.39 : 0.04 : 0.04 : 0.54
D[b]	THF	1 eq	2 eq	48%[c]	0%	0.24 : 0.10 : 0.17 : 0.50
E	THF	1 eq	0.95 eq	33%[c]	0%	0.33 : 0.07 : 0.07 : 0.54

a – the suspension of **25** in THF was sonicated for 5 min before addition of triphosgene and DMAP; *b* - suspension of **25** in THF was added dropwise to a solution of triphosgene and DMAP in THF; *c* - not isolated; *d* - ratio of initial amount of **25** converted to **35**, **38**, other unidentified products, or unreacted

Table 1 shows the calculated yields obtained for the desired urea product **35**. Also shown is the ratio of the initial amount of **25** converted to **35**, **38**, other unidentified

products or which remained as unreacted **25** - giving an indication of the progress of the first part of the reaction to form the isocyanate intermediate. During the original execution of the reaction (Method A) it was seen that RHS fragment **25** is sparingly soluble in THF and so the reaction was repeated in CH_2Cl_2 (Method B), in which the starting material indazole **25** is soluble. Unfortunately, although more of the starting material reacted, this seemed to promote the reaction of **25** with its isocyanate form to generate the homo-coupled product **38**. Rather than changing to a third solvent, Method C involved sonication of the suspension of **25** in the original solvent, THF, prior to addition of the other reagents, thus ensuring aggregation was minimised. This resulted in a reduction in the relative amount of **38** forming and an increase in the conversion of **36** to the desired urea coupled product **35**.

It was reasoned that further reduction of the amount of **38** formed may be achieved by reversing the order of addition of reagents in the first step, *i.e.* adding the sonicated suspension of **25** in THF to a solution of triphosgene and DMAP in THF. Dropwise addition would introduce amine **25** to an environment of excess coupling reagent, promoting the formation of the isocyanate over the homo-coupled product. Stirring the triphosgene with DMAP before addition of the amine may also activate the coupling reagent, again leading to preference of the amine to react with the coupling reagent over the isocyanate intermediate. The result shown for Method D contradicts our reasoning, suggesting that less of the starting material reacted either to form the isocyanate intermediate or the urea by-product **38**. This suggests a shortage of the coupling reagent and might be explained by the initial activation period stirring with DMAP in THF. Nucleophilic attack on triphosgene releases a molecule of phosgene, which is itself a coupling agent – hence the requirement of 0.35 eq. only of triphosgene. However, phosgene has a boiling point of *ca.* 8 °C and, despite the reaction mixture being kept at 0 °C, some may have evaporated leaving a shortage of the coupling reagent.

Valdivielso and co-workers reported that increasing the equivalence of dipeptide fragment **33** above 0.5 eq. did not improve the overall yield and only resulted in the unreacted starting material co-eluting with the desired urea coupled product. Investigating the effect of increasing the equivalence of the aliphatic amine seemed to confirm their findings with the ratio of unreacted **36** increasing in Method E.

It was decided to implement the conditions of Method C into a reaction involving an inexpensive dipeptide fragment to avoid the unnecessary potential loss of the more valuable RWJ-58259 LHS fragment **33**, and yet still generate an analogue of interest. During the development of RWJ-58259 Zhang and co-workers showed that analogues where the terminal aromatic moiety was switched with the primary amine functional group of the first amino acid position also demonstrated potent PAR-1 antagonist activity.[171] In some cases these analogues displayed an increased binding affinity to the receptor; although inferior selectivity for PAR-1 was also witnessed for some analogues. Encouraged by these reported findings, dipeptide fragment **44** (Scheme 13) was seen as a useful compound to be used to confirm the validity of the modified urea coupling reaction conditions.

tert-Butyl (2-aminoethyl)carbamate **39** was prepared in excellent yield by stirring di-*tert*-butyl dicarbonate with a 10-fold excess of ethylenediamine in $CHCl_3$ overnight, followed by evaporation of the unreacted amine under high vacuum. Amide coupling of amine **39** with Fmoc-Phe-OH **40** was achieved using coupling reagent HBTU with HOBt, and subsequent Fmoc deprotection gave amine **42**. The reaction sequence was repeated to install the second Phe residue giving the desired Boc protected dipeptide product **44** in 25**%** yield over 4 steps (Scheme 13).
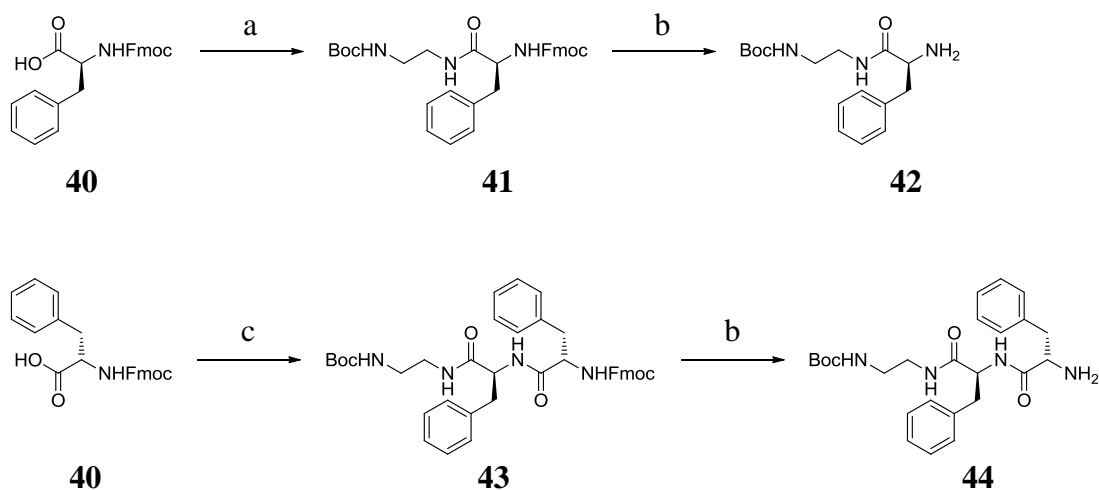

Reagents and conditions: a) HBTU, HOBt, DIPEA, DMF, **39**, 3 h; b) Et₂NH, DMF, 2 h; c) HBTU, HOBt, DIPEA, **42**, DMF, 16 h
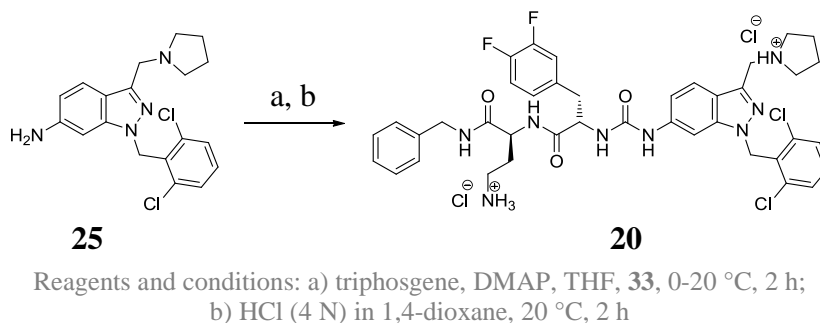
Scheme 13 LHS fragment analogue synthesis

Implementing the optimised conditions (Method C, Table 1), urea coupling of RHS fragment **25** with dipeptide fragment **44** and subsequent Boc deprotection was

achieved in 53% yield to give compound **45** as an analogue of RWJ-58259 (Scheme **14**).

Reagents and conditions: a) triphosgene, DMAP, THF, **44**, 0-20 °C, 2 h; b) HCl (4 N) in 1,4-dioxane, 20 °C, 2 h
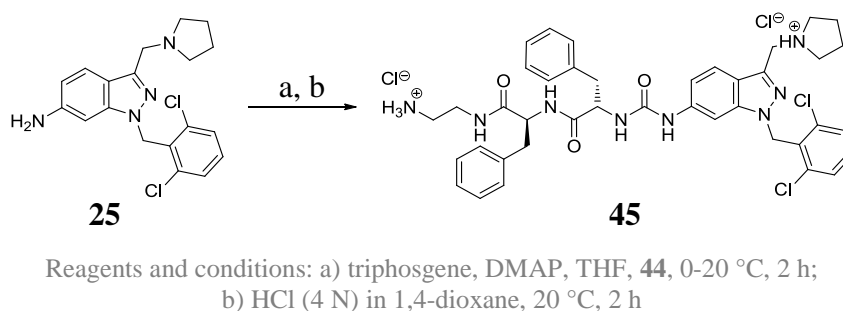
Scheme 14 Urea coupling and Boc deprotection with modified conditions

Despite still being far from the 91% yield reportedly achieved by Valdivielso and co-workers, a significant improvement had been achieved on previous attempts at the urea coupling reaction. As a result it was decided that the conditions could be utilised in the synthesis of a new batch of RWJ-58259 **20** (Scheme 15). Satisfyingly the coupling and deprotection steps were achieved in a 53% yield to furnish compound **20**.

Reagents and conditions: a) triphosgene, DMAP, THF, **33**, 0-20 °C, 2 h; b) HCl (4 N) in 1,4-dioxane, 20 °C, 2 h

Scheme 15 Urea coupling and deprotection to generate RWJ-58259

2.3 Biological evaluation

Having successfully synthesised a new batch of RWJ-58259 it was necessary to evaluate its antagonistic activity against PAR-1 and compare this with the previous batch of RWJ-58259 used in the Chambers research group. This was done to ensure continuity of the quality of data gathered, as well as to validate the structure of active compound in the previous batch. For clarity, despite having the same chemical structure, the batch synthesised during this research is referred to as compound **20** in the biological results, and the previous sample as RWJ-58259.

2.3.1 *In vitro* testing

As introduced in section 1.5.2 (page 20), activation of PAR-1 triggers various cell signalling pathways, with preference for the signal transduction route depending upon the nature of the activating ligand. Activation following cleavage by thrombin favours signalling *via* the $G\alpha_{12/13}$ sub-unit and further through the actions of RhoA. This small GTPase is a key component in the regulation of diverse cellular processes, including stress fiber formation, actin skeleton organisation, gene expression and cell division.[172]

Thrombin induced activation also triggers signalling *via* the $G\alpha_q$ sub-unit, albeit not to the extent to which the $G\alpha_{12/13}$ sub-unit is activated, which initiates the phospholipase C (PLC) signal transduction pathway. Activated PLC is responsible for the cleavage of phosphatidylinositol 4,5-biphosphate (PIP_2) into diacyl glycerol (DAG) and inositol 1,4,5-triphosphate (IP_3). IP_3 is released into the cytosol and interacts with IP_3 receptors on the smooth endoplasmic reticulum resulting in the release of Ca^{2+} ions into the cytosol, which bind to and activate various signal transduction components affecting numerous cellular processes. DAG remains bound to the cell membrane following cleavage of PIP_2 by PLC and activates protein kinase C (PKC), which goes on to phosphorylate and activate the Raf1 enzyme, among various other proteins, leading to the regulation of gene transcription *via* the mitogen-activated protein kinases/extracellular signal-regulated kinases (MAPK/ERK) signalling pathway.

In vitro testing methods employed previously within the Chambers research group have involved measurement of: i) phosphorylated ERK and total ERK, by western blot; and ii) the production and secretion of chemokine (C-C motif) ligand 2 (CCL2), by an enzyme-linked immunosorbent assay (ELISA). However, early efforts with the western blot assay revealed significant variability in results gathered. This made it difficult to compare accurately between compound **20** and RWJ-58259 other than to confirm that both are somewhat active against thrombin at low concentration. Therefore, initial *in vitro* testing to evaluate the biological activity of compound **20** was focussed on CCL2 ELISA.

On account of RWJ-58259 and analogues being compounds of interest in the development of PAR-1 antagonists to investigate pulmonary fibrosis, the primary choice of cell type employed for *in vitro* testing was the human lung fibroblast (HLF). This was due to the central role the fibroblast cell is believed to play in the progression of disease through proliferation and differentiation into the myofibroblast cell (see section 1.4.3.1). To confirm that results generated with the use of HLFs translate accordingly to murine lung fibroblasts (MLFs), since *in vivo* studies are done in mice, initial *in vitro* testing was done using both cell types.

2.3.1.1 CCL2 ELISA

Expression and secretion of chemokine CCL2 is known to be promoted in various cell types in response to exposure to thrombin. CCL2 plays an important role in the recruitment of various leukocytes to the site of injury through chemotaxis - the movement of cells in response to chemical stimulus. An ELISA was used to analyse the ability of compound **20** to inhibit CCL2 release in both HLFs and MLFs. Previous work done in the Chambers group by Deng and co-workers had shown thrombin signalling to be mediated exclusively *via* PAR-1 in MLFs, demonstrated by the ability of PAR-1 antagonist Q94 to arrest secretion of CCL2 triggered by high thrombin concentrations (10 nM).[173]

Interestingly, more recent work showed that treatment of HLFs with RWJ-58259 failed to fully prevent CCL2 release following exposure to even low concentrations of thrombin (0.3 nM); no inhibition was achieved at a concentration of 10 nM thrombin.[174] Incubating the HLFs with a combination of cleavage-blocking monoclonal antibodies (ATAP2 and WEDE15), prevented CCL2 release at low thrombin concentration; however, at high thrombin concentration these too were unable to achieve full inhibition. These findings suggest that CCL2 secretion results from thrombin signalling mediated by an alternative route in conjunction with PAR-1 in HLFs. This may be due to contribution from PAR-4 since this receptor's lower thrombin affinity would mean signalling *via* PAR-4 would be triggered at high thrombin concentrations.

Results from the ELISA for compound **20** corroborate previous findings in that CCL2 release is reduced but not fully inhibited in HLFs treated with 0.3 nM thrombin. However, in MLFs treated with 10 nM thrombin full inhibition of CCL2 secretion was achieved by compound **20** (Figure 16).
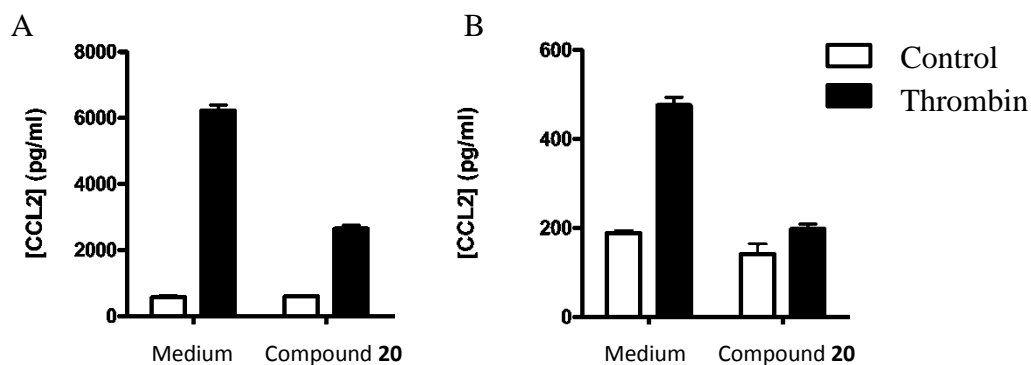

Figure 16 CCL2 release in response to thrombin following treatment with compound **20**; A - in HLFs, B - in MLFs

2.3.2 *In vivo* testing

In order to confirm that the positive *in vitro* testing results were transferrable to an *in vivo* setting, compound **20** was entered into a mouse model of LPS-induced lung inflammation. PAR-1 activation is known to regulate the release of pro-inflammatory mediators, including various cytokines and chemokines, which play an integral role in the orchestration of the response to injury.[175] One of the earliest responses is the recruitment of leukocytes to the area affected and first to respond are neutrophils, arriving within minutes of the injury occurring (see section 1.2.1, page 4). Recent work reported by the Chambers research group has demonstrated that blockade of PAR-1 activation by administration of RWJ-58259 in mice leads to a reduction in acute neutrophil recruitment at 3 h after LPS administration.[176] This reduction is also maintained, not simply delayed, since measurements at 6 h and 24 h also showed reduced numbers of neutrophils in bronchoalveolar lavage (BAL) fluid. Their work suggested an important role for chemokines CCL2 and CCL7 in the early recruitment of neutrophils, since their expression was identified as being up-regulated in cells in response to LPS, and PAR-1 antagonism with RWJ-58259 was shown to attenuate their levels of expression. Levels of macrophages recruited after LPS administration were shown not to be affected by neutralisation of CCL2 and CCL7 with anti-CCL2

and anti-CCL7 antibodies, respectively. This indicates an enhanced role for CCL2 and CCL7 in the recruitment of neutrophils in particular, whereas recruitment of other leukocytes depend on other chemokines and cytokines, the expressions of which are likely affected to a lesser degree by PAR-1 antagonism.

2.3.2.1 LPS-induce lung inflammation study

Compound **20** was administered intraperitoneally to mice 30 min after being challenged with LPS from *Streptococcus pneumoniae*. Mice were euthanized 24 h after exposure to LPS, lungs were lavaged and the cells counted.

The neutrophil cell counts were shown to be reduced as a result of treatment with compound **20** compared with those of mice not given the PAR-1 antagonist (A, Figure 17); this result unequivocally confirmed the quality of synthesised compound **20** to be equivalent to RWJ-58259. The ineffectiveness of anti-CCL2 and anti-CCL7 antibodies in the reduction of macrophage cell counts following LPS-induced lung inflammation was also verified through the failure of compound **20** to reduce the numbers of macrophages recruited to the site of injury (B, Figure **17**).
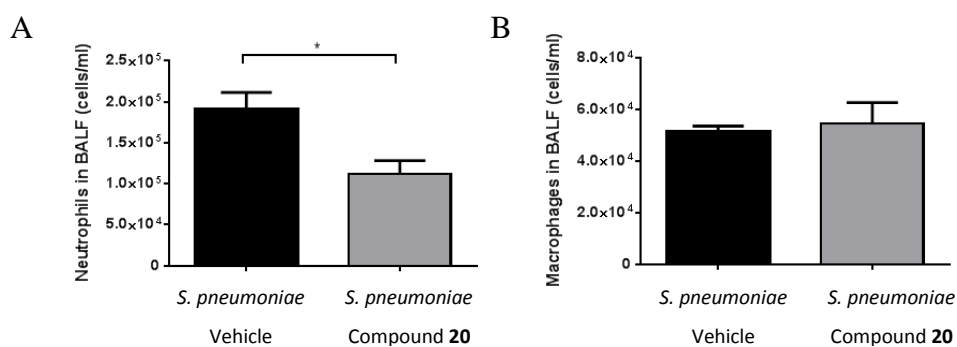

Figure 17 Cell counts in BAL fluid 24 h after LPS challenge;

A - neutrophils, B - macrophages

With the quality of compound **20** confirmed through thorough biological evaluation, work could begin on DM/PK profiling of the compound and the development of analogues for biological testing with compound **20** as the standard.

Chapter 3   Synthesis of Analogues

3.1 Introduction

Having confirmed the biological activity of compound **20** relative to that of results obtained for the original batch of RWJ-58259, the remainder of the investigation could commence with confidence in the identity and quality of the project lead compound.

The next stage of the investigation involved the synthesis of analogues of the project lead compound aimed at improving metabolic stability while maintaining PAR-1 antagonist activity. In order to decide on synthetic targets, information was required on the metabolic pathways to which RWJ-58259 is susceptible. DM/PK profiling of RWJ-58259 would provide information on the major metabolites formed; using this information in conjunction with the data obtained from literature regarding SAR studies of the RWJ-series of PAR-1 antagonists, analogues of interest could be selected as synthetic targets.

Synthesised compounds would then be tested in *in vitro* studies with HLFs against compound **20** as the standard. However, due to the variability experienced in the western blot assay and the time consumption associated with running the ELISA, it was decided that development of a new assay would be beneficial for the *in vitro* testing of analogues generated during the investigation. This decision was influenced by the accessibility to a Fluorescence Imaging Plate Reader (FLIPR) at UCL's Department of Physiology. As a consequence, assay development would be required in order to ensure accurate and relevant data is acquired.

3.2 DM/PK profiling

Very little information is reported in literature on the pharmacokinetics of RWJ-58259; only a short paragraph details the findings of Damiano and co-workers on the DM/PK profile of the compound following administration to rats and from a human liver microsomes assay.[166] They report a $t_{1/2}$ of 9.2 min for RWJ-58259 in a human liver microsome assay with oxidation and hydroxylation of the pyrrolidine

ring being the major metabolites formed - though no information is given on the precise location of metabolism or the degree of oxidation. The $t_{1/2}$ following intravenous administration of the compound in rats was measured as 0.32 h. Expanding the DM/PK profile of RWJ-58259 would benefit future *in vivo* investigations using this compound and assist in the development of metabolically stable analogues.

The intended approach for DM/PK profiling of compound **20** involved firstly an investigation into the chemical stability of the compound. The reported negligible bioavailability of RWJ-58259 following oral administration in rats may suggest a possible issue with the compound surviving conditions found in the digestive tract.[166] Therefore, it was decided an investigation into the hydrolytic stability of compound **20** over a range of pH may provide information as to why oral administration is unsuccessful.

Secondly, using the facilities made available through our collaborators at GSK, compound **20** would be submitted to a liver microsome assay, as described by Damiano and co-workers; however, since our *in vivo* studies are carried out in mice, microsomes from mouse as well as from human livers would be used. Data gathered from the liver microsome assay would give valuable information on the clearance time of compound **20** in mice, which would assist in future decisions on dosing regimen for *in vivo* studies. Analysis of the supernatant from the liver microsome assay by LC/MS/MS and by nanoscale ^{1}H NMR would allow for the identification of major metabolites generated during incubation of compound **20** with the microsomes.

Finally, using the analytical techniques described to analyse supernatant from the microsome assay (*vide supra*) and guided with information gained regarding clearance times of compound **20**, blood samples taken at time intervals from mice treated with the compound would be analysed to identify major metabolites generated *in vivo*. Based on the identity of these metabolites and those identified from the liver microsome assay, analogues of compound **20** could be synthesised with structural modifications aimed at blocking major sites of metabolism with the intention of generating a more metabolically stable analogue with potent PAR-1 antagonistic properties.

Unfortunately, due to complications at GSK during the time of this research, access to the facilities required for the running of the liver microsome assay was reduced. This resulted in a delay in obtaining the data required for deciding on approaches towards synthesising analogues of interest. Fortunately, however, access was provided to *in silico* screening facilities at GSK allowing calculated predictions to be made for the major sites of metabolism. This approach was adopted whilst access to the liver microsome assay was restricted.

3.2.1 Hydrolytic stability study

The hydrolytic stability of compound **20** was investigated to determine whether a range of physiologically relevant pH would be tolerated. Although the amide bond is considered relatively stable to hydrolysis under physiological conditions, the presence of neighbouring groups may, in some cases, act to catalyse the bond cleavage.[177] The chemical stability of compound **20** in an aqueous environment with pH ranging from 4 to 9 was investigated with the aim of identifying products of hydrolysis. The compound was incubated in buffered solutions of pH 4, 7.4 and 9, at 20 °C and 37 °C, and aliquots of the solutions were taken for analysis at time intervals of 10 min, 3 h, 24 h, 48 h.
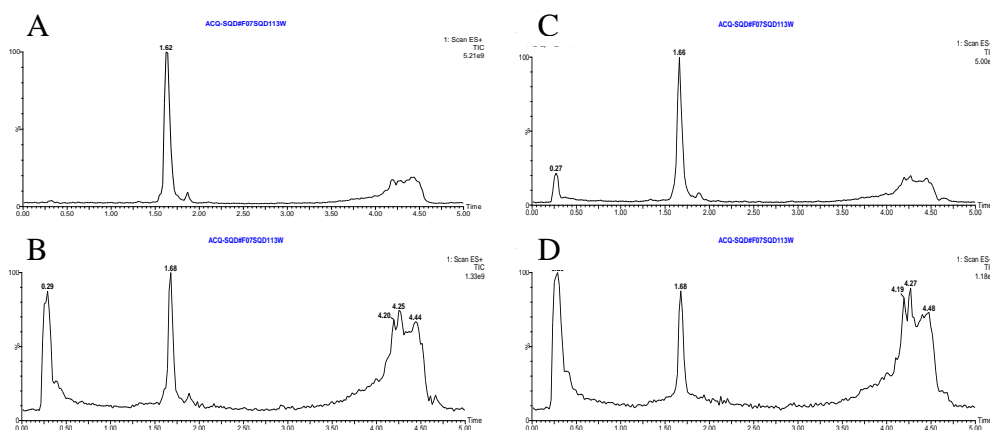
A - compound **20** standard; B - pH 7.4, 37 °C at 24 h; C - pH 4, 37 °C at 24 h; D - pH 9, 37 °C at 24 h. The peak at 20 s represents the injection peak and the broad peak at the end results from high organic solvent content in the mobile phase washing trace levels of organic compounds from the column.

Figure 18 ES+ TIC traces for hydrolytic study of compound **20**

Analysis of the samples by LC/MS showed only the parent compound present up to 48 h, suggesting compound **20** is stable to hydrolysis at physiologically relevant pH.

3.2.2 *In silico* screening

Compound **20** was submitted for *in silico* screening by our collaborators at GSK. The screening methodology used, MetaSite (Molecular Discovery Ltd), requires only the compound structure information in SMILES format. A 3D structure is generated, from which a set of descriptors is obtained for each atom using GRID probe pharmacophore recognition. Individual atoms are categorised based on their hydrophobic, hydrogen-bond donor, hydrogen-bond acceptor, or charge capabilities and their distances in space are binned and converted into clustered distances. A fingerprint for each atom in the substrate molecule is generated and compared with the fingerprint of human cytochrome P450s (CYPs) recognised as being responsible for major drug metabolism. Sets of descriptors based on GRID flexible molecular interaction fields are generated for CYP1A2, CYP2C9, CYP2C19, CYP2D6 and CYP3A4 to give the enzyme fingerprint. The fingerprint of both the substrate and the enzyme are compared and a score is calculated for the reactivity (R_i) and accessibility (E_i) of each atom (i) in the substrate. The site of metabolism is described as a probability function: $P_{SM},(i) = E_i R_i$.

The output gained from MetaSite ranks the probability of each atom to be a site of metabolism, calculated from its propensity to undergo the reaction type associated with the enzyme as well as the atom's accessibility towards the enzyme's heme group (Figure 19). The output also includes the likely products resulting from metabolism at the most probable sites (Figure 20, page 62).
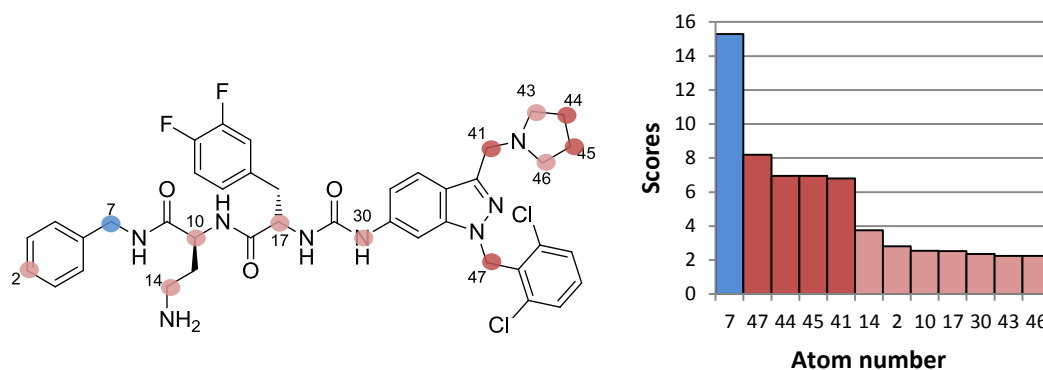
Graph shows scores calculated for the probability of an atom being a site of metabolism with colour coding identifying the position of the atom in the drawn structure; blue - most likely; dark red - likely; light red - less likely. See appendix for full MetaSite output.

Figure 19 Representation of MetaSite output for compound **20**
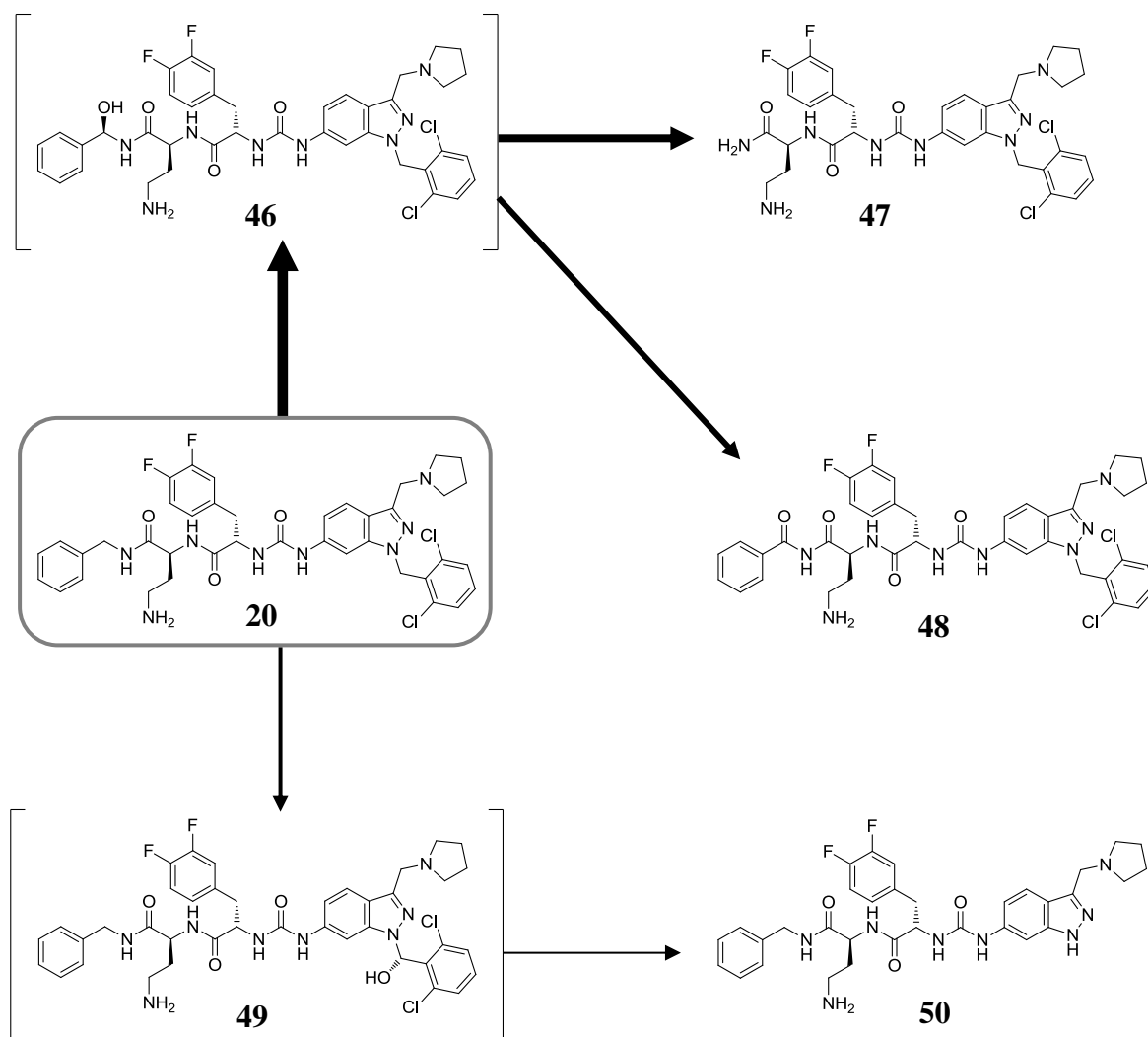
Figure 20 Predicted major metabolites formed from RWJ-58259

As demonstrated in Figure 19 Representation of MetaSite output for compound **20**, probability scores are displayed for each atom and major sites likely to be metabolised are highlighted. For compound **20**, MetaSite predicted the benzylic position of the terminal benzylamine (atom number 7) as the major site of metabolism, suggesting that this site is likely to undergo oxidation to form (*R*)-α-hydroxybenzylamine analogue **46** (Figure 20). This compound is likely to be an intermediate for both of the compounds predicted as the next most likely metabolites; cleavage of the C-N bond leaves primary amide **47** or further oxidation of **46** to yield benzamide derivative **48** (Figure 20).

The benzylic position of the *N*-(2,6-dichloro)benzyl substituent (atom number 47) was ranked as the atom next most probable to be a site of major metabolism, again

with oxidation leading to the (*S*)-α-hydroxy-(2,6-dichloro)benzyl analogue **49** and subsequent cleavage of the C-N bond to give 1*H*-indazole **50** (Figure 20).

Also highlighted as important sites to consider with regards to metabolism of compound **20** were the 3-position of the pyrrolidine ring (atom numbers 44 & 45) and the methylene substituent at the 3-position of the indazole (atom number 41). Taking this information into account, it is possible the hydroxylation and oxidation of the pyrrolidine ring, reported by Damiano and co-workers, occurs at the 3-position making it an important site to consider investigating.[166]

3.3 Analogue synthesis
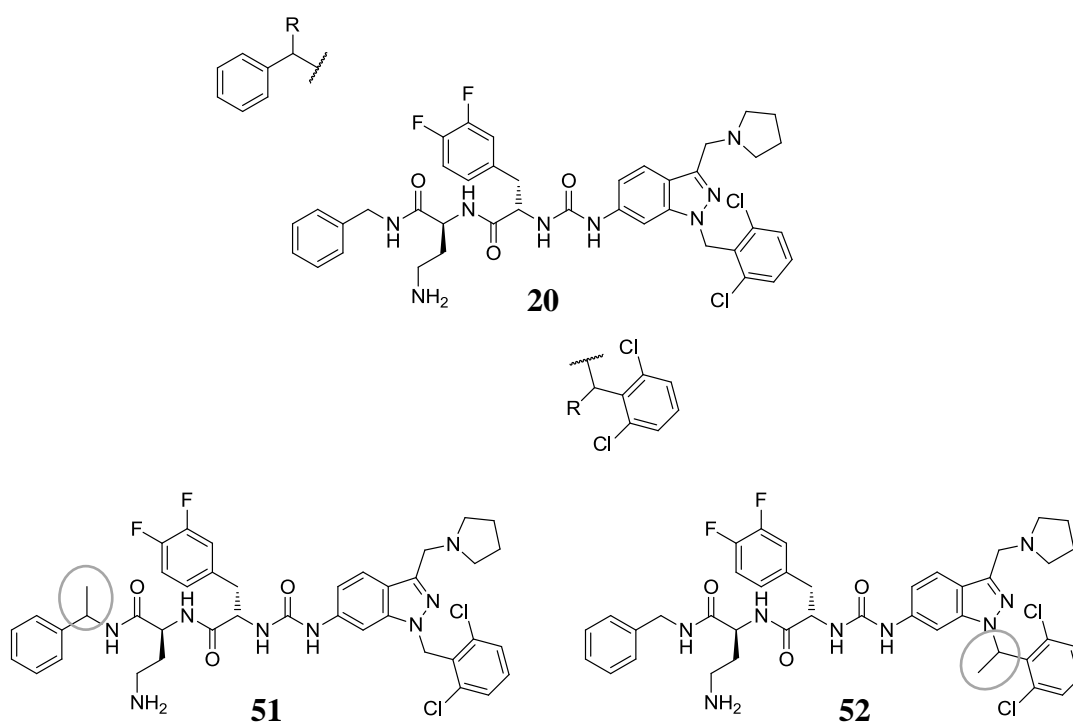
Figure 21 Primary synthetic targets

Guided by the information gained from the *in silico* screen of compound **20**, initial focus for the synthesis of analogues aimed at reducing metabolism centred on both benzylic positions highlighted in the MetaSite output (Figure 19). It was proposed that installation of a relatively small substituent at these positions would be enough to reduce their accessibility to the heme group at the active site of the metabolising enzyme, yet would be small enough so as not to interfere with vital interactions of the compound within the binding site of PAR-1, thus maintaining antagonist activity.

The methyl group was selected as the substituent to be introduced at these benzylic positions to investigate its effect on biological activity of the compound. Therefore, initial synthetic targets were decided as compounds **51** and **52** (Figure 21).

On account of the reported findings by Damiano and co-workers regarding metabolism of RWJ-58259, it was decided attention should also be directed towards the pyrrolidine ring of the molecule. Major metabolites identified from a human liver microsome assay involving RWJ-58259 were reported to result from hydroxylation and oxidation of the pyrrolidine ring. Taking this information in context with the data obtained from the MetaSite output, in which the 3-position of the ring was highlighted as a probable site of metabolism by CYPs, an analogue with a substituent at this position was proposed as a compound of interest (Figure 22).
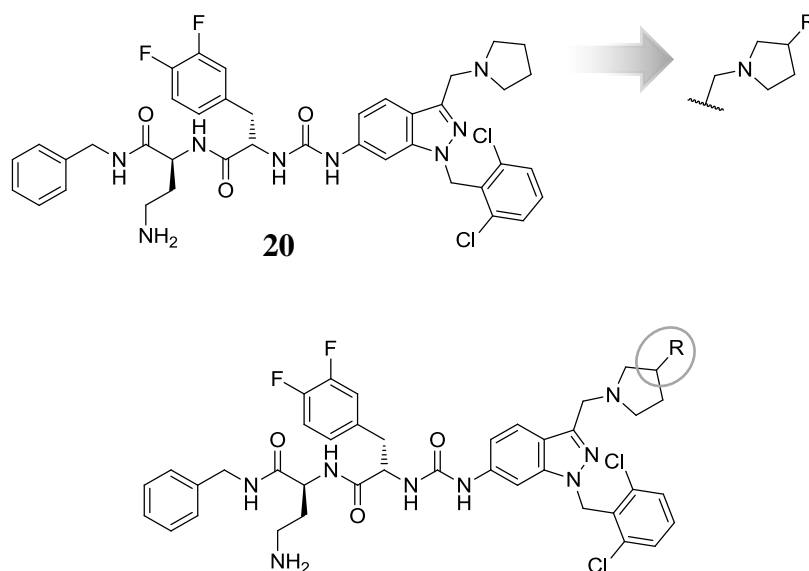
Figure 22 Secondary synthetic target

3.3.1 α-Methylbenzylamine analogue synthesis

Employing a convergent approach to the synthesis of compound **20** meant it was necessary only to alter the synthesis of the LHS fragment, followed by urea coupling to the previously synthesised RHS fragment **25**, in order to achieve analogue **51** (Figure 21).

