Structure-Guided Discovery of a Novel Steroid-Derived Family of Autotaxin Inhibitors Based on Endogenous Allosteric Modulators

Jennifer M. Clark

September 2020

STRUCTURE-GUIDED DISCOVERY OF A NOVEL STEROID-DERIVED FAMILY OF AUTOTAXIN INHIBITORS BASED ON ENDOGENOUS ALLOSTERIC MODULATORS

Thesis submitted to the University of Strathclyde in fulfilment of the requirements for the degree of Doctor of Philosophy.

By

Jennifer M. Clark

2020

Declaration of Copyright

This thesis is a result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date: 14 September 2020

Publication List

1. *"Structure-Guided Discovery of a Novel Family of Autotaxin Inhibitors Based on Endogenous Allosteric Modulators"* Jennifer M. Clark, Fernando Salgado-Polo, Simon. J. F. Macdonald, Tim Barrett, Anastassis Perrakis and Craig Jamieson. *Manuscript in Preparation.*

2. *"Development of Novel Autotaxin Inhibitors Derived From PF-8380"* Frances Potjewyd, Jennifer M. Clark, Fernando Salgado-Polo, Willem-Jan Keune, Emma L. Duffy, Diana Castagna, Lisa M. Miller, Louise C. Young, Simon J. F. Macdonald, Anastassis Perrakis, Allan. J. B. Watson and Craig Jamieson. *Manuscript in Preparation.*

Conference List

Oral Presentations

1. *"Development of a family of chemical probes for autotaxin based on endogenous allosteric inhibitors"* Jennifer M. Clark (2020), Oral Presentation for GSK Medicinal Chemistry Seminar (Virtual).

2. *"Design, synthesis and evaluation of novel autotaxin inhibitors"* Jennifer M. Clark (2020), Oral Presentation scheduled at American Chemical Society National Meeting and Expo, Philadelphia, USA, cancelled due to the ongoing COVID-19 pandemic.

3. *"Design, synthesis and evaluation of novel autotaxin inhibitors"* Jennifer M. Clark (2019), Oral Presentation at 13th BMCS Medicinal Chemistry Symposium, Cambridge, U.K.

Poster and Flash Presentations

1. 2019, 2nd RSC Anglo-Nordic Medicinal Chemistry Symposium, Snekkersten, Denmark.

2. 2019, 20th SCI/RSC Medicinal Chemistry Symposium, Cambridge, U.K.

3. 2019, Mastering Medicinal Chemistry V, Windlesham, U.K.

Abstract

Autotaxin (ATX) facilitates the hydrolysis of lysophosphatidyl choline to lysophosphatidic acid, a bioactive signalling lipid, which acts on its cognate receptors to facilitate a diverse range of cellular effects in multiple tissues. Abnormal LPA expression can lead to the progression of diseases such as cancer and fibrosis.

Modulation of ATX has been predominantly based around the generation of orthosteric or "Type I" inhibitors, which structurally mimic the natural substrate, LPA (**Figure 1**). However, despite being active, their therapeutic development has been limited. Progress in inhibitor diversification has recently led to the evolution of potent hybrid inhibitors binding in both the tunnel and pocket of ATX, which fall into the class of "Type IV" inhibitors.

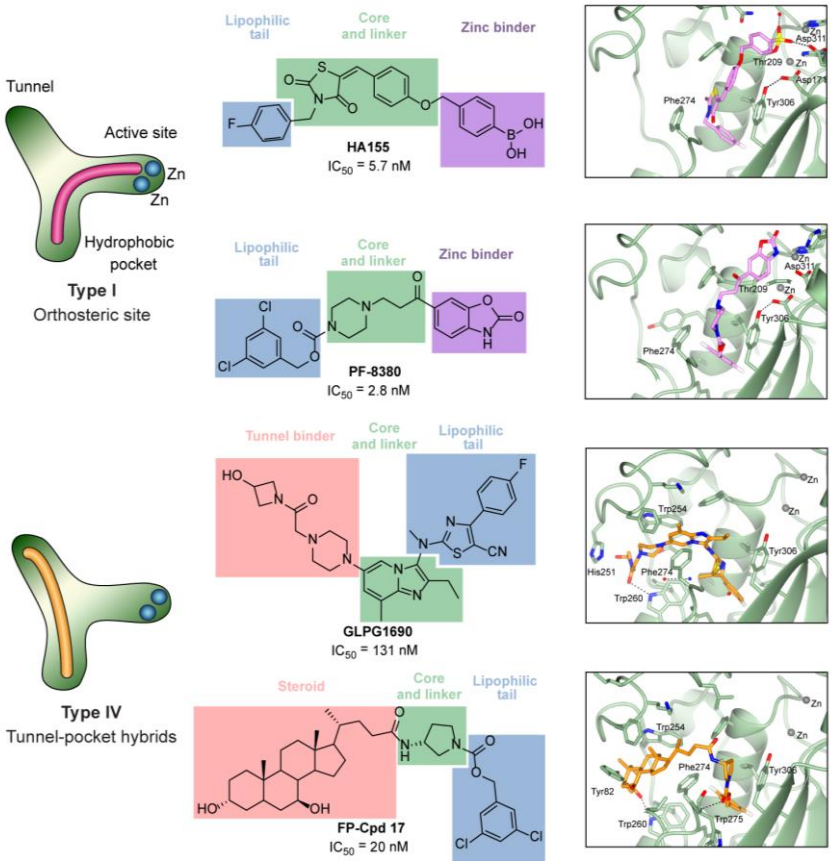


Figure 1 Two classes of highly potent ATX inhibitors – Type I - orthosteric binders and Type IV - tunnel-pocket hybrids with mixed allosteric-orthosteric inhibition.

Collaboration between Netherlands Cancer Institute and University of Strathclyde has facilitated the development of a novel class of inhibitors which can be confidently designated as first-in-class "Type V" inhibitors (**Figure 2**), the design, synthesis and biological evaluation of which will be described in this thesis.

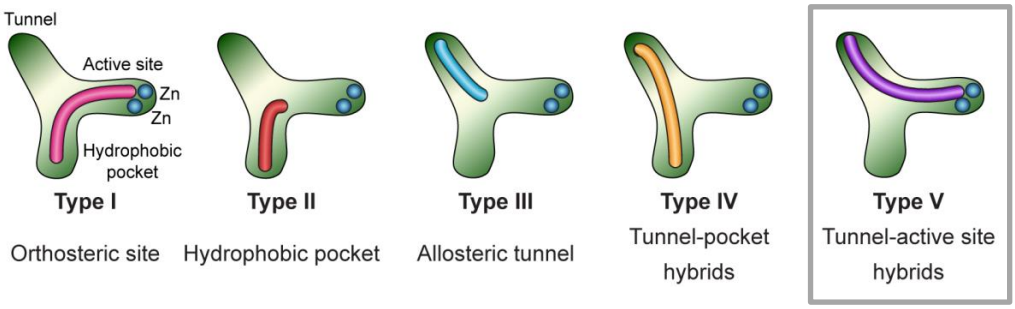


Figure 2 Designation of Type V inhibitors – tunnel-active site hybrids with mixed allosteric-orthosteric inhibition.

Amalgamation of key structural features from two comprehensive historical design hypotheses, based on both endogenous allosteric modulators and competitive orthosteric inhibitors, has led to the discovery of steroid-derived inhibitors armed with a key warhead pharmacophore, capable of interacting with the tunnel and active site of ATX.

Our binding hypothesis has been corroborated with three crystallographic structures of potent exemplar compounds bound to ATX, one of which we additionally disclose as a novel zinc binder with comparable potency to our lead analogue.

Further confirmatory studies have been undertaken through ATX kinetics, cell-based analysis and phenotypic studies in order to ascertain the fate of these novel compounds in the downstream signalling cascade and also in a disease-relevant context. Our lead compounds act through competitive inhibition of ATX and achieve modulation of LPA-mediated ATX allostery and inhibition of LPA-dependent signalling pathways in cells, which we believe serves as a solid baseline for further investigation into the impact of ATX in debilitating diseases, such as fibrosis and cancer (**Figure 3**).

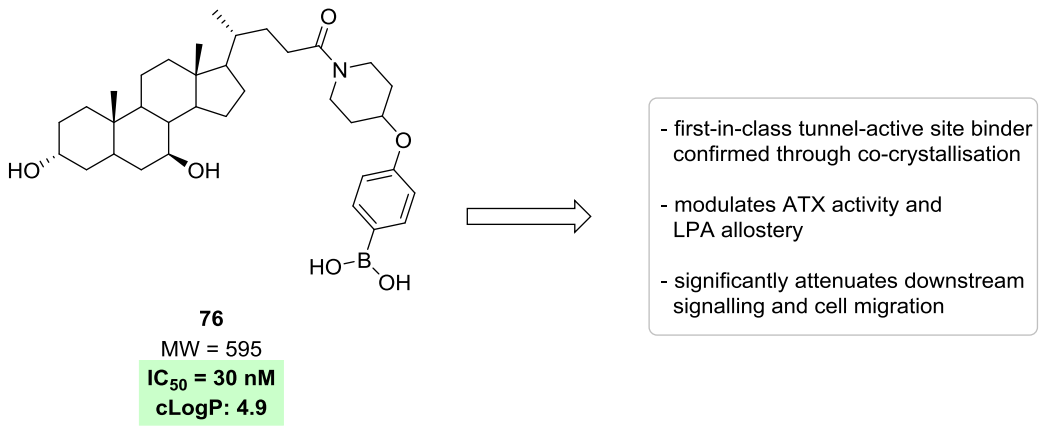


Figure 3 Lead analogue of Type V SAR library.

“Do not think of today’s failures, but of the success that may come tomorrow.

You have set yourself a difficult task, but you will succeed if you persevere and you will find joy in overcoming obstacles.”

Helen Keller

Acknowledgements

Firstly, I would like to thank my supervisor Dr Craig Jamieson for allowing me the opportunity to work on this research project, being a part of the autotaxin family has been such a pleasure. Thank you also for encouraging me to pursue the multi-disciplinary aspect and for supporting me in my conference and secondment ventures (even through all the bad luck streaks!).

I would also specifically like to thank Fernando Salgado-Polo and Professor Anastassis Perrakis at the Netherlands Cancer Institute in Amsterdam for sharing our interest in autotaxin and for your diligence with the biochemical and cell-based studies for this project. Your support and guidance throughout these last four years has been invaluable and I'll never be able to thank you enough for your hospitality and the knowledge you have gifted me during my two secondments. I would also like to acknowledge the Perrakis group and all of the groups on B8 for looking after me during my time in the Netherlands - thank you also for introducing me to Heineken at our Borrelen, the only beer I actually like!

I would like to extend my most sincere appreciation and gratitude to Medical Research Scotland for their financial support throughout my studentship. I have connected with many different people through their PhD programme and I will be forever grateful for the personal and professional opportunities that I have been afforded over the last four years. Thank you also to Professor Phillip Winn and Dr Alex Graham for investing your time and effort in us.

Thank you also to Dr Tim Barrett and Dr Simon Macdonald (GlaxoSmithKline, Stevenage) for your collective project input throughout the duration of my studentship.

A huge thank you to the technical and academic staff based in the Thomas Graham Building at University of Strathclyde, specifically, Craig Irving (NMR), Pat Keating and Alex Clunie (Mass Spec), my two favourite technicians Paul and Billy, stores staff - Margaret and Marc, the cleaners who kept me company during my early mornings and our resident office staff – Ros and the team, to name a few. You have been instrumental in my progression over the years and I would not be where I am without each and every one of you.

Special thanks to the past and present members of the CJAM group. I will miss being the lab mum - thank you for being a constant source of company in the lab throughout the last four years and for putting up with my love of dance music, may the JMC raves live on forever!

I would specifically like to thank my girls, Jennifer Duffy, Meghan Hamer and Jenna Mowat. Our friendship has been one of the highlights – thank you for the best daily support, chats and encouragement and for always being game for cocktails when required.

To my friends – you know who you are. Thank you for supporting me on my chemistry adventures, from a year in Frankfurt to secondments in Amsterdam to PhD life in Glasgow. To Calum, thank you for endless love, support and encouragement and for always having my favourite snacks and wine at the ready.

I would also like to thank the many outstanding and courageous women, both past and present, who have contributed to scientific research and continue to be positive role models for young women in STEM-related subjects. Their fortitude and strength will always be an inspiration for me and I am so thankful for the opportunity to follow in their footsteps.

Finally, to my mum, dad and little brother, Ross. Thank you from the bottom of my heart for everything you have done for me. I can't put into words how grateful I am for your time, patience, love, support and encouragement over the course of my learning years. Thank you for shaping me to be the person I am today and for encouraging me to follow my dreams. Most importantly, thank you for always believing in me ♥

2020 has been far from a normal year for most. I would like to take this opportunity to thank all of the key/essential workers, both in the UK and worldwide, from NHS staff to those working in food production or the transport sector and beyond, for their selflessness, strength and dedication in uniting us in this global fight against COVID-19.

Abbreviations

AC_{50}	Half-maximum activation concentration
ALT	Alanine aminotransferase
AKT	Protein kinase B
ASK-1	Apoptosis signal-regulating kinase 1
AST	Aspartate aminotransferase
ATX	Autotaxin
BCAA	Bicinchoninic acid
Bis-*p*NPP	Bis-(*p*-nitrophenyl) phosphate
BJeH	Human BJ foreskin fibroblasts
BSA	Bovine serum albumin
Boc	*tert*-Butyloxycarbonyl
CA	Cholic acid
Cbz	Carboxybenzyl
CDCA	Chenodeoxycholic acid
CDI	Carbonyldiimidazole
CMBP	Cyanomethylenetributylphosphorane
CO	Choline oxidase
DCM	Dichloromethane
DIAD	Diisopropyl azodicarboxylate
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethylsulfoxide

DOC	Sodium deoxycholate
DTT	Dithiothretiol
ECL	Enhanced chemiluminescent
EFL	EF-hand like
EDG	Endothelial differentiation gene
ENPP	Ectonucleotide pyrophosphatase/phosphodiesterase
ERK	Extracellular signal-regulated kinase
ex/em	Excitation/emission
FAF-BSA	Fatty acid free-bovine serum albumin
FCS	Foetal calf serum
FDA	Food and Drug Administration
FGFR	Fibroblast growth factor receptor
FVC	Forced vital capacity
GERD	Gastroesophageal reflux disease
GPCR	G protein-coupled receptor
GSK	GlaxoSmithKline
HRCT	High resolution computed tomography
HRMS	High resolution mass spectrometry
HRP	Horseradish peroxidase
HVA	Homovanillic acid
IC_{50}	Half-maximum inhibitory concentration
ILD	Interstitial lung disease
IPF	Idiopathic pulmonary fibrosis

LCA	Lithocholic acid
LPA	Lysophosphatidic acid
LPAR	Lysophosphatidic acid receptor
LPC	Lysophosphatidylcholine
NaAz	Sodium azide
NAFL	Non-alcoholic fatty liver
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NKI-AVL	Netherlands Cancer Institute-Antoni van Leewenhowk
PBS	Phosphate-buffered saline
PD	Pharmacodynamics
PDGFR	Platelet derived growth factor receptor
PEG	Polyethylene glycol
PI3K	Phosphatidylinositol-3-kinase
PK	Pharmacokinetics
*p*NP-TMP	*p*-Nitrophenyl-thymidine monophosphate
RIPA	Radio immunoprecipitation assay
RPM	Revolutions per minute
RPMI	Roswell Park Memorial Institute (Medium)
S1P	Sphingosine-1-phosphate
SAR	Structure-activity relationship
SCD-1	Stearoyl-CoA desaturase 1
SDS	Sodium dodecyl sulfate

SEM	Standard error of the mean
SLB	Surgical lung biopsy
SMB	Somatomedin-β
TGF-β1	Transforming growth factor-β1
THF	Tetrahydrofuran
TUDCA	Tauroursodeoxycholic acid
UDCA	Ursodeoxycholic acid
VEGFR	Vascular endothelial growth factor receptor
7-HCS	7-α-Hydroxycholesterol

Table of Contents

List of Figures, Schemes, and Tables

Figures

Tables

Schemes

1 Introduction

1.1 Ectonucleotide pyrophosphatase/phosphodiesterase family

1.1.1 Phosphodiesterases

Phosphodiester bonds represent a fundamental repeat unit in various biological environments, an example being in the construction of the sugar-phosphate backbone of DNA (**Figure 4**). The incorporation of nucleotides into a DNA strand results in the formation of a covalent bond through the joining of the phosphate group of one nucleotide via S_N2 reaction to the hydroxyl group of the adjacent nucleotide.