Due to the MetaSite prediction that oxidation would lead to the (R)-α-hydroxybenzylamine analogue **46** (Figure 20) stereoisomer preferentially, synthesis of both α-methyl stereoisomers at this benzylic position was explored to

see whether a difference in metabolic stability and biological activity would be observed. Employing the previously validated synthetic route, acid **28** was activated by stirring with HBTU/HOBt in the presence of DIPEA for a short period (3 min) before the introduction of either (*R*)- or (*S*)-1-phenylethanamine (Scheme 16). Following Fmoc deprotection, amine **54a/54b** was stirred with HBTU/HOBt activated acid **31** to achieve compounds **56a** and **56b** in good yield (59% and 57%, respectively) after a second Fmoc deprotection step.
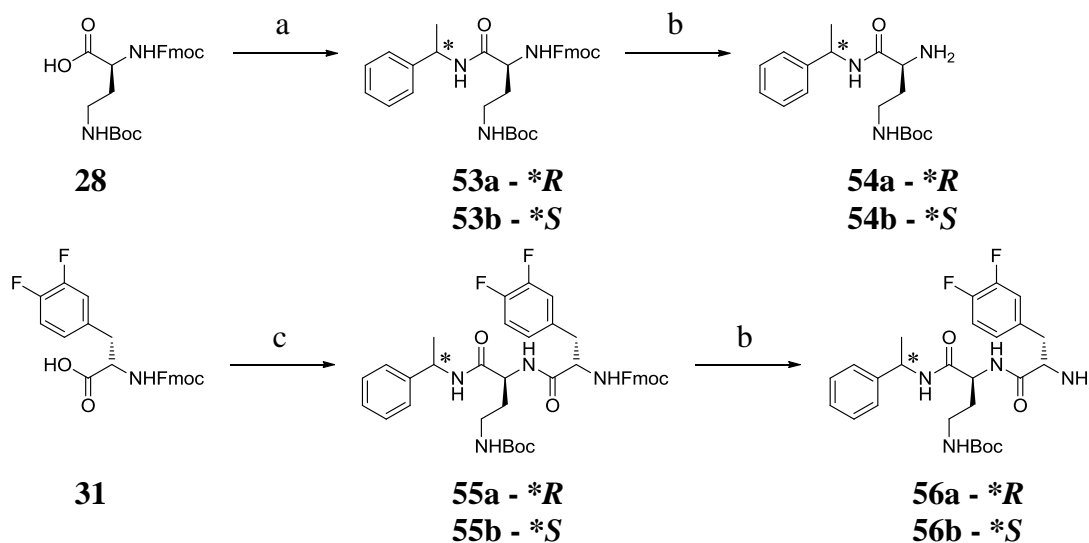
Reagents and conditions: a) HBTU, HOBt, DIPEA, DMF, (*R*)- or (*S*)-1-phenylethanamine, 3 h; b) Et₂NH, DMF, 2 h; c) HBTU, HOBt, DIPEA, **54a** or **54b**, DMF, 16 h

Scheme 16 Synthetic route to LHS fragment analogues

Urea coupling of amine **56a/56b** to amine **25** was achieved *via* the *in situ* isocyanate formation of the aniline amine by the optimised conditions described earlier (see section 2.2.3, page 47).
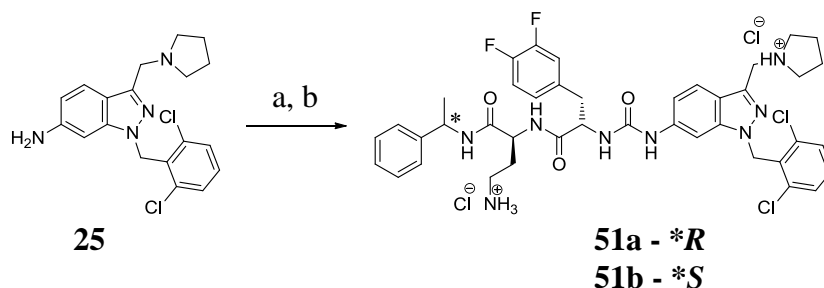
Reagents and conditions: a) triphosgene, DMAP, THF, **56a** or **56b**, 0-20 °C, 2 h; b) HCl (4 N) in 1,4-dioxane, 20 °C, 2 h

Scheme 17 Urea coupling and Boc deprotection

Compounds **51a** and **51b** were obtained in reasonable yield following Boc deprotection and purification by preparative HPLC.

3.3.2 α-Methyl-(2,6-dichloro)benzyl analogue synthesis

For the synthesis of compound **52** a slight modification of the RHS synthetic route was required to allow the installation of the methyl group at the benzylic position. The modified RHS fragment could be coupled to LHS fragment **33** to give the target compound.

Preparation of the alkylating reagent was required beforehand and was achieved, as shown in Scheme 18, in excellent yield. 2,6-Dichlorobenzaldehyde **57** was treated with methyl Grignard reagent under an inert atmosphere at 0 °C 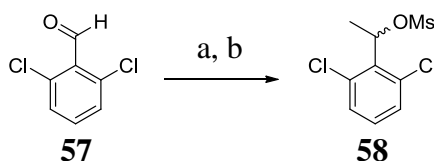for 90 min to give the secondary alcohol in quantitative yield as a clear yellow oil following the work up. The crude product was then stirred with triethylamine at 0 °C, to which was added methanesulfonyl chloride and after 90 min, purification of the reaction mixture gave the desired mesylate product **58** in 76% yield as a white crystalline solid.

Reagents and conditions: a) MeMgBr (3 M) in Et₂O, THF (dry), Argon, 0 °C, 1.5 h; b) MsCl, Et₃N, CH₂Cl₂, 0 °C, 1.5 h

Scheme 18 *N*-alkylating reagent synthesis

N-alkylation of indazole **23** with mesylate **58** (Scheme 19) proved to be more difficult than the equivalent reaction with the 2,6-dichlorobenzyl bromide, likely to be due to the added steric hindrance by the methyl group at the site of nucleophilic attack. The reaction would not proceed without heating, and even after stirring at 50 °C in THF for 24 h only 25% yield of the desired product **59** was isolated. It was noticed that regioisomer **60** was also generated in a ~5% yield, with *N*-alkylation occurring at the 2-position rather than the desired 1-position. In order to heat the reaction mixture at a higher temperature a solvent change was required; it was reasoned that using a more polar solvent might assist in stabilizing the anion formation and promote *N*-alkylation at the 1-position. Unfortunately, the ratio of compounds **59**:**60** was not improved by heating at 100 °C in DMF, however the

conversion of starting material to products was improved with compounds **59** and **60** being isolated in 55% and 14% yield, respectively. As a result, sufficient material was obtained to carry *N*-1-alkylated indazole **59** through to the next step. Reduction of the nitro group gave amine **61** in 84% yield.
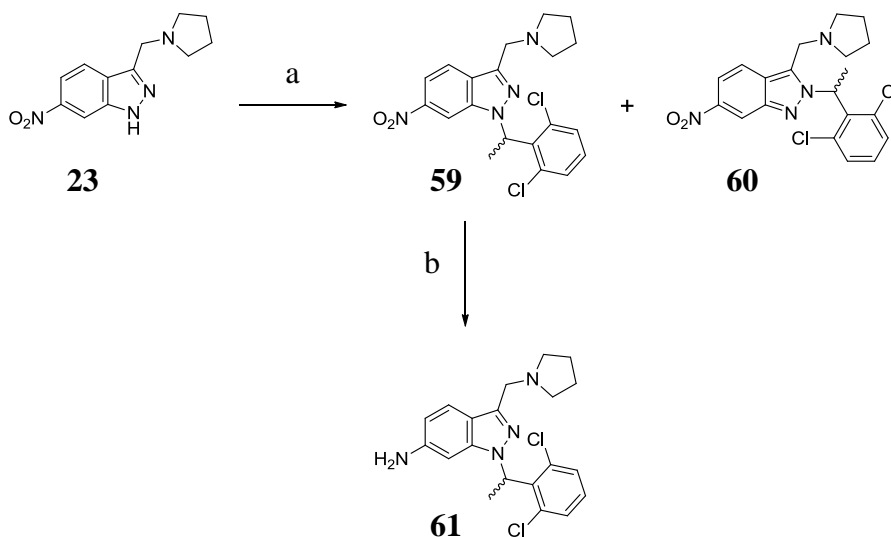
Reagents and conditions: a) CsCO$_3$, **58**, DMF, 100 °C, 20 h; b) FeCl$_3$.6H$_2$O, activated charcoal, Me$_2$NNH$_2$, MeOH, reflux, 2 h.

Scheme 19 Synthetic route to RHS fragment methylated analogue

Urea coupling with dipeptide fragment **33** followed by Boc deprotection provided target compound **52** in 35% yield (2 steps) after purification by preparative HPLC (Scheme 20).
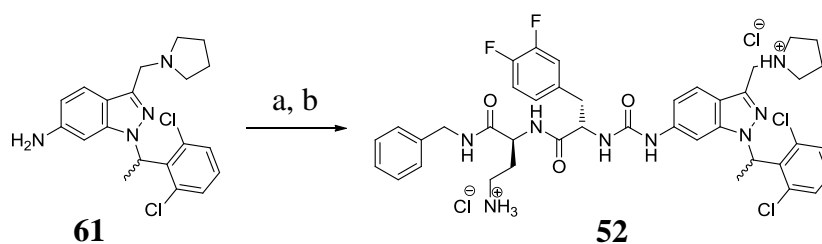
Reagents and conditions: a) triphosgene, DMAP, THF, **33**, 0-20 °C, 2 h; b) HCl (4 N) in 1,4-dioxane, 20 °C, 2 h

Scheme 20 Urea coupling and Boc deprotection

3.3.3 3-Fluoropyrrolidine analogue synthesis

When deciding on the substituent to introduce at the pyrrolidine 3-position, a review of the published literature surrounding the development and optimisation of

RWJ-58259 related compounds showed very little variation away from the pyrrolidine ring structure for this region of the molecule.[163, 164, 171] Structure-activity relationship (SAR) data published in patents filed, indicate that variation in this region of the molecule is not well tolerated (see section 1.9.6, page 39). A significant drop in antagonist activity is seen when bulkier substituents are introduced, as well as compounds incorporating a secondary rather than a tertiary amine. Therefore, it was proposed that introducing a fluorine atom at the 3-position (**62**, Figure 23) may reduce the likelihood of hydrogen atom abstraction occurring by CYPs, which leads to hydroxylation of the carbon atom.

62
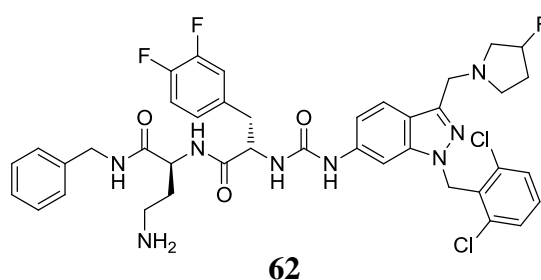
Figure 23 Secondary synthetic target

Fluorine is generally considered a good bioisostere for hydrogen in medicinal chemistry, since many structural and chemical features correspond closely.[178] Fluorine is the most electronegative atom in the periodic table and as a result its three lone pairs of electrons are held tightly by its highly positive nucleus. This renders the fluorine atom practically inert to interactions such as resonance or hydrogen bonding, unlike nitrogen or oxygen. Many consequences of its electronegativity make fluorine the atom/group closest resembling hydrogen in its structural properties. Fluorine's atomic radius is the nearest matching to that of hydrogen (F = 1.47 Å, *cf* H = 1.2 Å), as is its bond length to carbon (C-F = 1.35 Å, *cf* C-H = 1.09 Å). This means substitution of hydrogen with fluorine results in minimal steric influence. The carbon-fluorine bond is also the strongest bond in organic chemistry with regards to homolytic cleavage (bond dissociation energy for C-F = 105.4 kcal mol^{-1}, *cf* C-H = 98.8 kcal mol^{-1}), which again is a consequence of the high electronegativity of fluorine. The C-F bond is highly polarised with most of the electron density near the fluorine atom and can be considered more electrostatic in character than covalent. Such properties make fluorine a good substitution for hydrogen at likely sites of metabolism by CYPs, due to increased difficulty in breaking the C-F bond.

Therefore synthesis was initiated towards the production of RHS fragments **65a** and **65b** (Scheme 21), in which a fluorine atom was incorporated into the pyrrolidine 3-position as either the *R*- or *S*-stereoisomer. As with compounds **51a** and **51b**, both stereoisomers were synthesised to identify any difference in metabolic stability when submitted to liver microsome assays.

As with other RHS fragments, synthesis began from 6-nitroindole **21** and following rearrangement to indazole **22**, reductive amination of the aldehyde was carried out with either (*R*)- or (*S*)-3-fluoropyrrolidine to give **63a** and **63b** in yields of 62% and 59%, respectively. *N*-alkylation with 2,6-dichlorobenzyl bromide was achieved in good yield for both **64a** and **64b**; it was noted during purification that retention factors (Rf) for these fluorinated compounds on thin layer chromatography (TLC) plates were significantly higher than those seen for compounds **24** and **59** (Scheme 1, page 42 and Scheme 19, page 67, respectively). This suggested the extreme electronegativity of the fluorine atom may be having an effect on the basicity of the neighbouring tertiary amine, which made the compounds easier to purify by flash column chromatography. However, this raised concern as to what effect this may have on biological activity, since it is likely that the tertiary amine is involved in binding site interactions within PAR-1.
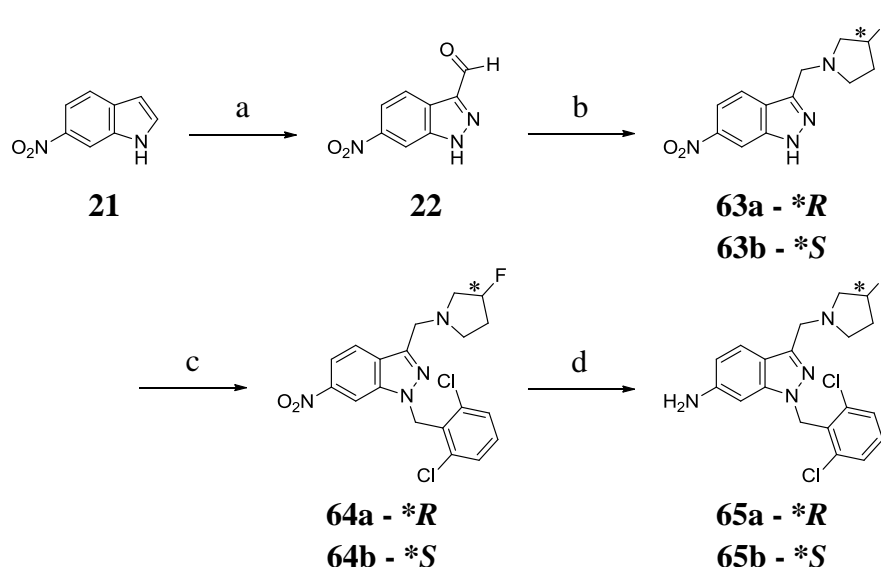
Reagents and conditions: a) NaNO$_2$, 6 N HCl, 5 h; b) (*R*)- or (*S*)-3-fluoropyrrolidine, NaB(OAc)$_3$H, CH$_2$Cl$_2$/DMF/AcOH (90:9:1), 3 h; c) 2,6-dichlorobenzyl bromide, CsCO$_3$, THF, 20 h; d) FeCl$_3$.6H$_2$O, activated charcoal, Me$_2$NNH$_2$, MeOH, reflux, 3.5 h.

Scheme 21 Synthetic route to RHS fragment 3-fluoropyrrolidine analogues

Nitro reduction was achieved, requiring a slightly longer reflux time, to give compounds **65a** and **65b** as pale yellow solids in 72% and 76% yields, respectively.

Fluorinated RHS fragments **65a** and **65b** were coupled to the original RWJ-58259 LHS fragment **33** by the optimised conditions to give the Boc protected intermediates in 58% and 64% yield, respectively (Scheme 22). Stirring with HCl (4 N) in 1,4-dioxane for 2 h then purification by preparative HPLC gave compounds **62a** and **62b** in good yield as white solids after repeated lyophilisation from 0.1 M HCl (aq).
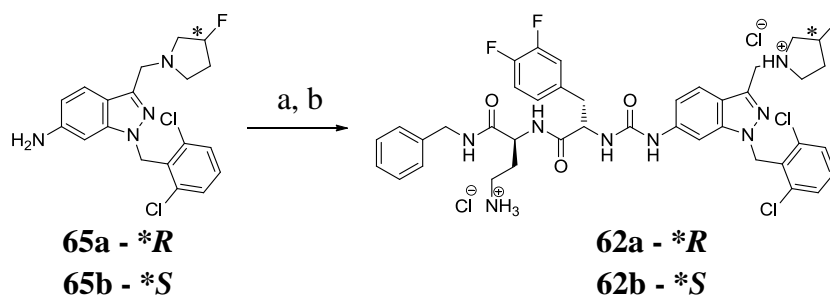
65a - *R*
65b - *S*

62a - *R*
62b - *S*

Reagents and conditions: a) triphosgene, DMAP, THF, **33**, 0-20 °C, 2 h; b) HCl (4 N) in 1,4-dioxane, 20 °C, 2 h

Scheme 22 Urea coupling and Boc deprotection

3.4 Metabolite generation

In conjunction with investigations into the development of analogues with improved metabolic stability, it was considered of interest to examine the biological activity of possible metabolites generated from RWJ-58259. Since the compound demonstrates efficacy in *in vivo* studies, and yet has been shown to have such a short $t_{1/2}$, it was reasoned that metabolism of RWJ-58259 may be generating metabolites that are also active against PAR-1, which may be prolonging the expected antagonistic activity of the parent compound. To this end it was reasoned that 3-pyrrolidinol analogue **66** (Figure 24) could be synthesised in much the same way as fluorinated analogues **62a** and **62b**, avoiding the requirement of developing an alternative synthetic route. Compound **66** was seen as the representative metabolite proposed by information gathered from the MetaSite output together with the findings reported by Damiano and co-workers.[166]
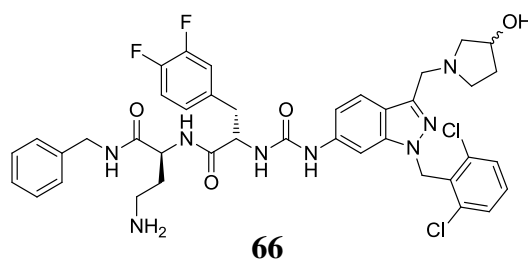
Figure 24 Metabolite target for synthesis

Also investigated, with assistance from our collaborators at GSK, was the use of biocatalysis as an approach to generate a range of possible metabolites directly from compound **20**. Interest in the use of biocatalysis in industry has increased in recent years as greater emphasis is put on environmental sustainability for manufacturing processes. No longer seen as the slow option, suitable only for academic curiosity, biocatalysis has seen rapid development recently. With the increasing number of protein structures being solved, growing understanding of protein functions in biosynthesis and development in computational methods, the process of engineering enzymes for specific conversions at precise locations on a substrate has become highly efficient. As a result, the use of biocatalysis in industry has escalated where employing engineered enzymes to perform chemical conversions, which would otherwise involve a number of synthetic steps, results in improved efficiency and ultimately cuts costs.

It was decided to investigate the use of a Codex® screening panel (Codexis) to screen a range of engineered microbial CYP enzymes for activity on compound **20**. Enzymes found to be actively metabolising compound **20** would then be incorporated in scaled up reactions with the aim of generating sufficient amounts of the metabolites in order to isolate and screen them in our *in vitro* studies for biological activity.

3.4.1 3-Pyrrolidinol analogue synthesis

The 3-pyrrolidinol, to be introduced by reductive amination of indazole **22**, was acquired as the *N*-Boc protected compound. As such, Boc deprotection was required and achieved in quantitative yield by stirring in diethyl ether whilst bubbling HCl (g) through the solution.

Reductive amination of aldehyde **22** with the Boc deprotected 3-pyrrolidinol was achieved in 39% yield over the two steps from indole **21**. Indazole **67** was then *N*-alkylated with 2,6-dichlorobenzyl bromide to give indazole **68** in 57% yield (Scheme 23). Nitro reduction was achieved after an extended period at reflux to give RHS fragment **69** in 63% yield.
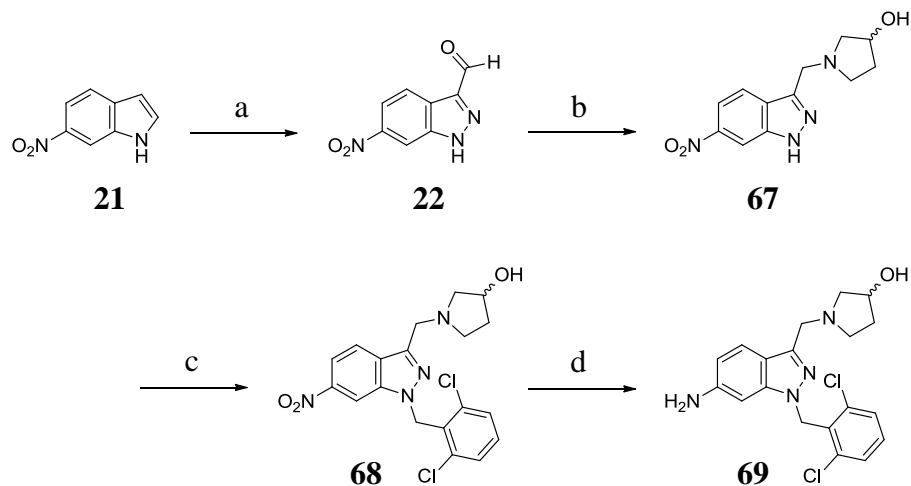
Reagents and conditions: a) $NaNO_2$, 6 N HCl, 5 h; b) 3-pyrrolidinol, $NaB(OAc)_3H$, CH_2Cl_2/DMF/AcOH (90:9:1), 3 h; c) 2,6-dichlorobenzyl bromide, $CsCO_3$, THF, 20 h; d) $FeCl_3.6H_2O$, activated charcoal, Me_2NNH_2, MeOH, reflux, 5 h.

Scheme 23 Synthetic route to RHS fragment 3-pyrrolidinol analogue

It was noticed during the synthesis that, similar to the equivalent fluorinated 6-nitroindazoles (**64a** and **64b**), the Rf for compound **68** on TLC plates was significantly raised compared to the original RWJ-58259 equivalent synthetic intermediate, compound **24**. However, introduction of an OH group could not be seen as reducing the polarity of the compound. In fact a cLogP for **68** was calculated at 3.44 compared with 4.83 for compound **24** (Figure 25), suggesting a significant reduction in lipophilicity with the introduction of an OH group, as would be expected.
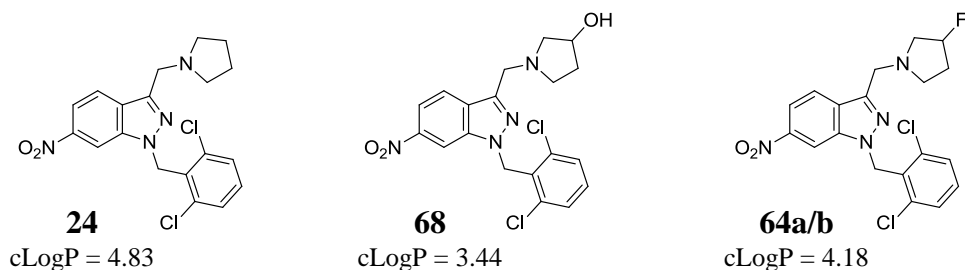
Figure 25 Calculated Log P values for intermediates

An alternative explanation may be that the hydrogen of the alcohol group is involved in weak hydrogen bonding with a π-bond in the aromatic indazole ring system. This may result in reduced interaction of the OH group with the silica during chromatography, leading to the increased Rf values witnessed.

Urea coupling of RHS fragment **69** with dipeptide fragment **33**, followed by Boc deprotection, gave target compound **66** after purification by preparative HPLC and lyophilisation from dilute HCl (aq) (Scheme 24); disappointingly only a 19% yield of product was achieved.
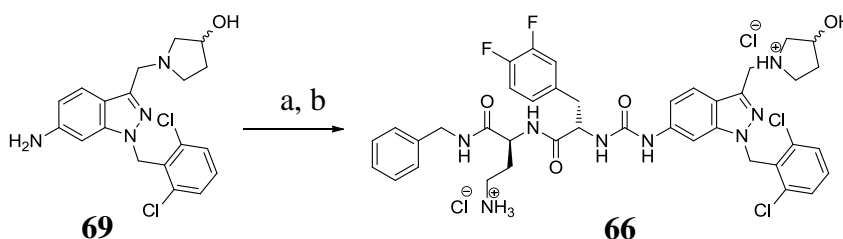
Reagents and conditions: a) triphosgene, DMAP, THF, **33**, 0-20 °C, 2 h; b) HCl (4 N) in 1,4-dioxane, 20 °C, 2 h

Scheme 24 Urea coupling and Boc deprotection

3.4.2 Biocatalysis

The Codex® panel is a 96-well plate with each well containing a different engineered microbial enzyme. These enzymes are selected for their range of reaction type activity and high tolerance for substrate variability. The panel is used to screen compounds against a wide range of enzyme classes in order to identify enzymes for which the compound is a substrate. The screen identifies the class or classes of enzyme(s) to be investigated further with the objective of discovering or developing an engineered enzyme that could be employed to carry out chemical conversions on the compound of interest.

Since the aim of this investigation was to generate possible metabolites of RWJ-58259 in order to screen them for PAR-1 activity, a range of engineered microbial CYPs were used. Compound **20** was dissolved in a 5% dextrose solution and added to the wells, which contained the enzymes dissolved in the reaction buffer solution, to give a final concentration of 0.1 mg/mL of compound **20**. A porous lid was applied to avoid cross contamination between wells and yet allow air to reach

the reaction mixture, and the plate was shaken thoroughly (250 rpm) at 30 °C for 16 h. The contents of each well were then diluted with acetonitrile causing the protein to precipitate out, the plate was centrifuged and the remaining solution analysed by LC/MS to identify any compounds of interest generated.

The results from LC/MS analysis showed that despite positive controls proving conditions to be optimal for enzyme activity, no reactions were seen to occur in wells containing compound **20**. Although this result seemed disappointing initially, the positive consequence of this outcome is the indication that compound **20** is not extensively metabolised. It is important to state that the Codex$^{®}$ panel enzymes are engineered based on genetic information from microbial CYPs, which are generally seen as being more active and stable than plant or animal enzymes, making them more suitable for industrial purposes. As a result their activity towards compound **20** is likely to differ from human or mouse CYPs. However, the fact that compound **20** proved to be inert to engineered enzymes selected for their activity and substrate tolerance would suggest that any metabolism occurring as a result of interaction with human or mouse CYPs may be restricted to very few pathways. This would prove to be an encouraging result since blocking metabolic pathways might require only very few modifications of the project lead compound to achieve the extended $t_{1/2}$ desired.

3.5 Biological evaluation of analogues

In order to be able to evaluate the compounds generated in the course of this research it was necessary to have a robust method of screening compounds for activity. Ideally this method would allow for the possibility of its use in a high throughput format for potential long ranging applications. Previous *in vitro* assays used within the Chambers research group to investigate PAR-1 inhibition have involved the phosphorylated ERK/total ERK western blot and the CCL2 ELISA; however, both assays are relatively time consuming and labour intensive and so an alternative assay was investigated.

3.5.1 Calcium ion mobilisation assay

Access to a Molecular Devices FLIPR Tetra High Throughput Cellular Screening System at the Department of Physiology, UCL, encouraged us to look into the

possibility of running a calcium ion mobilisation assay. PAR-1 activation by thrombin is known to lead to a quick response in the release of Ca^{2+} into the cytosol, which contributes to the activation of various components responsible for signal transduction (see section 2.3.1, page 54). Use of a dye, which upon Ca^{2+} complexation increases its fluorescence, allows for the real-time monitoring of PAR-1 activation and data derived from measurements of fluorescence detected are obtained immediately following the running of the assay. This method would reduce the experiment time by up to two days and use of a 96-well plate with a well-designed plate layout would maximise the number of compounds possible to test in each assay run.

3.5.1.1 Assay development

Use of the FLIPR was made to analyse the degree of PAR-1 activation in HLFs following exposure to thrombin. The Fluo-4 AM dye (Molecular Probes®) was used to complex with Ca^{2+} and HLFs, grown to confluence in 96-well plates, were incubated with either: i) the dye in media only, as control or, ii) dye in media with the compound of interest. Since Ca^{2+} mobilisation is a relatively quick response to PAR-1 activation, addition of the agonist was automated and added to each well simultaneously by the FLIPR. The cells were irradiated with light of wavelength (λ) 470-495 nm, which resulted in excitation of the Ca^{2+}-Fluo-4 AM complex and its emission of light at $\lambda = 515$-575 nm was detected. The degree of PAR-1 activation was determined by measuring the amount of fluorescence detected over a period of time.

Initially the assay was run in the absence of any compounds, with varying concentration of thrombin. This was in order to ensure the Ca^{2+} release response was measurable and to determine the effect of agonist concentration on degree of response.

The initial result was very promising as illustrated in Figure 26 where, in the absence of an antagonist, the time and extent of the Ca^{2+} mobilisation response is clearly related to the concentration of thrombin administered.
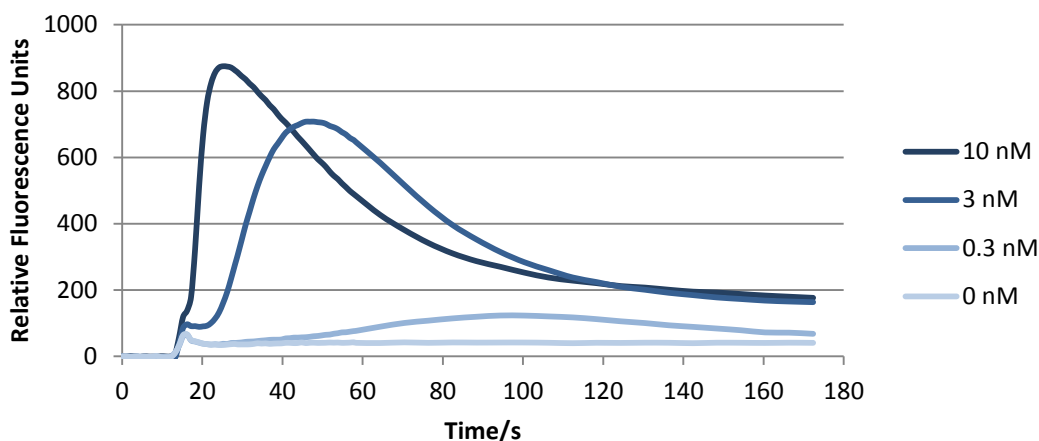
Figure 26 Ca^{2+} mobilisation responses over time in HLFs to varying concentrations of thrombin

The concentrations of thrombin used were chosen to represent low (0.3 nM), standard (3 nM) and high (10 nM) levels of thrombin. Physiological concentrations of thrombin vary depending upon the state of the surrounding tissue, since thrombin exists as its zymogen, prothrombin, until activation of the coagulation cascade leads to its cleavage by FXa (see Figure 2, page 4). In extravascular compartments a range of 0.05-0.13 nM is considered typical for the low thrombin concentrations seen following initiation of the coagulation cascade. This concentration rises to 1.2-2.4 nM by the time of clot formation (22-26 min), due to amplification of the initial activation through thrombin's positive feedback mechanisms.[179] Much higher concentrations of thrombin are possible for intravascular injuries, however, a concentration of 10 nM was considered high yet physiologically relevant with respect to the our interest in injuries occurring to the pulmonary interstitium.

The next stage was to validate the assay using compound **20** and RWJ-58259, both of which had demonstrated PAR-1 antagonist activity in the CCL2 ELISA and the *in vivo* LPS-induced lung inflammation study. The 96-well plate layout was designed as shown in Figure 27, allowing for six different compounds to be tested against four different concentrations of thrombin, with four experimental replicates for each condition. Compound **20** was tested alongside RWJ-58259 in order to compare directly between the different batches of the project lead compound. Also tested was the di-Phe analogue of RWJ-58259, compound **45**, synthesised as a model compound of interest in order to validate the synthetic route. Of the final three rows available one was filled with dye medium only, as a positive control, and another contained a

compound from a different research project, compound X. The remaining row was filled with the dye in media containing 0.1% DMSO; since compound stock solutions were made in DMSO it was necessary to demonstrate that the maximum amount of DMSO present following dilution did not affect the outcome.

Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45
Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45	Medium only	Medium + DMSO	Cmpd. X	RWJ-58259	Cmpd. 20	Cmpd. 45

Concentration of thrombin added = | 0 nM | 0.3 nM | 3 nM | 10 nM |

Figure 27 Plate layout for initial screening method

Results were plotted as the average change in relative fluorescence units (RFU) detected in the four repeats, *i.e.* maximum-minimum. As illustrated in Figure 28, the effect of 0.1% DMSO in the test medium is negligible and statistically both groups of data, from medium only and medium with DMSO, could not be said to have been from different sources (see appendix). Encouragingly, compounds **20, 45** and RWJ-58259, tested at a concentration of 3 μM, demonstrated antagonist activity against thrombin concentrations up to and including 10 nM. This result for novel compound **45** supports the reported findings of Zhang and co-workers that compounds in this series with a terminal ethyl- or propyl-diamine substituent coupled to an amino acid with an aromatic substituent demonstrate potent PAR-1 antagonist activity.
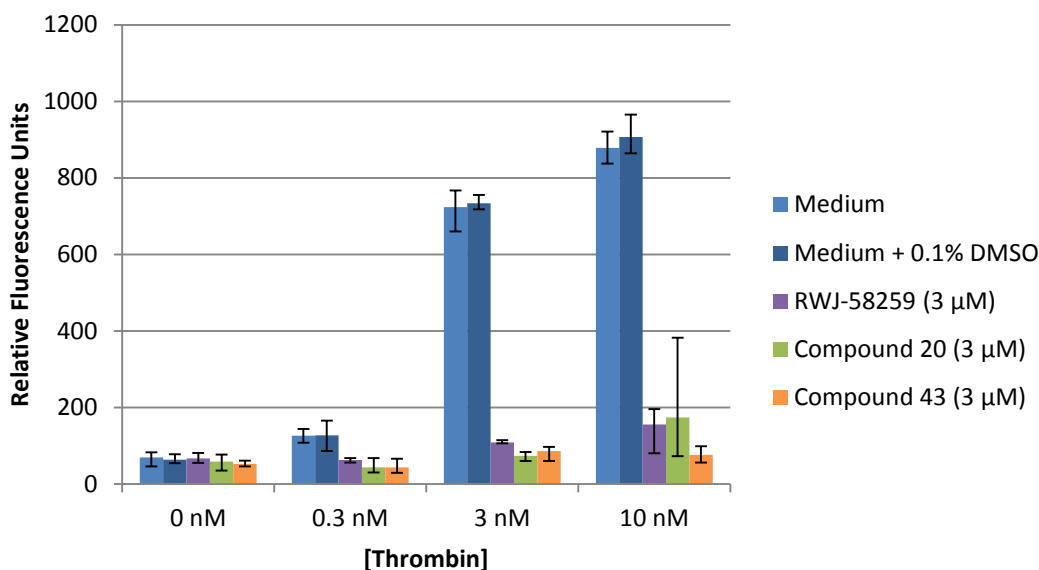
Figure 28 Change in fluorescence detected following exposure to thrombin

This initial screening method proved useful in identifying active compounds from compounds with weak or no PAR-1 antagonist activity; however, occasional variability in results motivated further development of the method. Error bars shown in Figure 28 represent the range of results given from the four repeated experiments on a single 96-well plate. At 10 nM thrombin, results for compound **20** give a much larger range than those for RWJ-58259. To be confident that variations in antagonist activity were measured accurately a method was required that would give a more detailed account of a compound's performance. Therefore, it was decided that the assay would be set up to collect data representing the effect varying antagonist concentration has on the cells' response to agonist exposure, a dose-response relationship.

Initial plate layout for the dose-response method involved only two compounds to ensure that four experimental replicates of each condition were performed (Figure 29). From the initial screening method results it was seen that a slight response was achieved at the highest thrombin concentration despite being incubated with the antagonists. Therefore, it was decided to test an antagonist concentration range of 0.1-10 μM against both 3 nM and 10 nM thrombin concentrations, to gather information on the performance window of the antagonists.