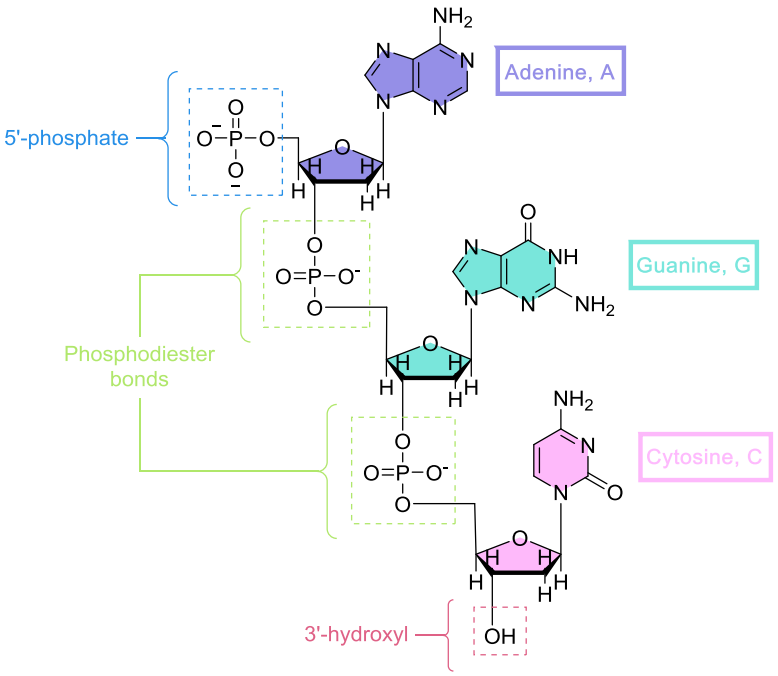


Figure 4 Primary structure of an example DNA strand showing key components: Sugar-phosphate backbone linked by phosphodiester bonds with respective nitrogenous bases.

Consequently, phosphodiester bonds can be hydrolytically cleaved by phosphodiesterase enzymes which release the respective phosphomonoesters. These enzymes can be categorised based on the nature of their substrate: lipid phosphodiesterases, for example,

sphingomyelinases,[1] cyclic nucleotide phosphodiesterases (PDE)[2] and ectonucleotide pyrophosphatases/phosphodiesterases (ENPP).[3]

1.1.2 Structure and Function of ENPP Enzymes

The mammalian ENPP family is comprised of seven highly conserved and structurally related ectoenzymes[4] which are numbered chronologically as they were discovered and were concomitantly named ENPP1-7. They function by hydrolysing phosphodiester and pyrophosphate bonds in both nucleotides and also in other substrates found extracellularly, for example, ENPP2, also known as autotaxin (ATX), exhibits lysoPLD activity[5] and is capable of facilitating the hydrolysis of lysophosphatidyl choline (LPC) to the bioactive signalling lipid, lysophosphatidic acid (LPA) as demonstrated in **Figure 5**.

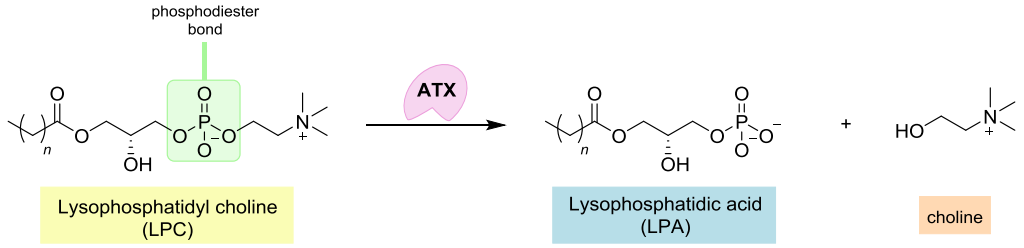


Figure 5 Phosphodiester bond hydrolysis of lysophosphatidyl choline to lysophosphatidic acid.

The tertiary structure of the ENPP family is modular in nature, wherein they all possess a catalytic domain which consists of around 400 residues, however the family can be sub-divided into two further categories: ENPP1-3 and ENPP4-7 wherein the respective structural composition at their termini demarcates the two categories.[6] The sequence composition of all seven proteins is exemplified in **Figure 6**.

ENPP4-7 are substantially less well-researched however, a putative Type I orientation has been assigned postulating that their catalytic domain is flanked by an N-terminal signal peptide and C-terminal transmembrane domain meaning they are retained within the plasma membrane.[6]

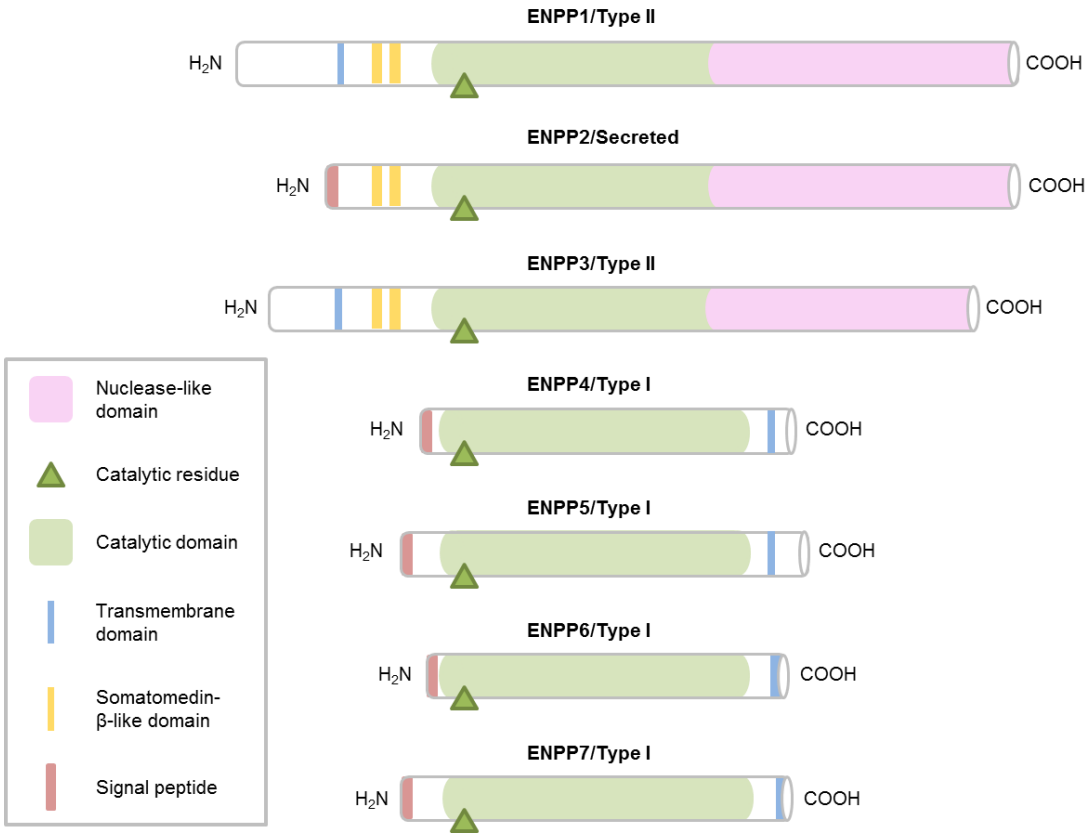
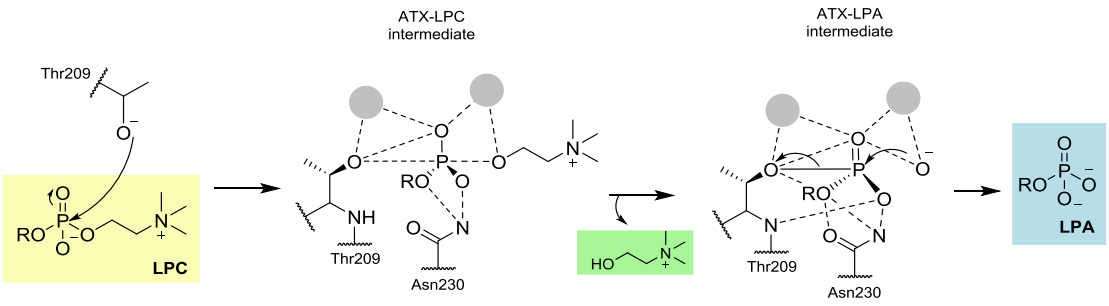


Figure 6 Cartoon representation of structural domain composition of ENPP isoenzymes – drawn generally, not specifically to scale.

ENPP1-3 are the most well-characterised of the seven proteins[4] and accordingly, have had their domain structures identified. ENPP1[7] and ENPP3[8] are confirmed as Type II transmembrane proteins and their structure is comprised of a C-terminal extracellular domain at one end which is linked to a single transmembrane domain and capped with an N-terminal intracellular domain.[9] ENPP2 is similar to ENPP1 and ENPP3 in that it consists of a C-terminal nuclease-like domain and two somatomedin-β (SMB) like domains located at the N-terminus, however ENPP2 is distinct in its own right. Surprisingly, it does not contain a transmembrane domain at its N-terminus[6] therefore is the only member of the ENPP family which is secreted. It shares a structural similarity with ENPP4-7 in that the sequence capping the N-terminus is a signal peptide therefore after removal of the signal peptide and subsequent trimming, ENPP2 is released into the secretory pathway.[10]

Focusing more specifically on ENPP1-3, the C-terminal domain can otherwise be known as the nuclease-like domain, and it is widely believed to be catalytically inactive for both ENPP1 and ENPP3 however, the presence of the nuclease domain is important in promoting the catalytic activity of the phosphodiesterase domain.[11] Reports suggest that the nuclease-like domain of ENPP2 could be responsible for its anti-adhesion properties.[6,12] At the N-terminus, all three isoenzymes contain two somatomedin-β like domains which were originally thought to be involved in homodimerisation through disulfide bridges,[13] however it was subsequently reported that neither ENPP2 nor extracellular ENPP1 form dimers.[14] When compared to the crystal structure of vitronectin,[15] it was apparent that the cysteine residues were in fact forming intra-chain disulfide bonds which is believed to be the same protein interactions involved in the somatomedin-β like domains of ENPP1-3.[6]

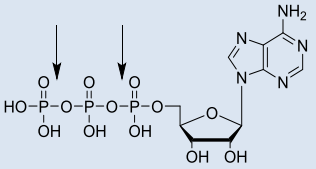
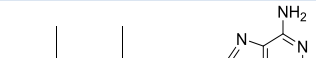
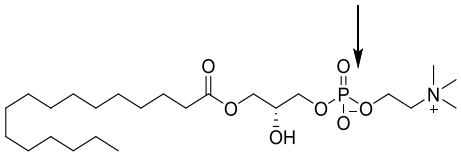
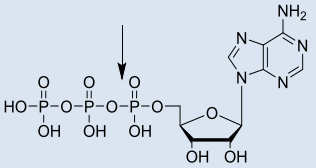
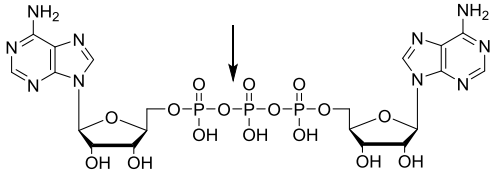
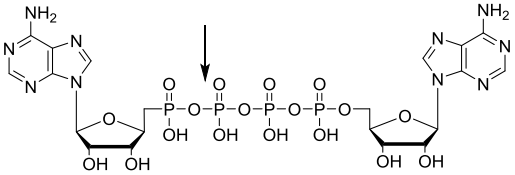
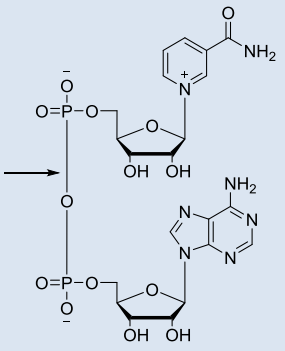
In terms of structure, the amino acid residues are conserved within the phosphodiesterase domain across all members of the ENPP family, wherein the catalytic residue threonine (serine in ENPP6) is stabilised by two zinc atoms in the active site. The catalytic mechanism proceeds similarly for all approaching substrates in a two-step fashion and is comparable to that observed in other phosphor/sulfo-coordinating enzymes such as the alkaline phosphatase family.[16] Using ENPP2 as an example, in the first step, the hydroxyl group of the catalytic threonine (or serine for ENPP6) is activated by one of the two zinc cations, initiating nucleophilic attack of the hydroxyl into the phosphodiester functionality of the inbound substrate (LPC) which, in turn, forms a new covalent bond, as demonstrated in **Scheme 1**.[17] A neighbouring water molecule is activated by the second zinc cation, initiating a second nucleophilic attack, which restores the respective catalytic residue and releases the cleaved product, in this case, LPA.[17]

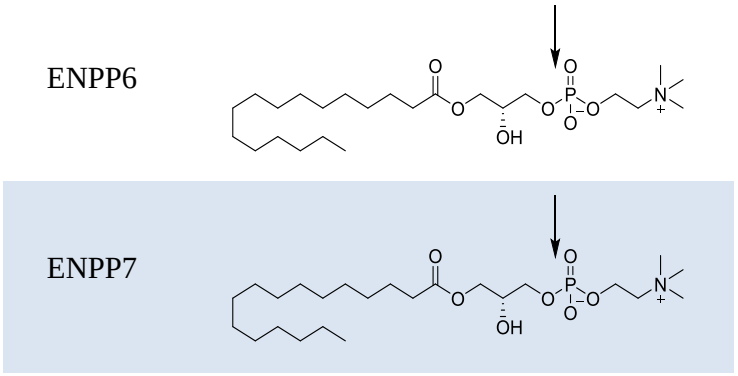


Scheme 1 Proposed catalytic mechanism initiated in phosphodiesterase domain of ENPP family (Ser for ENPP6).[17]

Significant differences exist between the independent isoenzymes in terms of substrate specificity at a binding site level and tissue distribution on an anatomical level. Firstly, with regard to specificity, little is known about the underlying preference for substrate binding. It appears that the small structural differences between each ENPP subtype have a significant influence on the outcome of substrate binding. For example, ENPP1 and ENPP3 have a specific preference for nucleotides whereas ENPP2/ATX is selective for phospholipids however, it can also accept nucleotide substrates (**Table 1**).[18] When compared to ENPP1, an interesting observation was noted in that ENPP1 has a segment of 18 amino acids which correspond with the hydrophobic pocket of ATX. Whilst ATX has a deletion in this region, it is believed that this additional section of amino acids in ENPP1 prevents binding of lysophospholipids and therefore explains why ENPP1 only hydrolyses nucleotides.[14]

Table 1 Isoenzyme substrate specificity

Isoenzyme	Substrate
ENPP1	
ENPP2	 
ENPP3	
ENPP4	 
ENPP5	



To add further complexity, it is possible to hydrolyse these substrates in different ways. Using LPC as an example, the phosphodiester bond can be hydrolysed either at one side which facilitates the release of LPA and choline by ENPP2, otherwise known as lysoPLD activity. It can be also be cleaved by ENPP6 at the other side to deliver monoacylglycerol and choline phosphate which is described as lysoPLC activity.[6] The substrate preference for ENPP4 and ENPP5 remains unclear despite studies elucidating their domain structures. Recently, the substrates for ENPP4 and ENPP5 have been identified as diadenosine polyphosphates Ap3A and Ap4A for ENPP4[19] and NAD^+ for ENPP5.[20]

The diversity associated with the ENPP family has inspired intensive research into their independent pathophysiological functions. However, ENPP2/ATX has remained at the forefront, predominantly due to its unique extracellular nature and distinctive role in a plethora of pathological and physiological modalities, which makes it an attractive and clinically relevant biomarker. In addition, being the only subtype with a hydrophobic pocket region, this structural distinction has been of particular interest over the years for exploratory drug discovery.