	0.1 μM	0.3 μM	1 μM	3 μM	10 μM		0.1 μM	0.3 μM	1 μM	3 μM	10 μM
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**

Concentration of thrombin added = 3 nM   10 nM

Figure 29 Initial plate layout for dose-response method

Results obtained from the assay clearly showed that the antagonist concentration span was too narrow and a broader range would be required in order to obtain a full dose-response curve. Also, from the two thrombin concentrations employed in the assay, the data plotted for 10 nM showed the beginnings of the curve representing a dose-response (B, Figure 30), whereas data points for 3 nM were all plotted at 100% inhibition (A, Figure 30). Therefore, it was decided the higher thrombin concentration would be better to achieve data for a full dose-response curve.
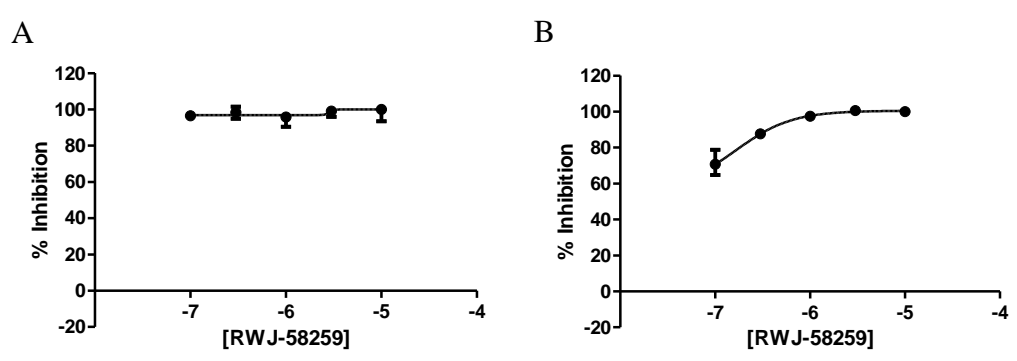
Results plotted are the mean of four repeated experiments, with error bars representing the range of data from individual experiments. A - [thrombin] = 3 nM; B - [thrombin] = 10 nM.

Figure 30 Initial dose-response results for RWJ-58259 *versus* thrombin

Based on the findings from the initial dose-response method the plate layout was adapted to cover an antagonist concentration range of 0.1 nM to 10 µM, to be tested against a thrombin concentration of 10 nM only (Figure 31).

	0.0001 µM	0.0003 µM	0.001 µM	0.003 µM	0.01 µM	0.03 µM	0.1 µM	0.3 µM	1 µM	3 µM	10 µM
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259	RWJ-58259
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**
Medium only	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**	Cmpd. **20**

Concentration of thrombin added =   10 nM

Figure 31 Modified plate layout for dose-response method

Data gathered from the Ca^{2+} mobilisation assay with the adapted plate layout were very pleasing. Percentage of full inhibition was calculated for each antagonist concentration and plotted to give full dose-response curves for RWJ-58259 and compound **20** (A and C, Figure 32). A second plate was also used to test compound **45** at the same time, which also achieved a full dose-response curve (D, Figure 32).

Analysis of the data using GraphPad Prism software calculated IC_{50} values of 320 nM and 279 nM for compound **20** and RWJ-58259, respectively. This result confirmed the quality of compound **20** compared with the earlier batch and gave confidence for its planned use as the standard against which analogues would be tested in further Ca^{2+} mobilisation assays. An IC_{50} of 362 nM was determined for compound **45** confirming this analogue too as a potent PAR-1 antagonist.
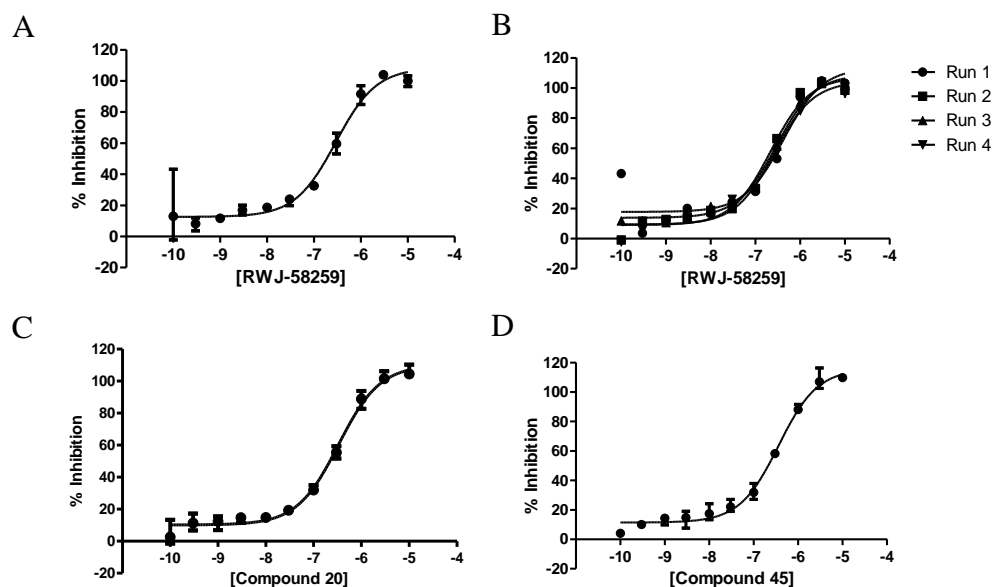
Graphs A, C and D plot mean values for data points at different concentrations for RWJ-58259, compound **20** and compound **45**, respectively, with error bars representing the range. Graph B plots each data point at different concentrations of RWJ-58259 with each horizontal row on the plate grouped into a run allowing for a representative dose-response curve to be generated for each experimental replicate (run).

Figure 32 Dose-response curves for compounds *versus* thrombin (10 nM) in HLFs

The error bars shown in graphs A, C and D (Figure 32) illustrate the range of data for percentage inhibition at any given antagonist concentration. Although, in general, the data is quite tight fitting, a broad range was seen for RWJ-58259 at a concentration of 1×10^{-10} M. However, plotting the four repeated experiments separately (B, Figure 32) shows that one data point is responsible for this wide range and although it cannot be ruled out statistically as an outlier, these plots would indicate it as an exception to the normal trend.

Since discrepancies have been reported in the responsiveness to thrombin of cells from different species, it was important to confirm that the performance of compound **20** in HLFs was repeated in MLFs. Therefore, dose-response data was gathered from the Ca^{2+} mobilisation assay using MLFs and showed compound **20** to have an increased level of potency (Figure 33). The dose-response curve shows a shift to the left by more than one log unit compared to those seen in Figure 32 for HLFs, and IC_{50} was measured at a value of 11 nM. This outcome suggests that MLFs are less responsive to thrombin, since 50% inhibition of the Ca^{2+} mobilisation response was achieve at almost a 30-fold lower concentration of compound **20**.
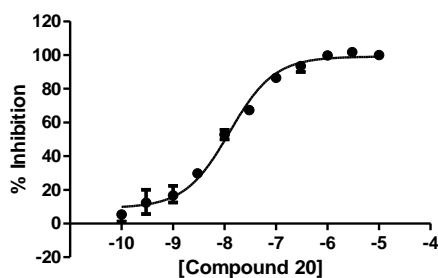
Figure 33 Dose-response curve for compound **20** *versus* thrombin (10 nM) in MLFs

During the course of this investigation further modifications were made to the plate layout in order to ensure the quality of the data was reliable and to optimise the assay with the objective of maximising the possible number of compounds tested simultaneously. Commonly encountered in cell-based assays is a phenomenon known as the temporal effect, whereby the performance of cells in wells located at the edges of a plate tends to vary unexplainably from those located away from the edges. The causes of such effects are likely to be numerous and complex, and screening departments often avoid the use of the peripheral wells to ensure variability of results is minimised. As a result of witnessing the edge effect on cells in the Ca^{2+} mobilisation assay on a few occasions, which resulted in the loss of data from the positive controls, an alternative plate layout was designed (Figure 34). Loss of the data from the positive controls, *i.e.* wells containing medium only, invalidates all other data gathered from the assay since there is no control measure of the cells' normal response to thrombin (10 nM) exposure.

This modified design to the plate layout ensured that plenty of wells representing the positive controls were placed away from the edges of the plate. An antagonist concentration range of 0.1nM to 10 μM was maintained whilst also introducing a column of wells that were not exposed to thrombin - as a negative control - in order to identify any partial agonist behaviour exhibited by compounds tested.

The number of repeat experiments for the standard (compound **20**) in each assay run was reduced to a single row due to the consistency observed in data obtained for this compound. This allowed for two other compounds to be tested, in three repeat experiments, on each 96-well plate and so maximising the throughput for the Ca^{2+} mobilisation assay.

82

10 µM	0.0001 µM	0.0003 µM	0.001 µM	0.003 µM	0.01 µM	0.03 µM	0.1 µM	0.3 µM	1 µM	3 µM	10 µM
Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a
Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a
Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a	Cmpd. 51a
Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only	Med. Only
Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20	Cmpd. 20
Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b
Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b
Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b	Cmpd. 51b

Concentration of thrombin added = 0 nM | 10 nM

Figure 34 Optimised plate layout for dose-response method

3.5.2 Biological evaluation of methylated analogues

Synthesised compounds **51a**, **51b** and **52** were evaluated for their PAR-1 antagonist activity in the Ca^{2+} mobilisation assay described earlier (see section 3.5.1, page 74). Dose-response data was obtained for each compound *versus* thrombin (10 nM) using compound **20** as the standard for setting full inhibition (Figure 35). This means the percentage inhibition determined for these compounds was calculated as a percentage of the value for full inhibition by compound **20**.

As can be seen from the dose-response curves in Figure 35 compounds **51a**, **51b** and **52** were all shown to perform as full antagonists of PAR-1. These compounds are able to block Ca^{2+} mobilisation in HLFs triggered by a high concentration of thrombin (10 nM) to the same extent as compound **20**. Compounds **51a** and **51b** were determined to have IC_{50} values of 524 nM and 548 nM, respectively. This indicates that the binding site pocket, into which RWJ-58259-like compounds sit, can accommodate substitution at the benzylic position of the LHS terminal benzylamine. This is an encouraging result since this benzylic position was predicted in the *in silico* screen as being the most probable site of metabolism (see section 3.2.2, page 61). Interestingly, data for compound **52** was calculated as giving an IC_{50} value

of 89 nM, suggesting substitution at this benzylic position results in a more potent PAR-1 antagonist.
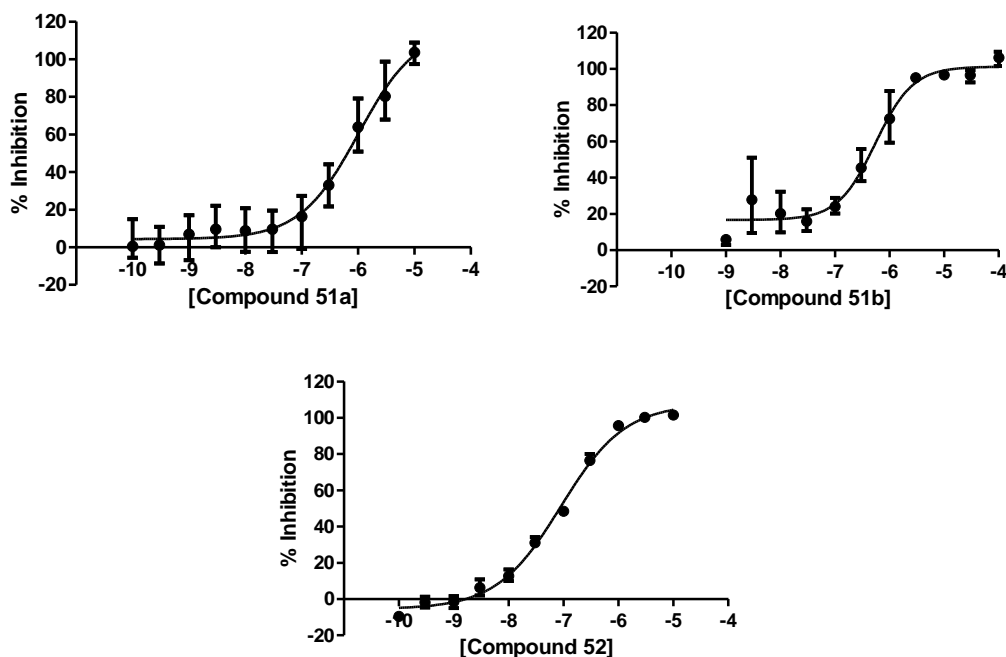
Figure 35 Dose-response curves for compounds *versus* thrombin (10 nM) in HLFs

3.5.3 Biological evaluation of fluorinated analogues

Following repeated attempts to run the Ca^{2+} mobilisation assay with fluorinated compounds **62a** and **62b** and failing to obtain reliable data, the compounds were tested at a higher concentration range. The 3-fluoropyrrolidine analogues were tested at a range of 3 nM to 300 μM against a thrombin concentration of 10 nM.

Gratifyingly, data was obtained that fitted a dose-response relationship between the antagonist and the receptor (Figure 36). Interestingly, however, it was noticed that both compound **62a** and **62b** did not achieve full inhibition of the cells' Ca^{2+} mobilisation response to thrombin (10 nM). Comparing these results with that of full antagonist compound **20**, tested on the same plate, compounds **62a** and **62b** were seen to achieve a maximum of 79% and 63% inhibition, respectively. Repeating the assay with an even higher concentration of antagonist, in the hope of achieving full inhibition, was deemed to be of no use since data obtained at a concentration of 300 μM was found not to be viable due to the baseline fluorescence detection being abnormally heightened before thrombin addition (see appendix). There is a

possibility that the antagonist might be interacting with the fluorescent dye at higher concentrations resulting in obscured results.
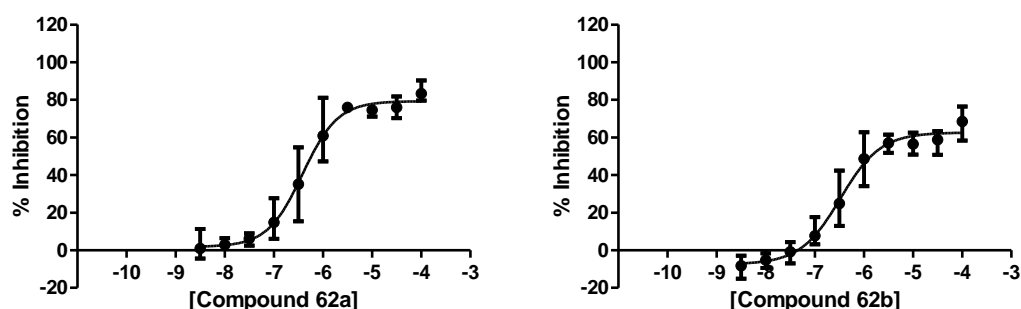
Figure 36 Dose-response curves for compounds *versus* thrombin (10 nM) in HLFs

IC_{50} values were calculated using the GraphPad Prism software, indicating the concentration at which 50% of the full pharmacological inhibitory effect of the compounds was achieved. Values of 384 nM and 332 nM were calculated for compounds **62a** and **62b**, respectively. This suggests that the pharmacologically active window for these compounds is similar to the project lead compound, RWJ-58259, despite being unable to fully inhibit PAR-1 activation in competition with high concentrations of thrombin.

3.5.4 Biological evaluation of 3-pyrrolidinol analogue

Following the interesting results obtained for fluorinated compounds **62a** and **62b**, compound **66** was tested under the suspicion that substitution of the pyrrolidine ring had detrimental effects on the PAR-1 activity of RWJ-58259-like compounds.

Tested against thrombin (10 nM) over a concentration range of 0.1 nM to 10 μM, data obtained for compound **66** was plotted and shown to follow a dose-response relationship (Figure 37). However, as with the 3-fluoropyrrolidine analogues, the 3-pyrrolidinol derivative **66** failed to achieve full inhibition of the Ca^{2+} mobilisation response in HLFs. Maximum inhibitory effect was measured at 54% of the full inhibition achieved with compound **20**. However, IC_{50} was calculated as being 266 nM, indicating the compound's pharmacological effects occur within a similar range to the other antagonists tested.
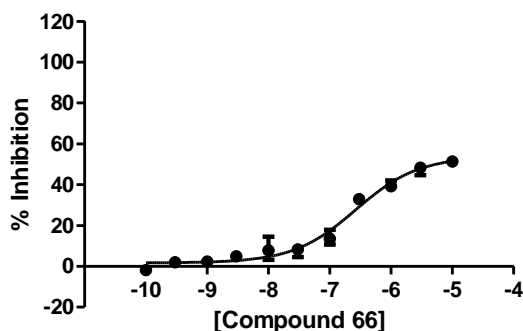
Figure 37 Dose-response curves for compound **66** *versus* thrombin (10 nM) in HLFs

3.6 Discussion

Utilising a Ca^{2+} mobilisation assay optimised to determine the PAR-1 inhibitory activity of synthesised analogues of PAR-1 antagonist RWJ-58259 *versus* thrombin, encouraging results were obtained. With the objective of developing analogues with improved metabolic stability and equivalent PAR-1 antagonist activity, compounds **51a**, **51b** and **52** were synthesised in which a methyl substituent was incorporated on the atoms predicted, by *in silico* screening, as the most probable sites of metabolism. Screening these compounds in the Ca^{2+} mobilisation assay demonstrated their PAR-1 antagonist activity to be comparable to, or even slightly improved upon, results obtained for the project lead, compound **20** (RWJ-58259). Compounds **51a** and **51b**, with a methyl group integrated at the benzylic position of the LHS fragment benzylamine, were determined to have IC_{50} values of 524 nM and 548 nM, respectively, compared with an IC_{50} of 320 nM for compound **20**. Since this position was predicted as the most probable site of metabolism by CYPs, confirmation that introduction of a substituent can be tolerated, with regards to receptor binding and inhibition, is a very encouraging result.

Promisingly, data gathered for compound **52** gave an IC_{50} value of 89 nM, improving on the potency of the project lead compound. This is an interesting result, possibly indicating a site for potential investigation into further development of analogues. SAR data published on the development of RWJ-58259 shows little variation from the *N*-1 benzyl substituent has been investigated, with substitutions on the benzene ring having been the main focus. This slight enhancement in potency suggests incorporating a substituent at this benzylic position leads to a more effective binding

of the receptor, requiring a lower concentration to restrict the conformational change promoted by the tethered ligand to activate the receptor. This could be as a result of the increased sterics reducing rotational entropy around the C-N bond of the benzyl substituent at the 1-position of the indazole, in effect locking it in a more favourable position to bind more effectively to the receptor. Alternatively, there may be a lipophilic pocket in this region of the receptor and any increase in lipophilicity of the molecule at this position would enhance the ligand receptor interaction. In this case, perhaps introducing a larger lipophilic group may improve potency further.

Methylated analogues **51a**, **51b** and **52**, along with compound **20**, have been submitted for screening in the liver microsomes assay discussed earlier (see section 3.2, page 58). Identification of differences in metabolism of these compounds compared with compound **20** will determine whether modifications made to the molecules have improved their metabolic stability. Improvement on the $t_{1/2}$ measured for compound **20**, combined with the proven antagonist activity, would present analogues **51a**, **51b** and **52** as enhanced tool compounds for use in *in vivo* investigations into PAR-1 antagonism. Alternatively, identification of additional sites of metabolism would direct synthesis of a second generation of RWJ-58259 analogues.

Biological evaluation of analogues **62a** and **62b**, incorporating a fluorine at the 3-position of the pyrrolidine ring, produced some interesting results. Despite showing activity over a similar range of concentrations as the full antagonists, these compounds failed to achieve full inhibition of thrombin induced PAR-1 activation, achieving only 63-79% of the degree of inhibition demonstrated by compound **20**. Similarly, proposed major metabolite compound **66** with a hydroxyl group at the 3-position also demonstrated a reduced ability to block receptor activation following exposure of the HLFs to thrombin. This latter result, as well as being of interest due to reduced degree of inhibition, is significant due to the observed retention of antagonist activity. Demonstrating PAR-1 antagonist activity for a likely metabolite of RWJ-58259 suggests that data from *in vivo* studies involving the parent compound may result from the biological activity of a combination of the parent compound and metabolites. This may provide a reason for the significant biological response achieved by dosing RWJ-58259 in mice despite its suspected short $t_{1/2}$. Activity may

be prolonged by the metabolites formed and so extending the desired therapeutic effect and, in the case of the mouse model of LPS-induced lung inflammation, reducing neutrophil recruitment up to 24h after dosing RWJ-58259.

With regards to the reduced antagonist effect, a common feature of these compounds is the presence of an electronegative substituent at the 3-position of the pyrrolidine ring. The proximity of such an electronegative substituent to the pyrrolidine nitrogen is likely to significantly reduce the basicity of the amine, which may affect the nature of any interactions involving this moiety as the compound sits in the binding pocket of the receptor. Precisely what effect is had on the nature of the binding and the resulting outcome that is witnessed from the receptor requires further investigation.

The pKa of a protonated pyrrolidine nitrogen can drop by more than two units as a result of introducing a fluorine atom at the 3-position and, though perhaps not quite to the same extent, a similar change could be expected with the incorporation of an oxygen atom.[180] This would reduce the amount of ionic interaction occurring with any negatively charged residues within the receptor binding site. The conformation of ring structures are also reported to be effected by the preference of fluorine substituents to be positioned co-axial with positively charged species to maximise dipole-dipole interactions.[178] A change in the conformation of the pyrrolidine ring system may be effecting interactions involving this region of the molecule. Alternatively, reduced basicity of the pyrrolidine amine may affect its capability of forming hydrogen bonding interactions within the receptor's binding pocket. Any hydrogen bonding interactions occurring with RWJ-58259 may be disrupted by the presence of a hydroxyl group at the 3-position, with the oxygen competing with the nitrogen as a hydrogen bond acceptor. This may lead to an interaction that creates a conformational change in the protein that is less capable of blocking activation. The hydroxyl group could alternatively act as a hydrogen bond donor, picking up a different interaction altogether.

The overall outcome resulting from the altered binding interaction is a reduction in the maximal inhibitory activity achieved on PAR-1 in response to thrombin exposure. However, from the biological data obtained for compounds **62a**, **62b** and **66** using the calcium ion mobilisation assay, it is difficult to determine precisely how this outcome is achieved.

From the predicted effect on binding interactions of introducing an electronegative substituent at the pyrrolidine 3-position, it would be reasonable to suggest that the compound's receptor binding affinity is reduced. This would mean an increased off-/on-rate (K_d, dissociation constant) for the compound binding to the receptor, resulting in a less competitive antagonist. However, the profiles of the dose-response curves obtained for these compounds do not match that which would be expected for an antagonist with reduced binding affinity. Reduced activity due to an increased K_d can be overcome by increasing the antagonist concentration, making the antagonist more competitive with the agonist for receptor binding by increasing the number of molecules competing. At a high enough concentration, full occupancy of the binding site is achieved by the antagonist, and thus full inhibition of receptor activation. Plotting the data would give a full dose-response curve, reaching 100% inhibition, only shifted to the right to give a lower IC_{50} value. From the plotted data for compounds **62a**, **62b** and **66** it can be seen that the curves plateau at 50-79%, meaning that increasing the concentration would have no effect on degree of inhibition achieved. Also, IC_{50} values calculated show that no significant shift to the right is observed for the dose-response curves.

A possible reason for the activity seen would be that the compounds behave as partial agonists. This would mean that at low concentration of the partial agonist the activity seen results from the full agonist, thrombin. However, as partial agonist concentration increases this compound would out-compete the full agonist for occupancy of the receptor activation site, and the resulting degree of receptor activation represents the reduced agonist activity achieved by the partial agonist. The data obtained from the calcium ion mobilisation assay could be seen to fit this reasoning, with 50-79% antagonism achieved being seen as 21-50% agonism achieved by the partial agonist out-competing the tethered ligand, uncovered by thrombin, for the receptor binding site.

In the calcium ion mobilisation assay the cells are incubated with the compound of interest for 1 hour before addition of thrombin in the FLIPR. This was done so that any stress responses triggered in receptors on the HLFs by the introduction of a foreign substance would have time to settle before addition of the thrombin and the beginning of fluorescence detection. If these compounds demonstrated some agonist

activity, upon exposure to the cells PAR-1 would have been activated, albeit to a limited extent, and calcium ion mobilisation triggered well before any measurements were recorded. What is unknown is how these compounds affect the PAR-1 surface population following addition to the cells. It is known that following activation by thrombin, or by TRAPs, PAR-1 is rapidly internalised by endocytosis and trafficked to lysosomes for degradation.[105] Alternatively, receptors are modified and recycled to the cell membrane to replenish the surface population; however, recycled PAR-1 has been shown to be modified such that activation by subsequent addition of thrombin does not occur.[181] If compounds **62a**, **62b** and **66** were performing as agonists, or partial agonists, and upon exposure to the cells activated PAR-1 leading to internalisation and modification, it might be expected that this would have consequences on the population of PAR-1 available for activation by thrombin, thus resulting in a reduced level of calcium ion mobilisation.

Another possible explanation would be that these compounds are negative allosteric modulators (NAMs).[182] Rather than binding at the site occupied by the tethered ligand upon activation, these compounds might bind at an allosteric site, which, when occupied, results in a conformational change at the orthosteric site *i.e.* the activating binding site. The conformational change affected by allosteric modulators can have a positive, negative or neutral effect on the interaction of ligands to the orthosteric site. The inhibitory effect these compounds have on the activation of PAR-1 as a result of cleavage by thrombin would make them NAMs. It could be conceived that all RWJ-58259-like compounds interact with this allosteric site and that receptor binding affinity at this site is unaffected by modifications to the pyrrolidine ring, hence the comparable IC_{50} values for the compounds tested. However, introduction of an electronegative substituent at the 3-position of the pyrrolidine ring may affect the compound's ability to achieve the conformational change required to alter the orthosteric site in a way that fully inhibits binding of the tethered ligand and activation of the receptor. This would result in a maximal effect of only partially blocking activation of the receptor following exposure to the agonist, as seen in the results obtained for compounds **62a**, **62b** and **66**.

Further *in vitro* biological testing of the synthesised compounds would assist in deciphering the reasons behind the partial inhibition witnessed in the calcium

mobilisation assay for compounds **62a**, **62b** and **66**. Establishing a baseline for all compounds, by adding them at a range of concentrations whilst measuring fluorescence in the FLIPR, would help identify any agonist response triggered. This would distinguish between a partial agonist and a NAM, since the latter would have no agonist activity yet would still only achieve partial inhibition in the conventional assay.

It is of interest to note that during the completion of this research the co-crystal structure of T4 lysozyme-stabilized PAR-1 bound to vorapaxar was solved and published by Zhang and co-workers.[159] The structure identifies interactions of the potent PAR-1 antagonist with various amino acid residues located on EC2, EC3 and transmembrane segments (TMs) 3-7. Preliminary investigations at GSK into docking of RWJ-58259 into the identified vorapaxar binding site indicated that interactions observed with vorapaxar were not picked up by RWJ-58259. This suggests the conformation of the protein may change when bound to RWJ-58259. Alternatively, this could support the notion of the compound binding allosterically and exacting its inhibitory effects by modifying the shape of the agonist binding site rather than directly blocking access to the binding site. In order to investigate the binding of RWJ-58259 to PAR-1 thoroughly a co-crystal structure of the receptor bound to the antagonist would be needed.

3.7 Conclusions and Further Work

The main objective of this research was to develop analogues of the project lead compound, RWJ-58259, maintaining its potent PAR-1 antagonist activity whilst improving on metabolic stability. The work involved investigation of the metabolic profile of RWJ-58259 to predict likely metabolites and identify possible *in vivo* pathways for their generation. *In silico* screening of the project lead using MetaSite software identified probable sites of metabolism and predicted likely metabolites formed as a result. A chemical stability study, looking at the susceptibility of RWJ-58259 to hydrolysis over a range of pH, as well as a biocatalysis approach, aiming to generate possible metabolites, was undertaken. The results suggested RWJ-58259 may not be extensively metabolised but rather a specific site on the molecule may be susceptible, suggesting that simple modifications may improve metabolic stability significantly.

The synthesis of novel analogues of RWJ-58259 was accomplished, incorporating either a methyl or fluorine substituent on atoms highlighted from the *in silico* screen as the most probable sites of metabolism. Development of a calcium ion mobilisation assay was achieved in order to evaluate biological activity of synthesised compounds on PAR-1 on HLFs with high efficiency. Encouragingly, compounds **51a**, **51b** and **52,** with a methyl substituent introduced at benzylic positions of the parent molecule, demonstrated full antagonist activity and were shown to maintain the high level of potency demonstrated by RWJ-58259. Analogues **62a** and **62b**, which incorporate a fluorine substituents at the pyrrolidine ring 3-position, were shown to have reduced antagonist activity, although the potency of these compounds was comparable to the full antagonists. This interesting result suggested an alternative activity profile for these compounds, possibly performing as partial agonists or NAMs.

Results are awaited for the synthesised compounds, as well as the project lead, from liver microsome studies carried out by collaborators at GSK. Should the data returned show a reduction in the clearance rate for the synthesised analogues compared with RWJ-58259, then an improved metabolic profile will have been achieved a new tool compound will have been generated for use in future *in vivo* studies.

Interestingly, this work has shown that a proposed major metabolite of RWJ-58259, in which the pyrrolidine ring 3-position is hydroxylated, also demonstrates PAR-1 antagonist activity at similar concentrations to the parent compound. Although activity resembles that of a partial agonist/NAM, this result is significant with regards to *in vivo* data obtained following dosing of RWJ-58259.

Future work in this area of research should establish the nature of the biological activity seen for the analogues proposed to be partial agonists/NAMs. Elucidating the reason for the reduced inhibition seen for these compounds may shed light on the true binding nature of RWJ-58259-like molecules and allow for the development of further compounds of interest. For example, should these compounds be shown to bind allosterically then, since PAR-1 has such a wide distribution in tissues throughout the body, it might be seen as more beneficial to develop a covalently bound NAM which does not fully inhibit PAR-1 activation. This may avoid possible side effects caused by full antagonists of PAR-1 with high binding affinity.

Further analogue synthesis may be guided by results from liver microsome assays identifying alternative sites of metabolism not considered in this research. Synthesis of a second generation of analogues of RWJ-58259 designed to block these newly identified sites may result in compounds with more enhanced metabolic stability. Also, synthesis of analogues investigating the space available at the benzylic position of the indazole N-1 substituent may uncover new binding interactions and improve on the potency demonstrated by RWJ-58259.

The FDA approval of vorapaxar as the first active ingredient to accomplish its therapeutic effect by inhibiting the activation of PAR-1 can be seen as a major endorsement of the significance of research in this area. Despite the receptor's widespread distribution throughout the body and varied role in many physiological processes, further investigation of the role of PAR-1 in various diseases may highlight a potential benefit for its targeting with therapeutics. Diseases such as pulmonary fibrosis, whereby topical application of a therapeutic to the afflicted area is possible, may benefit most from further investigations since such application would limit any side-effects resulting from inhibition of PAR-1 involved in normal physiological processes. The development of better tool compounds, such as those achieved during this research, may play an important role in assisting investigations

into the role of PAR-1 in various diseases and may lead to novel approaches in the treatment of some diseases.

Experimental

Organic Chemistry

General Experimental

Reactions carried out under dry conditions were done in oven-dried glassware under an argon atmosphere. Where addition of solids was required reactions were carried out in three-necked round-bottom flasks and the addition done under a positive pressure of argon. Reactions were stirred magnetically and monitored by thin-layer chromatography (TLC) using aluminium sheets coated with silica gel 60 F_{254} purchased from Merck, which were then viewed under UV irradiation ($\lambda = 254$ nm) and developed further with a potassium permanganate, ninhydrin or 4-nitrophenylhydrazine dip. Silica gel 60 (40-63 μm) from Merck was used to purify compounds by flash column chromatography. All solvents were of commercial grade, purchased from Fisher Scientific, VWR or Sigma-Aldrich and used without further purification. Petrol refers to petroleum ether (b.p. 40-60 °C). Reagents were purchased from Sigma-Aldrich, Alfa Aesar, Iris Biotech or Acros Organics and used as received from the supplier.

Instrumentation

Purification by preparative scale HPLC was carried out on a Varian Prostar (column: Phenomenex Jupiter 300 A C18, 250 x 4.6 mm, 10 μm; mobile phase gradient: 0-2 min (5%), 2-8 min (5-95%), 8-14 min (95%), 14-15 min (95-5%), 15-20 min (5%) MeCN in H_2O with 0.1% TFA). Melting points were measured using a Gallenkamp apparatus and are uncorrected. A Perkin Elmer Spectrum 100 FT-IR spectrometer or Bruker Alpha FT-IR spectrometer was used to obtain infrared spectra. ^{1}H NMR spectra of pure samples dissolved in either $CDCl_3$ or DMSO-d_6 were recorded at 400 MHz, 500 MHz and 600 MHz and ^{13}C NMR at 125 MHz and 150 MHz on a Bruker AMX400, Avance 500 and Avance III 600 spectrometer at 25 °C. Chemical shifts (δ) for ^{1}H and ^{13}C nuclei are reported relative to the residual solvent signal on a parts per million (ppm) scale. Coupling constants (*J* values) are reported in Hertz (Hz) and are reported as J_{H-H} couplings unless otherwise stated. High resolution mass spectra were

acquired for pure samples on either a Thermo Finnigan MAT900Xp (EI and CI) or Water LCT Premier XE (ES) mass spectrometer. LC/MS analysis was carried out using a Waters Acquity UPLC (column: C18 BEH 50 x 2.1 mm, 1.7 μm; mobile phase gradient (5 min): 5-95% MeCN in H_2O with 0.1% formic acid).

Experimental for chemical stability study

A solution of compound **20** (1.00 mg, 1.16 μmol) in DMSO (500 μL) was diluted 50 μL in 950 μL of buffered solutions of pH 4, 7.4 & 9. Eppendorf vials containing the different solutions were shaken at either room temperature (~20 °C) or 37 °C using a thermomixer and aliquots taken at 10 min, 3 h, 24 h, 48 h, 6 d and 7 d for analysis by LC/MS to identify degradation of the parent compound.

Experimental data

6-Nitro-3-(pyrrolidin-1-ylmethyl)-1*H*-indazole (23)[164]
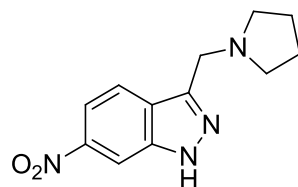
6-Nitroindole (500 mg, 3.08 mmol) was added to a stirred solution of sodium nitrite (2.13 g, 30.8 mmol) in H_2O (25 mL) and DMF (2 mL) at 18 °C. 6 M HCl (4.6 mL, 27.8 mmol) was added dropwise over 10 min and stirring continued for 3 h before diluting with EtOAc (30 mL) and extracting. Aqueous layer extracted further with EtOAc (2 x 15 mL). Organic extracts were combined and washed with H_2O (15 mL) and brine (15 mL), then dried ($MgSO_4$) and concentrated *in vacuo*. Residual dark brown solid was dissolved in CH_2Cl_2/DMF/AcOH (54/5.4/0.6 mL), to which was added pyrrolidine (548 mg, 0.643 mL, 7.70 mmol) and the reaction mixture was stirred at 18 °C for 20 min. Sodium triacetoxyborohydride (1.63 g, 7.70 mmol) was added portionwise over 10 min and stirring continued for 3 h before diluting with EtOAc (100 mL) and quenching with saturated aqueous $NaHCO_3$ (60 mL). Aqueous layer separated and extracted further with EtOAc (2 x 20 mL). Organic extracts were combined and washed with sat. $NaHCO_3$ (aq) (2 x 25 mL), H_2O (2 x 25 mL) and brine (2 x 25 mL) then dried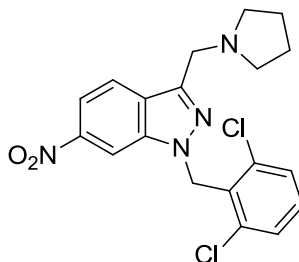 ($MgSO_4$) and concentrated *in vacuo* to leave dark brown viscous oil. Purification by flash column chromatography (0-2% MeOH/CH_2Cl_2) gave target compound (474 mg, 62% (88% crude[164])) as a dark brown oil; Rf 0.18 (10% MeOH/CH_2Cl_2); IR ν_{max} (thin film from $CHCl_3$) 3398 (N-H), 1649 (C=N), 1524 (NO_2), 1348 (NO_2), 1048-994 (C-N), 824 (ArC-H), 762 (ArC-H) cm^{-1}; ^{1}H NMR (500 MHz, DMSO-d$_6$) δ 13.49 (1H, bs, N*H*), 8.39 (1H, d, *J* = 1.3, CC*H*CNH), 8.06 (1H, d, *J* = 8.8, C*H*CHCNO$_2$), 7.91 (1H, dd, *J* = 8.8, 2.0, CHC*H*CNO$_2$), 3.99 (2H, s, C*H*$_2$CNN), 2.51-2.47 (4H, m, 2 x C*H*$_2$N(CH$_2$)CH$_2$C), 1.71-1.67 (4H, m, 2 x C*H*$_2$CH$_2$N(CH$_2$)CH$_2$); ^{13}C NMR (125 MHz, DMSO-d6) δ 146.3 (Ar*C*), 144.4 (Ar*C*), 140.0 (Ar*C*), 125.5 (Ar*C*), 122.3 (*C*HCHCNO$_2$), 114.8 (CH*C*HCNO$_2$), 107.4 (C*C*HCNO$_2$), 54.0 (2 x *C*H$_2$N(CH$_2$)CH$_2$C), 51.8 (*C*H$_2$CNN), 23.7 (2 x *C*H$_2$CH$_2$N(CH$_2$)CH$_2$C); m/z (FAB+) 247 (100%, [M+H]$^+$); HRMS $C_{12}H_{15}N_4O_2$ calcd. 247.1195, found 247.1198.

1-(2,6-Dichlorobenzyl)-6-nitro-3-(pyrrolidin-1-ylmethyl)-1*H*-indazole (24)[170]

To a stirred solution of **23** (468 mg, 1.87 mmol) in dry THF (20 mL) under argon was added 2,6-dichlorobenzyl bromide (449 mg, 1.87 mmol) followed by portionwise addition over 10 min of caesium carbonate (609 mg, 1.87 mmol). The reaction mixture was stirred at 16 °C for 24 h then diluted with H_2O (10 mL) and extracted with EtOAc (2 x 20 mL). Organic extracts were washed with H_2O (2 x 10 mL) then brine (10 mL), dried (MgSO$_4$) and concentrated *in vacuo* to give a dark brown solid residue. Purification by flash column chromatography (0-1% MeOH/CH$_2$Cl$_2$) gave target compound (453 mg, 60% (66%[170])) as a brown solid; mp 136-138 °C (138-140 °C[170]); Rf 0.25 (5% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 3081-2789 (C-H), 1521 (NO$_2$), 1438 (C-H), 1342 (NO$_2$), 1090 (C-N), 806-765 (ArC-H), 733 (C-Cl) cm^{-1}; ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.33 (1H, d, *J* = 1.6, CC*H*CNH), 8.00-7.96 (1H, d, *J* = 8.8, C*H*CHCNO$_2$ & 1H, dd, *J* = 8.8, 1.6, CHC*H*CNO$_2$), 7.41 (2H, d, *J* = 7.9, 2 x C*H*CCl), 7.29 (1H, m, C*H*CHCCl), 5.88 (2H, s, C*H*$_2$NN), 4.01 (2H, s, C*H*$_2$CNN), 2.59 (4H, m, 2 x C*H*$_2$N(CH$_2$)CH$_2$C), 1.79 (4H, m, 2 x C*H*$_2$CH$_2$N(CH$_2$)CH$_2$C); ^{13}C NMR (125 MHz, CDCl$_3$) δ 146.9 (Ar*C*), 144.5 (Ar*C*), 139.8 (Ar*C*), 137.3 (2 x *C*Cl), 131.2 (Ar*C*), 130.9 (*C*HCHCCl), 129.2 (2 x *C*HCCl), 126.9 (Ar*C*), 122.4 (*C*HCHCNO$_2$), 115.3 (CH*C*HCNO$_2$), 106.5 (C*C*HCNO$_2$), 54.6 (2 x *C*H$_2$N(CH$_2$)CH$_2$C), 52.3 (*C*H$_2$CNN), 49.1 (*C*H$_2$NN), 24.1 (2 x *C*H$_2$CH$_2$N(CH$_2$)CH$_2$C); m/z (CI+) 405 (100%, [^{35,35}M+H]$^+$), 407 (63%, [^{35,37}M+H]$^+$), 409 (12%, [^{37,37}M+H]$^+$); HRMS C$_{19}$H$_{19}$^{35}Cl$_2$N$_4$O$_2$ ([M+H]$^+$) calcd. 405.0885, found 405.0889.

1-(2,6-Dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1*H*-indazol-6-amine (25)[164]
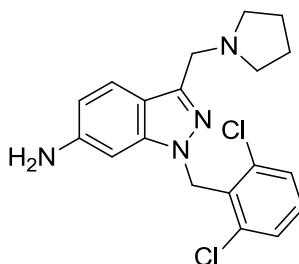
To a solution of **24** (539 mg, 1.33 mmol) in MeOH (40 mL) was added $FeCl_3.6H_2O$ (83.0 mg, 0.310 mmol) and activated charcoal (647 mg), followed by *N,N*-dimethylhydrazine (2.02 mL, 26.60 mmol). The reaction mixture was heated to reflux for 2 h then allowed to cool to ambient temperature before filtering through celite, washing through with a cold mixture of CH_2Cl_2/MeOH (90 mL, 4:1). The filtrate was concentrated *in vacuo* to leave a brown solid residue, which was purified by flash column chromatography (0-4% MeOH/CH_2Cl_2) to give the desired product (370 mg, 74% (60%[170])) as a light brown solid; mp 177-179 °C; Rf 0.1 (10% MeOH/CH_2Cl_2); IR v_{max} (neat) 3323-3214 (N-H), 2961-2480 (C-H), 1626 (C=N), 1581-1498 (ArC=C), 1438 (C-H), 1090 (C-N), 816-760 (ArC-H), 730 (C-Cl) cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 (1H, d, *J* = 8.6, C*H*CHCNH$_2$), 7.27 (2H, m, 2 x C*H*CCl), 7.15 (1H, m, C*H*CHCl), 6.53 (1H, dd, *J* = 8.6, 1.8, CHC*H*CNH$_2$), 6.49 (1H, d, *J* = 1.5, CC*H*CNH$_2$), 5.51 (2H, s, C*H$_2$*NN), 4.15 (2H, s, C*H$_2$*CNN), 4.03 (2H, bs, N*H$_2$*), 2.97 (4H, m, 2 x C*H$_2$*N(CH$_2$)CH$_2$C), 1.73 (4H, bs, 2 x C*H$_2$*CH$_2$N(CH$_2$)CH$_2$C); ^{13}C NMR (100 MHz, $CDCl_3$) δ 146.6 (Ar*C*), 142.4 (Ar*C*), 136.82 (Ar*C*), 136.78 (2 x *C*Cl), 131.5 (Ar*C*), 130.1 (*C*HCHCCl), 128.6 (2 x *C*HCCl), 120.9 (*C*HCHCNH$_2$), 117.4 (Ar*C*), 113.4 (CH*C*HCNH$_2$), 91.9 (C*C*HCNH$_2$), 52.3 (2 x *C*H$_2$N(CH$_2$)CH$_2$C), 48.7 (*C*H$_2$CNN), 47.9 (*C*H$_2$NN), 23.7 (2 x *C*H$_2$CH$_2$N(CH$_2$)CH$_2$C); m/z (CI+) 375 (27%, [^{35,35}M+H]$^+$), 377 (17%, [^{35,37}M+H]$^+$), 379 (3%, [^{37,37}M+H]$^+$), 84 (100%, [1-methylpyrrolidine]$^+$), 305 (8%, [M+H-C$_4$H$_8$N]$^+$); HRMS C$_{19}$H$_{21}$^{35}Cl$_2$N$_4$ ([M+H]$^+$) calcd. 375.1143, found 375.1137.

tert-Butyl {(S)-3-[(S)-2-amino-3-(3,4-difluorophenyl)propanamido]-4-(benzylamino)-4-oxobutyl}carbamate (33)
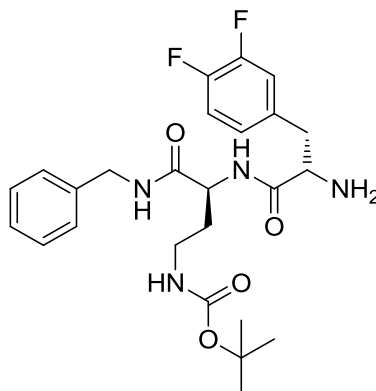
i) To a stirred solution of (S)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-4-[(tert-butoxycarbonyl)amino]butanoic acid (540 mg, 1.24 mmol), HBTU (430 mg, 1.14 mmol) and HOBt (152 mg, 1.14 mmol) in DMF (25 mL) was added DIPEA (0.20 mL, 1.14 mmol). After 3 min of stirring, benzylamine (0.11 mL, 1.04 mmol) was added and stirring continued at 18 °C for 5 h. The reaction mixture was concentrated *in vacuo* and the white residue taken up in CHCl$_3$/EtOAc (60 mL, 1:1) and washed with H$_2$O (15 mL), 10% citric acid solution (15 mL), sat. NaHCO$_3$ solution (15 mL) and brine (15 mL). The organic layer was concentrated *in vacuo* and the residue purified by flash column chromatography (25% EtOAc/Pet) to give the Fmoc-protected amide intermediate (Rf 0.13 (45% EtOAc/Pet)) as a colourless oil. This was dissolved in DMF (10 mL) and stirred with Et$_2$NH (1.28 mL, 12.4 mmol) at 18 °C for 3 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in CHCl$_3$/EtOAc (40 mL, 1:1) and washed with H$_2$O (2 x 15 mL), brine (15 mL) and sat. LiCl solution (4 x 15 mL). The organic layer was concentrated *in vacuo* and the residue purified by flash column chromatography (0-2% MeOH/CH$_2$Cl$_2$) to give the deprotected amide intermediate (259 mg, 81%) (Rf 0.2 (10% MeOH/CH$_2$Cl$_2$).

ii) To a stirred solution of (S)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-3-(3,4-difluorophenyl)propanoic acid (410 mg, 0.97 mmol), HBTU (338 mg, 0.89 mmol) and HOBt (120 mg, 0.89 mmol) in DMF (20 mL) was added DIPEA (0.16 mL, 0.89 mmol). After 3 min of stirring, a solution of the deprotected amide intermediate (250 mg, 0.81 mmol), from step i), in DMF (5 mL) was added and

stirring continued at 18 °C for 16 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in EtOAc (30 mL) and washed with H_2O (10 mL), 10% citric acid solution (10 mL), sat. $NaHCO_3$ solution (10 mL) and brine (10 mL). The organic layer was concentrated *in vacuo* and the residue purified by flash column chromatography (25-80% EtOAc/Pet) to give the Fmoc-protected title compound (Rf 0.19 (50% EtOAc/Pet)), which was dissolved in DMF (10 mL) and stirred with Et_2NH (1.00 mL, 9.69 mmol) at 18 °C for 3 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in EtOAc (30 mL) and washed with H_2O (2 x 10 mL), brine (10 mL) and sat. LiCl solution (4 x 15 mL). The organic layer was concentrated *in vacuo* and the residue purified by flash column chromatography (0-2% $MeOH/CH_2Cl_2$) to give the title compound (227 mg, 57% (84% crude[164])); as a white solid; mp 218-224 °C (decomp.); Rf 0.56 (10% $MeOH/CH_2Cl_2$); IR ν_{max} (neat) 3286 (N-H), 3062-2926 (C-H), 1678 (C=O), 1634 (C=O), 1515 (ArC=C), 1453-1365 (C-H), 1279 (C-O), 1165 (C-N), 1116 (C-F), 770-694 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 8.48 (1H, t, *J* = 6.0, N*H*CH$_2$Ph), 8.12 (1H, d, *J* = 7.5, CHN*H*COCH), 7.33-7.22 (7H, m, ArH), 7.04 (1H, m, ArH), 6.74 (1H. t, *J* = 5.7, N*H*CH$_2$CH$_2$), 4.32 (1H, m (obscured), C*H*NHCOCH), 4.29 (2H, d, *J* = 6.0, C*H$_2$*NHCO), 3.44 (1H, dd, *J* = 8.3, 4.9, C*H*NH$_2$), 2.96-2.86 (2H, m, C*H$_2$*CH$_2$CHNH & 1H, apparent dd, *J* = 13.6, 4.9, C(H)*H*CHNH$_2$), 2.64 (1H, dd, *J* = 13.6, 8.3, C(*H*)HCHNH$_2$), 1.88 (2H, bs, N*H$_2$*), 1.78 (1H, m, CH$_2$C(H)*H*CHNH), 1.63 (1H, m, CH$_2$C(*H*)HCHNH), 1.38 (9H, s, 3 x C*H$_3$*); ^{13}C NMR (150 MHz, DMSO-d$_6$) δ 174.1 (NH*C*OC), 171.2 (NH*C*OC), 155.5 (NH*C*OO), 149.9-148.9 (dd, *J* = 137.1, 12.5, *C*F), 148.3-147.3 (dd, *J* = 135.9, 12.5, *C*F), 139.2 (Ar*C*), 136.6 (Ar*C*), 128.3 (Ar*C*H), 127.1 (Ar*C*H), 126.9 (Ar*C*H), 126.1 (Ar*C*H), 118.3 (Ar*C*H), 118.7 (Ar*C*H), 116.9 (Ar*C*H), 116.8 (Ar*C*H), 77.7 (*C*(CH$_3$)$_3$), 55.8 (*C*HNH$_2$), 50.3 (*C*HNHCOCH), 42.0 (Ph*C*H$_2$NH), 39.8 (*C*H$_2$CHNH$_2$), 36.8 (*C*H$_2$CH$_2$CHNH), 32.9 (*C*H$_2$CHNHCOCH), 28.3 (3 x *C*H$_3$); m/z (CI+) 491 (20%, [M+H]$^+$), 391 (100%, [M+2H-Boc]$^+$); HRMS $C_{25}H_{33}F_2N_4O_4$ ([M+H]$^+$) calcd. 491.2470, found. 491.2467.

(2S,2'S)-Dimethyl 2,2'-[carbonylbis(azanediyl)]bis(3-phenylpropanoate) (37)
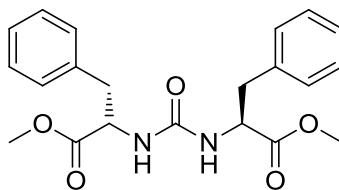
To a stirred solution of (S)-methyl 2-amino-3-phenylpropanoate (19.2 mg, 0.107 mmol) in THF (dry) (3 mL) at 0 °C was added triphosgene (2.70 mg, 8.80 μmol) followed by Et$_3$N (15.0 μL, 0.107 mmol). The reaction mixture was stirred at 0 °C for 5 min before allowing to reach ambient temperature (18 °C) and stirring for a further 1 h. The reaction mixture was diluted with EtOAc (30 mL) and washed with H$_2$O (2 x 15 mL) then brine (15 mL). Organics were dried (MgSO$_4$), concentrated *in vacuo* and the residue purified by flash column chromatography (33% EtOAc/Petrol) to give the title compound (9.30 mg, 90%); as a white solid; mp 158-159 °C (155-157 °C[183]); R$_f$ 0.38 (33% EtOAc/Petrol); IR ν$_{max}$ (neat) 3345 (N-H), 3059-2947 (C-H), 1739 (C=O), 1711 (C=O), 1631 (C=O), 1564 (ArC=C), 1496-1437 (C-H), 1174 (C-F), 749-696 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.27-7.07 (10H, m, ArH), 5.09 (2H, d, *J* = 8.1, 2 x N*H*), 4.77 (2H, dt, *J* = 7.9, 5.8, 2 x C*H*NH), 3.66 (6H, s, 2 x C*H$_3$*), 3.03 (4H, m, 2 x C*H$_2$*); ^{13}C NMR (150 MHz, CDCl$_3$) δ 173.1 (2 x *C*O$_2$CH$_3$), 156.2 (2 x *C*O), 136.1 (2 x Ar*C*), 129.5 (4 x Ar*C*H), 128.6 (4 x Ar*C*H), 127.1 (2 x Ar*C*H), 54.1 (2 x *C*HNH), 52.4 (2 x *C*H$_3$), 38.6 (2 x *C*H$_2$); m/z (CI+) 385 (100%, [M+H]$^+$); HRMS C$_{21}$H$_{25}$N$_2$O$_5$ ([M+H]$^+$) calcd. 385.1764, found. 385.1754.

(S)-methyl 2-{3-[1-(2,6-dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1H-indazol-6-yl]ureido}-3-phenylpropanoate (35)
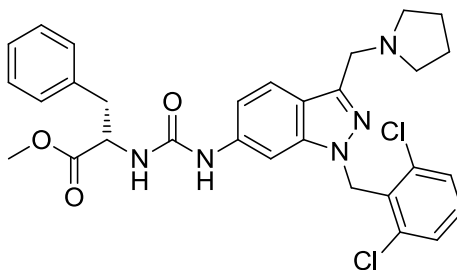
General procedure (Method A): to a stirred solution of **25** (20.0 mg, 53.3 μmol) in dry THF (3 mL) cooled to 0 °C was added, under dry conditions, triphosgene (5.40

mg, 17.6 μmol) and DMAP (13.4 mg, 107 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min under argon. The reaction mixture was cooled to 0 °C for 5 min before addition of a solution of (S)-methyl 2-amino-3-phenylpropanoate (4.80 mg, 26.8 μmol) in dry THF (0.5 mL). The reaction mixture was left to reach ambient temperature and stirred for 1 h under argon. The reaction mixture was diluted with EtOAc (8 mL) and washed with H_2O (2 x 3 mL) then brine (3 mL), dried ($MgSO_4$) and concentrated *in vacuo*. The crude material was analysed by 1D and 2D ^1H NMR spectroscopy to identify peaks for title compound as follows: ^1H NMR (600 MHz, $CDCl_3$) δ 7.75 (1H, bs, N*H*CONHCH), 7.68 (1H, s, CC*H*CNH), 7.49 (1H, d, *J* = 8.5, C*H*CHCNH), 7.34-7.11 (8H, m, ArH), 6.70 (1H, d, *J* = 8.7, CHC*H*CNH), 5.95 (1H, d, *J* = 7.5, CHN*H*CONH), 5.54 (2H, s, C*H₂*NN), 4.82 (1H, dt, *J* = 7.3, 6.2, C*H*NHCONH), 3.87 (2H, s, C*H₂*CNN), 3.68 (3H, s, C*H₃*), 3.08 (2H, ddd, *J* = 47.1, 13.7, 5.7, C*H₂*CHNH), 2.56 (4H, bs, C*H₂*N(CH₂)CH₂C), 1.69 (4H, bs, C*H₂*CH₂N(CH₂)CH₂C); m/z (ESI+) 580 (100%, [^{35,35}M+H]$^+$), 582 (60%, [^{35,37}M+H]$^+$), 584 (12%, [^{37,37}M+H]$^+$), 548 (80%, [M-OMe]$^+$); HRMS $C_{30}H_{32}{}^{35}Cl_2N_5O_3$ ([M+H]$^+$) calcd. 580.1882, found. 580.1873.

Method B: as general procedure using CH_2Cl_2 in place of THF.

Method C: as general procedure with sonication (5 min) of solution before addition of triphosgene and DMAP.

Method D: as general procedure except the sonicated (5 min) solution of **25** (20.0 mg, 53.3 μmol) in dry THF (2.5 mL) was added dropwise to a cooled (0 °C) solution of triphosgene (5.40 mg, 17.6 μmol) and DMAP (13.4 mg, 107 μmol) in dry THF (0.5 mL).

Method E: as general procedure adding (S)-methyl 2-amino-3-phenylpropanoate (10.0 mg, 56.0 μmol) - 1.05 equivalence in place of 0.5 equivalence.

1,3-Bis[1-(2,6-dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1*H*-indazol-6-yl]urea (37) - not isolated
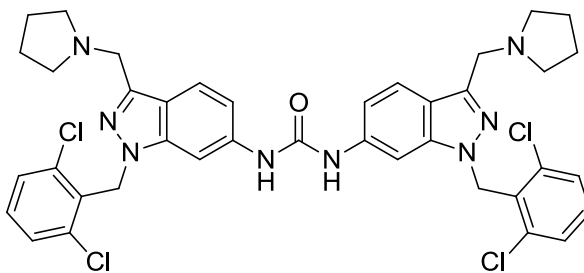
Identifiable peaks: ¹H NMR (600 MHz, CDCl₃) δ 7.92 (2H, s, 2 x CC*H*CNH), 7.42 (2H, obscured d, C*H*CHCNH), 6.84 (2H, d, *J* = 8.5, 2 x CHC*H*CNH).

***tert*-Butyl (2-aminoethyl)carbamate (39)**[184]
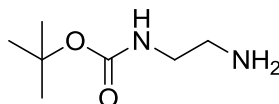
To a solution of ethylene diamine (3.35 mL, 50.0 mmol) in CHCl₃ (10 mL) stirred at 18 °C was added a solution of di-tert-butyl dicarbonate (1.09 g, 5.00 mmol) in CHCl₃ (30 mL) dropwise over 2 h. After stirring for 20 h the reaction mixture was concentrated *in vacuo* to leave a white solid, which was dissolved in sat. K₂CO₃ solution (50 mL) and extracted with CH₂Cl₂ (3 x 70 mL). Organics were combined, dried (MgSO₄), concentrated *in vacuo* and left to dry under vacuum for 6 h to give the title compound (710 mg, 87% (quant.[184])); as a slightly yellow oil; Rf 0.19 (10% MeOH/CH₂Cl₂); IR ν$_{max}$ (neat) 3358 (N-H), 2976 (C-H), 2933 (C-H), 1686 (C=O), 1517 (C=O), 1391 (C-H), 1365 (C-H), 1272 (C-O), 1249 (C-O), 1164 (C-N) cm⁻¹; ¹H NMR (600 MHz, CDCl₃) δ 4.95 (1H, bs, N*H*), 3.15 (2H, apparent dt, *J* = 5.8, 5.5, C*H₂*NH), 2.77 (2H, t, *J* = 5.8, C*H₂*NH₂), 1.42 (9H, s, 3 x C*H₃*), 1.17 (2H, bs, N*H₂*); ¹³C NMR δ 156.3 (*C*O), 79.3 (*C*(CH₃)₃), 43.5 (*C*H₂NH), 42.0 (*C*H₂NH₂), 28.5 (3 x *C*H₃); m/z (CI+) 161 (100%, [M+H]⁺); HRMS C₇H₁₇N₂O₂ ([M+H]⁺) calcd. 161.1290, found. 161.1289.

tert-Butyl (2-{(*S*)-2-[(*S*)-2-amino-3-phenylpropanamido]-3-phenylpropanamido}ethyl)carbamate (44)
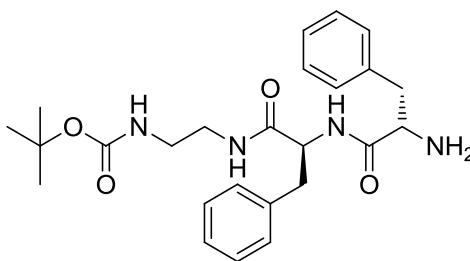
i) To a stirred solution of (*S*)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-3-phenylpropanoic acid (504 mg, 1.30 mmol) in MeCN (50 mL) was added HOBt (172 mg, 1.28 mmol) followed by DCC (526 mg, 2.55 mmol). After 5 min of stirring a solution of tert-butyl (2-aminoethyl)carbamate (200 mg, 1.25 mmol) in MeCN (2 mL) was added and stirring was continued at 18 °C for 3 h. The reaction mixture was cooled on ice, the precipitate filtered off and washed with cold MeCN (50 mL). The white solid collected was suspended in MeCN (50 mL) and stirred with Et$_2$NH (1.35 mL, 13.0 mmol) for 2.5 h. The reaction mixture was concentrated *in vacuo*, the residue taken up in EtOAc (25 mL) and washed with H$_2$O (3 x 10 mL). The organic layer was dried (MgSO$_4$), concentrated *in vacuo* and the residue purified by flash column chromatography (0-5% MeOH/CH$_2$Cl$_2$) to give the deprotected amide intermediate (190 mg, 49%) as a white solid (Rf 0.17 (10% MeOH/CH$_2$Cl$_2$)).

ii) To a stirred solution of (*S*)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-3-phenylpropanoic acid (127 mg, 0.33 mmol) in MeCN (30 mL) was added HOBt (43.0 mg, 0.32 mmol) followed by EDC (0.30 mL, 1.69 mmol). After 5 min of stirring a solution of the amide intermediate (101 mg, 0.33 mmol), from step i), in MeCN (2 mL) was added and stirring continued at 18 °C for 20 h. The reaction mixture was cooled on ice and the precipitate filtered and washed with cold MeCN (50 mL). The white solid was suspended in MeCN (30 mL) and stirred with Et$_2$NH (1.35 mL, 13.0 mmol) for 16 h. The reaction mixture was concentrated *in vacuo* and the residue triturated in Petrol until flaky pale solid achieved, which was then dissolved in CHCl$_3$ (10 mL), dried (MgSO$_4$) and concentrated *in vacuo*. The residue was purified by flash column chromatography (0-5% MeOH/CH$_2$Cl$_2$) to give the desired product (75.0 mg, 50%); as a yellow oil; Rf 0.08 (10% MeOH/CH$_2$Cl$_2$); IR v$_{max}$ (neat) 3295 (N-H), 2977 (C-H), 2933 (C-H), 1679 (C=O), 1649 (C=O), 1526

(C=O), 1366 (C-H), 1255 (C-O), 1170 (C-N), 743 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.82 (1H, d, J = 7.5, NHCOCHNH$_2$), 7.31-7.15 (10H, m, ArH), 6.79 (1H, s, NHCOO^{t}Bu), 4.88 (1H, s, NHCOCHNH), 4.65 (1H, m, CHNH), 3.61 (1H, m, CHNH$_2$), 3.31-3.19 (2H, m, CH_2NHCOO^{t}Bu), 3.17-3.01 (2H, m, CH_2NHCOCH, 2H, m, CH_2CHNH & 1H, m, C(H)HCHNH$_2$), 2.46 (1H, dd, J = 13.2, 10.0, C(H)HCHNH$_2$), 1.67 (2H, bs, NH_2), 1.39 (9H, s, 3 x CH_3); ^{13}C NMR (150 MHz, CDCl$_3$) δ 174.6 (COCHNH$_2$), 171.4 (COCHNH), 156.8 (COO^{t}Bu), 137.7 (ArC), 136.9 (ArC), 129.5 (ArCH), 129.4 (ArCH), 128.9 (ArCH), 128.7 (ArCH), 127.1 (ArCH), 127.0 (ArCH), 79.7 (C(CH$_3$)$_3$), 56.4 (CHNH$_2$), 54.2 (CHNH), 40.8 (CH$_2$CHNH$_2$), 40.7 (CH$_2$NHCOO^{t}Bu), 40.3 (CH$_2$NHCO), 38.3 (CH$_2$CHNH), 28.5 (3 x CH$_3$): m/z (ESI+) 477 (100%, [M+Na]$^+$), 355 (50%, [M+H-Boc]$^+$); HRMS C$_{25}$H$_{34}$N$_4$O$_4$Na ([M+Na]$^+$) calcd. 477.2478, found. 477.2484.

(S)-N-(2-aminoethyl)-2-({S}-2-{3-[1-(2,6-dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1H-indazol-6-yl]ureido}-3-phenylpropanamido)-3-phenylpropanamide dihydrochloride salt (45)
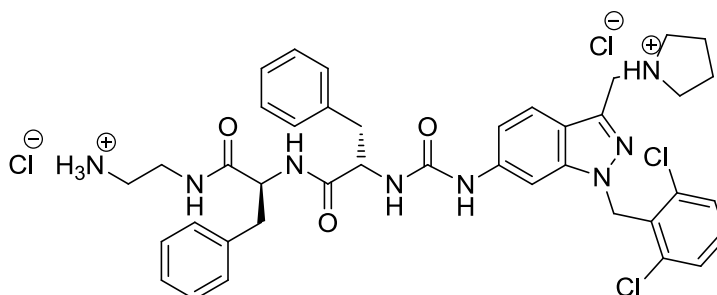
To a sonicated (5 min), then stirred solution of **25** (100 mg, 330 μmol) in dry THF (20 mL), under dry conditions and cooled to 0 °C, was added together triphosgene (28.0 mg, 116 μmol) and DMAP (66.0 mg, 660 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 1 h under argon. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **44** (75.0 mg, 165 μmol) in dry THF (10 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then diluted with EtOAc (100 mL) and washed with H$_2$O (2 x 20 mL) then brine (20 mL), dried (MgSO$_4$) and concentrated *in vacuo*. The pale orange residue was suspended in CH$_2$Cl$_2$ (10 mL), sonicated, filtered and rinsed with CH$_2$Cl$_2$ (10 mL) to give the Boc-protected intermediate as an off-white solid (75.0 mg, 53%). Solid (63.0 mg, 73.6 μmol) was

taken up in a solution of HCl (4N) in dioxane (2 mL), MeOH (2 drops) added to assist solvation and stirred for 2 h before concentrating *in vacuo*. The product was suspended in H_2O and lyophilised to give the title compound (41.0 mg, 67%); as a white solid; mp 164-168 °C (decomp.); Rf 0.06 (10% MeOH/CH_2Cl_2); IR v_{max} (neat) 3291 (N-H), 3057 (N-H), 2956 (C-H), 1637 (C=O), 1584 (C=O), 1535 (C=O), 1493 (C-H), 1435 (C-H), 1228 (C-N), 1160 (C-N), 740 (C-Cl), 697 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 10.66 (1H, bs, N$^+$*H*CH$_2$CNN), 9.25 (1H, s, N*H*CONHCH), 8.47 (1H, d, *J* = 8.1, CHN*H*COCH), 8.34 (1H, t, *J* = 5.7, N*H*CH$_2$CH$_2$N$^+$*H*$_3$), 8.10 (3H, bs, N$^+$*H*$_3$), 8.01 (1H, d, *J* = 0.8, CC*H*CNH), 7.84 (1H, d, *J* = 8.7, CC*H*CHCNH), 7.54 (2H, d, *J* = 8.1, 2 x C*H*CCl), 7.43 (1H, m, C*H*CHCCl), 7.27-7.12 (9H, m, ArH), 7.01 (1H, dd, *J* = 8.7, 1.3, CCHC*H*CNH), 6.55 (1H, d, *J* = 7.9, CHN*H*CONH), 5.66 (2H, s, C*H*$_2$NN), 4.59 (2H, s, C*H*$_2$CNN), 4.52 (1H, dt, *J* = 4.9, 8.3, C*H*NHCONH), 4.48 (1H, dt, *J* = 5.3, 8.7, C*H*NHCOCH), 3.36 (2H distorted by H$_2$O peak, m, 2 x C(H)*H*N$^+$H(CH$_2$)CH$_2$C), 3.34-3.26 (2H, m, CH$_2$C*H*$_2$NHCO), 3.11-3.01 (2H, m, 2 x C(*H*)HN$^+$H(CH$_2$)CH$_2$C & 1H, m, CC(H)*H*CHNHCOCH & 1H, m, CC(H)*H*CHNHCONH), 2.87 (1H, dd, *J* = 13.7, 9.2, CC(*H*)HCHNHCOCH), 2.83-2.74 (1H, m, CC(H)*H*CHNHCONH & 2H, m, C*H*$_2$CH$_2$NHCO), 1.91-1.85 (2H, m, 2 x C(H)*H*CH$_2$N$^+$H(CH$_2$)CH$_2$C), 1.83-1.76 (2H, m, 2 x C(*H*)HCH$_2$N$^+$H(CH$_2$)CH$_2$C); ^{13}C NMR (150 MHz, DMSO-d$_6$) δ 171.4 (NH*C*OCH), 154.7 (NH*C*OCH), 141.5 (NH*C*ONH), 139.8 (Ar*C*), 137.8 (Ar*C*), 137.6 (Ar*C*), 136.1 (2 x *C*Cl), 135.6 (Ar*C*), 131.5 (Ar*C*), 131.0 (*C*HCHCCl), 129.5 (Ar*C*H), 129.3 (Ar*C*H), 128.8 (2 x *C*HCCl), 128.1 (Ar*C*H), 126.3 (Ar*C*H), 120.7 (*C*HCHCNH), 117.7 (Ar*C*), 114.5 (CH*C*HCNH), 96.0 (C*C*HCNH), 54.3 (*C*HNHCOCH), 53.9 (*C*HNHCONH), 52.6 (2 x *C*H$_2$N$^+$H(CH$_2$)CH$_2$C), 47.7 (*C*H$_2$CNN), 47.3 (*C*H$_2$NN), 38.3 (*C*H$_2$CHNHCONH & *C*H$_2$CH$_2$NHCO), 37.6 (*C*H$_2$CHNHCOCH), 36.4 (CH$_2$*C*H$_2$NHCO), 22.7 (2 x CH$_2$*C*H$_2$N$^+$H(CH$_2$)CH$_2$C); m/z (ESI+) 755 (100%, [^{35,35}M+H]$^+$), 757 (77%, [^{35,37}M+H]$^+$), 759 (9%, [^{37,37}M+H]$^+$); HRMS C$_{40}$H$_{45}$^{35}Cl$_2$N$_8$O$_3$ ([M+H]$^+$) calcd. 755.2992, found. 755.2996.

(*S*)-4-Amino-*N*-benzyl-2-({*S*}-2-{3-[1-(2,6-dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1*H*-indazol-6-yl]ureido}-3-{3,4-difluorophenyl}propanamido)butanamide dihydrochloride salt (20)
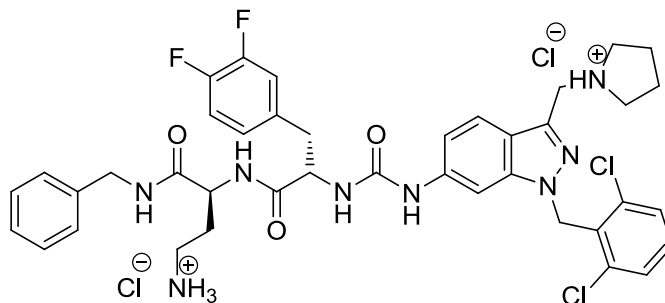
To a sonicated (5 min), then stirred solution of **25** (40.0 mg, 0.106 mmol) in dry THF (9 mL), under dry conditions and cooled to 0 °C, was added together triphosgene (11.0 mg, 37.1 μmol) and DMAP (26.0 mg, 212 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min under argon. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **33** (26.0 mg, 53.0 μmol) in dry THF (4 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then concentrated *in vacuo* and the residue purified by flash column chromatography (3-5% MeOH.NH$_3$/CH$_2$Cl$_2$) to give the Boc-protected intermediate (25.0 mg), Rf 0.35 (10% MeOH/CH$_2$Cl$_2$). The off-white solid was suspended in a solution of HCl (4N) in dioxane (2 mL) and stirred for 2.5 h before concentrating *in vacuo*, taking up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (20.0 mg, 53%); as a white solid; mp 176-184 °C; Rf 0.06 (10% MeOH/CH$_2$Cl$_2$); IR ν$_{max}$ (neat) 3280 (N-H), 3059 (N-H), 2926 (C-H), 1634 (C=O), 1585 (C=O), 1537 (C=O), 1435 (C-H), 1209 (C-N), 769 (C-Cl), 696 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 10.43 (1H, bs, N$^+$*H*CH$_2$CNN), 9.26 (1H, s, N*H*CONHCH), 8.64-8.63 (1H, m, N*H*CH$_2$Ph & 1H, m, CHN*H*COCH), 7.97 (1H, s, CC*H*CNH), 7.95 (3H, bs, N$^+$*H$_3$*), 7.80 (1H, d, *J* = 8.7, CC*H*CHCNH), 7.54 (2H, d, *J* = 8.1, 2 x C*H*CCl), 7.43 (1H, m, C*H*CHCCl), 7.35-7.21 (7H, m, ArH), 7.09-7.07 (1H, m, ArH), 7.02 (1H, dd, *J* = 8.7, 1.1, CHC*H*CNH), 6.67 (1H, d, *J* = 7.7, CHN*H*CONH), 5.61 (2H, s, C*H$_2$*NN), 4.60-4.55 (1H, m, C*H*NHCONH & 2H, m, C*H$_2$*CNN), 4.42 (1H, apparent dt, *J* = 6.0, 7.9, C*H*NHCOCH), 4.33 (2H, m, C*H$_2$*NHCO), 3.38 (2H distorted by H$_2$O peak, m, 2 x C(H)*H*N$^+$H(CH$_2$)CH$_2$C), 3.11-3.06 (2H, m, 2 x C(*H*)HN$^+$H(CH$_2$)CH$_2$C & 1H, m, C(H)*H*CHNHCONH), 2.89-2.78 (1H, m, C(*H*)HCHNHCONH & 2H, m,

C*H*₂CH₂CHNH), 2.11-1.76 (2H, m, C*H*₂CHNHCOCH & 4H, m, 2 x C*H*₂CH₂N⁺H(CH₂)CH₂C); ¹³C NMR (150 MHz, DMSO-d₆) δ 171.5 (NH*C*OCH), 170.5 (NH*C*OCH), 154.8 (NH*C*ONH), 148.2 (d, J_{CF} = 12.5, *C*F), 147.5 (d, J_{CF} = 12.5, *C*F), 141.4 (Ar*C*), 139.6 (Ar*C*), 139.1 (Ar*C*), 136.1 (2 x *C*Cl), 131.5 (Ar*C*), 131.0 (*C*HCHCCl), 128.8 (2 x *C*HCCl), 128.5 (Ar*C*), 128.3 (2 x Ar*C*H), 127.5 (Ar*C*), 127.1 (2 x Ar*C*H), 126.9 (Ar*C*H), 126.3 (Ar*C*H), 120.7 (*C*HCHCNH), 118.3 (d, *J* = 16.7, Ar*C*H), 117.7 (Ar*C*), 117.0 (d, *J* = 16.1, Ar*C*H), 114.5 (CH*C*HCNH), 96.2 (C*C*HCNH), 53.8 (*C*HNHCONH), 52.8 (2 x *C*H₂N⁺H(CH₂)CH₂C), 50.6 (*C*HNHCOCH), 48.0 (*C*H₂CNN), 47.2 (*C*H₂NN), 42.1 (*C*H₂NHCO), 37.5 (*C*H₂CHNHCONH), 36.2 (*C*H₂CH₂CHNH), 30.1 (*C*H₂CHNHCOCH), 22.7 (2 x *C*H₂CH₂N⁺H(CH₂)CH₂C); m/z (ESI+) 791 (100%, [³⁵,³⁵M+H]⁺), 793 (80%, [³⁵,³⁷M+H]⁺), 795 (9%, [³⁷,³⁷M+H]⁺); HRMS C₄₀H₄₃³⁵Cl₂F₂N₈O₃ calcd. 791.2803, found 791.2819.