1.1.3 ENPP2/Autotaxin

First isolated from A2058 melanoma cells in 1992 by Stracke *et al.*,[21] autotaxin, otherwise known as ENPP2, was originally proposed as an autocrine motility factor which was established to be the ATXα isoform. Alternative splicing of ATX mRNA is understood to form distinct, catalytically active isoforms, namely α, β, γ, δ and ε which vary in sequence on exons 12, 19 and 21, facilitating their differential expression.

Studies undertaken independently by Tokumura *et al.*[22] and Umezu-Goto *et al.*[23] established that ATX exhibited similar activity to that observed of plasma lysoPLD. Related studies reported that plasma lysoPLD was specifically identical to the β-isoform of ATX, and consequently ATXβ is universally recognised as the canonical isoform, forming the basis of most ATX-related research. After confirming that ATX and lysoPLD were, in fact, the same, it was later identified that ENPP2/ATX was the fundamental mediator in the release of bioactive signalling lipid, LPA through the cleavage of LPC.[24] Delivery of LPA to its cognate G-protein coupled receptors, $LPAR_{1-6}$, results in the initiation of a variety of downstream signalling processes, for example, cell proliferation, migration and survival.

ATX/ENPP2 exists as a 125 kDa multi-domain protein and carries a central phosphodiesterase catalytic domain flanked by a catalytically inactive C-terminal nuclease-like domain (NUC) and two cysteine-rich N-terminal SMB-like domains (**Figure 7**).[24] Both of the SMB domains are extensively involved in interactions with the catalytic domain and it is understood that they mainly exist to facilitate ATX-platelet interaction with β3 integrins. An integrin-binding sequence known as RGD was recognised as part of SMB2 which further corroborates the role of integrin-binding with studies suggesting that SMB domains also play a role in LPA binding and regulation of ATX activity.[25]

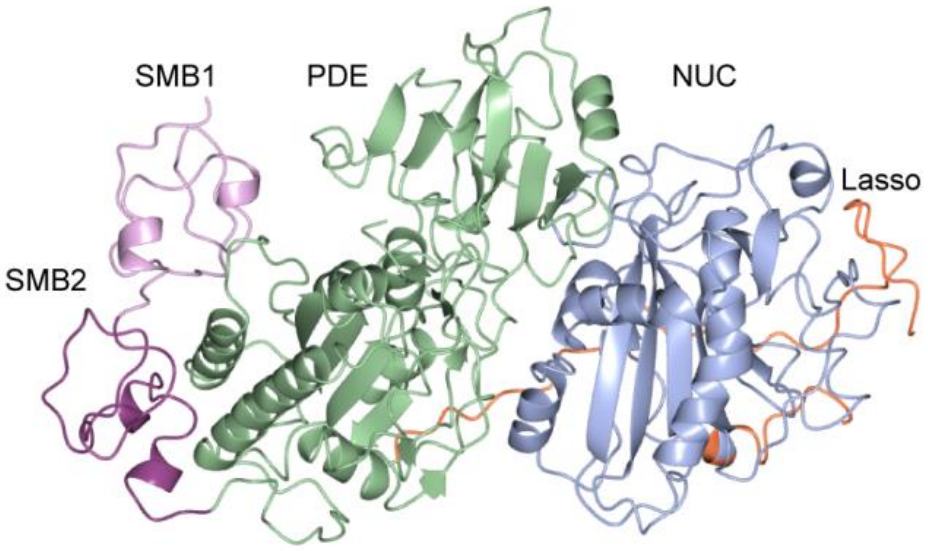


Figure 7 Crystal structure of ATX with key domains differentiated by colour (PDB ID: 2XR9).

The interface between the phosphodiesterase and nuclease domain is comprised primarily of an intermolecular Cys-bridge between residues 413 and 805 which is fundamental in preserving catalytic activity.[26] Additionally, a lasso, which consists of 50 amino acid residues, extends back around the length of the nuclease domain creating a tether between the far end of the catalytic domain and the nearside of the nuclease domain. A 13 residue-long EF-hand-like (EFL) functionality, which binds Ca^{2+} cations, is contained within the nuclease domain[26] and it is believed to be essential for activity in ENPP1 and ENPP3 but not for ATX.[27] The first residue of the EFL forms a salt bridge with Lys430 and interestingly, is the only residue outside the C-terminus of the nuclease domain that makes an interaction with the catalytic domain.

As discussed previously, the general construct of the phosphodiesterase domain is retained across all members of the ENPP family (apart from the catalytic residue in ENPP6) and can be sub-divided into three constituent components: the active site, hydrophobic pocket and hydrophobic tunnel. Catalytic activity takes place at the active site which contains two zinc cations and a threonine residue (Thr209 in *r*ATX and Thr210 in *h*ATX) (**Figure 8**).[26,28] The zinc cation closest to Thr209/210 is defined as the proximal zinc and is stabilised by the side chains of Asp171, Asp358 and His359. The second Zn^{2+} cation is termed the distal zinc and is coordinated by Asp311, His315 and His474. Another important feature is the Asn230 (Asn231 in *h*ATX) residue, positioned facing the two zinc atoms, which is understood to play a significant role in catalytic activity, in addition to being responsible for the predisposition of ENPP2 to phosphodiesterase activity. This has been confirmed through expression of a mutant ATX variant which did not contain Asn230, substantiating the hypothesis that it plays a necessary role in substrate hydrolysis.[29]

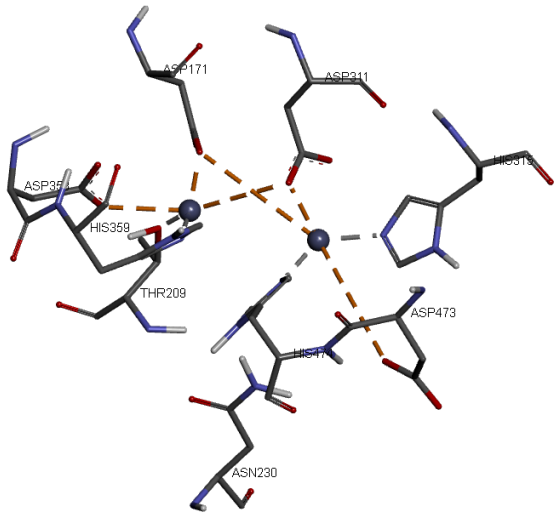


Figure 8 Key residues located in the active site of ATX (PDB ID: 2XR9),[18] viewed in DS Visualizer.

Directly following on from the active site, the hydrophobic pocket is exclusive to ENPP2 and contains several key residues, including Phe210, Leu213 and Tyr306, which have been confirmed as common interacting residues for most known LPC substrates.[28] LPA exists in circulation with differing acyl chain lengths and degrees of saturation, for example, LPA 14:0 has 14 carbon units and no unsaturation present whereas LPA 22:6 contains 22 carbon units and has 6 double bonds present in the acyl chain. Further discussion on LPA species can be found in later sections.

The final section of the phosphodiesterase domain is defined as the hydrophobic channel (**Figure 9**) and has long remained elusive in terms of its function despite being structurally characterised. Contributing polar and hydrophobic residues from the SMB-1 domain (Leu78, Tyr81 and Ser82) and the catalytic domain (Phe210, Tyr214, Lys248, Phe249, His251, Trp254, Pro258 and Trp260) form this solvent accessible interface which was previously believed to facilitate the shuttle of lipid-like molecules, between the active site and receptor surface.[28]

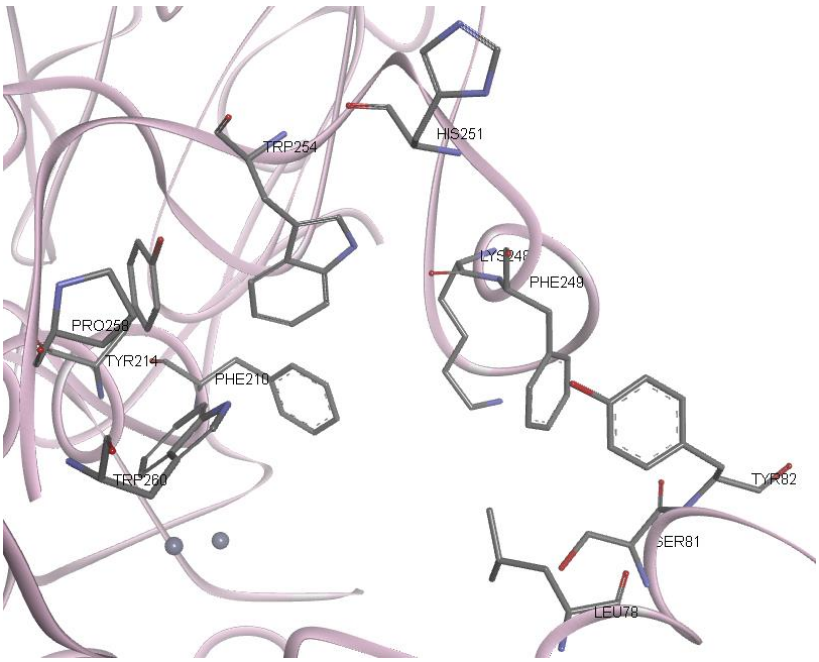


Figure 9 Key residues located in the hydrophobic tunnel of ATX (PDB ID: 2XR9),[18] viewed in DS Visualizer.

In order to provide a better understanding of the architecture of ATX, seminal work carried out by Nishimasu *et al.* implemented co-crystallography studies with mouse ATX involving individual LPA species[28] which provided an invaluable starting point for information on LPA production and delivery. In general, it was established that LPA interacts naturally with the active site via its phosphate group with important hydrogen bonds formed between the Thr209 and Asn230 residues. Electrostatic interactions were also noted between the phosphate group and the zinc atoms. The central ester and secondary alcohol functionalities straddle the boundary between the active site and hydrophobic tunnel with hydrogen bonding present with residues such as Asp171, Asp311 and Tyr306. These interactions anchor the core component and direct the phosphate group into the active site and their independent effects coalesce to increase affinity of the substrate for the catalytic site.

Interactions differed depending on the various lengths of LPA, however this highlighted the particular nature of the hydrophobic pocket and it was reasoned that there was a trend in affinity for LPA species. For example, it was determined that the lipid tail,

demonstrated in the ATX-LPA 14:0 complex (**Figure 10**), formed interactions with residues such as Ile167, Leu216, Ala217, Leu259, Phe273, Trp275 and Met512.[28]

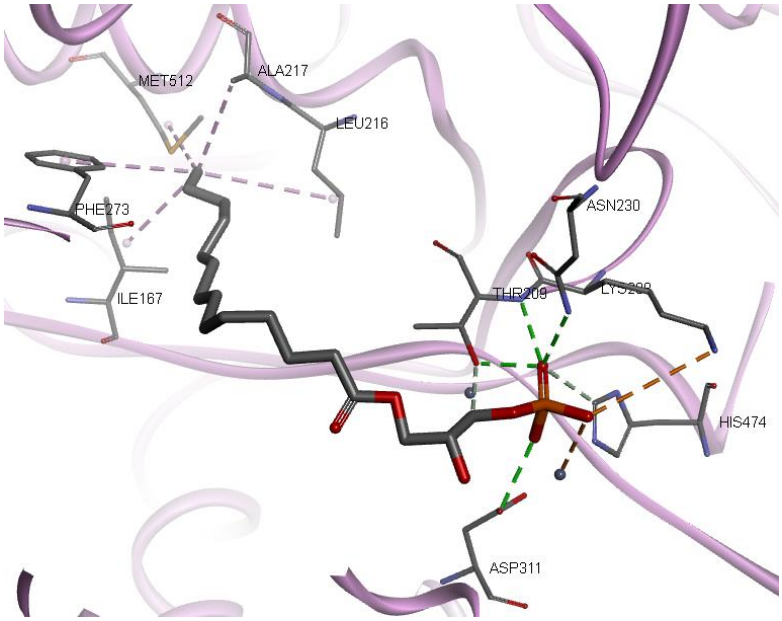


Figure 10 Co-crystal of LPA 14:0 and ATX showing interactions with key residues in the hydrophobic pocket and active site (PDB ID: 3NKN) viewed in DS Visualizer.

Interestingly however, upon considering unsaturation, the central construct of the tail is accommodated in a slightly different way. In the ATX-LPA 18:3 complex, shown in **Figure 11**, interactions were visible with Phe210, Leu213, Tyr214, Trp254 and Trp260 residues which are found on a different surface of the hydrophobic pocket.[28]

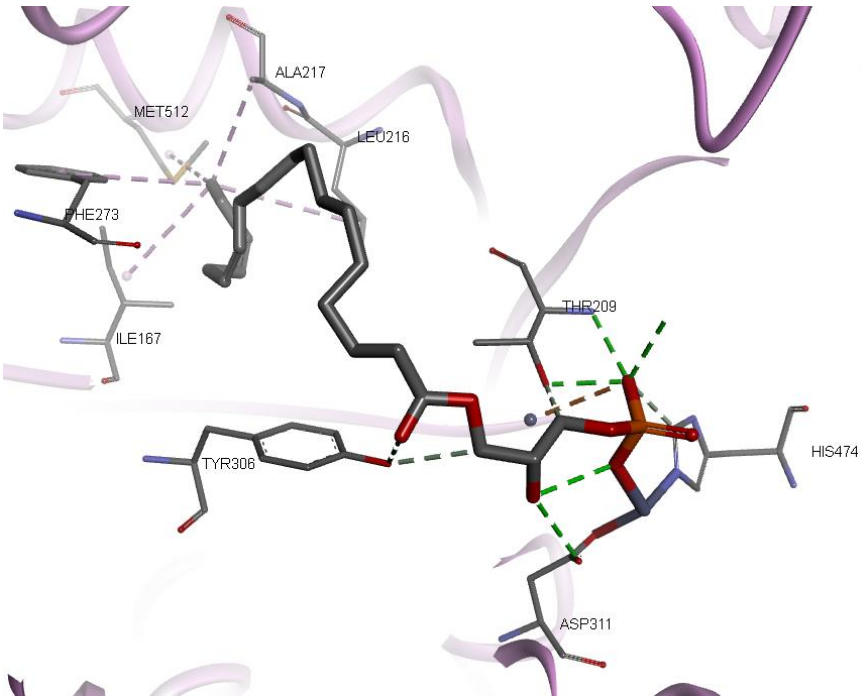


Figure 11 Co-crystal of LPA 18:3 and ATX showing interactions with key residues in the hydrophobic pocket and active site (PDB ID: 3NKQ) viewed in DS Visualizer.

This indicated that, despite not reaching the same depth of the hydrophobic pocket as lipids such as LPA 14:0, longer acyl chains with a higher degree of unsaturation can be accommodated. The trends established can be summarised to state that ATX has a chain length preference of 18:0 < 16:0 < 14:0 and an unsaturated bond preference of 18:0 < 18:1 < 18:3.[28]

It was concluded from this in-depth investigation of ATX that there were two pharmacophoric features present in the substrate that could be correlated to binding affinity. Subsequent structural compartmentalisation of the LPA scaffold revealed a zinc-binding head group and lipophilic tail, fused together by a central core component, as demonstrated in **Figure 12**, which reflect the hydrophobic nature of ATX and allow for its efficient delivery to cognate receptors, $LPAR_{1\text{-}6}$.