***tert*-Butyl ([*S*]-3-[(*S*)-2-amino-3-(3,4-difluorophenyl)propanamido]-4-oxo-4-{[(*R*)-1-phenylethyl]amino}butyl)carbamate (56a)**
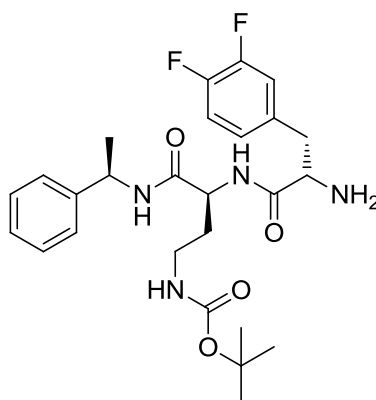
i) To a stirred solution of (*S*)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-4-[(tert-butoxycarbonyl)amino]butanoic acid (100 mg, 0.23 mmol), HBTU (80.0 mg, 0.21 mmol) and HOBt (28.0 mg, 0.21 mmol) in DMF (5 mL) was added DIPEA (37.0 μL, 0.21 mmol). After 3 min of stirring, (*R*)-1-phenylethanamine (24.0 μL, 0.19 mmol) was added and stirring continued at 18 °C for 5 h. The reaction mixture was concentrated *in vacuo* and the white residue taken up in CHCl₃/EtOAc (60 mL, 1:1) and washed with H₂O (15 mL), 10% citric acid solution (15 mL), sat. NaHCO₃ solution (15 mL) and brine (15 mL). The organic layer was concentrated *in vacuo* and the residue dissolved in DMF (5 mL) and stirred with Et₂NH (0.24 mL,

2.30 mmol) at 18 °C for 3 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in CHCl₃/EtOAc (40 mL, 1:1) and washed with H_2O (2 x 15 mL), brine (15 mL) and sat. LiCl solution (4 x 15 mL). The organic layer was concentrated *in vacuo* and the residue purified by flash column chromatography (0-2% MeOH/CH₂Cl₂) to give the deprotected amide intermediate (48.0 mg, 79%) (Rf 0.63 (10% MeOH/CH₂Cl₂).

ii) To a stirred solution of (*S*)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-3-(3,4-difluorophenyl)propanoic acid (74.0 mg, 0.17 mmol), HBTU (61.0 mg, 0.16 mmol) and HOBt (22.0 mg, 0.16 mmol) in DMF (5 mL) was added DIPEA (28.0 µL, 0.16 mmol). After 3 min of stirring, a solution of the deprotected amide intermediate (47.0 mg, 0.15 mmol), from step i), in DMF (5 mL) was added and stirring continued at 18 °C for 16 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in EtOAc (30 mL) and washed with H_2O (10 mL), 10% citric acid solution (10 mL), sat. NaHCO₃ solution (10 mL) and brine (10 mL). The organic layer was concentrated *in vacuo* and the residue dissolved in DMF (5 mL) and stirred with Et₂NH (0.18 mL, 1.74 mmol) at 18 °C for 3 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in EtOAc (30 mL) and washed with H_2O (2 x 10 mL), brine (10 mL) and sat. LiCl solution (4 x 15 mL). The organic layer was concentrated *in vacuo* and the residue purified by flash column chromatography (0-2% MeOH/CH₂Cl₂) to give the title compound (55.0 mg, 75%); as a white solid; mp 118-120 °C; Rf 0.38 (10% MeOH/CH₂Cl₂); IR $ν_{max}$ (neat) 3293 (N-H), 3061 (N-H), 2971-2871 (C-H), 1673 (C=O), 1634 (C=O), 1514 (C=O), 1446-1365 (C-H), 1280 (C-O), 1250 (C-N), 1209 (C-N), 1115 (C-F), 696 (ArC-H) cm⁻¹; ¹H NMR (600 MHz, DMSO-d₆) δ 8.46 (1H, d, *J* = 7.9, N*H*CH(CH₃)Ph), 8.07 (1H, d, *J* = 7.9, N*H*CHCH₂CH₂), 7.33-7.19 (7H, m, Ar*H*), 7.00 (1H, m, Ar*H*), 6.75 (1H, t, *J* = 5.7, N*H*CH₂CH₂), 4.90 (1H, quintet, *J* = 7.2, C*H*CH₃), 4.32 (1H, apparent dt, *J* = 6.8, 7.5, NHC*H*CH₂CH₂), 3.39 (1H, dd, *J* = 8.3, 4.9, C*H*NH₂), 2.94-2.87 (2H, m, NHCHCH₂C*H₂* & 1H, apparent dd, *J* = 13.6, 4.5, C(H)*H*CHNH₂), 2.60 (1H, dd, *J* = 13.6, 8.3, C(*H*)HCHNH₂), 1.81 (2H, bs, N*H₂*), 1.74 (1H, m, NHCHC(H)*H*CH₂), 1.61 (1H, m, NHCHC(*H*)HCH₂), 1.38 (9H, s, (C*H₃*)₃C), 1.34 (3H, d, *J* = 7.2, C*H₃*CH); ¹³C NMR (150 MHz, DMSO-d₆) δ 173.9 (NH*C*OC), 170.1 (NH*C*OC), 155.5 (NH*C*OO), 149.8-148.8 (dd, J_{C-F} = 137.7, 12.5, *C*F), 148.2-147.2 (dd, J_{C-F} = 136.5, 12.5, *C*F), 144.1 (Ar*C*), 136.6 (Ar*C*), 128.3 (2 x Ar*C*H), 126.8 (Ar*C*H),

126.1 (ArCH), 126.0 (2 x ArCH), 118.2 (d, $J_{\text{C-F}}$ = 16.7, CHCF), 116.8 (d, $J_{\text{C-F}}$ = 16.7, CHCF), 77.7 (C(CH$_3$)$_3$), 55.7 (CHNH$_2$), 50.2 (NHCHCH$_2$CH$_2$), 47.9 (CHCH$_3$), 40.1 (CH$_2$CHNH$_2$), 36.7 (CH$_2$CH$_2$NHCO), 33.2 (CH$_2$$CH_2$NHCO), 28.3 (C($CH_3$)$_3$), 22.4 ($CH_3$CH); m/z (CI+) 505 (88%, [M+H]$^+$), 405 (100%, [M+H-Boc]$^+$); HRMS C$_{26}$H$_{35}$F$_2$N$_4$O$_4$ ([M+H]$^+$) calcd. 505.2626, found. 505.2623.

tert-Butyl ([_S_]-3-[(_S_)-2-amino-3-(3,4-difluorophenyl)propanamido]-4-oxo-4-{[(_S_)-1-phenylethyl]amino}butyl)carbamate (56b)
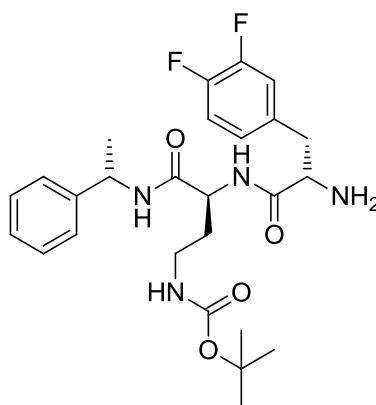
i) To a stirred solution of (_S_)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-4-[(tert-butoxycarbonyl)amino]butanoic acid (200 mg, 0.46 mmol), HBTU (160 mg, 0.42 mmol) and HOBt (56.0 mg, 0.42 mmol) in DMF (5 mL) was added DIPEA (74.0 μL, 0.42 mmol). After 3 min of stirring (_S_)-1-phenylethanamine (48.0 μL, 0.38 mmol) was added and stirring continued at 18 °C for 5 h. The reaction mixture was concentrated _in vacuo_ and purified by flash column chromatography (25% EtOAc/Pet) to give the Fmoc-protected amide intermediate (Rf 0.75 (50% EtOAc/Pet)), which was dissolved in DMF (5 mL) and stirred with Et$_2$NH (0.48 mL, 4.60 mmol) at 18 °C for 5 h. The reaction mixture was concentrated _in vacuo_ and the residue taken up in CHCl$_3$/EtOAc (40 mL, 1:1) and washed with H$_2$O (2 x 15 mL), brine (15 mL) and sat. LiCl solution (4 x 15 mL). The organic layer was concentrated _in vacuo_ and the residue purified by flash column chromatography (0-10% MeOH/CH$_2$Cl$_2$) to give the deprotected amide intermediate (90.0 mg, 74%) (Rf 0.25 (10% MeOH/CH$_2$Cl$_2$).

ii) To a stirred solution of (_S_)-2-({[(9H-fluoren-9-yl)methoxy]carbonyl}amino)-3-(3,4-difluorophenyl)propanoic acid (133 mg, 0.31 mmol), HBTU (111 mg, 0.29 mmol) and HOBt (38.0 mg, 0.29 mmol) in DMF (5 mL) was added DIPEA

(51.0 μL, 0.29 mmol). After 3 min of stirring, a solution of the deprotected amide intermediate (88.0 mg, 0.27 mmol), from step i), in DMF (2mL) was added and stirring continued at 18 °C for 16 h. The reaction mixture was concentrated *in vacuo* and the residue taken up in EtOAc (30 mL) and washed with H_2O (10 mL), 10% citric acid solution (10 mL), sat. $NaHCO_3$ solution (10 mL) and brine (10 mL). The organic layer was concentrated *in vacuo* and the residue dissolved in DMF (5 mL) and stirred with Et_2NH (0.33 mL, 3.14 mmol) at 18 °C for 5 h. The reaction mixture was concentrated *in vacuo* and the residue purified by flash column chromatography (0-3% $MeOH/CH_2Cl_2$) to give the title compound (106 mg, 77%); as a white solid; mp 106-112 °C; Rf 0.38 (10% $MeOH/CH_2Cl_2$; IR v_{max} (neat) 3291 (N-H), 3061 (N-H), 2973-2873 (C-H), 1676 (C=O), 1633 (C=O), 1515 (C=O), 1446-1365 (C-H), 1280 (C-O), 1249 (C-N), 1210 (C-N), 1116 (C-F), 696 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 8.50 (1H, d, J = 7.9, N*H*CH(CH$_3$)Ph), 8.06 (1H, d, J = 8.3, N*H*CHCH$_2$CH$_2$), 7.32-7.20 (7H, m, Ar*H*), 7.04 (1H, m, Ar*H*), 6.71 (1H, t, J = 5.7, N*H*CH$_2$CH$_2$), 4.88 (1H, quintet, J = 7.3, C*H*CH$_3$), 4.34 (1H, apparent dt, J = 6.0, 8.1, NHC*H*CH$_2$CH$_2$), 3.43 (1H, dd, J = 8.3, 4.9, C*H*NH$_2$), 2.91 (1H, dd, J = 13.6, 4.9, C(H)*H*CHNH$_2$), 2.88-2.79 (2H, m, NHCHCH$_2$C*H$_2$*), 2.63 (1H, dd, J = 13.6, 8.3, C(*H*)HCHNH$_2$), 1.79 (2H, bs, N*H$_2$*), 1.71-1.65 (1H, m, NHCHC(H)*H*CH$_2$), 1.57-1.51 (1H, m, NHCHC(*H*)HCH$_2$), 1.38 (9H, s, (C*H$_3$*)$_3$C), 1.34 (3H, d, J = 7.0, C*H$_3$*CH); ^{13}C NMR (150 MHz, DMSO-d$_6$) δ 174.0 (NH*C*OC), 170.1 (NH*C*OC), 155.4 (NH*C*OO), 149.9-148.9 (dd, J_{C-F} = 137.7, 12.5, *C*F), 148.3-147.3 (dd, J_{C-F} = 136.5, 12.5, *C*F), 144.5 (Ar*C*), 136.7 (Ar*C*), 128.3 (2 x Ar*C*H), 126.7 (Ar*C*H), 126.1 (Ar*C*H), 125.8 (2 x Ar*C*H), 118.2 (d, J_{C-F} = 16.7, *C*HCF), 116.8 (d, J_{C-F} = 16.7, *C*HCF), 77.7 (*C*(CH$_3$)$_3$), 55.7 (*C*HNH$_2$), 50.0 (NH*C*HCH$_2$CH$_2$), 47.9 (*C*HCH$_3$), 40.1 (*C*H$_2$CHNH$_2$), 36.7 (CH*C*H$_2$CH$_2$NH), 33.2 (CHCH$_2$*C*H$_2$NH), 28.3 (C(*C*H$_3$)$_3$), 22.6 (*C*H$_3$CH); m/z (ESI+) 505 (75%, [M+H]$^+$), 527 (100%, [M+Na]$^+$); HRMS C$_{26}$H$_{35}$F$_2$N$_4$O$_4$ ([M+H]$^+$) calcd. 505.2626, found. 505.2621.

(*S*)-4-Amino-2-({*S*}-2-{3-[1-(2,6-dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1*H*-indazol-6-yl]ureido}-3-{3,4-difluorophenyl}propanamido)-*N*-[(*R*)-1-phenylethyl]butanamide dihydrochloride salt (51a)
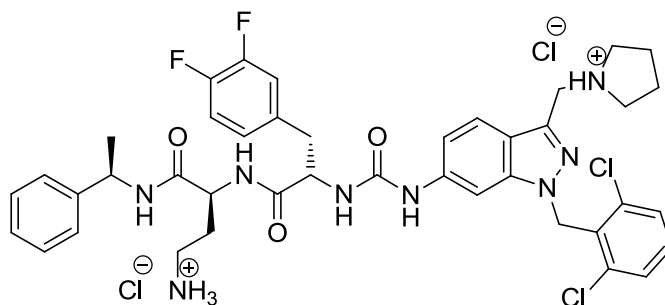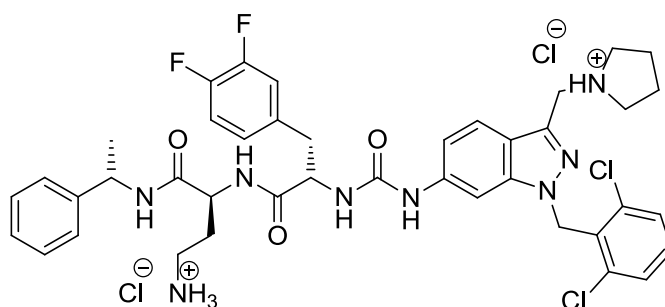

To a sonicated (5 min), then stirred, solution of **25** (50.0 mg, 133 μmol) in dry THF (5 mL), under dry conditions and cooled to 0 °C, was added together triphosgene (14.0 mg, 46.6 μmol) and DMAP (32.0 mg, 266 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min under argon. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **56a** (32.0 mg, 66.5 μmol) in dry THF (3 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then concentrated *in vacuo* and the residue purified by flash column chromatography (2-5% MeOH/CH$_2$Cl$_2$) to give the Boc-protected intermediate, Rf 0.31 (10% MeOH/CH$_2$Cl$_2$). The off-white solid was suspended in a solution of HCl (4N) in dioxane (2 mL) and stirred for 2.5 h before concentrating *in vacuo* and purifying by preparative RP-HPLC (retention time: 10.82 min). The isolated product was taken up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (3.70 mg, 6%[‡‡]); as a white solid; mp 158-163 °C (decomp.); Rf 0.06 (10% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 3274 (N-H), 3059 (N-H), 2973 (C-H), 1663 (C=O), 1647 (C=O), 1515 (C=O), 1436 (C-H), 1198 (C-N), 1177 (C-N), 1128 (C-F), 834-699 (ArC-H) cm^{-1}; ^{1}H NMR (DMSO-d$_6$, 600MHz) δ 9.98 (1H, bs, N$^+$*H*CH$_2$CNN), 9.08 (1H, s, N*H*CONHCH), 8.59 (1H, d, *J* = 8.3, N*H*CHCH$_2$CH$_2$), 8.49 (1H, d, *J* = 8.3, N*H*CH(CH$_3$)Ph), 7.97 (1H, d, *J* = 1.1, CC*H*CNH), 7.75-7.70 (1H, d, *J* = 8.7, C*H*CHCNH & 3H, bs, N$^+$*H$_3$*), 7.53 (2H, d, *J* = 8.3, 2 x C*H*CCl), 7.42 (1H, t, *J* = 8.3, C*H*CHCCl), 7.32-7.10 (7H, m, ArH), 7.01 (1H, dd, *J* = 8.7, 1.1, CHC*H*CNH), 6.97 (1H, m, ArH), 6.52 (1H, d, *J* = 7.9, CHN*H*CONH), 5.62 & 5.57 (2H, ABq, *J* = 14.5, C*H$_2$*NN), 4.95 (1H, dq, *J* = 7.5, 7.2,

---

[‡‡] Low yield was suspected to be a result of poor quality THF (dry)

C*H*CH$_3$), 4.61 (2H, d, *J* = 5.3, C*H*$_2$CNN), 4.57 (1H, apparent dt, *J* = 4.9, 7.9, C*H*NHCONH), 4.41 (1H, apparent dt, *J* = 7.5, 7.2, C*H*CH$_2$CH$_2$), 3.38 (2H distorted by H$_2$O peak, m, 2 x C(H)*H*N$^+$H(CH$_2$)CH$_2$C), 3.11-3.06 (2H, m, 2 x C(*H*)HN$^+$H(CH$_2$)CH$_2$C), 2.99 (1H, dd, *J* = 13.6, 4.1, C(H)*H*CHNHCONH), 2.83-2.77 (1H, m, C(*H*)HCHNHCONH & 2H, m, C*H*$_2$CH$_2$CHNH), 2.01-1.95 (1H, m, C(H)*H*CHNHCOCH), 1.93-1.84 (1H, m, C(*H*)HCHNHCOCH & 2H, m, 2 x C(H)*H*CH$_2$N$^+$H(CH$_2$)CH$_2$C), 1.81-1.74 (2H, m, 2 x C(*H*)HCH$_2$N$^+$H(CH$_2$)CH$_2$C), 1.37 (3H, d, *J* = 6.8, C*H*$_3$); ^{13}C NMR (150 MHz, DMSO-d$_6$) δ 171.3 (NH*C*OC), 169.4 (NH*C*OC), 157.8 (d, *J* = 31.6, Ar*C*), 154.6 (NH*C*ONH), 149.1 (Ar*C*), 147.5 (Ar*C*), 144.0 (Ar*C*), 141.4 (Ar*C*), 139.6 (Ar*C*), 136.1 (2 x *C*Cl), 135.6 (Ar*C*), 131.4 (Ar*C*), 131.0 (*C*HCHCCl), 128.8 (2 x *C*HCCl), 128.3 (2 x Ar*C*H), 126.8 (Ar*C*H), 126.2 (Ar*C*H), 126.0 (2 x Ar*C*H), 120.5 (C*C*HCHCNH), 118.2 (d, *J* = 16.7, Ar*C*H), 117.7 (Ar*C*), 116.9 (d, *J* = 16.7, Ar*C*H), 114.6 (CH*C*HCNH), 96.3 (C*C*HCNH), 53.6 (*C*HNHCONH), 53.1 (2 x *C*H$_2$N$^+$H(CH$_2$)CH$_2$C), 50.4 (*C*HCH$_2$CH$_2$), 48.0 (*C*H$_2$CNN), 47.95 (*C*HCH$_3$), 47.3 (*C*H$_2$NN), 37.6 (*C*H$_2$CHNHCONH), 36.3 (*C*H$_2$CH$_2$CHNH), 30.2 (*C*H$_2$CHNHCOCH), 22.6 (2 x *C*H$_2$CH$_2$N$^+$H(CH$_2$)CH$_2$C), 22.3 (*C*H$_3$); m/z (ESI+) 805 (100%, [^{35,35}M+H]$^+$), 807 (78%, [^{35,37}M+H]$^+$), 809 (12%, [^{37,37}M+H]$^+$); HRMS C$_{41}$H$_{45}$^{35}Cl$_2$F$_2$N$_8$O$_3$ ([M+H]$^+$) calcd. 805.2960, found. 805.2922.

(*S*)-4-amino-2-({*S*}-2-{3-[1-(2,6-dichlorobenzyl)-3-(pyrrolidin-1-ylmethyl)-1*H*-indazol-6-yl]ureido}-3-{3,4-difluorophenyl}propanamido)-*N*-[(*S*)-1-phenylethyl]butanamide dihydrochloride salt (51b)


To a sonicated (5 min), then stirred, solution of **25** (60.0 mg, 133 μmol) in dry THF (8 mL), under dry conditions and cooled to 0 °C, was added together triphosgene (16.0 mg, 54.0 μmol) and DMAP (37.0 mg, 306 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min under

argon. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **56b** (39.0 mg, 77.0 μmol) in dry THF (3 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then concentrated *in vacuo*. The residue was suspended in a solution of HCl (4N) in dioxane (2 mL) and stirred for 2.5 h before concentrating *in vacuo* and purifying by preparative RP-HPLC (retention time: 8.56 min). The isolated product was taken up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (22.5 mg, 33%); as a white solid; mp 174-178 °C (decomp.); Rf 0.06 (10% MeOH/CH$_2$Cl$_2$); IR ν$_{max}$ (neat) 3275 (N-H), 3061 (N-H), 2972 (C-H), 1665 (C=O), 1633 (C=O), 1514 (C=O), 1436 (C-H), 1198 (C-N), 1179 (C-N), 1129 (C-F), 834-699 (ArC-H) cm^{-1}; ^{1}H NMR (DMSO-d$_6$, 600MHz) δ 10.11 (1H, bs, N$^+$*H*CH$_2$CNN), 9.17 (1H, s, N*H*CONHCH), 8.58-8.55 (overlapping 1H, d, *J* = 7.9, N*H*CHCH$_2$CH$_2$ & 1H, d, *J* = 7.9, N*H*CH(CH$_3$)Ph), 8.01 (1H, d, *J* = 0.9, CC*H*CNH), 7.78-7.76 (3H, bs, N$^+$*H$_3$* & 1H, d, *J* = 8.5, C*H*CHCNH), 7.55 (2H, d, *J* = 8.3, 2 x C*H*CCl), 7.44 (1H, m, C*H*CHCCl), 7.36-7.21 (7H, m, Ar*H*), 7.07-7.03 (1H, m, Ar*H* & 1H, dd, *J* = 8.9, 1.5, CHC*H*CNH), 6.63 (1H, d, *J* = 7.7, CHN*H*CONH), 5.63 (1H, s, C*H$_2$*NN), 4.94 (1H, overlapping dq, *J* = 7.0, 7.3, C*H*CH$_3$), 4.63-4.58 (2H m, C*H$_2$*CNN & 1H, m, C*H*NHCONH), 4.43 (1H, overlapping dt, *J* = 6.2, 7.9, C*H*CH$_2$CH$_2$), 3.39 (2H, obscured by H$_2$O peak, 2 x C(H)*H*N$^+$H(CH$_2$)CH$_2$C), 3.10 (2H, m, 2 x C(*H*)HN$^+$H(CH$_2$)CH$_2$C), 3.04 (1H, dd, *J* = 13.9, 4.7, C(H)*H*CHNHCONH), 2.88 (1H, dd, *J* = 13.9, 8.1, C(*H*)HCHNHCONH), 2.77-2.67 (2H, m, C*H$_2$*CH$_2$CHNH), 1.98-1.88 (2H, m, 2 x C(H)*H*CH$_2$N$^+$H(CH$_2$)CH$_2$C & 1H C(H)*H*CHNHCOCH), 1.87-1.75 (2H, m, 2 x C(*H*)HCH$_2$N$^+$H(CH$_2$)CH$_2$C & 1H C(*H*)HCHNHCOCH), 1.37 (3H, d, *J* = 7.0, C*H$_3$*); ^{13}C NMR (150 MHz, DMSO-d$_6$) δ 171.3 (NH*C*OC), 169.4 (NH*C*OC), 158.0 (q, *J* = 30.8, Ar*C*), 154.7 (NH*C*ONH), 149.9-149.2 (dd, *J* = 104.5, 13.2, *C*F), 148.2-147.5 (dd, *J* = 103.4, 13.2, *C*F), 144.4 (Ar*C*), 141.5 (Ar*C*), 139.7 (Ar*C*), 136.1 (2 x *C*Cl), 135.6 (Ar*C*), 131.5 (Ar*C*), 131.0 (*C*HCHCCl), 128.8 (2 x *C*HCCl), 128.4 (2 x Ar*C*H), 126.8 (Ar*C*H), 126.2 (Ar*C*H), 125.9 (2 x Ar*C*H), 120.5 (C*C*HCHCNH), 118.3 (d, *J* = 17.6, *C*HCF), 117.7 (Ar*C*), 117.0 (d, *J* = 16.5, *C*HCF), 114.6 (CH*C*HCNH), 96.3 (C*C*HCNH), 53.6 (*C*HNHCONH), 53.1 (2 x *C*H$_2$N$^+$H(CH$_2$)CH$_2$C), 50.3 (*C*HCH$_2$CH$_2$), 48.2 (*C*H$_2$CNN), 48.05 (*C*HCH$_3$), 47.3 (*C*H$_2$NN), 37.6 (*C*H$_2$CHNHCONH), 36.2 (*C*H$_2$CH$_2$CHNH), 30.4 (*C*H$_2$CHNHCOCH), 22.7 (2 x *C*H$_2$CH$_2$N$^+$H(CH$_2$)CH$_2$C), 22.6 (*C*H$_3$); m/z (ESI+)

805 (100%, [^{35,35}M+H]$^+$), 807 (64%, [^{35,37}M+H]$^+$), 809 (11%, [^{37,37}M+H]$^+$); HRMS C$_{41}$H$_{45}$^{35}Cl$_2$F$_2$N$_8$O$_3$ ([M+H]$^+$) calcd. 805.2960, found. 805.2968.

Racemic mixture of (*S*)-1-(2,6-dichlorophenyl)ethyl methanesulfonate & (*R*)-1-(2,6-dichlorophenyl)ethyl methanesulfonate (58)
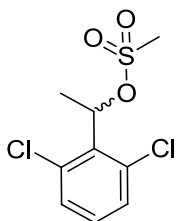

To a solution of 2,6-dichlorobenzaldehyde (2.00 g, 11.4 mmol) in THF (dry) (50 mL) at 0 °C under dry conditions was added a solution of methyl magnesium bromide (3 M) in Et$_2$O (11.4 mL, 34.3 mmol) in two portions. The reaction mixture was stirred for 1.5 h before quenching with cold sat. NH$_4$Cl (aq) solution (100 mL) and extracting with Et$_2$O (3 x 80 mL). Organics were combined, dried (MgSO$_4$) and concentrated *in vacuo* to leave a clear yellow oil, which was taken up in CH$_2$Cl$_2$ (30 mL) and stirred at 0 °C. To this solution was added triethylamine (3.25 mL, 22.2 mmol) followed by methanesulfonyl chloride (1.17 mL, 14.4 mmol). The reaction mixture was stirred for 1.5 h before quenching with H$_2$O (30 mL) and extracting with CH$_2$Cl$_2$ (3 x 20 mL). Organics were combined and washed with H$_2$O (20 mL), HCl (1 M) (20 mL), sat. NaHCO$_3$ (aq) solution (20 mL) and brine (20 mL), then dried (MgSO$_4$) and concentrated *in vacuo* to leave clear, colourless oil. Purified by flash column chromatography (5% Et$_2$O/Pet) and dried under high vacuum to give target compound (2.72 g, 89%); as a white solid; mp 49-50 °C; Rf 0.06 (20% Et$_2$O/Pet); IR ν_{max} (neat) 3025-2934 (C-H), 1580 (ArC=C), 1562 (ArC=C), 1435 (C-H), 1349 (SO$_2$), 1171 (SO$_2$), 1084 (C-O), 906-763 (ArC-H), 518 (C-Cl) cm^{-1}; H NMR (600 MHz, CDCl$_3$) δ 7.35 (2H, d, *J* = 8.3, C*H*CCl), 7.22 (1H, t, *J* = 8.3, C*H*CHCCl), 6.49 (1H, q, *J* = 6.8, C*H*CH$_3$), 2.86 (3H, s, C*H$_3$*S), 1.84 (3H, d, *J* = 6.8, C*H$_3$*CH); ^{13}C NMR δ 134.8 (2 x *C*Cl), 133.8 (Ar*C*), 130.4 (*C*HCHCCl), 129.7 (*C*HCCl), 75.9 (*C*HCH$_3$), 38.5 (*C*H$_3$S), 19.6 (*C*H$_3$CH); m/z (CI+) 268 (2%, ^{35,35}M$^+$), 173 (100%, [^{35,35}M-OMs]$^+$) , 175 (66%, [^{35,37}M-OMs]$^+$) , 177 (11%, [^{37,37}M-OMs]$^+$); HRMS C$_9$H$_{10}$^{35}Cl$_2$O$_3$S (M$^+$) calcd. 267.9728, found. 267.9727.

Racemic mixture (*S*)-1-[1-(2,6-dichlorophenyl)ethyl]-6-nitro-3-(pyrrolidin-1-ylmethyl)-1H-indazole & (*R*)-1-[1-(2,6-dichlorophenyl)ethyl]-6-nitro-3-(pyrrolidin-1-ylmethyl)-1H-indazole (59)
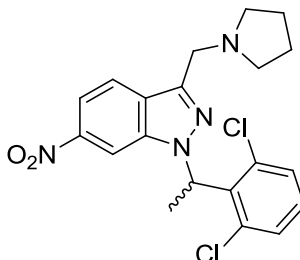

To a stirred solution of **23** (183 mg, 0.743 mmol) in DMF (20 mL), under dry conditions, was added **58** (200 mg, 0.743 mmol) and caesium carbonate (242 mg, 0.743 mmol). The reaction mixture was stirred under argon at 100 °C for 20 h before allowing to reach ambient temperature then diluted with H_2O (10 mL) and extracted with EtOAc (3 x 20 mL). The organics were combined and washed with H_2O (10 mL) then brine (2 x 10 mL), dried (MgSO$_4$) and concentrated *in vacuo* to leave an oily brown residue. Purification by flash column chromatography (0-1.5% MeOH in CH$_2$Cl$_2$) gave target compound (171 mg, 55%); as a light brown solid; mp 92-96 °C; Rf 0.44 (10% MeOH in CH$_2$Cl$_2$); IR ν_{max} (neat) 3063-2783 (C-H), 1579 (ArC=C), 1561 (ArC=C), 1521 (NO$_2$), 1438 (C-H), 1339 (NO$_2$), 1210 (C-N), 1122 (C-N), 1067 (C-N), 870-734 (ArC-H), 593 (C-Cl) cm^{-1}; ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.91 (1H, dd, *J* = 8.8, 0.8, C*H*CHCNO$_2$), 7.85 (1H, ddd, *J* = 8.8, 2.0, 0.8, CHC*H*CNO$_2$), 7.68 (1H, dd, *J* = 1.8, 0.8, CC*H*CNO$_2$), 7.26 (2H, d, *J* = 8.1, 2 x C*H*CCl), 7.13 (1H, m, C*H*CHCCl), 6.45 (1H, q, *J* = 7.1, C*H*CH$_3$), 4.07 & 3.98 (2H, ABq, *J* = 13.5, C*H*$_2$CNN), 2.55 (4H, m, 2 x C*H*$_2$N(CH$_2$)CH$_2$C), 2.15 (3H, d, *J* = 7.1, C*H*$_3$), 1.73 (4H, m, 2 x C*H*$_2$CH$_2$N(CH$_2$)CH$_2$C); ^{13}C NMR (100 MHz, CDCl$_3$) δ 146.1 (Ar*C*), 142.5 (Ar*C*), 138.8 (Ar*C*), 135.2 (Ar*C*), 134.6 (2 x *C*Cl), 129.9 (2 x *C*HCCl), 129.7 (*C*HCHCCl), 127.4 (Ar*C*), 121.9 (*C*HCHCNO$_2$), 115.1 (CH*C*HCNO$_2$), 105.9 (C*C*HCNO$_2$), 56.2 (*C*HCH$_3$), 54.1 (2 x *C*H$_2$N(CH$_2$)CH$_2$C), 52.0 (*C*H$_2$CNN), 23.7 (2 x *C*H$_2$CH$_2$N(CH$_2$)CH$_2$C), 17.0 (*C*H$_3$); m/z (ESI+) 419 (100%, [^{35,35}M+H]$^+$), 421 (95%, [^{35,37}M+H]$^+$), 423 (19%, [^{37,37}M+H]$^+$); HRMS C$_{20}$H$_{21}$^{35}Cl$_2$N$_4$O$_2$ ([M+H]$^+$) calcd. 419.1042, found. 419.1046.

Racemic mixture of (*S*)-1-[1-(2,6-dichlorophenyl)ethyl]-3-(pyrrolidine-1-ylmethyl)-1*H*-indazol-6-amine & (*R*)-1-[1-(2,6-dichlorophenyl)ethyl]-3-(pyrrolidine-1-ylmethyl)-1*H*-indazol-6-amine (61)
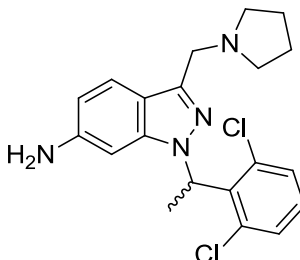
To a solution of **59** (138 mg, 0.33 mmol) in MeOH (30 mL) was added $FeCl_3.6H_2O$ (21.0 mg, 0.077 mmol) and activated charcoal (168 mg), followed by *N,N*-dimethylhydrazine (0.50 mL, 6.58 mmol). The reaction mixture was heated to reflux for 2 h then allowed to cool to ambient temperature before filtering through celite, washing through with a cold mixture of CH_2Cl_2/MeOH (60 mL, 4:1). The filtrate was concentrated *in vacuo* to leave a brown solid residue. Purification by flash column chromatography (0-5% MeOH/CH_2Cl_2) gave the desired product (107 mg, 84%); as a light brown solid; mp 196-199 °C (decomp.); Rf 0.25 (10% MeOH in CH_2Cl_2); IR v_{max} (neat) 3310-3202 (N-H), 2953-2789 (C-H), 1619 (C=N), 1560-1514 (ArC=C), 1434 (C-H), 1086 (C-N), 812-778 (ArC-H), 608 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.62 (1H, d, *J* = 8.6, C*H*CHCNH$_2$), 7.30 (2H, d, *J* = 8.1, 2 x C*H*CCl), 7.16 (1H, t, *J* = 8.1, C*H*CHCCl), 6.54 (1H, dd, *J* = 8.6, 1.8, CHC*H*CNH$_2$), 6.29 (1H, q, *J* = 7.1, C*H*CH$_3$), 5.98 (1H, d, *J* = 1.8, CC*H*CNH), 4.24 (1H, d, *J* = 13.6, C(H)*H*CCC), 4.14 (1H, d, *J* = 13.6, C(*H*)HCCC), 3.73 (2H, s, N*H*$_2$), 2.90 (4H, bs, 2 x C*H*$_2$NCH$_2$C), 2.14 (3H, d *J* = 7.1, C*H*$_3$), 1.87 (4H, bs, 2 x C*H*$_2$CH$_2$N); ^{13}C NMR (150 MHz, CDCl$_3$) δ 179.1 (Ar*C*), 145.4 (Ar*C*), 141.8 (Ar*C*), 136.0 (Ar*C*), 134.6 (2 x *C*Cl), 129.7 (2 x *C*HCCl), 129.0 (*C*HCHCCl), 121.4 (*C*HCHCNH$_2$), 118.3 (Ar*C*), 112.6 (CH*C*HCNH$_2$), 92.2 (C*C*HCNH$_2$), 55.4 (*C*HCH$_3$), 53.0 (2 x *C*H$_2$NCH$_2$), 50.4 (C*C*H$_2$N), 23.7 (2 x *C*H$_2$CH$_2$N), 17.0 (*C*H$_3$); m/z (ESI+) 389 (100%, [^{35,35}M+H]$^+$), 391 (91%, [^{35,37}M+H]$^+$), 393 (15%, [^{37,37}M+H]$^+$); HRMS C$_{20}$H$_{23}$^{35}Cl$_2$N$_4$ ([M+H]$^+$) calcd. 389.1300, found. 389.1302.

Racemic mixture of (*S*)-4-amino-*N*-benzyl-2-[(*S*)-2-(3-{1-[(*R*)-1-(2,6-dichlorophenyl)ethyl]-3-[pyrrolidin-1-ylmethyl]-1*H*-indazol-6-yl}ureido)-3-(3,4-difluorophenyl)propanamido]butanamide dihydrochloride salt **&** (*S*)-4-amino-*N*-benzyl-2-[(*S*)-2-(3-{1-[(*S*)-1-(2,6-dichlorophenyl)ethyl]-3-[pyrrolidin-1-ylmethyl]-1*H*-indazol-6-yl}ureido)-3-(3,4-difluorophenyl)propanamido]butanamide dihydrochloride salt (**52**)
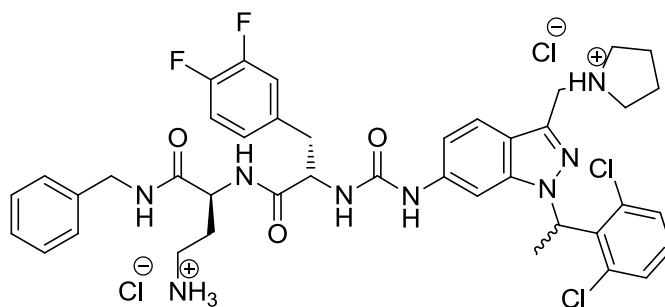
To a sonicated (5 min), then stirred, solution of **61** (50.0 mg, 129 μmol) in dry THF (5 mL), under dry conditions and cooled to 0 °C, was added together triphosgene (13.0 mg, 45.0 μmol) and DMAP (31.0 mg, 257 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min under nitrogen. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **33** (32.0 mg, 64.3 μmol) in dry THF (3 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under nitrogen then concentrated *in vacuo*. The residue was suspended in a solution of HCl (4 N) in dioxane (2 mL) and stirred for 2.5 h before concentrating *in vacuo* and purifying by preparative RP-HPLC (retention time: 8.92 min). The isolated product was taken up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (22.5 mg, 33%); as a white solid; mp 172-175 °C (decomp.); Rf 0.06 (10% MeOH/CH$_2$Cl$_2$); IR ν$_{max}$ (neat) 3258 (N-H), 3084 (N-H), 2920-2850 (C-H), 1649 (C=O), 1547 (C=O), 1512 (C=O), 1434 (C-H), 1158 (C-N), 1116 (C-F), 1080 (C-N), 770 (ArC-H), 576 (C-Cl) cm^{-1}; ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 10.88 (1H, bs, N$^+$*H*CH$_2$CNN), 9.10 (1H, s, N*H*CONHCH), 8.64-8.57 (1H, m, N*H*CH$_2$Ph & 1H, m, CHN*H*COCH), 8.05 (3H, bs, N$^+$*H$_3$*), 7.87 (1H, d, *J* = 8.8, C*H*CHCNH), 7.47 (2H, m, 2 x C*H*CCl), 7.40-7.14 (9H, ArH), 7.10 (1H, ddd, *J* = 8.8, 4.3, 1.5, CHC*H*CNH), 7.05-7.00 (1H, m, ArH), 6.54 (1H, t, *J* = 7.8-8.1, CHN*H*CONH), 6.35 (1H, quintet, *J* = 7.1-7.6, C*H*CH$_3$), 4.78-4.64 (2H, m, C*H$_2$*CNN), 4.56 (1H, m, C*H*NHCONH), 4.42 (1H, apparent dt, *J* = 6.0, 7.9, C*H*NHCOCH), 4.32 (2H, m, C*H$_2$*NHCO), 3.47 (2H, m, 2 x C(H)*H*N$^+$H(CH$_2$)CH$_2$C,

3.21 (2H, m, 2 x C(*H*)HN⁺H(CH₂)CH₂C), 3.03 (1H, dt, *J* = 13.9, 4.0, C(H)*H*CHNHCONH), 2.84-2.78 (1H, m, C(*H*)HCHNHCONH & 2H, m, C*H₂*CH₂CHNH), 2.07-1.89 (3H, dd, *J* = 7.1, 2.0, C*H₃*CH & 2H, m, C*H₂*CHNHCOCH & 4H, m, 2 x C*H₂*CH₂N⁺H(CH₂)CH₂C); ¹³C NMR (150 MHz, DMSO-d₆) δ 171.5 (NH*C*OCH), 171.3 (Ar*C*), 170.5 (NH*C*OCH), 154.5 (NH*C*ONH), 154.4 (Ar*C*), 148.2 (Ar*C*), 147.5 (Ar*C*), 140.5 (Ar*C*), 139.4 (Ar*C*), 139.1 (Ar*C*), 134.8 (Ar*C*), 133.7 (2 x *C*Cl), 130.4 (*C*HCHCCl), 130.1 (2 x *C*HCCl), 128.34 (Ar*C*H), 128.3 (Ar*C*H), 127.2 (Ar*C*H), 127.1 (Ar*C*H), 126.9 (Ar*C*H), 126.3 (Ar*C*H), 120.8 (*C*HCHCNH), 118.8 (Ar*C*), 118.3 (d, *J* = 16.5, Ar*C*H), 116.9 (d, *J* = 16.5, Ar*C*H), 114.4 (CH*C*HCNH), 95.7 (C*C*HCNH), 55.3 (*C*HCH₃), 53.5 (*C*HNHCONH), 52.8 (2 x *C*H₂N⁺H(CH₂)CH₂C), 50.5 (*C*HNHCOCH), 48.1 (*C*H₂CNN), 42.1 (*C*H₂NHCO), 37.8 (*C*H₂CHNHCONH), 36.1 (*C*H₂CH₂CHNH), 30.1 (*C*H₂CHNHCOCH), 22.8 (2 x *C*H₂CH₂N(CH₂)CH₂C), 17.1 (*C*H₃CH); m/z (ESI+) 805 (100%, [³⁵,³⁵M+H]⁺), 807 (88%, [³⁵,³⁷M+H]⁺), 809 (22%, [³⁷,³⁷M+H]⁺); HRMS C₄₁H₄₅³⁵Cl₂F₂N₈O₃ ([M+H]⁺) calcd. 805.2960, found. 805.2903.

3-{[(3*R*)-3-Fluoropyrrolidin-1-yl]methyl}-6-nitro-1*H*-indazole (63a)
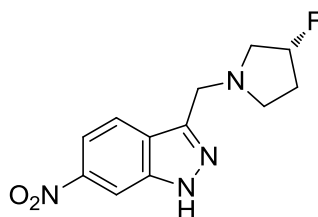
6-Nitroindole (500 mg, 3.08 mmol) was added to a stirred solution of sodium nitrite (2.13 g, 30.8 mmol) in H₂O (25 mL) and DMF (2 mL) at 18 °C. 6 M HCl (4.6 mL, 27.8 mmol) was added dropwise over 10 min and stirring continued for 3 h before diluting with EtOAc (30 mL) and extracting. Aqueous layer extracted further with EtOAc (2 x 15 mL). Organic extracts combined and washed with H₂O (15 mL) and brine (15 mL), then dried (MgSO₄) and concentrated *in vacuo*. Residual dark brown solid was dissolved in CH₂Cl₂/DMF/AcOH (54/5.4/0.6 mL), to which was added 3-*R*-(-)-fluoropyrrolidine.HCl (968 mg, 7.71 mmol) and the reaction mixture was stirred at 18 °C for 20 min. Sodium triacetoxyborohydride (1.63 g, 7.70 mmol) was added portionwise over 10 min and stirring continued for 3 h before diluting with EtOAc (100 mL) and quenching with sat. aqueous NaHCO₃ (60 mL). Aqueous layer separated and extracted further with EtOAc (2 x 20 mL). Organic extracts combined

and washed with sat. NaHCO$_3$ (aq) (2 x 25 mL), H$_2$O (2 x 25 mL) and brine (2 x 25 mL) then dried (MgSO$_4$) and concentrated *in vacuo* to leave dark brown oil. Purified by flash column chromatography (0.75% MeOH/CH$_2$Cl$_2$) gave target compound (505 mg, 62%); as a dark brown oil; Rf 0.19 (5% MeOH/CH$_2$Cl$_2$); [α]$_D^{28}$ +5.7 (c 0.35, MeOH); IR v_{max} (thin film from CHCl$_3$) 3290 (N-H), 3059-2921 (C-H), 1664 (C=N), 1548 (NO$_2$), 1345 (NO$_2$), 1238 (C-N), 1114 (C-F), 1086 (C-N), 762-697 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 8.41 (1H, d, *J* = 1.3, CC*H*CNO$_2$), 8.04-8.00 (1H, d, *J* = 8.9, C*H*CHCNO$_2$ & 1H, dd, *J* = 8.9, 1.5, CHC*H*CNO$_2$), 5.20 (1H, m, C*H*F), 4.15 & 4.12 (2H, ABq, *J* = 13.7, C*H$_2$*CNN), 3.02-2.94 (1H, m, NC(H)*H*CHF & 1H, m, NC(H)*H*CH$_2$CHF), 2.90-2.82 (1H, ddd, *J* = 29.9, 11.5, 5.1, NC(*H*)HCHF), 2.59 (1H, dt, *J* = 5.8, 8.3, NC(*H*)HCH$_2$CHF), 2.25-2.04 (2H, m, NCH$_2$C*H$_2$*CHF); ^{13}C NMR (150 MHz, CDCl$_3$) δ 147.1 (Ar*C*), 144.9 (Ar*C*), 140.0 (Ar*C*), 125.7 (Ar*C*), 122.0 (*C*HCHCNO$_2$), 115.7 (CH*C*HCNO$_2$), 106.8 (C*C*HCNO$_2$), 94.2-93.1 (d, *J* = 176.1, *C*HF), 60.7 (d, *J* = 23.1, N*C*H$_2$CHF), 52.5 (N*C*H$_2$CH$_2$CHF), 51.9 (*C*H$_2$CNN), 33.0 (d, *J* = 22.0, NCH$_2$*C*H$_2$CHF); m/z (ESI+) 265 (30%, [M+H]$^+$), 217 (100%, [M+H+MeCN-C$_4$H$_7$FN]$^+$; HRMS C$_{12}$H$_{14}$FN$_4$O$_2$ ([M+H]$^+$) calcd. 265.1101, found 265.1091.
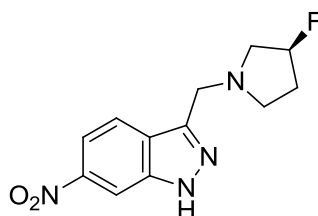
3-{[(3*S*)-3-Fluoropyrrolidin-1-yl]methyl}-6-nitro-1*H*-indazole (63b)

6-Nitroindole (500 mg, 3.08 mmol) was added to a stirred solution of sodium nitrite (2.13 g, 30.8 mmol) in H$_2$O (25 mL) and DMF (2 mL) at 18 °C. 6 M HCl (4.6 mL, 27.8 mmol) was added dropwise over 10 min and stirring continued for 3 h before diluting with EtOAc (30 mL) and extracting. Aqueous layer extracted further with EtOAc (2 x 15 mL). Organic extracts combined and washed with H$_2$O (15 mL) and brine (15 mL), then dried (MgSO$_4$) and concentrated *in vacuo*. Residual dark brown solid was dissolved in CH$_2$Cl$_2$/DMF/AcOH (54/5.4/0.6 mL), to which was added 3-S-(+)-fluoropyrrolidine.HCl (968 mg, 7.71 mmol) and the reaction mixture was stirred at 18 °C for 20 min. Sodium triacetoxyborohydride (1.63 g, 7.70 mmol) was added portionwise over 10 min and stirring continued for 3 h before diluting with

EtOAc (100 mL) and quenching with sat. aqueous NaHCO$_3$ (60 mL). Aqueous layer separated and extracted further with EtOAc (2 x 20 mL). Organic extracts combined and washed with sat. NaHCO$_3$ (aq) (2 x 25 mL), H$_2$O (2 x 25 mL) and brine (2 x 25 mL) then dried (MgSO$_4$) and concentrated *in vacuo* to leave dark brown oil. Purified by flash column chromatography (0.75% MeOH/CH$_2$Cl$_2$) gave target compound (483 mg, 59%); as a dark brown oil; Rf 0.19 (5% MeOH/CH$_2$Cl$_2$); $[\alpha]_D^{24}$ -12.5 (c 0.64, MeOH); IR ν_{max} (thin film from CHCl$_3$) 3343 (N-H), 2947-2848 (C-H), 1710 (C=N), 1522 (NO$_2$), 1345 (NO$_2$), 1296-1173 (C-N), 1111 (C-F), 1079 (C-N), 794-696 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 8.42 (1H, d, J = 1.5, CC*H*CNO$_2$), 8.05-8.01 (1H, d, J = 8.9, C*H*CHCNO$_2$ & 1H, dd, J = 8.9, 1.7, CHC*H*CNO$_2$), 5.20 (1H, m, C*H*F), 4.15 & 4.12 (2H, ABq, J = 13.8, C*H$_2$*CNN), 3.01-2.93 (1H, m, NC(H)*H*CHF & 1H, m, NC(H)*H*CH$_2$CHF), 2.89-2.82 (1H, ddd, J = 30.1, 11.7, 5.1, NC(*H*)HCHF), 2.59 (1H, dt, J = 5.8, 8.5, NC(*H*)HCH$_2$CHF), 2.25-2.04 (2H, m, NCH$_2$C*H$_2$*CHF); ^{13}C NMR (150 MHz, CDCl$_3$) δ 147.1 (Ar*C*), 145.0 (Ar*C*), 140.0 (Ar*C*), 125.7 (Ar*C*), 122.0 (*C*HCHCNO$_2$), 115.7 (CH*C*HCNO$_2$), 106.7 (C*C*HCNO$_2$), 94.2-93.1 (d, J = 176.1, *C*HF), 60.7 (d, J = 23.1, N*C*H$_2$CHF), 52.5 (N*C*H$_2$CH$_2$CHF), 51.9 (*C*H$_2$CNN), 33.0 (d, J = 22.0, NCH$_2$*C*H$_2$CHF); m/z (ESI+) 265 (32%, [M+H]$^+$), 217 (100%, [M+H+MeCN-C$_4$H$_7$FN]$^+$); HRMS C$_{12}$H$_{13}$FN$_4$O$_2$ ([M+H]$^+$) calcd. 265.1101, found 265.1111.

1-(2,6-Dichlorobenzyl)-3-{[(3*R*)-3-fluoropyrrolidin-1-yl]methyl}-6-nitro-1*H*-indazole (64a)
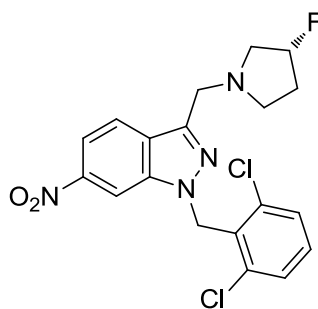
To a stirred solution of **63a** (505 mg, 1.91 mmol) in dry THF (30 mL) under argon was added 2,6-dichlorobenzyl bromide (454 mg, 1.89 mmol) followed by portionwise addition over 10 min of caesium carbonate (616 mg, 1.89 mmol). The reaction mixture was stirred at 16 °C for 24 h then diluted with H$_2$O (10 mL) and extracted with EtOAc (2 x 20 mL). Organic extracts were washed with H$_2$O (2 x 10

mL) then brine (10 mL), dried (MgSO$_4$) and concentrated *in vacuo.* The residue was purified by flash column chromatography (0-0.25% MeOH/CH$_2$Cl$_2$) gave target compound (630 mg, 79%); as a pale yellow solid; mp 134-136 °C; Rf 0.63 (5% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 2957-2812 (C-H), 1578 (C=N), 1515 (NO$_2$), 1427 (C-H), 1347 (NO$_2$), 1202 (C-N), 1139 (C-F), 1085 (C-N), 821-627 (ArC-H), 500 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 8.34 (1H, d, *J* = 1.9, CC*H*CNO$_2$), 7.98 (1H, d, *J* = 8.7, C*H*CHCNO$_2$), 7.96 (1H, dd, *J* = 8.7, 1.9, CHC*H*CNO$_2$), 7.39 (1H, d, *J* = 7.9, C*H*CCl), 7.27 (1H, t, *J* = 8.3, C*H*CHCCl), 5.85 (2H, s, C*H*$_2$NN), 5.20-5.09 (1H, m, C*H*F), 4.04 (2H, d, *J* = 1.1, NC*H*$_2$CN), 2.93-2.84 (1H, m, NC(*H*)HCHF & 1H, m, NC(*H*)HCH$_2$), 2.83-2.75 (1H, ddd, *J* = 29.7, 11.7, 4.9, NC(H)*H*CH$_2$), 2.52 (1H, dt, *J* = 5.7, 8.7, NC(H)*H*CHF), 2.19-1.98 (2H, m, CH$_2$C*H*$_2$CHF); ^{13}C NMR (150 MHz; CDCl$_3$) δ 146.7 (Ar*C*), 143.5 (Ar*C*), 139.5 (Ar*C*), 137.0 (2 x *C*Cl), 130.9 (Ar*C*), 130.6 (*C*HCHCCl), 129.0 (2 x *C*HCCl), 126.4 (Ar*C*), 122.1 (*C*HCHCNO$_2$), 115.2 (CH*C*HCNO$_2$), 106.3 (C*C*HCNO$_2$), 94.2-93.1 (d, J_{C-F} = 175.8, *C*HF), 60.5-60.4 (d, J_{C-F} = 23.3, N*C*H$_2$CHF), 52.3 (N*C*H$_2$CH$_2$CHF), 51.6 (N*C*H$_2$CN), 48.8 (*C*H$_2$NN), 33.1-32.9 (d, J_{C-F} = 22.1, CH$_2$*C*H$_2$CHF); m/z (CI+) 423 (100%, [^{35,35}M+H]$^+$), 425 (62%, [^{35,37}M+H]$^+$), 427 (10%, [^{37,37}M+H]$^+$); HRMS C$_{19}$H$_{18}$^{35}Cl$_2$FN$_4$O$_2$ ([M+H]$^+$) calcd. 423.0791, found. 423.0793.
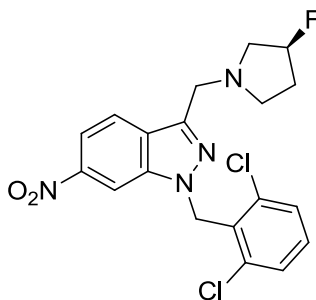
1-(2,6-Dichlorobenzyl)-3-{[(3*S*)-3-fluoropyrrolidin-1-yl]methyl}-6-nitro-1*H*-indazole (64b)

To a stirred solution of **63b** (483 mg, 1.83 mmol) in dry THF (30 mL) under argon was added 2,6-dichlorobenzyl bromide (454 mg, 1.89 mmol) followed by portionwise addition over 10 min of caesium carbonate (616 mg, 1.89 mmol). The reaction mixture was stirred at 16 °C for 24 h then diluted with H$_2$O (10 mL) and extracted with EtOAc (2 x 20 mL). Organic extracts were washed with H$_2$O (2 x 10 mL) then brine (10 mL), dried (MgSO$_4$) and concentrated *in vacuo.* The residue was

purified by flash column chromatography (0-0.25% MeOH/CH$_2$Cl$_2$) gave target compound (536 mg, 69%); as a yellow solid; mp 134-136 °C; Rf 0.63 (5% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 2955-2811 (C-H), 1562 (C=N), 1515 (NO$_2$), 1427 (C-H), 1346 (NO$_2$), 1202 (C-N), 1139 (C-F), 1085 (C-N), 802-627 (ArC-H), 500 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 8.34 (1H, dd, J = 1.7, 0.8, NCCHC), 7.99 (1H, dd, J = 8.7, 0.8, NCCHCHC), 7.97 (1H, dd, J = 8.7, 1.7, NCCHCHC), 7.39 (2H, d, J = 8.1, 2 x CHCCl), 7.27 (1H, dd, J = 8.3, 7.7, CHCHCCl), 5.85 (2H, s, CH_2NN), 5.20-5.09 (1H, m, CHF), 4.04 (2H, d, J = 1.5, NCH_2CN), 2.93-2.84 (1H, m, NC(H)HCHF & 1H, m, C(H)HCH$_2$CHF), 2.83-2.75 (1H, ddd, J = 29.7, 11.7, 5.1, NC(H)HCHF), 2.52 (1H, dt, J = 5.7, 8.7, C(H)HCH$_2$CHF), 2.19-1.99 (2H, m, CH$_2$CH_2CHF); ^{13}C NMR (150 MHz, CDCl$_3$) δ 146.7 (ArC), 143.5 (ArC), 139.5 (ArC), 137.0 (2 x CCl), 130.9 (ArC), 130.6 (CHCHCCl), 129.0 (2 x CHCCl), 126.4 (ArC), 122.1 (CHCHCNO$_2$), 115.2 (CHCHCNO$_2$), 106.3 (CCHCNO$_2$), 94.2-93.1 (d, J_{C-F} = 175.8, CHF), 60.5 (d, J_{C-F} = 23.3, NCH$_2$CHF), 52.3 (NCH$_2$CH$_2$CHF), 51.7 (NCH$_2$CN), 48.8 (CH$_2$NN), 33.0 (d, J_{C-F} = 22.1, CH$_2$$CH_2$CHF); m/z (ESI+) 423 (100%, [^{35,35}M+H]$^+$), 425 (75%, [^{35,37}M+H]$^+$), 427 (15%, [^{37,37}M+H]$^+$); HRMS C$_{19}$H$_{18}$^{35}Cl$_2$FN$_4$O$_2$ ([M+H]$^+$) calcd. 423.0791, found 423.0810.

(*R*)-1-(2,6-Dichlorobenzyl)-3-[(3-fluoropyrrolidin-1-yl)methyl]-1H-indazol-6-amine (65a)

To a solution of **64a** (610 mg, 1.44 mmol) in MeOH (40 mL) was added FeCl$_3$.6H$_2$O (78.0 mg, 0.290 mmol) and activated charcoal (784 mg), followed by *N,N*-dimethylhydrazine (2.19 mL, 28.8 mmol). The reaction mixture was heated to reflux for 2 h then allowed to cool to ambient temperature before filtering through celite, washing through with a cold mixture of CH$_2$Cl$_2$/MeOH (90 mL, 4:1). The filtrate was concentrated *in vacuo* to leave a brown solid residue, which was purified by flash column chromatography (0-1% MeOH/CH$_2$Cl$_2$) to give the desired product (408 mg,

72%); as a pale yellow solid; mp 158-160 °C; Rf 0.09 (5% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 3438-3209 (N-H), 2956-2786 (C-H), 1625 (C=N), 1564-1496 (ArC=C), 1435 (C-H), 1292-1088 (C-N), 803-688 (ArC-H), 594 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.57 (1H, d, J = 8.7, CHCHCNH$_2$), 7.34 (2H, d, J = 8.3, 2 x CHCCl), 7.21 (1H, t, J = 7.9, CHCHCCl), 6.52 (1H, dd, J = 8.7, 1.9, CHCHCNH$_2$), 6.45 (1H, d, J = 1.9, CCHCNH$_2$), 5.64 (2H, s, CH_2NN), 5.18-5.07 (1H, m, CHF), 3.95 (2H, s, NCH_2CN), 3.80 (2H, bs, NH_2), 2.91-2.79 (2H, m, NCH_2CHF & 1H, m, NC(H)HCH$_2$), 2.54 (1H, m, NC(H)HCH$_2$), 2.17-1.95 (2H, m, CH$_2$CH_2CHF); ^{13}C NMR (150 MHz, CDCl$_3$) δ 145.7 (ArC), 142.4 (ArC), 142.3 (ArC), 137.0 (2 x CCl), 131.9 (ArC), 129.9 (CHCHCCl), 128.8 (2 x CHCCl), 121.9 (CHCHCNH$_2$), 117.6 (ArC), 112.1 (CHCHCNH$_2$), 93.3-94.4 (d, $J_{C\text{-}F}$ = 175.2, CHF), 92.5 (CCHCNH$_2$), 60.3 (d, $J_{C\text{-}F}$ = 23.3, NCH$_2$CHF), 52.0 (NCH$_2$CH$_2$CHF), 51.4 (NCH$_2$CN), 48.4 (CH$_2$NN), 33.0 (d, $J_{C\text{-}F}$ = 22.1, CH$_2$$CH_2$CHF); m/z (ESI+) 415 (20%, [^{35,35}M+Na]$^+$), 304 (100%, [^{35,35}M-C$_4$H$_7$FN]$^+$) , 306 (70%, [^{35,37}M-C$_4$H$_7$FN]$^+$) , 308 (13%, [^{37,37}M-C$_4$H$_7$FN]$^+$); HRMS C$_{19}$H$_{19}$^{35}Cl$_2$FN$_4$Na ([M+Na]$^+$) calcd. 415.0869, found 415.0851.
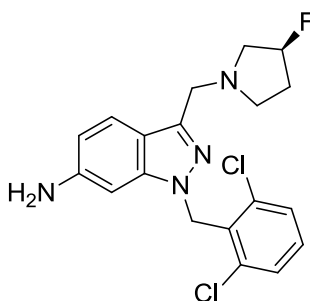
(*S*)-1-(2,6-Dichlorobenzyl)-3-[(3-fluoropyrrolidin-1-yl)methyl]-1H-indazol-6-amine (65b)

To a solution of **64b** (520 mg, 1.23 mmol) in MeOH (40 mL) was added FeCl$_3$.6H$_2$O (68.0 mg, 0.250 mmol) and activated charcoal (669 mg), followed by *N,N*-dimethylhydrazine (1.87 mL, 24.6 mmol). The reaction mixture was heated to reflux for 2 h then allowed to cool to ambient temperature before filtering through celite, washing through with a cold mixture of CH$_2$Cl$_2$/MeOH (90 mL, 4:1). The filtrate was concentrated *in vacuo* to leave a brown solid residue, which was purified by flash column chromatography (0-1% MeOH/CH$_2$Cl$_2$) to give the desired product (375 mg, 78%); as a pale yellow solid; mp 158-160 °C; Rf 0.09 (5% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 3438-3209 (N-H), 2956-2786 (C-H), 1625 (C=N), 1564-1496 (ArC=C), 1435

(C-H), 1292-1087 (C-N), 803-689 (ArC-H), 594 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.57 (1H, d, J = 8.7, C*H*CHCNH$_2$), 7.34 (2H, d, J = 7.9, 2 x C*H*CCl), 7.21 (1H, t, J = 8.1, C*H*CHCCl), 6.52 (1H, dd, J = 8.5, 1.7, CHC*H*CNH$_2$), 6.45 (1H, d, J = 1.5, CC*H*CNH$_2$), 5.64 (2H, s, C*H$_2$*NN), 5.18-5.07 (1H, m, C*H*F), 3.95 (2H, s, NC*H$_2$*CN), 3.81 (2H, bs, N*H$_2$*), 2.92-2.80 (2H, m, NC*H$_2$*CHF & 1H, m, NC(*H*)HCH$_2$), 2.56 (1H, m, NC(H)*H*CH$_2$), 2.16-1.95 (2H, m, CH$_2$C*H$_2$*CHF); ^{13}C NMR (150 MHz, CDCl$_3$) δ 145.7 (Ar*C*), 142.4 (Ar*C*), 142.3 (Ar*C*), 137.0 (2 x *C*Cl), 131.9 (Ar*C*), 129.9 (*C*HCHCCl), 128.8 (2 x *C*HCCl), 121.9 (*C*HCHCNH$_2$), 117.6 (Ar*C*), 112.1 (*C*HCHCNH$_2$), 94.5-93.3 (d, J_{C-F} = 175.2, *C*HF), 92.5 (C*C*HCNH$_2$), 60.3 (d, J_{C-F} = 22.7, N*C*H$_2$CHF), 52.0 (N*C*H$_2$CH$_2$CHF), 51.4 (N*C*H$_2$CN), 48.4 (*C*H$_2$NN), 33.0 (d, J_{C-F} = 21.5, CH$_2$*C*H$_2$CHF); m/z (ESI+) 415 (25%, [^{35,35}M+Na]$^+$), 304 (100%, [^{35,35}M-C$_4$H$_7$FN]$^+$) , 306 (87%, [^{35,37}M-C$_4$H$_7$FN]$^+$) , 308 (8%, [^{37,37}M-C$_4$H$_7$FN]$^+$); HRMS C$_{19}$H$_{19}$^{35}Cl$_2$FN$_4$Na ([M+Na]$^+$) calcd. 415.0869, found 415.0865.

(S)-4-Amino-N-benzyl-2-{[S]-2-[3-(1-{2,6-dichlorobenzyl}-3-{[(R)-3-fluoropyrrolidin-1-yl]methyl}-1H-indazol-6-yl)ureido]-3-[3,4-difluorophenyl]propanamido}butanamide dihydrochloride salt (62a)
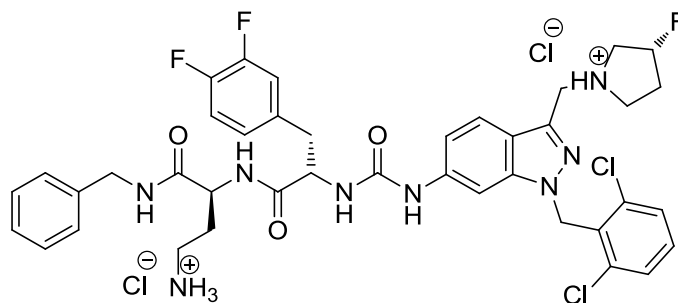
To a sonicated (5 min), then stirred solution of **65a** (50.0 mg, 0.127 mmol) in dry THF (5 mL), under dry conditions and cooled to 0 °C, was added triphosgene (13.0 mg, 44.5 μmol) and DMAP (31.0 mg, 254 μmol). After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min under argon. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **33** (31.0 mg, 63.6 μmol) in dry THF (1 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then concentrated *in vacuo* and the residue purified by flash column chromatography (0-2% MeOH/CH$_2$Cl$_2$) to give the Boc-protected intermediate (42.0 mg), Rf 0.31 (10% MeOH/CH$_2$Cl$_2$). The off-white solid was suspended in a solution of HCl (4 N) in dioxane and stirred for 2.5 h

before concentrating *in vacuo* and purifying by preparative RP-HPLC (retention time: 8.68 min). The isolated product was taken up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (6.70 mg, 17%); as a white solid; mp 169-172 °C; Rf 0.06 (10% MeOH/CH_2Cl_2); IR ν_{max} (neat) 3273 (N-H), 3060 (N-H), 2926 (C-H), 1633 (C=O), 1585 (C=N), 1536 (C=O), 1512 (C=O), 1434 (C-H), 1209 (C-N), 1162 (C-F), 1027 (C-N), 770-696 (ArC-H), 608 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.94-10.75 (1H, m, $CH_2N^+H(C_4H_7F)$), 9.20 (1H, s, N*H*CONHCH), 8.63 (1H, d, *J* = 7.9, N*H*CHCH$_2$CH$_2$), 8.58 (1H, t, *J* = 5.9, N*H*CH$_2$Ph), 8.00 (1H, s, CC*H*CNH), 7.84 (3H, bs, N$^+$H$_3$), 7.75 (1H, d, *J* = 8.7, CC*H*CHCNH), 7.54 (2H, d, *J* = 8.1, 2 x C*H*CCl), 7.44 (1H, m, C*H*CHCCl), 7.32-7.22 (7H, m, ArH), 7.06 (1H, m, ArH) 7.03 (1H, dd, *J* = 8.9, 1.5, CCHC*H*CNH), 6.69 (1H, d, *J* = 7.7, N*H*CONHC), 5.62 (2H, s, C*H$_2$*NN), 5.39 (1H, m, C*H*F), 4.67 (2H, bs, C*H$_2$*N$^+$H(C_4H_7F)), 4.58 (1H, dt, *J* = 4.9, 8.1, C*H*NHCONH), 4.42 (1H, dt, *J* = 6.2, 7.7, C*H*CH$_2$CH$_2$), 4.33 (2H, m, NHC*H$_2$*Ph), 3.80-3.25 (4H distorted by H$_2$O peak, m, 2 x C*H$_2$*N$^+$H(CH$_2$)CH$_2$C), 3.06 (1H, dd, *J* = 13.9, 4.5, C(H)*H*CHNHCONH), 2.89-2.77 (1H, m, C(*H*)HCHNHCONH & 2H, m, C*H$_2$*CH$_2$CHNH), 2.40-2.24 (1H distorted by solvent peak, m, CHFC(H)*H*CH$_2$), 2.18-2.05 (1H, m, CHFC(*H*)HCH$_2$), 2.05-1.86 (2H, m, CH$_2$C*H$_2$*CHNH); ^{13}C NMR (600 MHz, DMSO-d_6) δ 171.5 (NH*C*OC), 170.4 (NH*C*OC), 158.5-157.9 (q, *J* = 31.9, Ar*C*), 154.8 (NH*C*ONH), 149.9-149.1 (dd, *J* = 105.6, 12.1, *C*F), 148.3-147.5 (dd, *J* = 105.6, 12.1, *C*F), 141.4 (Ar*C*), 139.7 (Ar*C*), 139.1 (Ar*C*), 136.1 (2 x *C*Cl), 135.3 (Ar*C*), 131.4 (Ar*C*), 131.0 (*C*HCHCCl), 128.8 (2 x *C*HCCl), 128.3 (Ar*C*), 127.1 (Ar*C*), 126.9 (Ar*C*), 126.2 (Ar*C*), 120.5 (C*C*HCHCNH), 118.3 (d, *J* = 17.6, Ar*C*), 117.8 (Ar*C*), 117.0 (d, *J* = 16.5, Ar*C*), 114.6 (CC*H*CHCNH), 96.3 (C*C*HCNH), 92.3-91.2 (d, J_{C-F} = 175.2, *C*HF), 58.9 (NH$^+$*C*H$_2$CHF), 53.8 (*C*HCH$_2$C), 51.5 (N$^+$H*C*H$_2$CH$_2$CHF), 50.5 (*C*HCH$_2$CH$_2$), 48.5 (*C*H$_2$N$^+$H(C_4H_7F)), 47.3 (*C*H$_2$NN); 42.1 (*C*H$_2$Ph), 37.5 (CH*C*H$_2$C), 36.3 (*C*H$_2$N$^+$H$_3$), 30.6 (CHF*C*H$_2$CH$_2$), 30.1 (CHCH$_2$*C*H$_2$N$^+$H$_3$); m/z (ESI+) 809 (100%, [^{35,35}M+H]$^+$), 811 (55%, [^{35,37}M+H]$^+$), 813 (5%, [^{37,37}M+H]$^+$); HRMS $C_{40}H_{42}{}^{35}Cl_2F_3N_8O_3$ ([M+H]$^+$) calcd. 809.2709, found. 809.2719.

(S)-4-Amino-N-benzyl-2-{[S]-2-[3-(1-{2,6-dichlorobenzyl}-3-{[(S)-3-fluoropyrrolidin-1-yl]methyl}-1H-indazol-6-yl)ureido]-3-[3,4-difluorophenyl]propanamido}butanamide dihydrochloride salt (62b)
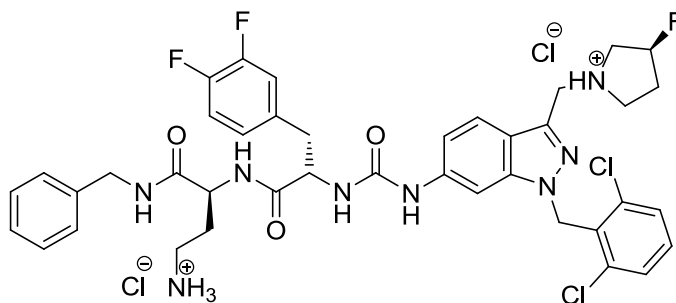
To a sonicated (5 min), then stirred, solution of **65b** (60.0 mg, 0.153 mmol) in dry THF (8 mL), under dry conditions and cooled to 0 °C, was added triphosgene (16.0 mg, 54.0 μmol) and DMAP (37.0 mg, 306 μmol) under argon. After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **33** (38.0 mg, 77.0 μmol) in dry THF (4 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then concentrated *in vacuo* and the residue purified by flash column chromatography (0-3% MeOH/CH$_2$Cl$_2$) to give the Boc-protected intermediate (45.0 mg), Rf 0.31 (10% MeOH/CH$_2$Cl$_2$). The off-white solid was suspended in a solution of HCl (4 N) in dioxane and stirred for 2.5 h before concentrating *in vacuo* and purifying by preparative RP-HPLC (retention time: 8.84 min). The isolated product was taken up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (20.0 mg, 46%) as a white solid; mp 154-156 °C; Rf 0.06 (10% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 3267 (N-H), 3061 (N-H), 2943 (C-H), 1633 (C=O), 1585 (C=N), 1537 (C=O), 1512 (C=O), 1434 (C-H), 1209 (C-N), 1161 (C-F), 1025 (C-N), 769-695 (ArC-H), 607 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 10.86-10.60 (1H, m, CH$_2$N$^+$H(C$_4$H$_7$F)), 9.15 (1H, s, NHCONHCH), 8.63 (1H, d, J = 7.9, NHCHCH$_2$CH$_2$), 8.58 (1H, t, J = 6.0, NHCH$_2$Ph), 8.00 (1H, s, CCHCNH), 7.78 (3H, bs, N$^+$H$_3$), 7.75 (1H, d, J = 8.7, CCHCHCNH), 7.55 (2H, d, J = 8.3, 2 x CHCCl), 7.44 (1H, t, J = 8.7, CHCHCCl), 7.32-7.22 (7H, m, ArH), 7.05 (1H, m, ArH), 7.03 (1H, dd, J = 9.0, 1.5, CCHCHCNH), 6.62 (1H, d, J = 7.9, NHCONHC), 5.62 (2H, s, CH$_2$NN), 5.39 (1H, dd, J = 53.8, 25.6, CHF), 4.70 (2H, d, J = 25.2, CH$_2$N$^+$H(C$_4$H$_7$F)), 4.60 (1H, dt, J = 4.9, 8.3, CHCH$_2$C), 4.42 (1H, apparent dt, J = 6.4, 7.9, CHCH$_2$CH$_2$), 4.35 (2H, ddd,

J = 21.1, 15.4, 6.0, NHCH_2Ph), 3.80-3.25 (4H distorted by H$_2$O peak, m, 2 x CCH_2N$^+$HCH_2), 3.06 (1H, dd, J = 13.9, 4.5, CHC(H)HC), 2.85 (1H, dd, J = 13.9, 8.3, CHC(H)HC), 2.83-2.74 (2H, m, CH_2N$^+$H$_3$), 2.37-2.23 (1H distorted by solvent peak, m, CHFC(H)HCH$_2$), 2.13-2.07 (1H, m, CHFC(H)HCH$_2$), 2.05-1.86 (2H, m, CHCH_2CH$_2$N$^+$H$_3$); ^{13}C NMR (150 MHz, DMSO-d$_6$) δ 171.5 (NHCOC), 170.4 (NHCOC), 158.4-157.7 (dd, J = 63.8, 31.6, ArC), 154.7 (NHCONH), 149.9-149.1 (dd, J = 104.9, 11.9, CF), 148.3-147.5 (dd, J = 104.3, 12.5, CF), 141.4 (ArC), 139.7 (ArC), 139.1 (ArC), 136.1 (2 x CCl), 135.2 (ArC), 131.4 (ArC), 131.0 (CHCHCCl), 128.8 (2 x CHCCl), 128.3 (ArC), 127.2 (ArC), 126.9 (ArC), 126.2 (ArC), 120.5 (CCHCHCNH), 118.3 (d, J = 16.7, ArC), 117.8 (ArC), 117.0 (d, J = 16.7, ArC), 114.6 (CCHCHCNH), 96.3 (CCHCNH), 92.3-91.2 (d, J_{C-F} = 175.2, CHF), 58.9 (d, J_{C-F} = 22.7 NH$^+$$CH_2$CHF), 53.7 ($CHCH_2$C), 51.5 (N$^+HCH_2CH_2$CHF), 50.5 ($CHCH_2CH_2$), 48.5 ($CH_2N^+$H(C$_4H_7$F)), 47.3 ($CH_2$NN); 42.1 ($CH_2$Ph), 37.5 (CH$CH_2$C), 36.3 ($CH_2N^+H_3$), 30.6 (d, J_{C-F} = 21.5, CHFCH$_2$CH$_2$), 30.1 (CHCH$_2$$CH_2N^+H_3$); m/z (ESI+) 809 (100%, [^{35,35}M+H]$^+$), 811 (40%, [^{35,37}M+H]$^+$), 813 (6%, [^{37,37}M+H]$^+$); HRMS C$_{40}$H$_{42}$^{35}Cl$_2$F$_3$N$_8$O$_3$ ([M+H]$^+$) calcd. 809.2709, found. 809.2718.