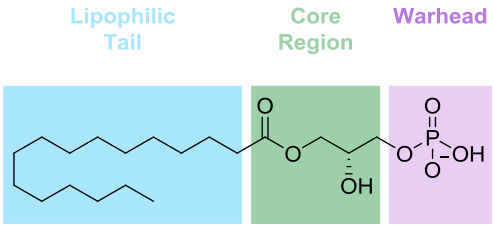
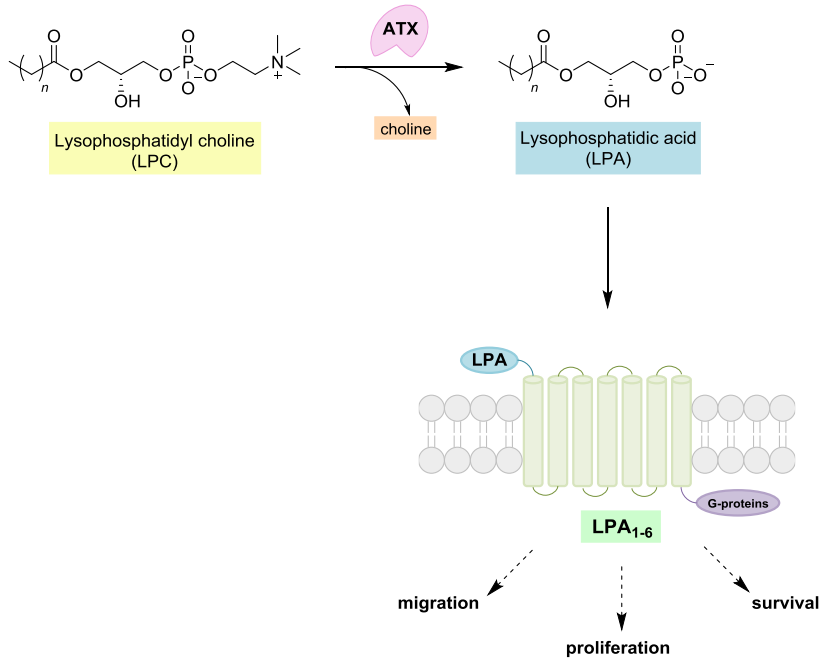


Figure 12 Key structural components of LPA essential for activity.

1.1.4 LysoPLD Activity

Understanding the connection between ATX and phospholipase D could perhaps be classed as one of the milestone markers in the history of ENPP2. Both were established independently with ATX originally isolated as a tumour-motility stimulating protein.[21] Isolation, purification and sequencing of bovine phospholipase D highlighted similarities in sequence identity to ATX and further investigation concluded that the two entities were, in fact, structurally identical.[22,23] This was apparent from studies displaying the ability of ATX to stimulate cell motility through LPA production[30] and the capacity of lysoPLD to facilitate the formation of LPA.[22]

Cell motility or migration is one of many downstream signalling processes mediated by the production of LPA. Others include, cell proliferation, cell apoptosis and cell invasion. The signalling pathway itself is initiated by cleavage of LPC by ATX to form LPA and its biological function is believed to be mediated through specific G protein-coupled receptors (GPCR), known as $LPAR_{1\text{-}6}$. (**Scheme 2**).[31]



Scheme 2 Simplified schematic of lysoPLD activity in ATX-LPA signalling pathway.

As discussed earlier, a plethora of LPA species exist in circulation which is a contributing factor to the complexity of ATX and are synthesised via different biological mechanisms. Structurally, the differences between LPA species can be defined by two attributes: hydrocarbon number and the degree of unsaturation present in the acyl chain, an example of which is shown in **Figure 13**. The common notation for differentiation between LPA forms is a ratio which represents the number of hydrocarbons to the number of unsaturated bonds in the acyl alkyl chain. Those most frequently referred to in the literature are LPA 14:0, 16:0, 18:0, 18:1, 18:3 and 22:6 and these are accommodated differently in terms of their binding to the enzyme which in turn directly correlates to their rate of hydrolysis.

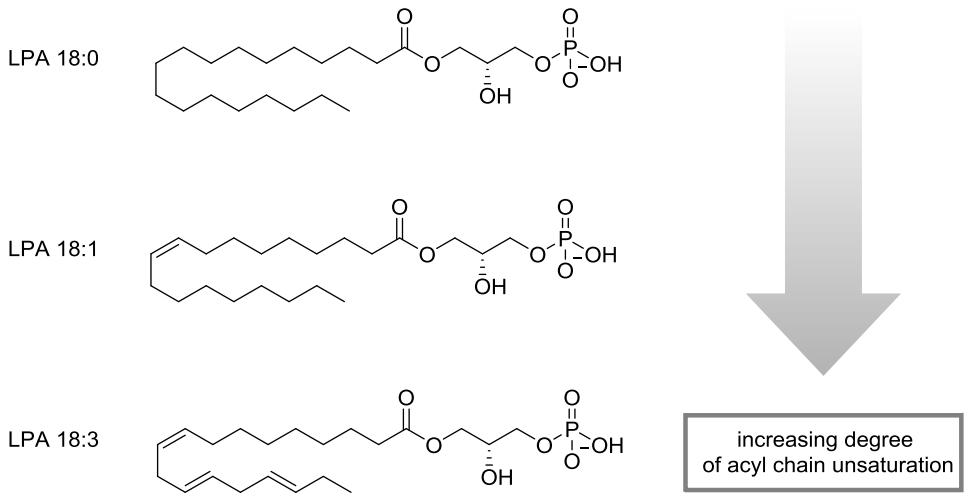


Figure 13 Structural examples from LPA 18:X family depicting increasing degree of unsaturation of one chain length.

In terms of binding, Nishimasu *et al.* determined the preference of ATX for chain lengths and degrees of unsaturation,[28] however supplementary kinetic studies on LPC hydrolysis were undertaken by a group based at Novartis, and the findings suggested that unsaturated LPC species have a higher turnover rate, e.g. 18:1 0.8 $min^{-1}\mu M^{-1}$ and 20:4 0.7 $min^{-1}\mu M^{-1}$, than their saturated counterparts e.g. 14:0 0.71 $min^{-1}\mu M^{-1}$ and 20:0 0.27 $min^{-1}\mu M^{-1}$ when analysed using the native substrate assay. The correlation between LPC species and their downstream effect on LPA_{1-6} remains unclear which simply highlights the complexity associated with this pathway and highlights the difficulty in investigating its potential as a therapeutic target.

With regard to means of synthesis, LPA is predominantly generated as a direct result of LPC hydrolysis by ATX or lysoPLD activity, however other pathways also exist albeit contributing minorly to bulk production (**Figure 14**).[32] In the first pathway, phospholipids are converted directly into their respective lysophospholipids either in platelets or in plasma, where ATX then converts LPC to LPA. In the second pathway, phosphatidic acid (PA) is produced from lysophospholipids or diacylglycerol and is then converted directly into LPA.

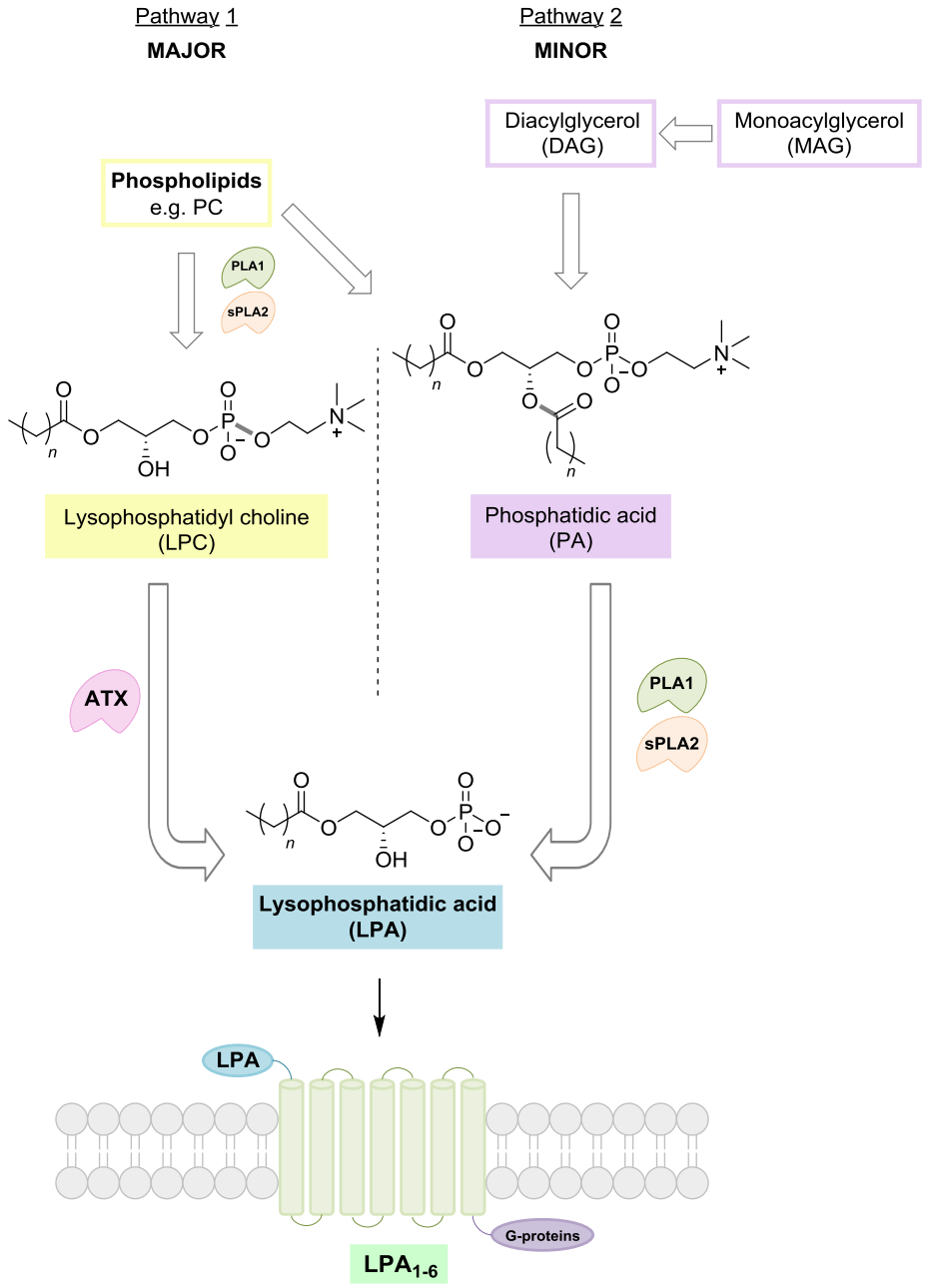


Figure 14 Schematic depicting two main pathways of LPA formation.

As discussed previously, the mechanism of lysoPLD activity has been studied extensively by Hausmann[26] through the use of various probe molecules and metal complexes which have elucidated the ATX-substrate complexes generated through the concerted S_N2 process that is believed to take place. Delivery of LPA to its receptors is understood to be facilitated through the tunnel region of ATX which is located at the interface between the phosphodiesterase domain and SMB domains. This movement of

LPA is further assisted by recognition of ATX by integrins present on the exterior wall of cells[25] and this can be attributed to the RGD binding sequences present in the SMB domains of ATX which are known to function in integrin recognition. Through binding of ATX to the integrin, a conformational change takes place in the tunnel and this is understood to place LPA in an optimal position for delivery to its receptors.[26,33]

At present, there are six established and characterised LPA receptors,[34] with LPA_6 having only been discovered and structurally validated in recent years. Historically, it was believed that LPA exerted its role through three subtypes of the endothelial differentiation gene (EDG) family however these were later identified as being synonymous with LPA_{1-3}.[35] On identification of a fourth receptor LPA_4,[36] it was clear that it belonged to a different subtype which was not structurally homologous with $LPAR_{1-3}$. Subsequent discovery of LPA_5[37] and LPA_6,[38] which also appeared to follow this same subtype structure, prompted designation of the "non-EDG family" terminology for LPA_{4-6}.

LPA is delivered to its cognate receptors $LPAR_{1-6}$, wherein each individual receptor couples to at least two of the four heterotrimeric Gα subunits ($G_{\alpha 12/13}$, $G_{\alpha q/11}$, $G_{\alpha i/o}$ and $G_{\alpha s}$).[39] The biological role of LPA is predominantly driven by the various extracellular responses elicited on activation of the various G-proteins which are illustrated in **Figure 15**. The breadth of signalling pathways that LPA can affect are very diverse, for example, activation of $LPAR_1$ facilitates signal transduction through the RHO, PLC, RAS and PI3K pathways, and these confer activity in cellular mechanisms, such as proliferation and migration. Dysregulation of this signalling process leads to the formation of fibrotic tissue in proliferative diseases and is also understood to contribute towards tumour progression in certain cancers,[31] however it is possible to attenuate this through screening of novel inhibitors.

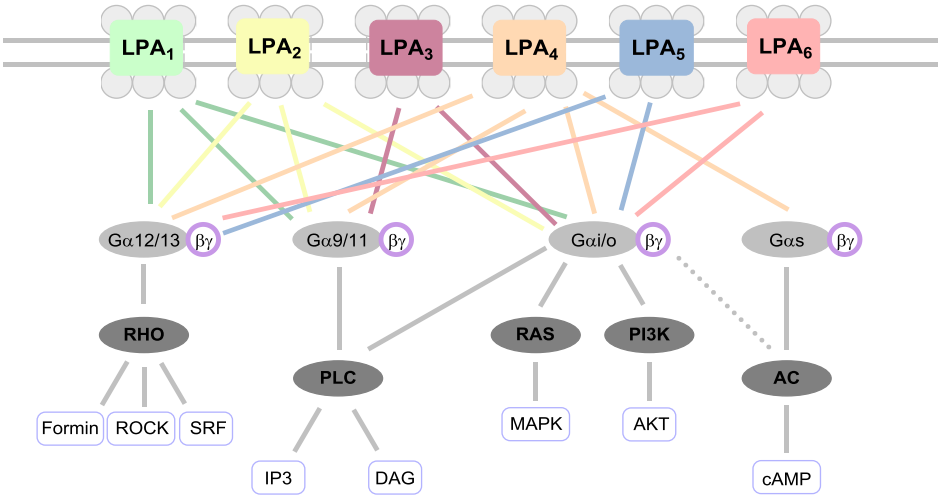


Figure 15 LPAR signalling cascade.

The other LPA receptors associate with different combinations of G-protein subtypes and this varied expression pattern results in both physiological and pathophysiological effects in almost all organs and cell environments in the body[39] as shown in **Table 2**.

$LPAR_1$ and $LPAR_2$ have been linked to both immune responses and also cellular responses such as migration, survival and proliferation, which are associated with autoimmune disorders[40] and neuropathic pain[41] in addition to cancer invasion[42] and hydrocephalus.[43] $LPAR_4$ has been connected to vascular, lymphatic and bone development which is believed to be a result of irregularity in cell aggregation, adhesion, survival and proliferation.[35] Interestingly, $LPAR_6$, the most recently identified member of the LPA receptor family, and therefore relatively poorly researched in comparison to the others, is believed to affect cell morphology and is possibly linked to hair maintenance which links to its expression in follicular cells, e.g. skin, scalp and eyebrow.[44]

Table 2 LPA receptor signalling responses and biological functions[39]

Receptor	Signalling Responses	Biological Function	Expression Pattern
LPA_1 + LPA_2	Migration, survival, proliferation, immune responses	Cancer invasion, autoimmune disorders, hydrocephalus, neuropathic pain	$LPAR_1$: Widely expressed – brain, placenta, small intestine, colon $LPAR_2$: Pancreas, ovary, prostate
LPA_3	Survival, proliferation	Fertility, reproductive regulation	Pancreas, ovary, prostate
LPA_4	Aggregation, adhesion, survival, proliferation	Vascular, lymphatic, bone development	Ovary
LPA_5	Adhesion, immune response	Neuropathic pain	Spleen, heart, platelets, GI lymphocytes, dorsal root ganglia, developing brain
LPA_6	Morphology	Hair maintenance	Leukocytes and follicular cells

Differentiation in signalling responses with respect to each receptor directly correlates to the range of biological functions which result from receptor activation. LPA_1 is expressed most widely throughout the body, with high mRNA levels reported in the brain, placenta, small intestine and colon. Comparatively, $LPAR_2$ and $LPAR_3$ exhibit a more minimal expression pattern which is localised to the pancreas, ovary and prostate. Low levels of $LPAR_4$ have been detected in most mammalian tissues however, it has been indicated that higher than baseline levels are present in ovarian tissue. $LPAR_5$ is expressed in the spleen, heart, platelets, gastrointestinal lymphocytes, dorsal root ganglia and the developing brain which correlates to its postulated involvement in pain signalling. Finally, due to the novelty of $LPAR_6$, work is still underway to examine the

full extent of its expression pattern however, emerging studies suggest that high levels of $LPAR_6$ are expressed in leukocytes and follicular cells.