Racemic mixture of (*S*)-1-[(6-nitro-1*H*-indazol-3-yl)methyl]pyrrolidin-3-ol & (*R*)-1-[(6-nitro-1*H*-indazol-3-yl)methyl]pyrrolidin-3-ol (67)
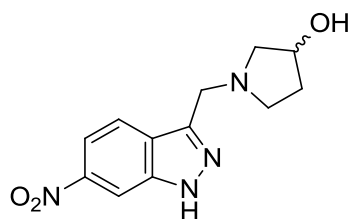
i) 6-Nitroindole (500 mg, 3.08 mmol) was added to a stirred solution of sodium nitrite (2.13 g, 30.8 mmol) in H$_2$O (25 mL) and DMF (2 mL) at 18 °C. 6 M HCl (4.6 mL, 27.8 mmol) was added dropwise over 10 min and stirring continued for 3 h before diluting with EtOAc (30 mL) and extracting. Aqueous layer extracted further with EtOAc (2 x 15 mL). Organic extracts combined, washed with H$_2$O (15 mL) and brine (15 mL), then dried (MgSO$_4$) and concentrated *in vacuo* to leave the indazole-3-carbaldehyde intermediate as a dark brown solid.

ii) HCl (g), generated by the dropwise addition of conc. H_2SO_4 (65 mL, 1.20 mol) to NaCl (s) (58.0 g, 1.07 mol), was bubbled through a stirred solution of 1-Boc-3-pyrrolidinol (1.00 g, 5.34 mmol) in Et_2O (50 mL) for 3 h. Stirring was then continued for 20 h before concentrating *in vacuo* and drying further under high vacuum to leave the hydrochloride salt of 3-pyrrolidinol as a pale pink solid.

iii) The indazole-3-carbaldehyde intermediate was dissolved in CH_2Cl_2/DMF/AcOH (45/4.5/0.5 mL), to which was added the 3-pyrrolidinol (660 mg, 5.36 mmol) and the reaction mixture was stirred at 18 °C for 20 min. Sodium triacetoxyborohydride (1.63 mg, 7.70 mmol) was added portionwise over 10 min and stirring continued for 3 h before diluting with EtOAc (100 mL) and quenching with sat. aqueous $NaHCO_3$ (60 mL). Aqueous layer separated and extracted further with EtOAc (2 x 30 mL). Organic extracts combined and washed with sat. $NaHCO_3$ (aq) (2 x 25 mL), H_2O (2 x 25 mL) and brine (2 x 25 mL) then dried ($MgSO_4$) and concentrated *in vacuo* to leave brown residue. Purification by flash column chromatography (0-10% MeOH/CH_2Cl_2) gave target compound (406 mg, 50%); as a brown solid; mp 166-168 °C; Rf 0.13 (10% MeOH/CH_2Cl_2); IR ν_{max} (neat) 3122 (O-H), 3068 (N-H), 2959-2835 (C-H), 1513 (NO_2), 1341 (NO_2), 1125-1059 (C-N), 874-730 (ArC-H) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 12.09 (1H, bs, NH), 8.40 (1H, d, J = 1.9, CCHC), 7.94 (1H, dd, J = 1.9, 9.0, CHCHCNO$_2$), 7.87 (1H, d, J = 9.0, CHCHCNO$_2$), 4.40 (1H, m, CHOH); 4.10 (2H, s, CCH_2N); 3.07 (1H, td, J = 8.7, 4.1, NCHHCH$_2$), 2.93 (1H, d, J = 10.2, NCHHCHOH), 2.63 (1H, dd, J = 10.2-10.5, 5.3-5.7, NCHHCHOH), 2.41 (1H, q, J = 8.3-8.7, NCHHCH$_2$), 2.29-2.23 (1H, m, NCH$_2$CHH), 1.84-1.79 (1H, m, NCH$_2$CHH); ^{13}C NMR (150 MHz, CDCl$_3$) δ 147.0 (ArC), 144.1 (ArC), 140.0 (ArC), 125.4 (ArC), 121.2 (CHCHCNO$_2$), 115.5 (CHCHCNO$_2$), 107.5 (CCHC), 71.5 (CHOH), 63.4 (NCH$_2$CHOH), 53.1 (NCH$_2$CH$_2$), 51.5 (CH$_2$CNN), 35.2 (NCH$_2$$CH_2$); m/z (CI+) 263 (100%, [M+H]$^+$); HRMS $C_{12}H_{15}N_4O_3$ ([M+H]$^+$) calcd. 263.1144, found. 263.1143.

Racemic mixture of (*S*)-1-{[1-(2,6-dichlorobenzyl)-6-nitro-1*H*-indazol-3-yl]methyl}pyrrolidin-3-ol & (*R*)-1-{[1-(2,6-dichlorobenzyl)-6-nitro-1*H*-indazol-3-yl]methyl}pyrrolidin-3-ol (68)
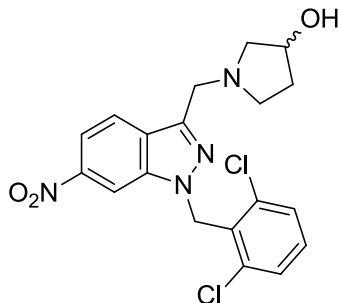
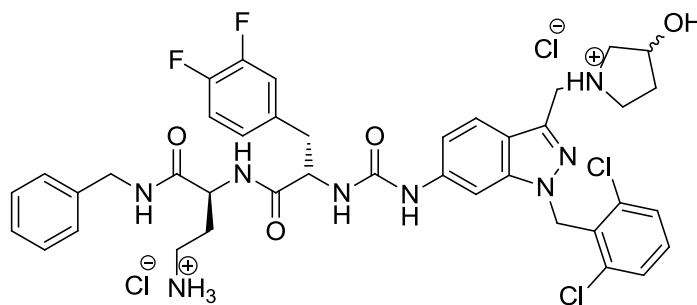To a stirred solution of **67** (319 mg, 1.22 mmol) in dry THF (30 mL) under argon was added 2,6-dichlorobenzyl bromide (293 mg, 1.22 mmol) followed by portionwise addition over 10 min of caesium carbonate (396 mg, 1.22 mmol). The reaction mixture was stirred at 16 °C for 20 h then diluted with H_2O (10 mL) and extracted with EtOAc (2 x 20 mL). Organic extracts were washed with H_2O (2 x 10 mL) then brine (10 mL), dried (MgSO$_4$) and concentrated *in vacuo*. The residue was purified by flash column chromatography (0-1.5% MeOH/CH$_2$Cl$_2$) gave target compound (292 mg, 57%); as a light brown solid; mp 161-163 °C; Rf 0.45 (10% MeOH/CH$_2$Cl$_2$); IR ν_{max} (neat) 3083 (O-H), 2963-2839 (C-H), 1579 (ArC=C), 1524 (NO$_2$), 1437 (C-H), 1345 (NO$_2$), 1298-1089 (C-N), 874-732 (ArC-H), 666 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 8.34 (1H, d, *J* = 1.9, NCC*H*C), 7.97 (1H, dd, *J* = 8.7, 1.9, NCC*H*CHC), 7.94 (1H, d, *J* = 8.7, NCCHC*H*C), 7.40 (2H, d, *J* = 7.9, 2 x C*H*CCl), 7.28 (1H, t, *J* = 7.9, C*H*CHCCl), 5.85 (2H, s, C*H*$_2$NN), 4.32 (1H, m, C*H*OH), 4.02 (2H, s, NC*H*$_2$CN), 2.90 (1H, td, *J* = 8.7, 5.3, C*H*HCH$_2$CHOH), 2.69 (1H, d, *J* = 10.2, NC*H*HCHOH), 2.62 (1H, dd, *J* = 10.2, 4.9, NCH*H*CHOH), 2.40 (1H, dt, *J* = 6.0-6.4, 8.7-9.4, CH*H*CH$_2$CHOH), 2.20-2.14 (1H, m, CH$_2$C*H*HCHOH), 1.93 (1H, bs, CHO*H*), 1.76-1.71 (1H, m, CH$_2$CH*H*CHOH); ^{13}C NMR (150 MHz, CDCl$_3$) δ 146.7 (Ar*C*), 143.6 (Ar*C*), 139.5 (Ar*C*), 137.0 (2 x *C*Cl), 130.8 (Ar*C*), 130.6 (*C*HCHCCl), 129.0 (2 x *C*HCCl), 126.5 (Ar*C*), 121.9 (NCCH*C*HC), 115.2 (NC*C*HCHC), 106.3 (NC*C*HC), 71.7 (*C*HOH), 62.9 (N*C*H$_2$CHOH), 52.4 (*C*H$_2$CH$_2$CHOH), 51.5 (N*C*H$_2$CN), 48.8 (N*C*H$_2$CCCl), 35.2 (CH$_2$*C*H$_2$CHOH); m/z (CI+) 421 (100%, [^{35,35}M+H]$^+$), 423 (55%, [^{35,37}M+H]$^+$); HRMS C$_{19}$H$_{19}$^{35}Cl$_2$N$_4$O$_3$ ([M+H]$^+$) calcd. 421.0834, found. 421.0831.

Racemic mixture of (*S*)-1-{[6-amino-1-(2,6-dichlorobenzyl)-1*H*-indazol-3-yl]methyl}pyrrolidin-3-ol & (*R*)-1-{[6-amino-1-(2,6-dichlorobenzyl)-1*H*-indazol-3-yl]methyl}pyrrolidin-3-ol (69)
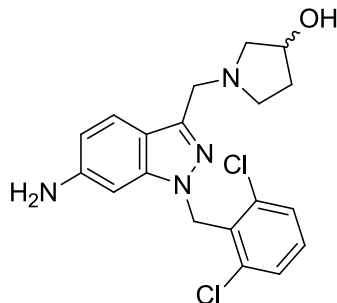
To a solution of **68** (206 mg, 0.489 mmol) in MeOH (30 mL) was added $FeCl_3.6H_2O$ (31.0 mg, 0.114 mmol) and activated charcoal (248 mg), followed by *N,N*-dimethylhydrazine (0.74 mL, 9.78 mmol). The reaction mixture was heated to reflux for 5 h then allowed to cool to ambient temperature before filtering through celite, washing through with a cold mixture of CH_2Cl_2/MeOH (90 mL, 4:1). The filtrate was concentrated *in vacuo* to leave a brown oily residue, which was purified by flash column chromatography (0-6% MeOH/CH_2Cl_2) to give the desired product (120 mg, 63%); as a light brown/pink solid; mp 148-152 °C; Rf 0.25 (20% MeOH/CH_2Cl_2); IR v_{max} (neat) 3432 (O-H), 3344-3230 (N-H), 2958-2848 (C-H), 1626 (C=N), 1564-1495 (ArC=C), 1434 (C-H), 1292-1090 (C-N), 767-691 (ArC-H), 631 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.51 (1H, d, *J* = 8.7, C*H*CHCNO$_2$), 7.33 (2H, d, *J* = 8.3, 2 x C*H*CCl), 7.20 (1H, t, *J* = 8.3, C*H*CHCCl), 6.51 (1H, dd, *J* = 8.7, 1.9, CHC*H*CNO$_2$), 6.45 (1H, d, *J* = 1.9, CC*H*C), 5.62 (2H, s, NNC*H$_2$*C), 4.26 (1H, m, C*H*OH), 3.95 (2H, s, C*H$_2$*CNN), 3.42 (1H, s, CHO*H*), 2.97 (1H, dt, *J* = 5.7, 8.7, C*H*HCH$_2$CHOH), 2.78 (1H, d, *J* = 10.5, NC*H*HCHOH), 2.62 (1H, dd, *J* = 10.4, 5.3, NCH*H*CHOH), 2.40 (1H, dt, *J* = 6.4, 9.0, CH*H*CH$_2$CHOH), 2.13-2.08 (1H, m, CH$_2$C*H*HCHOH), 1.74-1.69 (1H, m, CH$_2$CH*H*CHOH); ^{13}C NMR (150 MHz, CDCl$_3$) δ 145.9 (Ar*C*), 142.4 (Ar*C*), 141.7 (Ar*C*), 136.9 (2 x *C*Cl), 131.8 (Ar*C*), 130.0 (*C*HCHCCl), 128.8 (2 x *C*HCCl), 121.6 (*C*HCHCNO$_2$), 117.5 (Ar*C*), 112.3 (CH*C*HCNO$_2$), 92.5 (C*C*HC), 71.5 (*C*HOH), 62.5 (N*C*H$_2$CHOH), 52.0 (*C*H$_2$CH$_2$CHOH), 50.9 (*C*H$_2$CNN), 48.3 (*C*H$_2$NN), 35.1 (CH$_2$*C*H$_2$CHOH); m/z (CI+) 391 (100%, [^{35,35}M+H]$^+$), 393 (64%, [^{35,37}M+H]$^+$), 395 (9%, [^{37,37}M+H]$^+$); HRMS C$_{19}$H$_{21}$^{35}Cl$_2$N$_4$O ([M+H]$^+$) calcd. 391.1092, found. 391.1091.

Racemic mixture of (*S*)-4-amino-*N*-benzyl-2-{[*S*]-2-[3-(1-{2,6-dichlorobenzyl}-3-{[(*S*)-3-hydroxypyrrolidin-1-yl]methyl}-1*H*-indazol-6-yl)ureido]-3-[3,4-difluorophenyl]propanamido}butanamide dihydrochloride salt & (*S*)-4-amino-*N*-benzyl-2-{[*S*]-2-[3-(1-{2,6-dichlorobenzyl}-3-{[(*R*)-3-hydroxypyrrolidin-1-yl]methyl}-1*H*-indazol-6-yl)ureido]-3-[3,4-difluorophenyl]propanamido}butanamide dihydrochloride salt (66)

To a sonicated (5 min), then stirred solution of **69** (60.0 mg, 0.153 mmol) in dry THF (8 mL), under dry conditions and cooled to 0 °C, was added triphosgene (16.0 mg, 54.0 μmol) and DMAP (37.0 mg, 306 μmol) under argon. After 5 min the reaction mixture was allowed to reach ambient temperature and stirred for 30 min. The reaction mixture was cooled again to 0 °C for 5 min and a solution of **33** (38.0 mg, 77.0 μmol) in dry THF (4 mL) was added. The reaction mixture was left to reach ambient temperature and stirred for 2 h under argon then concentrated *in vacuo* and the residue purified by flash column chromatography (0-3% MeOH/CH$_2$Cl$_2$) to give the Boc-protected intermediate (27.0 mg). The off-white solid was suspended in a solution of HCl (4 N) in dioxane and stirred for 2.5 h before concentrating *in vacuo* and purifying by preparative RP-HPLC (retention time: 7.81 min). The isolated product was taken up in a 0.1 M HCl (aq) solution (2 mL) and lyophilised to give the title compound (12.0 mg, 19%); as a white solid; mp 182-186 °C (decomp.); Rf 0.06 (10% MeOH/CH$_2$Cl$_2$); IR ν$_{max}$ (neat) 3273 (O-H), 3080 (N-H), 2958-2848 (C-H), 1634 (C=O), 1584 (C=N), 1539-1493 (ArC=C), 1434 (C-H), 1280-1086 (C-N), 769-696 (ArC-H), 608 (C-Cl) cm^{-1}; ^{1}H NMR (600 MHz, DMSO-d$_6$) δ 10.72 (1H, m, N$^+$*H*CH$_2$CNN), 9.35 & 9.32 (1H, 2 x s, N*H*CONHCH), 8.68-8.66 (1H, m, N*H*CHCH$_2$CH$_2$ & 1H, m, N*H*CH$_2$Ph), 8.03 (3H, m, N$^+$*H*$_3$), 7.99 & 7.97 (1H, 2 x s, CC*H*CNH), 7.83 & 7.80 (1H, 2 x d, *J* = 8.9 & 8.7, CC*H*CHCNH), 7.54 (2H, m, 2 x C*H*CCl), 7.43 (1H, m, C*H*CHCCl), 7.37-7.17 (7H, m, ArH), 7.09 (1H, m, ArH), 7.02 & 7.01 (1H, 2 x d, *J* = 8.9 & 8.7, CCHC*H*CNH), 6.75 & 6.72 (1H, 2 x d, *J* = 7.7 &

7.7, N*H*CONHC), 5.61 (2H, m, C*H₂*NN), 4.65-4.54 (2H, m, C*H₂*CNN & 1H, m, C*H*NHCONH), 4.43 (1H, dt, J = 6.0, 7.9, C*H*CH₂CH₂), 4.36-4.29 (2H, m, NHC*H₂*Ph & 1H, m C*H*OH), 3.48 & 3.20 (0.5H & 0.5H, m, Pyr-C(5)*H₂*), 3.47 & 3.01 (0.5H & 0.5H, m, Pyr-C(2)*H₂*), 3.41 & 3.26 (0.5H & 0.5H, m, Pyr-C(5')*H₂*), 3.19 (0.5H & 0.5H, m, Pyr-C(2')*H₂*), 3.07 (1H, m, C(H)*H*CHNHCONH), 2.86-2.78 (1H, m, C(*H*)HCHNHCONH & 2H, m C*H₂*CH₂CHNH), 2.08-1.91 (2H, m, CH2CHNHCONH), 2.02 & 1.74 (0.5H & 0.5H, m, Pyr-C(4)*H₂*), 1.94 & 1.84 (0.5H & 0.5H, m, Pyr-C(4')*H₂*); ^{13}C NMR (150 MHz, DMSO-d₆) δ 171.5 (NH*C*OC), 170.5 (NH*C*OC), 154.82 & 154.81 (NH*C*ONH), 149.9-149.1 (dd, J = 104.5, 12.1, *C*F), 148.2-147.5 (dd, J = 104.5, 12.1, *C*F), 141.5 & 141.4 (Ar*C*), 139.7 (Ar*C*), 139.1 (Ar*C*), 136.1 (2 x *C*Cl), 135.5 & 135.3 (Ar*C*), 131.5 & 131.4 (Ar*C*), 131.0 (*C*HCHCCl), 128.83 (*C*HCCl), 128.80 (*C*HCCl), 128.3 (2 x Ar*C*H), 127.1 (2 x Ar*C*H), 126.9 (Ar*C*H), 126.3 (Ar*C*H), 120.8 & 120.7 (C*C*HCHCNH), 118.4 (d, J = 17.6, Ar*C*), 117.9 (Ar*C*), 117.6 (Ar*C*), 117.0 (d, J = 16.5, Ar*C*), 114.5 (CH*C*HCHCNH), 96.1 (C*C*HCNH), 68.3 (Pyr-*C*(3)HOH), 68.1 (Pyr-*C*(3')HOH), 60.3 (Pyr-*C*(2')H₂), 59.8 (Pyr-*C*(2)H₂), 53.9 (*C*HNHCONH), 51.8 (Pyr-*C*(5')H₂), 51.5 (Pyr-*C*(5)H₂), 50.6 (*C*HCH₂CH₂), 49.5 & 48.3 (*C*H₂CNN), 47.33 & 47.26 (*C*H₂NN), 42.1 (*C*H₂Ph), 37.5 (*C*H₂CHNHCONH), 36.2 (*C*H₂CH₂CH), 33.1 (Pyr-*C*(4)H₂), 32.5 (Pyr-*C*(4')H₂), 30.1 (*C*H₂CHNHCOCH); m/z (ESI+) 807 (100%, [^{35,35}M+H]⁺), 809 (71%, [^{35,37}M+H]⁺), 811 (15%, [^{37,37}M+H]⁺), 829 (20%, [M+Na]⁺).

Biology

Biological testing of compounds was done in collaboration with Prof Rachel Chambers' research group at the Centre for Inflammation and Tissue Repair (UCL, Division of Medicine). Calcium mobilisation assay work was assisted by Dr Natalia Smoktunowicz and CCL2 ELISA work was also carried out by Natalia. *In vivo* work was carried out by Dr Ricardo José.

Materials and methods

Reagents:

Thrombin, extracted from human plasma, was purchased from Calbiochem (Merck Biosciences, UK). PAR-1 agonist peptide and reverse peptide were supplied by Bachem. The original RWJ-58259 sample used as standard was a gift from Dr Claudia Derian (Johnson & Johnson Pharmaceutical Research & Development, USA).

Cell culture:

Primary human lung fibroblasts (pHLFs), grown from explant cultures of normal adult lung tissue, were a gift from Dr Robin J McAnulty (University College London). Cells were maintained in DMEM at 37 °C (10% CO_2) supplemented with glutamine (4 mM), penicillin, streptomycin, and 10% (v/v) FBS (all from Invitrogen) and were below passage 10 when used for experiments.

Measurement of intracellular Ca^{2+} levels:

pHLFs were seeded in clear-bottom, black, 96-well microplates at a density of 10,000 cells/well for 48 h and then quiesced for 24 h. Intracellular Ca^{2+} levels were assessed using the Fluo-4 NW kit (Invitrogen, UK) according to the manufacturer's instructions. The dye was dissolved in assay buffer (20 mM HEPES in HBSS) supplemented with 2.5 mM probenecid. Wells were aspirated and solutions of test compounds (3mM in DMSO), diluted in the dye buffer solution to give a final concentration of 3 μM, were added to the relevant wells (100 μL per well). For positive controls, the equivalent volume of DMSO only was diluted in the dye buffer

solution and added to the relevant wells. The cells were then incubated again at 37 °C (10% CO_2) for 30 min and at ambient conditions, in the dark, for 30 min. Cells were then stimulated by addition of a solution of thrombin in assay buffer (50 µL, 10 nM/3 nM/0.3 nM/0 nM) and changes in intracellular Ca^{2+} concentration were monitored in real-time using a fluorescent imaging plate reader (FLIPR® Tetra System, Molecular Devices, Inc). Measurements were taken every 1 sec for 1 min, then every 6 sec for a further 2 min. Experiments were performed in three replicates on the same plate and the data plotted as change in relative fluorescence units (RFU).[174]

Human CCL2 ELISA:

Cells were seeded at 10 x 10^4 cells/well, grown overnight to subconfluence and quiesced in serum-free medium for 24 h before stimulation or siRNA treatment. Wells of a 96-well ELISA plate were coated with 50 µL of 2 µg/mL MAB679 (R&D Systems Europe Ltd., Abingdon, UK) in bicarbonate coating buffer, left for 5 h at room temperature and then washed. 100 µL of PBS with 1% bovine serum albumin was added to each well and plates were left overnight at 4 °C. 50 µL of either CCL2 standard or sample was added to each well and left for 2 h at room temperature. After washing, 50 µL of 100 ng/mL BAF279 (R&D Systems) was added to each well and left for 2 h at room temperature. The plate was washed and 50 µL of a 2 µg/mL dilution of streptavidin-horseradish peroxidase (Dako UK Ltd, Ely, UK) was added for 30 min. After washing, 50 µL of 3,3',5,5'-tetramethylbenzidine substrate (BD Biosciences, San Diego, CA) was added to each well. After 30 min, colour development was quenched with 0.19 M H_2SO_4, and absorbance was measured on a plate reader at 450 nm. Alternatively, the human CCL2 ELISA Duoset kit DY279 (R&D Systems) was used according to the manufacturer's instructions.[174]

LPS-induced pulmonary inflammation model:

Experiments were conducted with local ethical approval in accordance with the Home Office, UK. Female Balb/c mice (6-8 wk old; Charles River, Margate, UK) were anesthetized (5% isoflurane) and challenged with LPS in sterile saline (125 µg/kg, 50 µL intranasally; *Escherichia coli* 0127:B8; Sigma Aldrich, Gillingham, UK) or sterile saline only. Treated mice were injected intraperitoneally with

compound **20** (5 mg/kg), 30 min after LPS administration. After 24 h, animals were euthanized (urethane, intraperitoneally, 20 g/kg), and BAL performed (3 x 0.5 mL of PBS). Total and differential cell counts were quantified after cytospin. Alternatively, BAL fluid was isolated and whole lungs were removed and homogenized.[176]

References

1. T. A. Wynn, *Journal of Pathology*, 2008, **214**, 199-210.
2. G. C. Gurtner, S. Werner, Y. Barrandon and M. T. Longaker, *Nature*, 2008, **453**, 314-321.
3. B. N. Brown and S. F. Badylak, *Translational Research*, 2014, **163**, 268-285.
4. P. F. Mercer and R. C. Chambers, *Biochimica et Biophysica Acta*, 2013, **1832**, 1018-1027.
5. V. Hernandez-Gea and S. L. Friedman, in *Annual Review of Pathology: Mechanisms of Disease, Vol 6*, eds. A. K. Abbas, S. J. Galli and P. M. Howley, Annual Reviews, Palo Alto, 2011, vol. 6, pp. 425-456.
6. S. L. Friedman, D. Sheppard, J. S. Duffield and S. Violette, *Science Translational Medicine*, 2013, **5**.
7. W. K. Stadelmann, A. G. Digenis and G. R. Tobin, *American Journal of Surgery*, 1998, **176**, 26s-38s.
8. T. A. Wynn, *Journal of Clinical Investigation*, 2007, **117**, 524-529.
9. L. G. Licari and J. P. Kovacic, *Journal of Veterinary Emergency and Critical Care*, 2009, **19**, 11-22.
10. M. Ueno, M. Kodali, A. Tello-Montoliu and D. J. Angiolillo, *Journal of Atherosclerosis and Thrombosis*, 2011, **18**, 431-442.
11. W. Fiers, R. Beyaert, W. Declercq and P. Vandenabeele, *Oncogene*, 1999, **18**, 7719-7730.
12. C. Hetz, *Nature Reviews Molecular Cell Biology*, 2012, **13**, 89-102.
13. B. N. Chau, C. Xin, J. Hartner, S. Ren, A. P. Castano, G. Linn, J. Li, P. T. Tran, V. Kaimal, X. Huang, A. N. Chang, S. Li, A. Kalra, M. Grafals, D. Portilla, D. A. MacKenna, S. H. Orkin and J. S. Duffield, *Science Translational Medicine*, 2012, **4**.
14. A. Leask, *Circulation Research*, 2010, **106**, 1675-1680.
15. G. R. Grotendorst, *Cytokine and Growth Factor Reviews*, 1997, **8**, 171-179.
16. B. C. Willis and Z. Borok, *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 2007, **293**, L525-L534.
17. P. Martin and S. J. Leibovich, *Trends in Cell Biology*, 2005, **15**, 599-607.
18. B. Hinz, S. H. Phan, V. J. Thannickal, A. Galli, M.-L. Bochaton-Piallat and G. Gabbiani, *American Journal of Pathology*, 2007, **170**, 1807-1816.
19. C. J. Scotton and R. C. Chambers, *Chest*, 2007, **132**, 1311-1321.
20. V. Dugina, L. Fontao, C. Chaponnier, J. Vasiliev and G. Gabbiani, *Journal of Cell Science*, 2001, **114**, 3285-3296.
21. J. J. Tomasek, G. Gabbiani, B. Hinz, C. Chaponnier and R. A. Brown, *Nature Reviews Molecular Cell Biology*, 2002, **3**, 349-363.
22. G. Serini and G. Gabbiani, *Experimental Cell Research*, 1999, **250**, 273-283.
23. A. Pardo and M. Selman, *Proceedings of the American Thoracic Society*, 2006, **3**, 383-388.
24. B. Deconinck, J. Verschakelen, J. Coolen, E. Verbeken, G. Verleden and W. Wuyts, *Lung*, 2013, **191**, 19-25.
25. O. J. Dempsey, K. M. Kerr, H. Remmen and A. R. Denison, *British Medical Journal*, 2010, **340**.
26. W. Osler, *The Principles and practice of medicine c.2*, D. Appleton, 1892.

27. M. J. Cushley, A. G. Davison, R. M. du Bois, J. Egan, C. D. R. Flower, G. J. Gibson, A. P. Greening, N. B. N. Ibrahim, I. D. A. Johnston, D. M. Mitchell, C. A. C. Pickering and G. Diffuse Parenchymal Lung Dis, *Thorax*, 1999, **54**, S1-S30.

28. R. P. Baughman, E. E. Lower and R. M. du Bois, *Lancet*, 2003, **361**, 1111-1118.

29. C. Ferri, G. Valentini, F. Cozzi, M. Sebastiani, C. Michelassi, G. La Montagna, A. Bullo, M. Cazzato, E. Tirri, F. Storino, D. Giuggioli, G. Cuomo, M. Rosada, S. Bombardieri, S. Todesco, G. Tirri and I. Systemic Sclerosis Study Grp, *Medicine*, 2002, **81**, 139-153.

30. P. I. Latsi and A. U. Wells, *Current Opinion in Rheumatology*, 2003, **15**, 748-755.

31. U. Costabel, R. M. Du Bois and J. J. Egan, *Diffuse Parenchymal Lung Disease*, Karger, 2007.

32. J. G. Scadding, *Thorax*, 1974, **29**, 271-281.

33. A. L. A. Katzenstein and J. L. Myers, *American Journal of Respiratory and Critical Care Medicine*, 1998, **157**, 1301-1315.

34. C. Agusti, *American Journal of Respiratory and Critical Care Medicine*, 2002, **166**, 426-426.

35. W. D. Travis, U. Costabel, D. M. Hansell, T. E. King, D. A. Lynch, A. G. Nicholson, C. J. Ryerson, J. H. Ryu, M. Selman, A. U. Wells, J. Behr, D. Bouros, K. K. Brown, T. V. Colby, H. R. Collard, C. R. Cordeiro, V. Cottin, B. Crestani, M. Drent, R. F. Dudden, J. Egan, K. Flaherty, C. Hogaboam, Y. Inoue, T. Johkoh, D. S. Kim, M. Kitaichi, J. Loyd, F. J. Martinez, J. Myers, S. Protzko, G. Raghu, L. Richeldi, N. Sverzellati, J. Swigris, D. Valeyre and A. E. C. O. Idiopathic, *American Journal of Respiratory and Critical Care Medicine*, 2013, **188**, 733-748.

36. V. Navaratnam, K. M. Fleming, J. West, C. J. Smith, R. G. Jenkins, A. Fogarty and R. B. Hubbard, *Thorax*, 2011, **66**, 462-467.

37. P. F. Mercer and R. C. Chambers, *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 2013, **304**, L466-L468.

38. S. Amer Thoracic and S. European Resp, *American Journal of Respiratory and Critical Care Medicine*, 2000, **161**, 646-664.

39. G. Raghu, D. Weycker, J. Edelsberg, W. Z. Bradford and G. Oster, *American Journal of Respiratory and Critical Care Medicine*, 2006, **174**, 810-816.

40. T. E. King, A. Pardo and M. Selman, *Lancet*, 2011, **378**, 1949-1961.

41. M. Selman, T. E. King and A. Pardo, *Annals of Internal Medicine*, 2001, **134**, 136-151.

42. B. Ley, H. R. Collard and T. E. King, *American Journal of Respiratory and Critical Care Medicine*, 2011, **183**, 431-440.

43. H. R. Collard, B. B. Moore, K. R. Flaherty, K. K. Brown, R. J. Kaner, T. E. King, J. A. Lasky, J. E. Loyd, I. Noth, M. A. Olman, G. Raghu, J. Roman, J. H. Ryu, D. A. Zisman, G. W. Hunninghake, T. V. Colby, J. J. Egan, D. M. Hansell, T. Johkoh, N. Kaminski, D. S. Kim, Y. Kondoh, D. A. Lynch, J. Muller-Quernheim, J. L. Myers, A. G. Nicholson, M. Selman, G. B. Toews, A. U. Wells, F. J. Martinez and C. Idiopathic Pulmonary Fibrosis, *American Journal of Respiratory and Critical Care Medicine*, 2007, **176**, 636-643.

44. J. W. Song, S. B. Hong, C. M. Lim, Y. Koh and D. S. Kim, *European Respiratory Journal*, 2011, **37**, 356-363.

45. M. Selman, G. Carrillo, A. Estrada, M. Mejia, C. Becerril, J. Cisneros, M. Gaxiola, R. Perez-Padilla, C. Navarro, T. Richards, J. Dauber, T. E. King, A. Pardo and N. Kaminski, *Plos One*, 2007, **2**.

46. V. Cottin, J. Le Pavec, G. Prevot, H. Mal, M. Humbert, G. Simonneau, J. F. Cordier and O. P. Germ, *European Respiratory Journal*, 2010, **35**, 105-111.

47. M. Mejia, G. Carrillo, J. Rojas-Serrano, A. Estrada, T. Suarez, D. Alonso, E. Barrientos, M. Gaxiola, C. Navarro and M. Selman, *Chest*, 2009, **136**, 10-15.

48. C. J. Lettieri, S. D. Nathan, S. D. Barnett, S. Ahmad and A. F. Shorr, *Chest*, 2006, **129**, 746-752.

49. C. D. Fell and F. J. Martinez, *Chest*, 2007, **131**, 641-643.

50. K. Hamada, S. Nagai, S. Tanaka, T. Handa, M. Shigematsau, T. Nagao, M. Mishima, M. Kitaichi and T. Izumi, *Chest*, 2007, **131**, 650-656.

51. H. F. Nadrous, P. A. Pellikka, M. J. Krowka, K. L. Swanson, N. Chaowalit, P. A. Decker and J. H. Ryu, *Chest*, 2005, **128**, 2393-2399.

52. D. Bouros, K. Hatzakis, H. Labrakis and K. Zeibecoglou, *Chest*, 2002, **121**, 1278-1289.

53. M. Hironaka and M. Fukayama, *Pathology International*, 1999, **49**, 1060-1066.

54. G. Raghu, H. R. Collard, J. J. Egan, F. J. Martinez, J. Behr, K. K. Brown, T. V. Colby, J. F. Cordier, K. R. Flaherty, J. A. Lasky, D. A. Lynch, J. H. Ryu, J. J. Swigris, A. U. Wells, J. Ancochea, D. Bouros, C. Carvalho, U. Costabel, M. Ebina, D. M. Hansell, T. Johkoh, D. S. Kim, T. E. King, Y. Kondoh, J. Myers, N. L. Muller, A. G. Nicholson, L. Richeldi, M. Selman, R. F. Dudden, B. S. Griss, S. L. Protzko, H. J. Schunemann and A. E. J. A. Comm, *American Journal of Respiratory and Critical Care Medicine*, 2011, **183**, 788-824.

55. W. A. Wuyts, C. Agostini, K. M. Antoniou, D. Bouros, R. C. Chambers, V. Cottin, J. J. Egan, B. N. Lambrecht, R. Lories, H. Parfrey, A. Prasse, C. Robalo-Cordeiro, E. Verbeken, J. A. Verschakelen, A. U. Wells and G. M. Verleden, *European Respiratory Journal*, 2013, **41**, 1207-1218.

56. J. S. Allam and A. H. Limper, *Current Opinion in Pulmonary Medicine*, 2006, **12**, 312-317.

57. C. H. M. van Moorsel, M. F. M. van Oosterhout, N. P. Barlo, P. A. de Jong, J. J. van der Vis, H. J. T. Ruven, H. W. van Es, J. M. M. van den Bosch and J. C. Grutters, *American Journal of Respiratory and Critical Care Medicine*, 2010, **182**, 1419-1425.

58. R. Chibbar, F. Shih, M. Baga, E. Torlakovic, K. Ramlall, R. Skomro, D. W. Cockcroft and E. G. Lemire, *Modern Pathology*, 2004, **17**, 973-980.

59. V. Pulkkinen, S. Bruce, J. Rintahaka, U. Hodgson, T. Laitinen, H. Alenius, V. L. Kinnula, M. Myllarniemi, S. Matikainen and J. Kere, *Faseb Journal*, 2010, **24**, 1167-1177.

60. J. J. Egan, J. P. Stewart, P. S. Hasleton, J. R. Arrand, K. B. Carroll and A. A. Woodcock, *Thorax*, 1995, **50**, 1234-1239.

61. J. P. Stewart, J. J. Egan, A. J. Ross, B. G. Kelly, S. S. Lok, P. S. Hasleton and A. A. Woodcock, *American Journal of Respiratory and Critical Care Medicine*, 1999, **159**, 1336-1341.

62. Y. W. Tang, J. E. Johnson, P. J. Browning, R. A. Cruz-Gervis, A. Davis, B. S. Graham, K. L. Brigham, J. A. Oates, J. E. Loyd and A. A. Stecenko, *Journal of Clinical Microbiology*, 2003, **41**, 2633-2640.