It is pertinent here to draw attention to the detrimental disruption that abnormal LPA signalling can have on most major organs within the body. The variety in biological functions of LPA highlight the extremely complex and diverse nature of the ATX-LPA signalling pathway. As a result, dysregulation of this pathway has been linked to a range of diseases including cancer,[45] fibrosis,[46] obesity,[47] autoimmune disorders[48] and inflammatory conditions.[49] For example, one of the first biological implications of LPA in cancer was established from the formative discovery that increased LPA levels corresponded to cell motility and invasion. Studies have now linked this increase in LPA levels to ovarian cancer in addition to prostate, head and neck, bowel and thyroid cancers.[50]

Interest in the ATX-LPA signalling cascade has increased dramatically in recent years[51] due to the vast array of diseases that abnormal LPA levels can be responsible for. More importantly however, the heterogeneous nature of proliferative diseases, such as fibrosis and cancer, has made their target a particular challenge.

1.2 Fibroproliferative Diseases

The term fibrosis describes the excessive and uncontrolled build-up of fibrous connective tissue which occurs in response to recurrent tissue inflammation or damage.[52] Fibroproliferative diseases can affect a range of organs and tissues in the body,[53] including but not limited to, the lungs,[54] skin,[55] liver[56] and kidneys[57] (**Figure 16**). The physiological evolution of diseases of this nature remains largely uncertain[58] and as a result, there are limited therapeutic options for their treatment which ultimately cause high morbidity and mortality.[53] Unfortunately, most clinical treatment solutions attenuate disease progression as opposed to offering a panacea and consequently, nearly 45% of all deaths in developed countries can be attributed to some variety of fibroproliferative disease.[59]

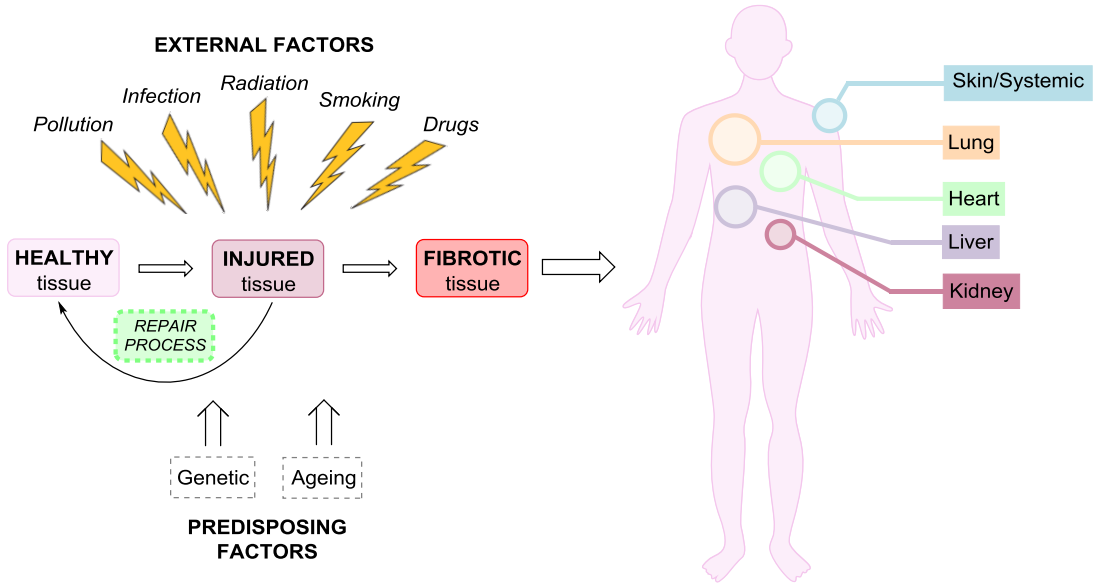


Figure 16 Schematic depicting progression of healthy tissue damage through organ-specific fibroproliferative diseases.

Under normal circumstances, exposure of healthy tissue to external factors such as, infection, smoking, or radiation causes small trauma which, on a cellular level, initiates a repair process and regenerates the site of injury. Generally speaking, fibrosis occurs incrementally as a result of excessive or repetitive damage to injured tissue over a

prolonged period of time, which ultimately, leads to organ failure. Other contributing factors can also be attributed to pre-disposing factors such as genetics[60] and ageing.[61]

All forms of fibrosis are notoriously difficult to diagnose and treat however, two which have been at the forefront of research efforts are Non-Alcoholic Fatty Liver Disease (NAFLD) and Idiopathic Pulmonary Fibrosis (IPF). Both NAFLD and IPF are of a poorly understood aetiology[62,63] meaning that there is an extremely blurred clinical and scientific understanding of the pathology of these diseases, which is a contributing factor to the difficulty in finding successful treatment methods.

1.2.1 Clinical Phenotype of Non-Alcoholic Fatty Liver Disease

Non-Alcoholic Fatty Liver Disease (NAFLD) is a progressive condition of the liver which exists due to the accumulation of fatty tissue through means other than excessive alcohol consumption.[62,64,65] Two main classifications of NAFLD exist: Non-Alcoholic Fatty Liver (NAFL) and Non-Alcoholic Steatohepatitis (NASH), both of which can lead to prolonged inflammation of the liver tissue resulting in fibrosis, which, without appropriate treatment, can progress to liver cirrhosis and end-stage liver failure.[65]

NAFLD occurs in four main stages as exemplified in **Figure 17**, the first of which being relatively benign. Non-alcoholic fatty liver (NAFL), by description, is a relatively harmless build-up of fat around the liver tissue. This initial stage tends to be asymptomatic, and is often diagnosed when the patient is seeking advice for another health condition.[66] NAFL is treatable with lifestyle and diet changes.[67] On progression of NAFLD, Non-Alcoholic Steatohepatitis (NASH) is the second phase of the disease where the patient presents with an inflamed liver and a build-up of fat at a more advanced stage than NAFL. Fibrosis of liver follows wherein prolonged inflammation results in excessive scarring of the liver tissue and surrounding blood vessels. Cirrhosis is the final and most severe stage of NAFLD which occurs after many years of repeated inflammation and liver injury. The liver shrinks substantially from its normal size and the scarring becomes extremely significant. At this point, the damage is permanent and irreversible which invariably leads to liver failure and possibly, liver cancer.

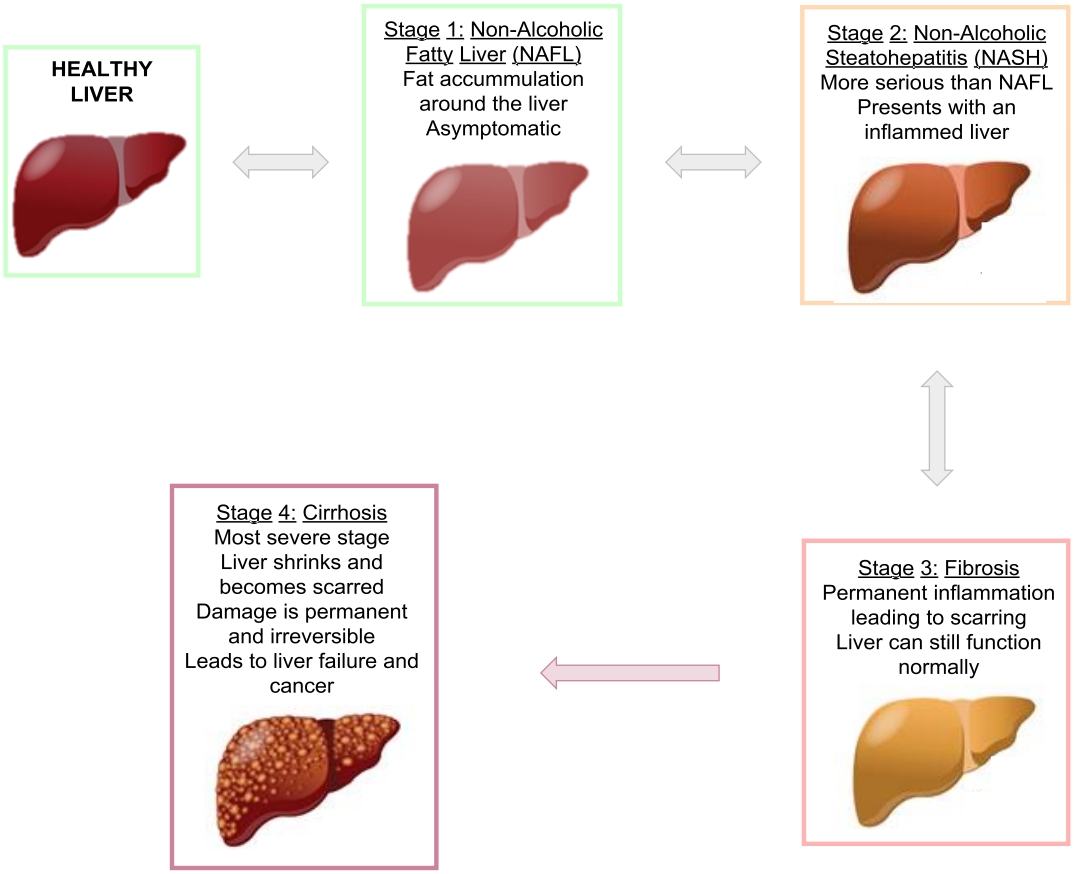


Figure 17 Progression of liver fibrosis and liver failure.

The "at risk" category for NAFLD tends to affect people that fall into a variety of clinical groupings. An individual, who is over the age of fifty, obese or overweight, suffers from Type II diabetes, presents with high blood pressure or high cholesterol, or smokes is most likely to develop symptoms of NAFLD/NASH. These risk factors highlight the reason for diet and lifestyle changes being recommend as the initial treatment phase for those diagnosed.

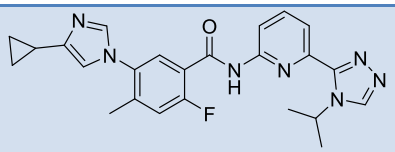
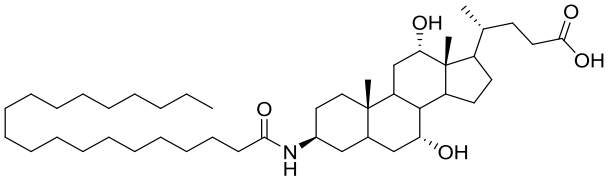
1.2.2 Current Treatment Methods for Non-Alcoholic Fatty Liver Disease

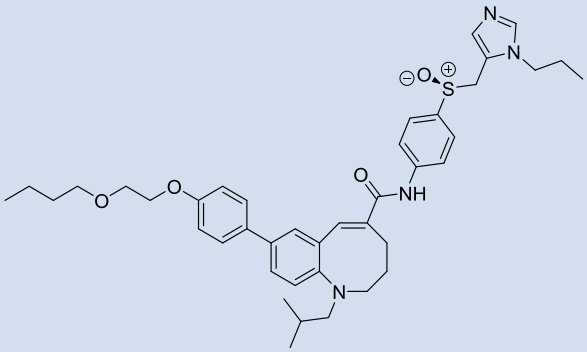
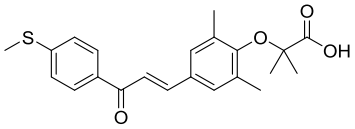
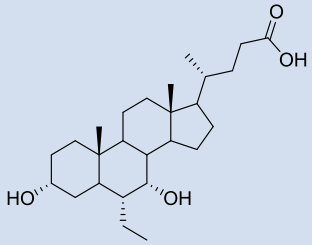
The chronic nature and increasing prevalence of NAFLD presents a large socioeconomic problem for the healthcare system, estimated at 24% globally,[68] which is reflective of the increase in diseases such as obesity and diabetes.[69] In the US alone, NAFLD is now at epidemic levels and tops the list of most common liver disease.[62]

Despite these overwhelming statistics, in terms of medication, there are currently no clinically approved methods to treat the underlying cause of NAFLD.[70] As mentioned above, the medical advice tends to be diet or lifestyle changes[70] as opposed to medication, regardless of the stage, which highlights the pressing demand for treatment options.

To put this into perspective, only recently, two notable compounds undergoing clinical trials have been considered for conditional FDA approval after positive preliminary results and these are detailed in **Table 3**. A recent major clinical candidate for NASH was Selonsertib, an orally-dosed apoptosis signal-regulating kinase 1 (ASK-1) inhibitor pioneered by Gilead, which was reported to prevent pro-inflammatory cytokine production therefore decreasing gene expression and suppressing uncontrolled apoptosis and cellular proliferation.[71] Unfortunately, in February 2019, long awaited results from Gilead's Phase III STELLAR-4 study disappointingly reported that Selonsertib had failed to reach its primary target,[72,73] with patients receiving Selonsertib showing no significant improvement over those who were given the placebo.

Table 3 Phase III clinical candidates for the treatment of NASH

Company	Name	Structure
Gilead	Selonsertib	
Galmed Pharmaceuticals	Aramchol	

Allergan	Cenicriviroc	
Genfit	Elafibranor	
Intercept Pharmaceuticals	Ocaliva	

Conversely, following success in Phase IIb studies, Galmed Pharmaceuticals are currently planning a Phase III/IV clinical trial for their novel fatty acid bile acid conjugate, Aramchol, which acts through stearoyl-CoA desaturase (SCD-1) biosynthesis pathway and is shown to reduce liver fat and fibrosis in patients with NASH.[74,75] A notable observation from this success is the steroidal functionality of the drug – an interesting pursuit for the treatment of any liver disease.

At current, three companies have candidates participating in Phase III clinical trials for NASH, as shown above in **Table 3**, with interim results expected to be released in due course. One of particular note is Ocaliva, or known chemically as obeticholic acid, a farnesoid-X (FXR) agonist[76] which was positioned to be the first FDA-approved drug for NASH.[77] Despite being approved for primary biliary cholangitis, the FDA rejected Ocaliva in July 2020 for the treatment of liver fibrosis resulting from NASH citing an uncertain histopathological endpoint which does not outweigh the potential risks[78] however, Intercept appear to be committed to pursuing the drug.

It is clear that the push to develop clinical candidates for the treatment of NASH is progressing however, notwithstanding the milestone successes; there are still significant

roadblocks in crossing the threshold of FDA approval. Unfortunately, NAFLD is not unique in the struggle to develop treatments. Other fibroproliferative diseases such as IPF have also faced immense challenges in terms of treatment options[59] and the common theme continues to be uncertainty surrounding disease pathophysiology and the heterogeneity of disease progression,[79–81] with these issues sometimes only coming to the forefront in a clinical environment with human subjects.

1.2.3 Clinical Phenotypes for Idiopathic Pulmonary Fibrosis

Idiopathic pulmonary fibrosis falls into the clinical category of an interstitial lung disease (ILD) and those suffering from IPF tend to exhibit chronic and progressive symptoms characteristic of most fibrotic diseases.[82] Due to its unknown aetiology, pulmonary fibrosis is described as an idiopathic disease state[83] however it is postulated that IPF symptoms are caused by repetitive and/or severe injury to cells within the lung cavity. Perpetual scarring of the lung tissue then follows which renders the lungs unable to exchange gas efficiently with the bloodstream.[84]

Physical symptoms characteristic of IPF tend to manifest themselves gradually, most commonly as shortness of breath (dyspnea), particularly during periods of physical exertion, and a non-productive cough persisting for longer than six months.[85] These symptoms can be directly attributed to this congregation of fibrotic scar tissue in the lung cavity preventing the exchange of essential gases, e.g. oxygen, between the alveoli and the bloodstream.

Due to the generality of these symptoms, IPF is often overlooked as ageing or misdiagnosed as another, similar lung disease, for example, asthma or bronchitis.[54] As a direct result of this, IPF sufferers can remain unaware of their condition for a prolonged period of time[86] before a definitive clinical diagnosis is made. In addition to this, the generality of these symptoms also renders an accurate clinical diagnosis extremely difficult. Recent advances in medical imaging, such as high resolution computed tomography (HRCT), are more diagnostically specific,[87] particularly in cases of suspected IPF, than the more invasive surgical lung biopsy meaning a faster and more confident diagnosis.

Other symptoms of IPF which can develop over time as the disease progresses include clubbing, or widening of the fingertips and toes, weight loss, breathing difficulties and

fatigue.[85] Eventually, organ shutdown prevails eventually with the likely outcome being death of the individual.[88]

The prognosis for IPF is generally extremely poor: the average survival rate post-diagnosis has been previously estimated to range between two and five years as recently as 2018,[89] with median survival being only three years[90] and the general five-year survival rate has been reported to be significantly lower than most cancers, excluding cancers of the lung and pancreas (**Figure 18**).[91]

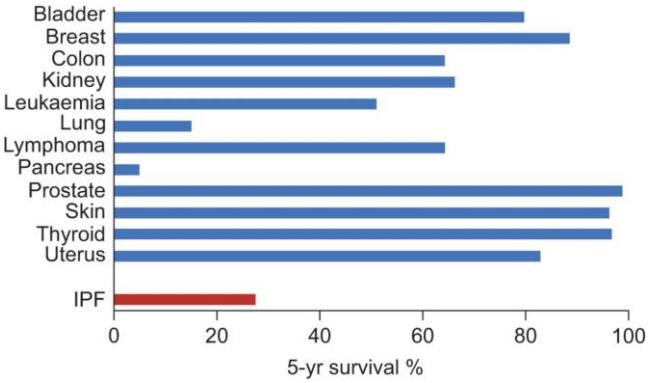


Figure 18 5-year survival rate (%).[91]

As is the case with most fibrotic diseases, the cause of IPF is largely unknown; however it is believed that symptoms of IPF occur as a result of a variety of stimuli (**Figure 19**), most likely, chronic inflammation or environmental contributors, i.e. asbestos or silica, drug toxicity, cigarette smoking or infection.[90,92] External factors tend to be those which are inhaled and cause direct damage to the lung parenchyma. Another potential factor which may influence the progression of IPF, and relates directly to its clinical description, is a genetic predisposition to connective tissue diseases.[93]

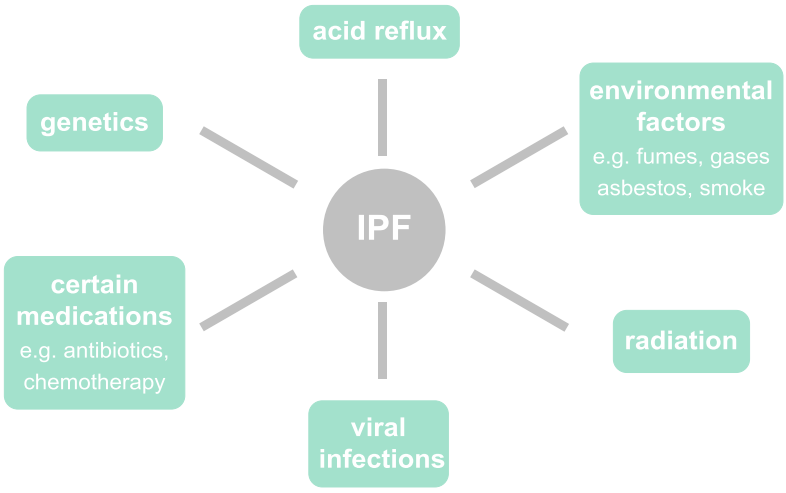


Figure 19 Possible risk factors/causes of IPF.

Due to lack of suitable treatment options and generally poor survival rate, IPF has attracted attention as a therapeutic target over recent years.[59,94,95] Studies have shown that males are slightly more prone to developing the disease than females,[86] IPF is non-specific to ethnic backgrounds[96] and the most common age bracket likely to suffer from IPF are adults between the ages of 60 and 70.[86] The occurrence of IPF is estimated between 14 to 43 people per 100,000[90] and there is evidence to support that the incidence of IPF is increasing worldwide[54] and so, due to the aging population living longer, there is a high chance that IPF diagnoses could become more common. Therefore, it is of pressing clinical importance to investigate the possibility of an effective and efficacious treatment which targets the cause of the disease rather than managing the symptoms.

Understanding the pathogenesis of IPF has been a long-standing effort in research, with full knowledge of the area still incomplete. The normal lung remodelling process in response to injury includes regeneration of cell and tissue functions and the disappearance of myofibroblasts. However, in IPF, and as a result of severe or repetitive injury, the normal tissue repair mechanism can become dysregulated and manifests itself as a progressively irreversible scarring response.[97]

It is believed that most interstitial lung diseases start with a primary epithelial injury in the lung tissue, predominantly due to oxidative or mechanical stresses, which results in either cell death or cell reprogramming.[98]

At this point, pro-inflammatory cytokines are activated, which encourages an immune response, and macrophage cells migrate towards the injury site. Vascular leakage occurs[99] in conjunction with extravascular coagulation, leading to fibroblast recruitment, followed by invasion and proliferation, attenuated by tyrosine kinases such as platelet-derived growth factor (PDGR) and vascular endothelial growth factor (VEGF).[100] These fibroblasts differentiate into myofibroblasts, a process which is assisted by multiple cytokines and growth factors, for example, transforming growth factor-β (TGF- β), $a_v\beta_6$, lysophosphatidic acid (LPA) and autotaxin (ATX), chemokines and enzymes which transform epithelial cells into myofibroblasts.[101]

Myofibroblasts also produce these cytokines and growth factors which attenuates fibrogenesis.[102] Additionally, collagen deposition is a natural and typically reversible process in normal wound healing however collagen build-up in the extracellular matrix (ECM) is a major contributory factor in scarring of the organ tissue due to cross-linking.

Accumulation of these processes, e.g. collagen deposition, fibroblast and myofibroblast formation and extravascular coagulation, results in cell structure collapse and irreversibly forms a new layer of epithelial cells which do not achieve their function, leading to permanent scarring and ultimately organ failure.

1.2.4 Current Treatment Methods for Idiopathic Pulmonary Fibrosis

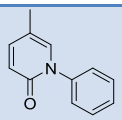
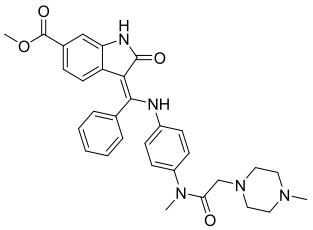
There are no current treatment methods available to treat the underlying cause of IPF, other than lung transplantation or enrolment in clinical trials.[103] Patient suitability for lung transplantation can vary drastically depending on health, age and many other variables and therefore is frequently not a solution to treating IPF in many cases. Accessibility to clinical trials can vary depending on geographical location and patients may not fit the eligibility criteria required to take part in a clinical trial[104] and there is also no guarantee that involvement in the clinical trial would necessarily mean administration of the active compound. Therefore, neither of these methods are a reliable or sustainable means of treating IPF. There are approaches available to manage and alleviate symptoms, for example, oxygen therapy, pulmonary rehabilitation and certain medications which can suppress inflammation, such as treatment with steroids.

In terms of past treatment strategies, it was originally believed that IPF was more inflammatory in nature than fibrotic due to the lack of concrete and in-depth research in

the field.[105] Therefore, the drug discovery community focused on targeting the early stages of the disease by minimising inflammation with the aim of preventing the progression to fibrosis. Drug candidates put forward for clinical trials in IPF have a notoriously high attrition rate and it is widely believed that this is due to a number of contributing factors.[106] Firstly, for example, there is a noticeable lack of suitable pre-clinical models which accurately depict and translate disease progression between animals and human subjects. Pre-clinical studies, for the most part, take the form of bleomycin-induced mice models however these are normally administered to young mice when IPF predominantly affects the elderly. Additionally, there is little consideration for the heterogeneity of fibroproliferative disease progression in the induction of fibrosis in pre-clinical models[97,107] and also discrepancies in clinical trial endpoints.

Despite this, a breakthrough in IPF drug discovery was noted in 2014 when two therapies were approved in quick succession by the FDA **(Table 4**) – pirfenidone (Esbriet®) and nintedanib (Ofev®).[108] Both drugs were deemed "first-in-class"[109] and successful clinical trials provided hope for patients suffering from a disease with no viable cure.

Table 4 Current FDA-approved drugs for the treatment of IPF

Company	Name	Structure
Roche	Pirfenidone (Esbriet®)	
Boehringer-Ingelheim	Nintedanib (Ofev®)	

Pirfenidone (5-methyl-1-phenyl-2-[1*H*]-pyridone) is an orally active, small molecule drug developed by Roche and was the original FDA-approved therapy for IPF. Although

its mechanism of action still remains unclear,[110] it continues to be the therapy of choice for patients due to its ability to significantly decrease the rate of IPF progression.[111] Existing studies suggest that pirfenidone acts by regulating levels of varying cytokines, growth factors and chemokines within the lung parenchyma.[112] Convincing evidence suggests that pirfenidone has the capacity to considerably reduce fibroblast proliferation and collagen production – two highly important cytokine-mediated pathways involved in the prognosis of IPF.[113]

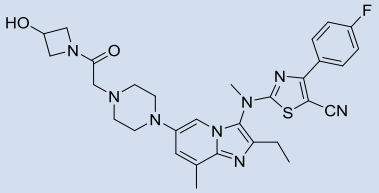
Nintedanib, or BIBF-1120, marketed by Boehringer-Ingleheim, is a small molecule implicated in tyrosine-kinase inhibition. In contrast to pirfenidone, the molecular targets of nintedanib are more widely understood and it is believed that the drug targets a range of receptors which consist of vascular endothelial growth factor receptor (VEGFR), fibroblast growth factor receptor (FGFR) and platelet derived growth factor receptor (PDGFR).[114] All of the aforementioned receptors have been connected to the underlying pathogenesis of IPF, and so it is believed that nintedanib acts by blocking these signalling pathways therefore decelerating the rate of deterioration in lung function.

A major caveat to both of these clinical compounds is their adverse effects. Both present with a variety of unfortunate side-effects, from gastrointestinal complications, e.g. dyspepsia/indigestion, gastroesophageal reflux disease (GERD) and nausea, to skin photosensitivity and hepatic dysfunction.[115] More seriously, pirfenidone is contraindicated in patients suffering from severe hepatic impairment due to its ability to increase enzyme levels in the liver, making this a highly unsuitable and risky option for those patients suffering from both IPF and severe hepatic function.[115]

Despite the adverse side effects associated with these candidates, both exhibit a significant effect in impeding the progression of IPF, an area which was drastically underdeveloped in terms of suitable treatment methods. Unfortunately, they continue to remain as progression-free survival strategies and do not present a disease-modifying approach for the treatment of IPF.[116]

More recently, several emerging compounds have demonstrated positive results in late-stage clinical trials and interestingly, target different modes of action associated with IPF. Some notable examples are shown below in **Table 5**.

Table 5 Recent clinical candidates for the treatment of IPF

Company	Phase	Name	Type	Structure
Galapagos	III	GLPG1690	Small Molecule	
Promedior /Roche	II	PRM-151	Small Molecule	Structure undisclosed
FibroGen	III	Pamrevulumab	Monoclonal Antibody	Structure undisclosed

GLPG1690 (Ziritaxestat), an autotaxin inhibitor developed by Galapagos which will be discussed in more detail in a later section, is currently in Phase III clinical studies and has been granted Orphan Drug Designation by the FDA in IPF. Additionally, Promedior, overseen by Roche, have demonstrated significant long-term safety and efficacy data from Phase II studies of their recombinant human pentraxin-2 candidate, PRM-151[117] with Phase III commencing in due course. Thirdly, Pamrevlumab, pioneered by FibroGen, is a monoclonal antibody which inhibits the activity of connective tissue growth factor (CTGF) in IPF and has demonstrated a promising safety and tolerability profile thus far.[118] It has also received both Orphan Drug and Fast Track Designation from the FDA. Early-stage clinical trials are also underway for a large number of other compounds, including BBT-877,[119] another ATX inhibitor recently disclosed as showing promise with dose-proportional suppression of LPA production in Phase I trials for IPF, however further discussion will not be included here as the structure of this asset has not been disclosed.

It is evident that whilst endeavours are underway for an FDA-approved treatment, there is still an unmet clinical demand for further validation of therapeutic targets that contribute to, or are associated with, the development and progression of fibrotic diseases, such as NASH and IPF.

1.3 Autotaxin as a Therapeutic Target

1.3.1 Validation in Cancer

From the mid 1990's, the interplay between the ATX-LPA signalling pathway and cancer has been apparent and has been consequently well-documented. Based on the effect that dysregulation of this pathway has on fundamental cellular processes, it is no surprise that overexpression of both ATX and LPA has been reported in tumour malignancies,[49] for example, non-small cell lung cancer. LPA receptor activation is directly related to various signalling pathways with roles in cell-cycle progression and tumour cell migration and invasion.[31] It is now believed that the interaction between cell-surface integrins and ATX positions locally-produced LPA close to its receptors. This suggests that ATX may facilitate a tumour microenvironment, therefore indirectly contributing to the augmentation of the metastatic process.[31] A study has also listed ATX in the top 40 genes found to be upregulated in metastatic cancers, including breast, prostate and colon cancer.[120]

LPA has been demonstrated to cover all ten hallmark processes involved in cancer,[121,122] from the stimulation of proliferation to resisting cell death through regulation of cell apoptosis. Over one thousand papers have reported a link between LPA and a multitude of different cancers[50] including, but not limited to, pancreatic, cervical, ovarian, breast, bladder and leukaemia-derived cancers, which has ultimately resulted in the validation of LPA as a therapeutic target for cancer treatment.[45]

Interestingly, conclusions have been drawn which implicate the ATX-LPA pathway in chemotherapy and radiotherapy resistance. This is more than likely a repeating cycle due to the increase in ATX secretion, which results in high concentrations of circulating LPA being aberrantly delivered to its receptors promoting uncontrolled signalling ultimately leading to aggressive tumour proliferation.

Despite the overwhelming evidence, there are currently no ATX or LPA-related clinical candidates for the treatment of cancer which is perhaps an unfortunate combination of the multi-faceted nature of the signalling pathway in addition to complexity of cancer itself. Comparatively more success has been established in fibroproliferative diseases such as IPF and NASH, at least in pre-clinical models.

1.3.2 Validation in Fibrosis

Validation of ATX as a therapeutic target for IPF was first established by Oikonomou *et al.* wherein it was reported that genetic knockout of ATX in mice within bronchial epithelial cells and macrophages/neutrophils resulted in a marked decrease of bleomycin induced lung fibrosis.[123] It could be deduced from this data that circulating ATX, present in bronchial epithelial cells, has the capacity to contribute towards IPF pathogenesis. An inhibitor of ATX was then tested which showed a decline in disease development, suggesting that ATX is a viable target for fibroproliferative diseases, specifically IPF. Additionally, increased ATX levels were also identified in analysed bronchoalveolar lavage fluid (BALF).

Since then, high levels of both LPA and ATX have been reported in both the BALF and breath condensate of patients presenting with IPF which categorically demonstrates the link between ATX and IPF. Numerous reviews and publications have described the challenges in treating IPF[95,103,124] and the efforts made in unravelling the complex pathology of the disease itself.