63. R. Meliconi, P. Andreone, L. Fasano, S. Galli, A. Pacilli, R. Miniero, M. Fabbri, L. Solforosi and M. Bernardi, *Thorax*, 1996, **51**, 315-317.
64. J. E. G. Pearson and R. S. E. Wilson, *Thorax*, 1971, **26**, 300-305.
65. G. Raghu, T. D. Freudenberger, S. Yang, J. R. Curtis, C. Spada, J. Hayes, J. K. Sillery, C. E. Pope and C. A. Pellegrini, *European Respiratory Journal*, 2006, **27**, 136-142.
66. C. Tcherakian, V. Cottin, P. Y. Brillet, O. Freynet, N. Naggara, Z. Carton, J. F. Cordier, M. Brauner, D. Valeyre and H. Nunes, *Thorax*, 2011, **66**, 226-231.
67. M. Korfei, C. Ruppert, P. Mahavadi, I. Henneke, P. Markart, M. Koch, G. Lang, L. Fink, R. M. Bohle, W. Seeger, T. E. Weaver and A. Guenther, *American Journal of Respiratory and Critical Care Medicine*, 2008, **178**, 838-846.
68. T. H. Sisson, M. Mendez, K. Choi, N. Subbotina, A. Courey, A. Cunningham, A. Dave, J. F. Engelhardt, X. M. Liu, E. S. White, V. J. Thannickal, B. B. Moore, P. J. Christensen and R. H. Simon, *American Journal of Respiratory and Critical Care Medicine*, 2010, **181**, 254-263.
69. B. C. Willis, J. M. Liebler, K. Luby-Phelps, A. G. Nicholson, E. D. Crandall, R. M. du Bois and Z. Borok, *American Journal of Pathology*, 2005, **166**, 1321-1332.
70. A. Moeller, S. E. Gilpin, K. Ask, G. Cox, D. Cook, J. Gauldie, P. J. Margetts, L. Farkas, J. Dobranowski, C. Boylan, P. M. O'Byrne, R. M. Strieter and M. Kolb, *American Journal of Respiratory and Critical Care Medicine*, 2009, **179**, 588-594.
71. Y. Miyazaki, K. Araki, C. Vesin, I. Garcia, Y. Kapanci, J. A. Whitsett, P. F. Piguet and P. Vassalli, *Journal of Clinical Investigation*, 1995, **96**, 250-259.
72. G. Raghu, K. K. Brown, U. Costabel, V. Cottin, R. M. du Bois, J. A. Lasky, M. Thomeer, J. P. Utz, R. K. Khandker, L. McDermott and S. Fatenejad, *American Journal of Respiratory and Critical Care Medicine*, 2008, **178**, 948-955.
73. J. S. Zhao, W. Shi, Y. L. Wang, H. Chen, P. Bringas, M. B. Datto, J. P. Frederick, X. F. Wang and D. Warburton, *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 2002, **282**, L585-L593.
74. V. Ruiz, R. M. Ordonez, J. Berumen, R. Ramirez, B. Uhal, C. Becerril, A. Pardo and M. Selman, *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 2003, **285**, L1026-L1036.
75. T. Okuma, Y. Terasaki, K. Kaikita, H. Kobayashi, W. A. Kuziel, M. Kawasuji and M. Takeya, *Journal of Pathology*, 2004, **204**, 594-604.
76. M. Gharaee-Kermani, R. E. McCullumsmith, I. F. Charo, S. L. Kunkel and S. H. Phan, *Cytokine*, 2003, **24**, 266-276.
77. L. Steinman, *Nature Medicine*, 2007, **13**, 139-145.
78. Y. Yogo, S. Fujishima, T. Inoue, F. Saito, T. Shiomi, K. Yamaguchi and A. Ishizaka, *Respiratory Research*, 2009, **10**.
79. M. P. Keane, *European Respiratory Review*, 2008, **17**, 151-156.
80. W. E. Lawson, P. F. Crossno, V. V. Polosukhin, J. Roldan, D. S. Cheng, K. B. Lane, T. R. Blackwell, C. Xu, C. Markin, L. B. Ware, G. G. Miller, J. E. Loyd and T. S. Blackwell, *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 2008, **294**, L1119-L1126.
81. M. Waghray, Z. B. Cui, J. C. Horowitz, I. M. Subramanian, F. J. Martinez, G. B. Toews and V. J. Thannickal, *Faseb Journal*, 2005, **19**, 854-+.

82. R. C. Chambers, P. Leoni, O. P. Blanc-Brude, D. E. Wembridge and G. J. Laurent, *Journal of Biological Chemistry*, 2000, **275**, 35584-35591.
83. R. G. Jenkins, X. Su, G. Su, C. J. Scotton, E. Camerer, G. J. Laurent, G. E. Davis, R. C. Chambers, M. A. Matthay and D. Sheppard, *Journal of Clinical Investigation*, 2006, **116**, 1606-1614.
84. P. F. Mercer, R. H. Johns, C. J. Scotton, M. A. Krupiczojc, M. Konigshoff, D. C. J. Howell, R. J. McAnulty, A. Das, A. J. Thorley, T. D. Tetley, O. Eickelberg and R. C. Chambers, *American Journal of Respiratory and Critical Care Medicine*, 2009, **179**, 414-425.
85. O. P. Blanc-Brude, F. Archer, P. Leoni, C. Derian, S. Bolsover, G. J. Laurent and R. C. Chambers, *Experimental Cell Research*, 2005, **304**, 16-27.
86. R. C. Chambers, K. Dabbagh, R. J. McAnulty, A. J. Gray, O. P. Blanc-Brude and G. J. Laurent, *Biochemical Journal*, 1998, **333**, 121-127.
87. G. S. Bogatkevich, E. Tourkina, R. M. Silver and A. Ludwicka-Bradley, *Journal of Biological Chemistry*, 2001, **276**, 45184-45192.
88. C. J. Scotton, M. A. Krupiczojc, M. Konigshoff, P. F. Mercer, Y. C. G. Lee, N. Kaminski, J. Morser, J. M. Post, T. M. Maher, A. G. Nicholson, J. D. Moffatt, G. J. Laurent, C. K. Derian, O. Eickelberg and R. C. Chambers, *Journal of Clinical Investigation*, 2009, **119**, 2550-2563.
89. R. C. Chambers, *Br J Pharmacol*, 2008, **153 Suppl 1**, S367-378.
90. D. C. J. Howell, R. H. Johns, J. A. Lasky, B. Shan, C. J. Scotton, G. J. Laurent and R. C. Chambers, *American Journal of Pathology*, 2005, **166**, 1353-1365.
91. L. B. Chen and J. M. Buchanan, *Proceedings of the National Academy of Sciences of the United States of America*, 1975, **72**, 131-135.
92. T. K. H. Vu, D. T. Hung, V. I. Wheaton and S. R. Coughlin, *Cell*, 1991, **64**, 1057-1068.
93. L. W. Liu, T. K. H. Vu, C. T. Esmon and S. R. Coughlin, *Journal of Biological Chemistry*, 1991, **266**, 16977-16980.
94. T. J. Rydel, K. G. Ravichandran, A. Tulinsky, W. Bode, R. Huber, C. Roitsch and J. W. Fenton, *Science*, 1990, **249**, 277-280.
95. R. M. Scarborough, M. A. Naughton, W. Teng, D. T. Hung, J. Rose, T. K. H. Vu, V. I. Wheaton, C. W. Turck and S. R. Coughlin, *Journal of Biological Chemistry*, 1992, **267**, 13146-13149.
96. S. Nystedt, K. Emilsson, A. K. Larsson, B. Strombeck and J. Sundelin, *European Journal of Biochemistry*, 1995, **232**, 84-89.
97. H. Ishihara, A. J. Connolly, D. W. Zeng, M. L. Kahn, Y. W. Zheng, C. Timmons, T. Tram and S. R. Coughlin, *Nature*, 1997, **386**, 502-506.
98. M. L. Kahn, Y. W. Zheng, W. Huang, V. Bigornia, D. W. Zeng, S. Moff, R. V. Farese, C. Tam and S. R. Coughlin, *Nature*, 1998, **394**, 690-694.
99. W. F. Xu, H. Andersen, T. E. Whitmore, S. R. Presnell, D. P. Yee, A. Ching, T. Gilbert, E. W. Davie and D. C. Foster, *Proceedings of the National Academy of Sciences of the United States of America*, 1998, **95**, 6642-6646.
100. M. Nakanishi-Matsui, Y. W. Zheng, D. J. Sulciner, E. J. Weiss, M. J. Ludeman and S. R. Coughlin, *Nature*, 2000, **404**, 609-+.
101. J. N. McLaughlin, M. M. Patterson and A. B. Malik, *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**, 5662-5667.
102. R. Ramachandran, F. Noorbakhsh, K. Defea and M. D. Hollenberg, *Nature Reviews Drug Discovery*, 2012, **11**, 69-86.

103. K. Defea, *British Journal of Pharmacology*, 2008, **153**, S298-S309.
104. S. Rajagopal, K. Rajagopal and R. J. Lefkowitz, *Nature Reviews Drug Discovery*, 2010, **9**, 373-386.
105. V. S. Ossovskaya and N. W. Bunnett, *Physiological Reviews*, 2004, **84**, 579-621.
106. A. Russo, U. J. K. Soh, M. M. Paing, P. Arora and J. Trejo, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 6393-6397.
107. U. J. K. Soh, M. R. Dores, B. X. Chen and J. Trejo, *British Journal of Pharmacology*, 2010, **160**, 191-203.
108. M. Riewald and W. Ruf, *Proceedings of the National Academy of Sciences of the United States of America*, 2001, **98**, 7742-7747.
109. V. Trivedi, A. Boire, B. Tchemychev, N. C. Kaneider, A. J. Leger, K. O'Callaghan, L. Covic and A. Kuliopulos, *Cell*, 2009, **137**, 332-343.
110. K. Oikonomopoulou, K. K. Hansen, M. Saifeddine, N. Vergnolle, I. Tea, M. Blaber, S. I. Blaber, I. Scarisbrick, E. P. Diamandis and M. D. Hollenberg, *Biological Chemistry*, 2006, **387**, 817-824.
111. J. N. McLaughlin, L. X. Shen, M. Holinstat, J. D. Brooks, E. DiBenedetto and H. E. Hamm, *Journal of Biological Chemistry*, 2005, **280**, 25048-25059.
112. B. L. Wolfe, A. Marchese and J. Trejo, *Journal of Cell Biology*, 2007, **177**, 905-916.
113. J. R. Hamilton, P. B. Nguyen and T. M. Cocks, *Circulation Research*, 1998, **82**, 1306-1311.
114. A. A. Laniyonu and M. D. Hollenberg, *British Journal of Pharmacology*, 1995, **114**, 1680-1686.
115. K. Hirano and M. Hirano, *Journal of Pharmacological Sciences*, 2010, **114**, 127-133.
116. D. T. Hung, T. K. H. Vu, V. I. Wheaton, K. Ishii and S. R. Coughlin, *Journal of Clinical Investigation*, 1992, **89**, 1350-1353.
117. N. Vergnolle, M. D. Hollenberg and J. L. Wallace, *British Journal of Pharmacology*, 1999, **126**, 1262-1268.
118. N. W. Bunnett, *Semin Thromb Hemost*, 2006, **32 Suppl 1**, 39-48.
119. W. R. Ferrell, J. C. Lockhart, E. B. Kelso, L. Dunning, R. Plevin, S. E. Meek, A. J. H. Smith, G. D. Hunter, J. S. McLean, F. McGarry, R. Ramage, L. Jiang, T. Kanke and J. Kawagoe, *Journal of Clinical Investigation*, 2003, **111**, 35-41.
120. A. C. Chin, N. Vergnolle, W. K. MacNaughton, J. L. Wallace, M. D. Hollenberg and A. G. Buret, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 11104-11109.
121. N. Vergnolle, *Gut*, 2005, **54**, 867-874.
122. J. L. K. Wee, Y. T. Chionh, G. Z. Ng, S. N. Harbour, C. Allison, C. N. Pagel, E. J. Mackie, H. M. Mitchell, R. L. Ferrero and P. Sutton, *Gastroenterology*, 2010, **138**, 573-582.
123. F. Mule, M. C. Baffi and M. C. Cerra, *British Journal of Pharmacology*, 2002, **136**, 367-374.
124. A. Kawabata, R. Kuroda, N. Nagata, N. Kawao, T. Masuko, H. Nishikawa and K. Kawai, *British Journal of Pharmacology*, 2001, **133**, 1213-1218.
125. N. G. Arizmendi, M. Abel, K. Mihara, C. Davidson, D. Polley, A. Nadeem, T. El Mays, B. F. Gilmore, B. Walker, J. R. Gordon, M. D. Hollenberg and H. Vliagoftis, *Journal of Immunology*, 2011, **186**, 3164-3172.

126. T. M. Cocks, B. Fong, J. M. Chow, G. P. Anderson, A. G. Frauman, R. G. Goldie, P. J. Henry, M. J. Carr, J. R. Hamilton and J. D. Moffatt, *Nature*, 1999, **398**, 156-160.

127. P. J. Vaughan, C. J. Pike, C. W. Cotman and D. D. Cunningham, *Journal of Neuroscience*, 1995, **15**, 5389-5401.

128. F. M. Donovan, C. J. Pike, C. W. Cotman and D. D. Cunningham, *Journal of Neuroscience*, 1997, **17**, 5316-5326.

129. Y. F. Wang, W. B. Luo, R. Stricker and G. Reiser, *Journal of Neurochemistry*, 2006, **98**, 1046-1060.

130. M. Thiyagarajan, J. A. Fernandez, S. M. Lane, J. H. Griffin and B. V. Zlokovic, *Journal of Neuroscience*, 2008, **28**, 12788-12797.

131. C. E. Junge, T. Sugawara, G. Mannaioni, S. Alagarsamy, P. J. Conn, D. J. Brat, P. H. Chan and S. F. Traynelis, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 13019-13024.

132. L. A. Boven, N. Vergnolle, S. D. Henry, C. Silva, Y. Imai, J. Holden, K. Warren, M. D. Hollenberg and C. Power, *Journal of Immunology*, 2003, **170**, 2638-2646.

133. J. I. Zwicker, B. C. Furie and B. Furie, *Critical Reviews in Oncology Hematology*, 2007, **62**, 126-136.

134. F. R. Rickles, *Pathophysiology of Haemostasis and Thrombosis*, 2006, **35**, 103-110.

135. A. Boire, L. Covic, A. Agarwal, S. Jacques, S. Sherifl and A. Kuliopulos, *Cell*, 2005, **120**, 303-313.

136. A. J. Ramsay, Y. Dong, M. L. Hunt, M. Linn, H. Samaratunga, J. A. Clements and J. D. Hooper, *Journal of Biological Chemistry*, 2008, **283**, 12293-12304.

137. V. Gratio, C. Loriot, G. D. Virca, K. Oikonomopoulou, F. Walker, E. P. Diamandis, M. D. Hollenberg and D. Darmoul, *American Journal of Pathology*, 2011, **179**, 2625-2636.

138. V. Gratio, N. Beaufort, L. Seiz, J. Maier, G. D. Virca, M. Debela, N. Grebenchtchikov, V. Magdolen and D. Darmoul, *American Journal of Pathology*, 2010, **176**, 1452-1461.

139. S. C. Even-Ram, M. Maoz, E. Pokroy, R. Reich, B. Z. Katz, P. Gutwein, P. Altevogt and R. Bar-Shavit, *Journal of Biological Chemistry*, 2001, **276**, 10952-10962.

140. S. Even-Ram, B. Uziely, P. Cohen, S. Grisaru-Granovsky, M. Maoz, Y. Ginzburg, R. Reich, I. Vlodavsky and R. Bar-Shavit, *Nature Medicine*, 1998, **4**, 909-914.

141. M. D. Hollenberg, M. Saifeddine, B. Al-Ani and A. Kawabata, *Canadian Journal of Physiology and Pharmacology*, 1997, **75**, 832-841.

142. Y. Kato, Y. Kita, M. Nishio, Y. Hirasawa, K. Ito, T. Yamanaka, Y. Motoyama and J. Seki, *European Journal of Pharmacology*, 1999, **384**, 197-202.

143. Y. Kato, Y. Kita, Y. Hirasawa-Taniyama, M. Nishio, K. Mihara, K. Ito, T. Yamanaka, J. Seki, S. Miyata and S. Mutoh, *European Journal of Pharmacology*, 2003, **473**, 163-169.

144. J. C. Barrow, P. G. Nantermet, H. G. Selnick, K. L. Glass, P. L. Ngo, M. B. Young, J. M. Pellicore, M. J. Breslin, J. H. Hutchinson, R. M. Freidinger, C. Condra, J. Karczewski, R. A. Bednar, S. L. Gaul, A. Stern, R. Gould and T.

M. Connolly, *Bioorganic & Medicinal Chemistry Letters*, 2001, **11**, 2691-2696.

145. P. G. Nantermet, J. C. Barrow, G. F. Lundell, J. M. Pellicore, K. E. Rittle, M. Young, R. M. Freidinger, T. M. Connolly, C. Condra, J. Karczewski, R. A. Bednar, S. L. Gaul, R. J. Gould, K. Prendergast and H. G. Selnick, *Bioorganic & Medicinal Chemistry Letters*, 2002, **12**, 319-323.

146. S. D. Tomasello, D. J. Angiolillo and S. Goto, *Expert Opinion on Investigational Drugs*, 2010, **19**, 1557-1567.

147. H. P. Zhao, H. M. Jiang and B. R. Xiang, *Expert Opinion on Investigational Drugs*, 2013, **22**, 1437-1451.

148. H. S. Ahn, L. Arik, G. Boykow, D. A. Burnett, M. A. Caplen, M. Czarniecki, M. S. Domalski, C. Foster, M. Manna, A. W. Stamford and Y. S. Wu, *Bioorganic & Medicinal Chemistry Letters*, 1999, **9**, 2073-2078.

149. G. D. Barry, G. T. Le and D. P. Fairlie, *Current Medicinal Chemistry*, 2006, **13**, 243-265.

150. S. Chackalamannil, Y. Xia, W. J. Greenlee, M. Clasby, D. Doller, H. Tsai, T. Asberom, M. Czarniecki, H. S. Ahn, G. Boykow, C. Foster, J. Agans-Fantuzzi, M. Bryant, J. Lau and M. Chintala, *Journal of Medicinal Chemistry*, 2005, **48**, 5884-5887.

151. S. Chackalamannil, Y. G. Wang, W. J. Greenlee, Z. Y. Hu, Y. Xia, H. S. Ahn, G. Boykow, Y. S. Hsieh, J. Palamanda, J. Agans-Fantuzzi, S. Kurowski, M. Graziano and M. Chintala, *Journal of Medicinal Chemistry*, 2008, **51**, 3061-3064.

152. M. P. Bonaca, B. M. Scirica, M. A. Creager, J. Olin, H. Bounameaux, M. Dellborg, J. M. Lamp, S. A. Murphy, E. Braunwald and D. A. Morrow, *Circulation*, 2013, **127**, 1522-+.

153. S. Leonardi, P. Tricoci, H. D. White, P. W. Armstrong, Z. Huang, L. Wallentin, P. E. Aylward, D. J. Moliterno, F. Van de Werf, E. Chen, L. Providencia, J. E. Nordrehaug, C. Held, J. Strony, T. L. Rorick, R. A. Harrington and K. W. Mahaffey, *European Heart Journal*, 2013, **34**, 1723-1731.

154. E. M. Ruda, M. C. Scrutton, P. W. Manley and D. P. Tuffin, *Biochemical Pharmacology*, 1990, **39**, 373-381.

155. E. M. Ruda, A. Petty, M. C. Scrutton, D. P. Tuffin and P. W. Manley, *Biochemical Pharmacology*, 1988, **37**, 2417-2426.

156. D. F. McComsey, L. R. Hecker, P. Andrade-Gordon, M. F. Addo and B. E. Maryanoff, *Bioorganic & Medicinal Chemistry Letters*, 1999, **9**, 255-260.

157. M. A. Ceruso, D. F. McComsey, G. C. Leo, P. Andrade-Gordon, M. F. Addo, R. M. Scarborough, D. Oksenberg and B. E. Maryanoff, *Bioorganic & Medicinal Chemistry*, 1999, **7**, 2353-2371.

158. M. P. Beavers, B. E. Maryanoff, D. Nguyen and B. D. Blackhart, in *Advances in Medicinal Chemistry*, eds. B. E. Maryanoff and A. B. Reitz, JAI Press Inc., Stamford, Connecticut, 1999, vol. 4, pp. 245-271.

159. C. Zhang, Y. Srinivasan, D. H. Arlow, J. J. Fung, D. Palmer, Y. Zheng, H. F. Green, A. Pandey, R. O. Dror, D. E. Shaw, W. I. Weis, S. R. Coughlin and B. K. Kobilka, *Nature*, 2012, **492**, 387-392.

160. M. S. Bernatowicz, C. E. Klimas, K. S. Hartl, M. Peluso, N. J. Allegretto and S. M. Seiler, *Journal of Medicinal Chemistry*, 1996, **39**, 4879-4887.

161. W. J. Hoekstra, B. L. Hulshizer, D. F. McComsey, P. Andrade-Gordon, J. A. Kauffman, M. F. Addo, D. Oksenberg, R. M. Scarborough and B. E. Maryanoff, *Bioorganic & Medicinal Chemistry Letters*, 1998, **8**, 1649-1654.

162. D. F. McComsey, M. J. Hawkins, P. Andrade-Gordon, M. F. Addo, D. Oksenberg and B. E. Maryanoff, *Bioorganic & Medicinal Chemistry Letters*, 1999, **9**, 1423-1428.

163. P. Andrade-Gordon, B. E. Mayanoff, C. K. Derian, H. C. Zhang, M. F. Addo, A. L. Darrow, A. J. Eckardt, W. J. Hoekstra, D. F. McComsey, D. Oksenberg, E. E. Reynolds, R. J. Santulli, R. M. Scarborough, C. E. Smith and K. B. White, *Proceedings of the National Academy of Sciences of the United States of America*, 1999, **96**, 12257-12262.

164. H. C. Zhang, C. K. Derian, P. Andrade-Gordon, W. J. Hoekstra, D. F. McComsey, K. B. White, B. L. Poulter, M. F. Addo, W. M. Cheung, B. P. Damiano, D. Oksenberg, E. E. Reynolds, A. Pandey, R. M. Scarborough and B. E. Maryanoff, *Journal of Medicinal Chemistry*, 2001, **44**, 1021-1024.

165. P. Andrade-Gordon, C. K. Derian, B. E. Maryanoff, H. C. Zhang, M. F. Addo, W. M. Cheung, B. P. Damiano, M. R. D'Andrea, A. L. Darrow, L. De Garavilla, A. J. Eckardt, E. C. Giardino, B. J. Haertlein and D. F. McComsey, *Journal of Pharmacology and Experimental Therapeutics*, 2001, **298**, 34-42.

166. B. P. Damiano, C. K. Derian, B. E. Maryanoff, H. C. Zhang and P. A. Gordon, *Cardiovascular Drug Reviews*, 2003, **21**, 313-326.

167. H. C. Zhang, D. F. McComsey, K. B. White, M. F. Addo, P. Andrade-Gordon, C. K. Derian, D. Oksenberg and B. E. Maryanoff, *Bioorganic & Medicinal Chemistry Letters*, 2001, **11**, 2105-2109.

168. G. Buchi, G. C. M. Lee, D. Yang and S. R. Tannenbaum, *Journal of the American Chemical Society*, 1986, **108**, 4115-4119.

169. S. R. Boothroyd and M. A. Kerr, *Tetrahedron Letters*, 1995, **36**, 2411-2414.

170. A. M. Valdivielso, M. T. Garcia-Lopez and R. Herranz, *Arkivoc*, 2008, 287-294.

171. H.-C. Zhang, K. B. White, D. F. McComsey, M. F. Addo, P. Andrade-Gordon, C. K. Derian, D. Oksenberg and B. E. Maryanoff, *Bioorganic & Medicinal Chemistry Letters*, 2003, **13**, 2199-2203.

172. I. M. Zohn, S. L. Campbell, R. Khosravi-Far, K. L. Rossman and C. J. Der, *Oncogene*, 1998, **17**, 1415-1438.

173. X. Deng, P. F. Mercer, C. J. Scotton, A. Gilchrist and R. C. Chambers, *Mol Biol Cell*, 2008, **19**, 2520-2533.

174. A. Ortiz-Stern, X. Deng, N. Smoktunowicz, P. F. Mercer and R. C. Chambers, *Journal of Cellular Physiology*, 2012, **227**, 3575-3584.

175. S. L. Tressel, N. C. Kaneider, S. Kasuda, C. Foley, G. Koukos, K. Austin, A. Agarwal, L. Covic, S. M. Opal and A. Kuliopulos, *Embo Molecular Medicine*, 2011, **3**, 370-384.

176. P. F. Mercer, A. E. Williams, C. J. Scotton, R. J. Jose, M. Sulikowski, J. D. Moffatt, L. A. Murray and R. C. Chambers, *American Journal of Respiratory Cell and Molecular Biology*, 2014, **50**, 144-157.

177. K. H. Glusenkamp, C. Mengede, W. Drosdziok, E. Jahde and M. F. Rajewsky, *Bioorganic & Medicinal Chemistry Letters*, 1998, **8**, 285-288.

178. D. O'Hagan, *Chemical Society Reviews*, 2008, **37**, 308-319.

179. M. A. Shuman and S. P. Levine, *Journal of Clinical Investigation*, 1978, **61**, 1102-1106.

180. C. Huang, W. J. Moree, S. Zamani-Kord, B. F. Li, F. C. Tucci, S. Malany, J. Wen, H. Wang, S. R. Hoare, C. Yang, A. Madan, P. D. Crowe and G. Beaton, *Bioorg Med Chem Lett*, 2011, **21**, 947-951.
181. L. F. Brass, S. Pizarro, M. Ahuja, E. Belmonte, N. Blanchard, J. M. Stadel and J. A. Hoxie, *Journal of Biological Chemistry*, 1994, **269**, 2943-2952.
182. L. T. May, K. Leach, P. M. Sexton and A. Christopoulos, in *Annual Review of Pharmacology and Toxicology*, 2007, vol. 47, pp. 1-51.
183. B. Gabriele, G. Salerno, R. Mancuso and M. Costa, *The Journal of Organic Chemistry*, 2004, **69**, 4741-4750.
184. A. Eisenführ, P. S. Arora, G. Sengle, L. R. Takaoka, J. S. Nowick and M. Famulok, *Bioorganic & Medicinal Chemistry*, 2003, **11**, 235-249.
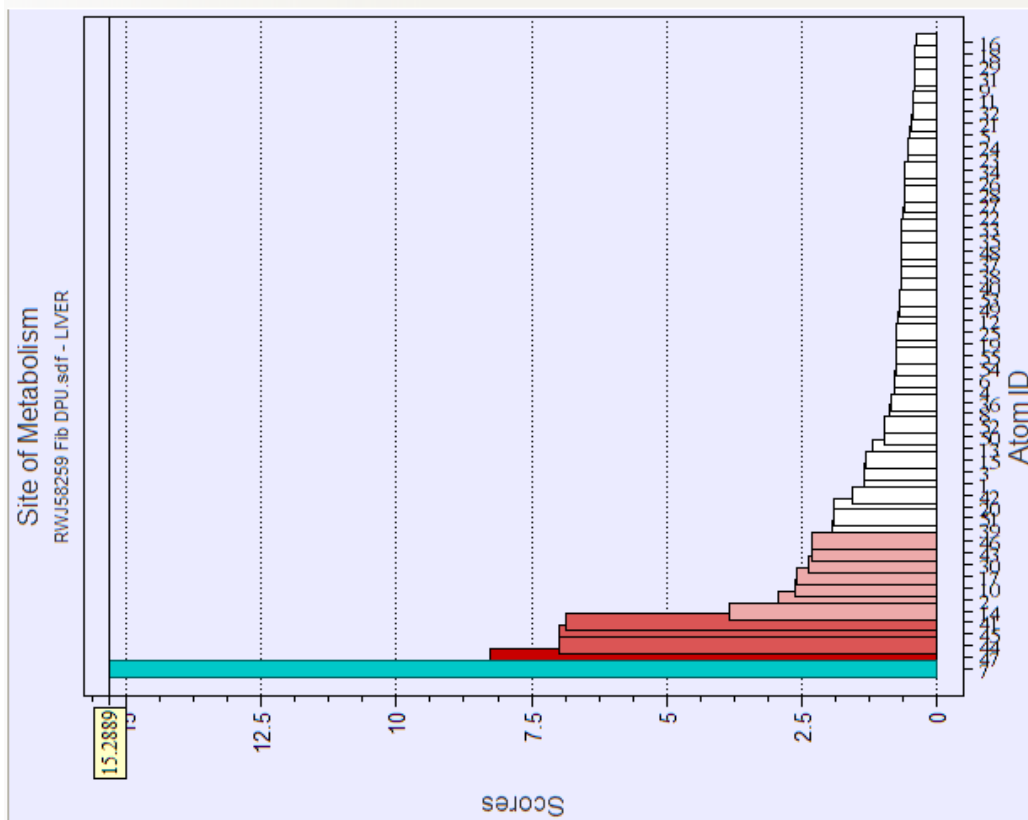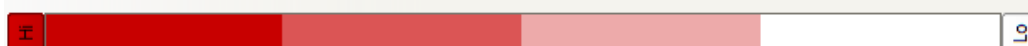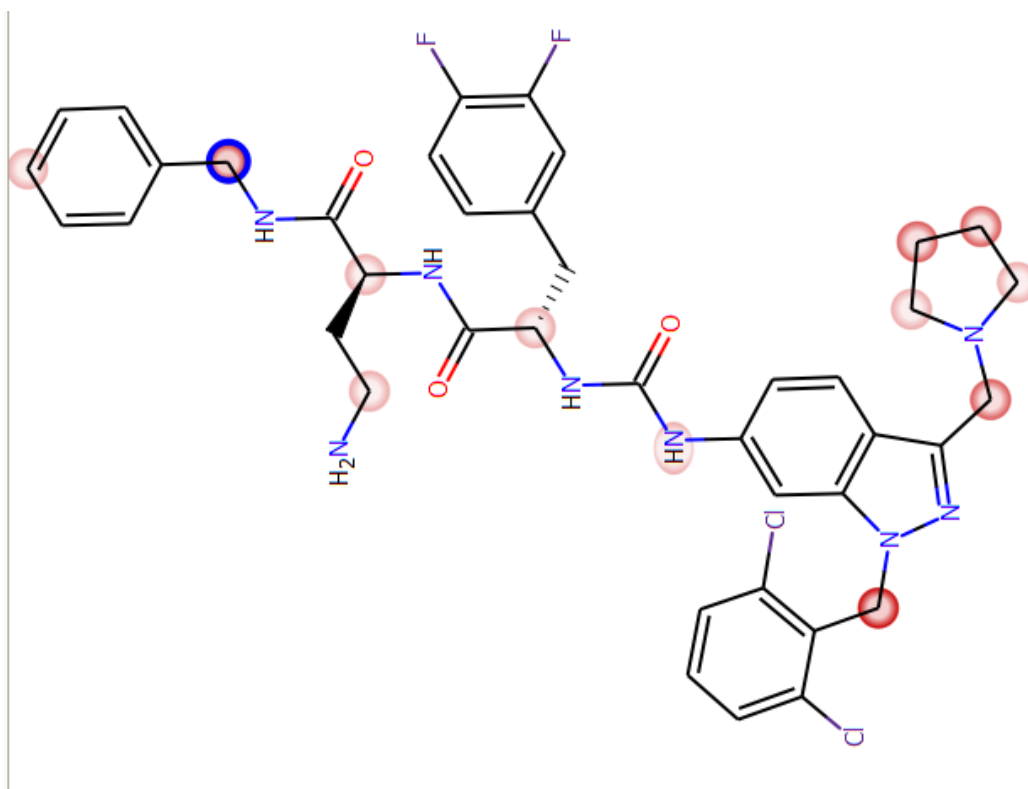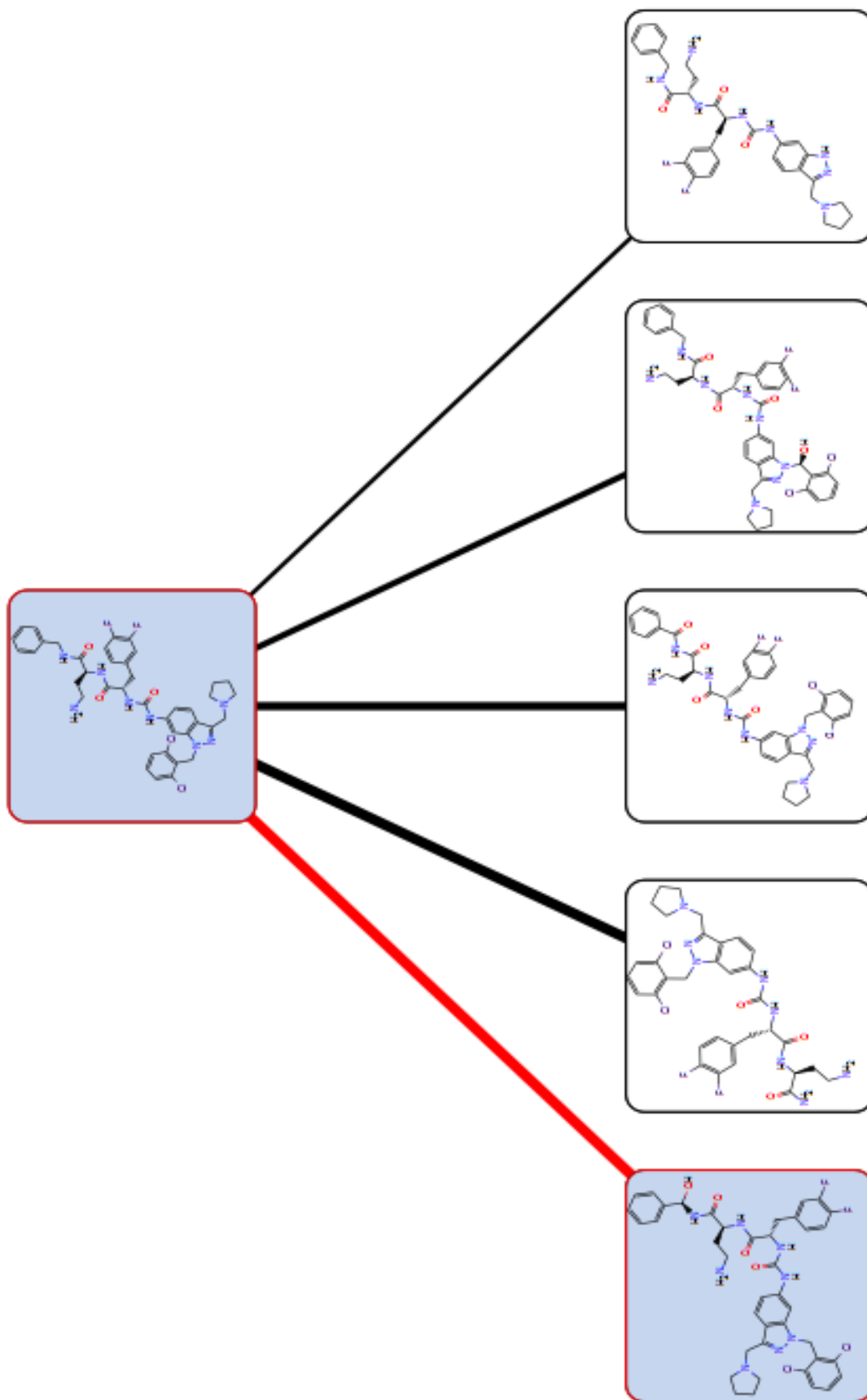
Appendix

MetaSite output

Statistical analysis of means from medium only *versus* medium with 0.1% DMSO

FLIPR traces for compounds **62a** & **62b**

MetaSite Output:

Student's T-test:

To determine if difference between mean values is significant (requires $p < 0.05$)

		[Thrombin]			
		0 nM	0.3 nM	3 nM	10 nM
Medium	Experimental replicate 1	70	108	767	868
	Experimental replicate 2	83	144	756	886
	Experimental replicate 3	46	115	713	921
	Experimental replicate 4	79	140	660	837
	Mean value	69.5	126.75	724	878
Med+DMSO	Experimental replicate 1	78	86	755	864
	Experimental replicate 2	54	142	730	908
	Experimental replicate 3	67	115	718	891
	Experimental replicate 4	54	166	733	965
	Mean value	63.25	127.25	734	907
	t test score	0.559351	0.980311	0.708511	0.334408

Therefore, difference is not significant and values can be said to be collect from same population, *i.e.* there is no significant difference caused by the presence of 0.1% DMSO.

FLIPR traces for compounds 62a & 62b:
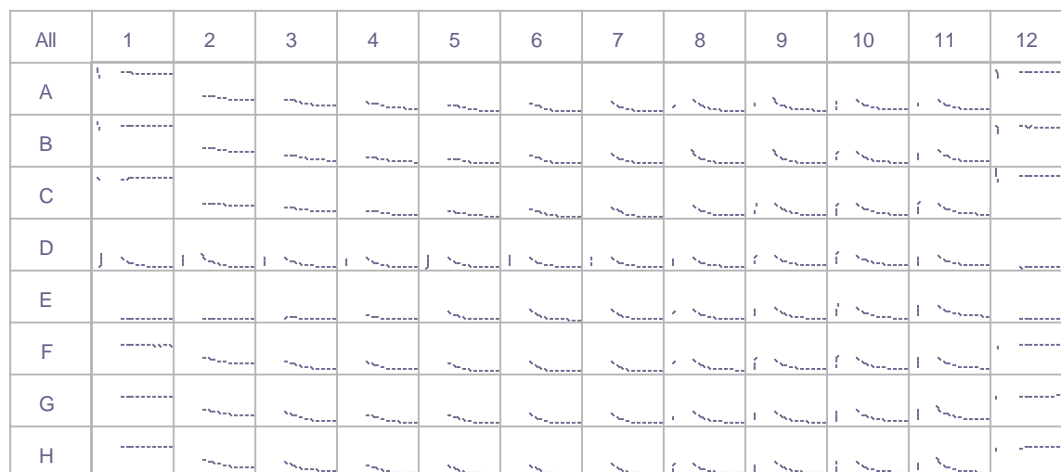
Rows A-C: Compound **62a** (Col 1-11, conc. 300 μM-3 nM; Col 12, conc. 300 μM)

Row D: Medium only

Row E: Compound **20** (Col 1-11, conc. 10 μM-0.1 nM; Col 12, conc. 10 μM)

Rows F-H: Compound **62b** (Col 1-11, conc. 300 μM-3 nM; Col 12, conc. 300 μM)

Columns 1-11: 10 nM thrombin

Column 12: 0 nM thrombin