NAFLD is somewhat less well-progressed than IPF in terms of its validation of ATX however in recent years, an increasing number of connections have been made between the ATX-LPA signalling pathway and liver fibrosis.[125] A collaborative effort between PharmAkea and Cayman Chemical[126] has confirmed the anti-fibrotic capacity of ATX in a high-fat diet mouse model of NASH and has shown that inhibition of ATX may provide a route towards treating liver fibroses, including NASH.

As will be discussed later, the most relevant validation for ATX in the treatment of pulmonary fibrosis has been in the recent clinical development of small molecule ATX inhibitor, GLPG1690, which showed efficacy in diminishing bleomycin-induced pulmonary fibrosis in pre-clinical studies.[127] Since then, it has demonstrated positive results in Phase I and II clinical trials for IPF, and it has recently been disclosed that it is also in clinical development for the treatment of systemic scleroderma, otherwise known as skin fibrosis.[51] This demonstrates human proof-of-concept for ATX inhibition and represents a milestone in the evolution of ATX inhibitors for the treatment of fibrosis.

1.4 Development and Evolution of Autotaxin Inhibitors

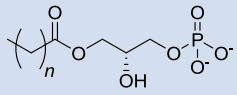
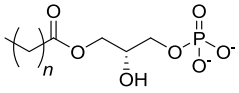
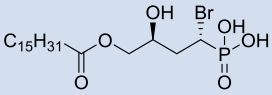
A plethora of ATX inhibitors have been reported in the literature setting targeting different areas of the catalytic domain,[128–131] however currently, there have been no therapeutic candidates which have successfully completed clinical trials and thus, no FDA-approvals for reversing IPF or NASH, however it is possible that this might change in the very near future with the apparent success of GLPG1690.

1.4.1 Substrate- and Lipid-based Inhibitors

Initially, the evolution of ATX inhibitors was closely aligned with the design of LPA antagonists[132] and accordingly, lipid-like molecules were designed that were structurally similar to the endogenous product itself, LPA. Both LPA (**Table 6**), and a related phospholipid, sphingosine-1-phopshate (S1P, **Table 6**), have been reported in the literature as "mixed" inhibitors of ATX showing significant activity.[133] Utilising an endogenous ligand as an inhibitor is described as feedback inhibition. Although LPA and S1P were efficient at minimising local LPA production, their ability to decrease LPA production *in vivo* was poor. Despite this, several attempts have been made at synthesising lipid-like analogues in an effort to decrease downstream signalling effects through ATX inhibition. Multiple examples exist in the literature with significant results, with one of the most notable being BrP-LPA (**Table 6**).

BrP-LPA is a potent inhibitor of ATX with chiral character and shows an ability to significantly reduce tumour volume in mouse *in vivo*. Despite this promising data, BrP-LPA has extremely poor physicochemical properties in addition to possessing a known toxicophoric functionality,[134] an alkyl bromide. Two common themes with most lipid-like candidates are their poor selectivity and notoriously unsatisfactory physicochemical properties. The poorly soluble nature of the phosphonate head group renders them unattractive for further therapeutic development.

Table 6 Lipid-like inhibitors of ATX.

Name	Structure	Activity
S1P		K_i = 50 nM
LPA		K_i = 110 nM
BrP-LPA		IC_{50} = 22 nM

1.4.2 Binding Modes of Autotaxin

Over the last ten years, with the assistance of co-crystal structures, research has moved away from designing substrate-based or lipid-like molecules and into the development of small molecule inhibitors of ATX. A significant body of small molecules have been reported to inhibit autotaxin, with modulation derived through interaction with different parts of the catalytic domain and have been neatly reviewed periodically by Castagna,[132] Nikolau[135] and Aidinis.[136]

Advancement in understanding the true potential of autotaxin has acted as springboard for chasing inhibitor diversification. The tripartite binding site of ATX has allowed for various avenues to be explored leading to the selective design of many different "types" of compounds. These have recently been eloquently categorised by Salgado-Polo based on their known crystal structures (**Figure 20**).[137]

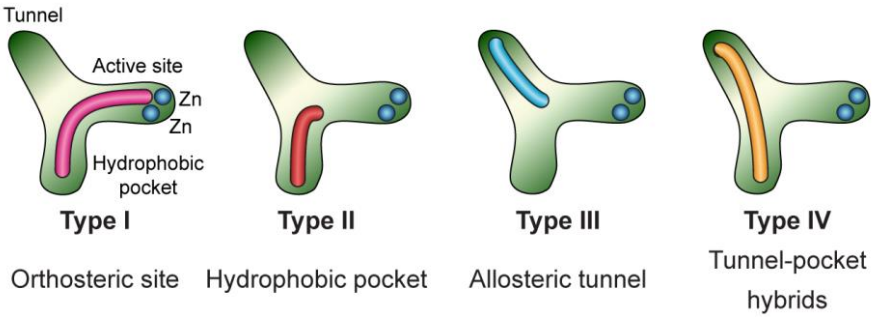
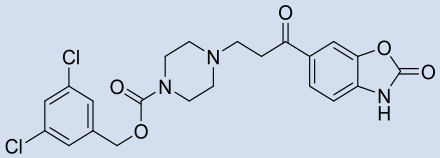
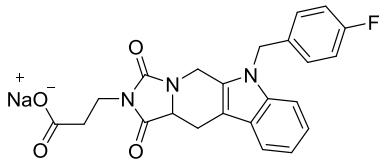
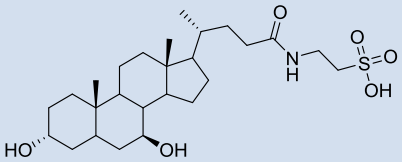
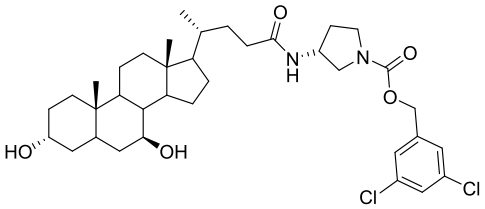


Figure 20 Classification of ATX inhibitor types based on their respective crystal structures.

Representative examples of these distinct binding modes are summarised below (**Table 7**). A significant fraction of inhibitors fall into the category of orthosteric site modulators (Type I), as they function to competitively block substrate binding in the active site and hydrophobic pocket as demonstrated with PF-8380 (**Table 7, Entry 1**).[129] In addition, hydrophobic binders (Type II) compete with substrate binding without possessing a warhead component which is targeted for the active site as Pharmakea exemplified with their analogue PAT-352 (**Table 7, Entry 2**).[138] In contrast, allosteric tunnel binders (Type III) owe their effect to the non-competitive nature of their binding and are often described as weak endogenous modulators, for example, tauroursodeoxycholic acid (TUDCA) (**Table 7, Entry 3**).[139] Finally, and most recently, hybrid inhibitors have been identified as tunnel-hydrophobic pocket binders (Type IV) acting allosterically but switching to competitive inhibition of ATX demonstrated by steroid hybrid **4** (**Table 7, Entry 4**).[140]

Table 7 Representative structures of four classes of ATX inhibitors

Entry	Name	Type	Binding Mode	Representative Examples
1	PF-8380	Type I	Orthosteric site	

2	PAT-352	Type II	Hydrophobic pocket	
3	TUDCA	Type III	Allosteric tunnel	
4	FP-Cpd 17	Type IV	Tunnel-pocket hybrids	

For the purpose of this thesis, only Types I, III and IV will be discussed in detail in the following sections.

1.4.3 Type I Inhibitors – Orthosteric Site

The insight gained from the lipid-like libraries provided instrumental information on key interactions and structural features necessary for recognition and binding. Combining the pursuit of more drug-like structures with the knowledge gained from these libraries, small molecule inhibitors were designed to mimic the natural product, LPA and consequently, are classed as orthosteric inhibitors. Most of these inhibitors exhibit a classic general structure with three distinct pharmacophoric elements that can be loosely related to LPA: a lipophilic tail component, core and linker region and zinc binder (**Figure 21**).

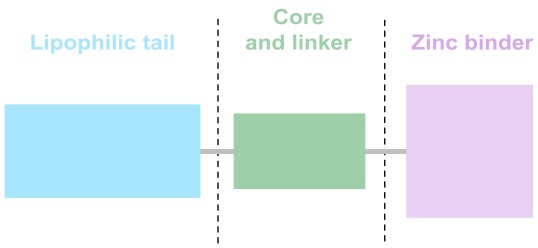


Figure 21 Building block construction approach of orthosteric ATX inhibitors.

The first small molecule Type I inhibitor was from the thiazolidinedione family. A classic SAR approach led to the discovery of HA155 which has been reported as an extremely active ATX inhibitor (**5**, IC_{50} = 5.7 nM, LPC) with a particularly noteworthy boronic acid functionality which forms a reversible covalent bond with the hydroxyl group on the Thr209 residue of ATX (**Figure 22**). The original hit compound from the screening process contained a carboxylic acid functionality which was considered to be acting as a phosphate group mimic, interacting with the Thr209 oxygen in the active site. Therefore, the next logical step was to consider bioisosteres of carboxylic acids, one of which being a boronic acid. This proved to be successful wherein replacement of the carboxylic acid (IC_{50} = 2.5 µM, LPC) with a boronic acid functionality provided a 400-fold increase in potency.

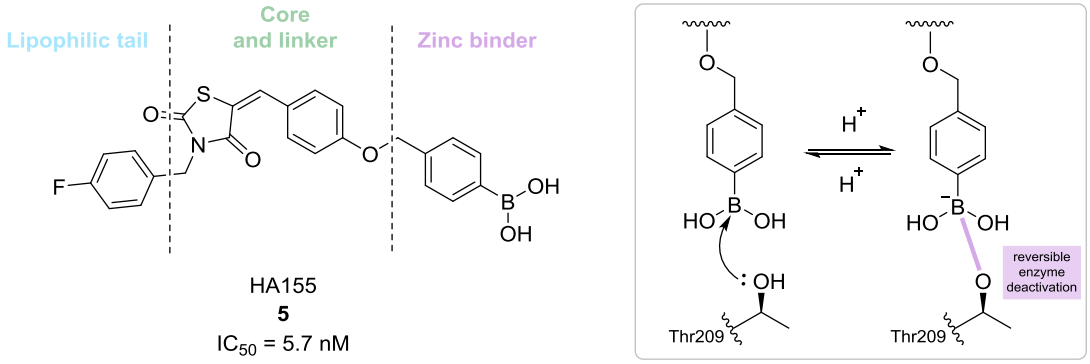


Figure 22 Structure of literature compound HA155, **5**.

Co-crystallisation of HA155 with ATX revealed a plethora of interactions with Thr209, Asp311, His359 and His474 at the active site in addition to interactions with Leu213 and Phe274 in the hydrophobic pocket. The covalent interaction between the boronic acid and Thr209 was of interest and consequently, mode of inhibition studies confirmed that HA155 and related active analogues furnished with a similar aryl boronic acid motif were indeed reversible inhibitors of ATX.

A second noteworthy orthosteric inhibitor is PF-8380 which continues to predominate the literature as one of the most potent inhibitors of ATX *in vitro* (**1**, IC_{50} = 1.7 nM, LPC). First reported by Pfizer in 2010 (**Figure 23**), design of the compound library was influenced by an earlier patent published by Merck KGaA.[129] Unfortunately, despite

their extremely desirable potency, PF-8380 and related analogues have never progressed in a therapeutic setting due to their sub-optimal physicochemical properties. Regardless, PF-8380 has been an extremely useful tool compound for further structural and biochemical elucidation.

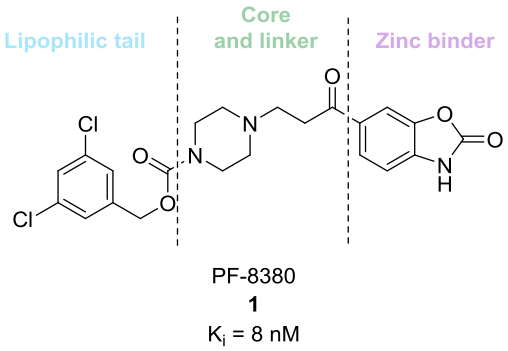


Figure 23 Structure of literature compound PF-8380, **1**.

Despite recent advancements in IPF research with regards to ATX, our understanding of the pathophysiological role of ATX needs to be improved. It has taken many years to reach this stage yet there still remain many unanswered questions.

With the knowledge that PF-8380 was a potent ATX inhibitor with known activity *in vivo*, attempts have been made to improve other physicochemical properties such as solubility, ligand efficiency and cLogD. Co-crystallisation studies of PF-8380 and ATX by Jones *et al.* allowed for a better understanding of key interactions that were required to be maintained whilst probing alterations in other parts of the framework to gain more favourable physicochemical parameters.

Extensive work by the Jamieson group[141,142] took each of three pharmacophoric elements of the scaffold in turn and worked on building a structure-activity relationship (SAR) profile in order to synthesise a library of compounds with more desirable physicochemical properties (**Figure 24**).

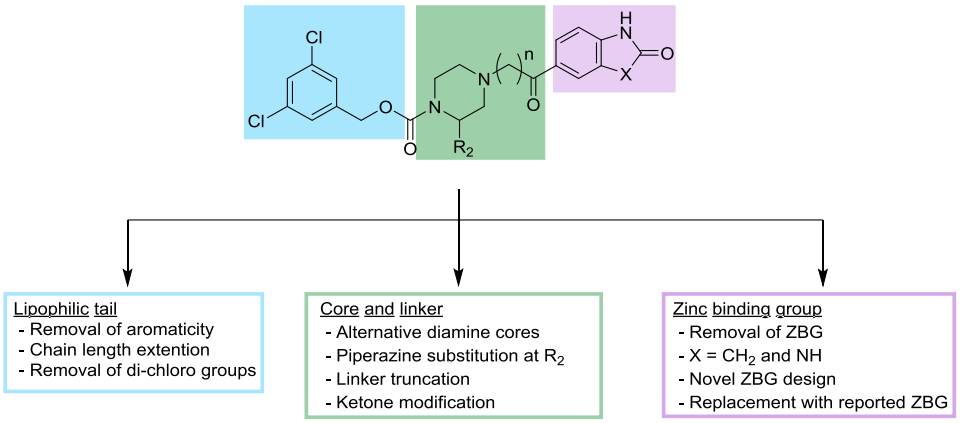
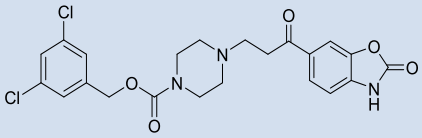
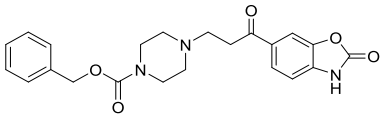
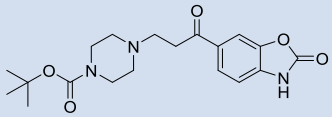
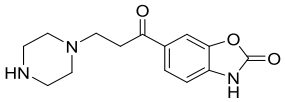
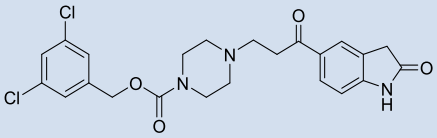
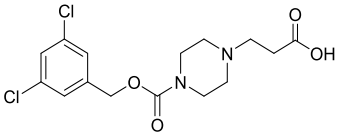
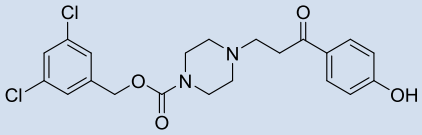
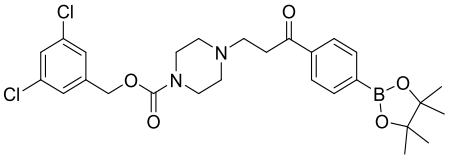


Figure 24 PF-8380 SAR exploration carried out within our laboratory[141,142]

The SAR exploration primarily focussed on determining the changes that are tolerated in the lipophilic tail region (**Table 8, Entries 6-8**). It had already been identified from the crystal structure that the 3,5-dichlorobenzylcarbamate moiety fills the pocket however, it was of interest to probe the extent of variation possible. Multiple variables were investigated, for example, replacement of the aryl component with an aliphatic moiety **7** and removal of the lipophilic tail **8** entirely. However it was the 3,5-dichlorobenzyl carbamate tail, as seen on PF-8380 (**1**), which was indeed the most favourable in terms of potency although, the unsubstituted benzyl carbamate **6** exhibited more desirable physicochemical properties.

The subsequent phase of the study investigated the effect of varying the benzo[*d*]oxazol-2(3*H*)-one zinc binding group (**Table 8, Entries 9-12**). It was established that removal of the ZBG, for example, **10**, was detrimental to potency and this was attributed to the fact that there were no hydrogen bond donors or acceptors present in the vicinity of the active site to provide interaction with the zinc atoms. Various modifications were made to the hydrogen bond donors and acceptors, such as, replacing the benzo[*d*]oxazol-2(3*H*)-one with a phenol functionality **11** which was detrimental to potency however, a notable result from this sub-series was that of the pinacol boronate ester **12** (K_i = 27 nM, bis-*p*NPP).

Table 8 Overview of key observations from SAR investigation around PF-8380.

	Compound	Structure	K_i (μM)	cLogD
Lipophilic Tail Diversification	**1**		0.008	3.62
	6		0.053	2.42
	7		> 30	1.75
	8		> 30	-1.23
Zinc Binding Group Diversification	**9**		0.022	3.36
	10		> 30	-0.40
	11		> 30	5.60
	12		0.027	5.74

PF-8380 and HA155 are perhaps two of the most pertinent tool compounds in the literature to date and have continued to serve as robust starting points for novel design strategies towards new inhibitor libraries.

Other orthosteric ATX inhibitors have been reported over the years, however it was apparent that orthosteric inhibition has not had much success in translation from pre-clinical models through the clinic. However, this hypothesis may well be disproved on

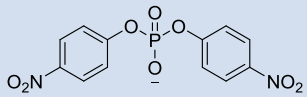
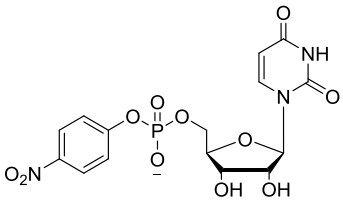
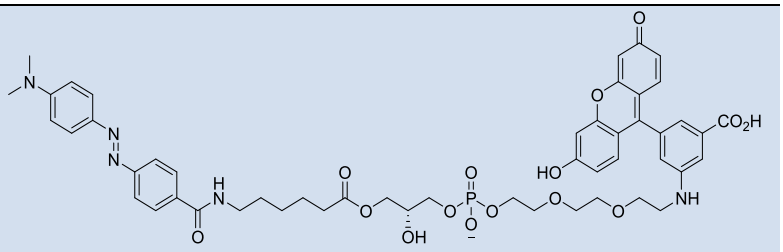
structural disclosure of BBT-877[119] pending announcement of results from the Phase II trial.

It is possible that, in the past, this has been due to the lack of robust preclinical models for fibrosis in general, or also the fact that there is more to understand about the role that autotaxin and LPA play in the pathogenesis of these conditions.

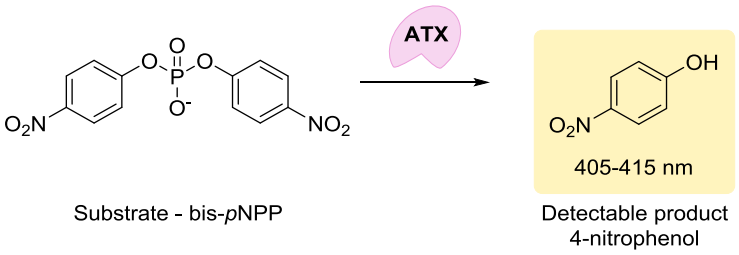
1.4.4 In vitro Biological Evaluation of Autotaxin Inhibitors

In the early stages of orthosteric inhibitor design, assays were developed to test activity against ATX in order to compliment SAR progression. To this day, there continues to be two well-defined *in vitro* biochemical assays used to determine ATX activity and these either engage an unnatural substrate, such as, bis-(*p*-nitrophenyl) phosphate (bis-*p*NPP), *p*-nitrophenyl-thymidine monophosphate (*p*NP-TMP), FS-3, or the natural substrate, LPC (**Table 9**).[143]

Table 9 Unnatural substrates for ATX biochemical assays.

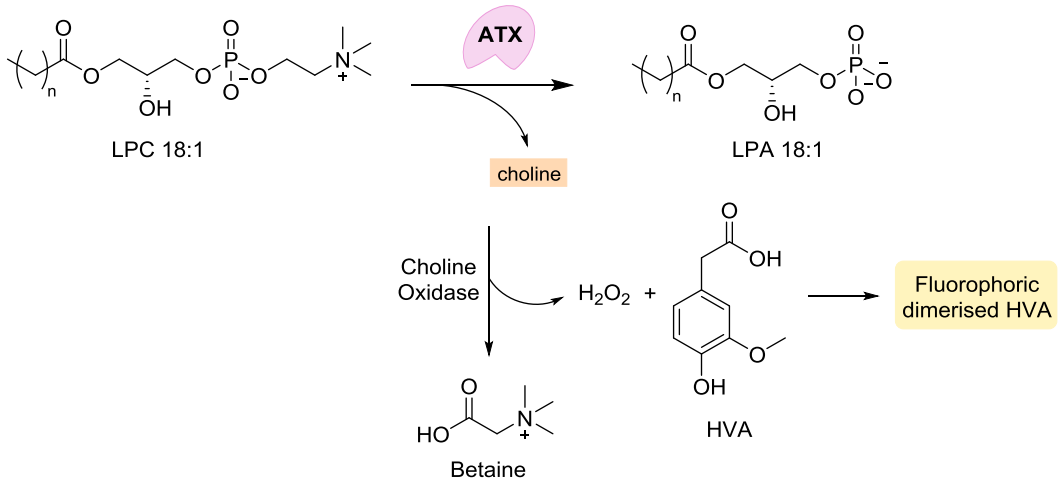
Name	Structure
bis-*p*NPP	
*p*NP-TMP	
FS-3	

As an example, the bis-*p*NPP assay measures absorbance facilitated by cleavage of the unnatural substrate into a fluorescent by-product, 4-nitrophenol (**Scheme 3**). As a cost-effective, preliminary measure, biochemical assays utilising an unnatural substrate are implemented as an initial screen in order to triage synthesised compounds from SAR libraries and is a technique implemented in our own laboratories. The protocol is straightforward however it is most effective for compounds targeting the active site.



Scheme 3 Cleavage of bis-*p*NPP by ATX forming fluorescent by-product, 4-nitrophenol.

If a compound is deemed active from the initial bis-*p*NPP screen, it may then be tested in the LPC assay which provides a more accurate measure of potency. The natural substrate assay measures choline release through a cascade reaction which eventually concludes with the formation of a fluorescent dimer (**Scheme 4**).



Scheme 4 Choline release assay mechanism.

The hydrolysis of LPC to LPA is catalysed by ATX which generates choline as a by-product. This release of choline can be monitored in a two-step enzymatic colouring reaction. The choline by-product is oxidised to betaine in the presence of choline oxidase (CO) generating hydrogen peroxide. Horseradish peroxidase (HRP) then facilitates the reaction between homovanillic acid (HVA) and hydrogen peroxide resulting in the formation of dimerised HVA which is fluorophoric in nature and therefore, can be used to monitor assay progression.

A major caveat associated with the choline release assay is that false positives are possible due to unwanted reactions between the inhibitor and the colouring reagents.

1.4.5 Type III Inhibitors – Allosteric Tunnel

From the outset, a large proportion of compounds were designed to interact with the hydrophobic pocket and the active site in the same way as the endogenous ligand, therefore acting as competitive inhibitors. This was, in part, due to the limited knowledge of the function of the tunnel region residing in the catalytic domain. Early structural work into autotaxin,[29] including crystal structure 2XR9, noted residual electron density in the tunnel region however this was not structurally identified. It was generally believed to be lipid-related,[144] possibly acquired from growth medium. Re-examination and refinement of the electron maps in subsequent years identified a quad-ring system with two axial methyl groups – characteristic of steroidal architecture. Further optimisation determined the structure to be 7-α-hydroxycholesterol (7HCS), which has been reported to be present in purified mammalian ATX structures but on analysis, it does not register any measurable activity.[139]

This crucial finding led to an extensive investigation into the tunnel region by our collaborators at the Netherlands Cancer Institute (NKI) which provided invaluable data exemplifying the previously unreported interplay between bile salts and LPC production.[139]

It was established that 7-HCS and other known steroid moieties such as testosterone, dexamethasone and prednisolone were completely ineffective in mediating LPC hydrolysis. However, a SAR screen focused on tail conjugates and hydroxyl substitution pattern on the ring construct determined that conjugated bile salts such as TUDCA and

UDCA bind ATX allosterically and have the ability to inhibit LPA production (**Figure 25**).

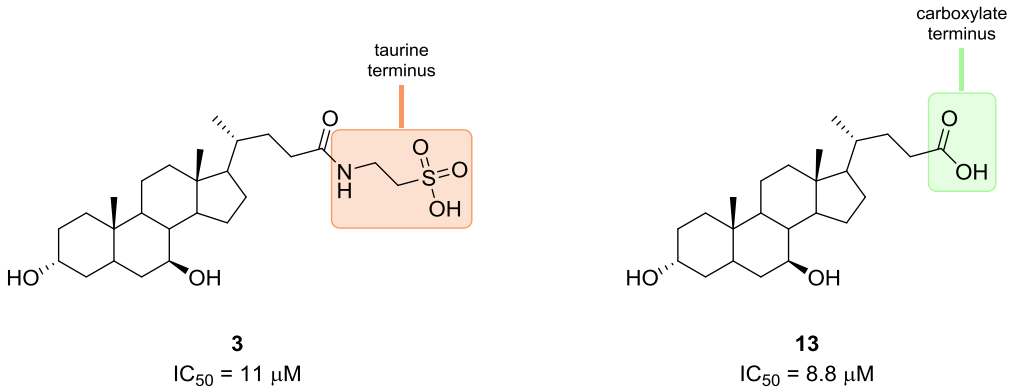


Figure 25 Structures of tauroursodeoxycholic acid (TUDCA) and ursodeoxycholic acid (UDCA) with their respective termini.

It was believed that co-crystallisation studies involving TUDCA and ATX (**Figure 26**) provided substantiation to the hypothesis that the tunnel acts as an LPA exit route, most likely by decreasing the rate of product release, however this has since been revisited by further studies undertaken by the Perrakis group.[145] Despite this, bile salts/allosteric modulators are representative examples of Type III inhibition.[137]

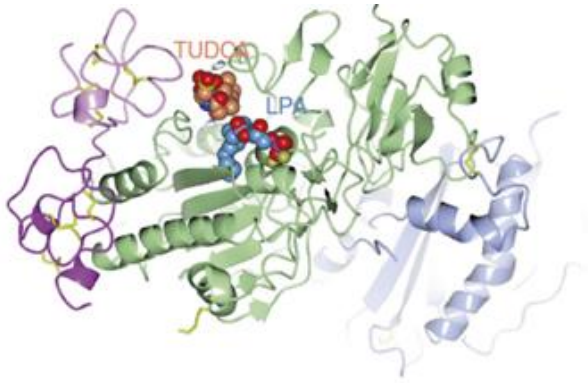


Figure 26 ATX structure bound with TUDCA **3** (orange carbons) in the tunnel and LPA (blue carbons) in the pocket, both shown as space filling models.[146]

Following on from this, it was proposed that these steroidal motifs could represent a functional handle to design compounds which would act as partial inhibitors. Synthetic manipulation of TUDCA would be limited due to its associated taurine terminus however, the synthetic tractability of UDCA with its carboxylate terminus was considered more promising, providing potential compounds which could be more closely aligned with the intrinsic functionalities of PF-8380 analogues.

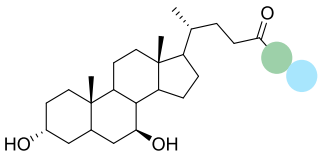
1.4.6 Type IV Inhibitors – Tunnel-Pocket Hybrids

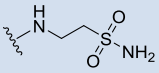
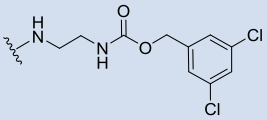
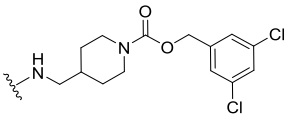
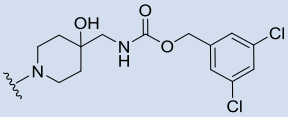
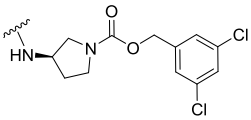
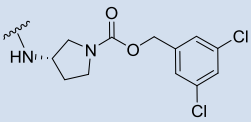
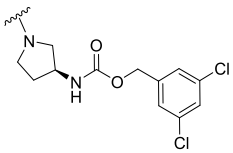
Type IV inhibitors have only recently emerged in the literature however they represent an extremely important direction for validity of autotaxin as a therapeutic target.

A library of allosteric modulators was designed by the Jamieson group combining structural knowledge of the binding mode of bile salts and previously reported competitive inhibitors of ATX.[140] The design hypothesis was built on the idea that it was possible to merge the steroidal framework with the lipophilic "tail" of a tool compound from within the PF-8380 series in order to generate previously unreported allosteric "Type IV" modulators of ATX.

With the knowledge that the 3,5-dichlorobenzyl carbamate tail was an optimal anchor in the hydrophobic pocket, the area for optimisation remained in the core region as it was necessary that the compounds vector in the correct way to accommodate both the key steroidal and lipophilic tail interactions. Based on co-crystallisation studies from the PF-8380 series, it was known that the piperazine core functionality had a degree of flexibility due to its ability to adopt different conformations. A formal evaluation of various core scaffolds (**Table 10**) found that 3-amino pyrrolidine **4** provided the best activity (**4**, IC_{50} = 20 nM, LPC). Further biochemical analysis concluded that the lead compound was in fact a competitive inhibitor of ATX. This report was the first example of a hybrid tunnel-hydrophobic pocket ATX inhibitor.[140]

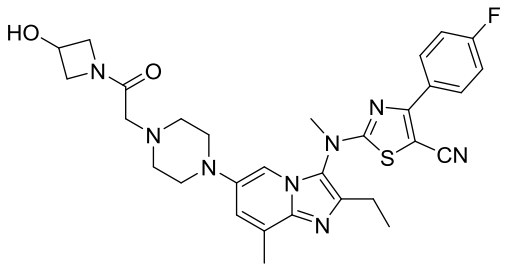
Table 10 Diamine core SAR for synthesis of hybrid compounds.



Entry		LPC IC_{50} (µM)
3		10.4
13		8.80
14		0.20
15		2.50
16		0.10
17		0.72
4		0.02
18		0.50

In addition, recent unpublished work in collaboration with NKI and the University of Twente has described, through histological studies and gene expression profiling that steroid hybrid **4** attenuates fibrosis, steatosis and inflammation post-treatment in established mouse NASH models. Liver function tests were also carried out which exemplified that on dosing **4**, a significant reduction in alanine aminotransferase (ALT), aspartate aminotransferase (AST), cholesterol and triglyceride plasma levels were detected. Elevated levels of ALT, AST, cholesterol and triglycerides are definitive markers of liver disease therefore; these *in vivo* observations contribute to the validation of ATX inhibition in the treatment of NAFLD and NASH.

However, the most significant advancement in the field of autotaxin inhibitors has been from the serendipitous discovery of GLPG1690, **19** (**Figure 27**).[127,147] In 2018, Galapagos reported a potent and selective inhibitor of ATX, **19**, which was shown to reduce plasma LPA levels *ex vivo* and *in vivo* in addition to exhibiting a favourable safety profile, good plasma pharmacokinetics (PK) and robust pharmacodynamics (PD) in Phase I clinical trials.



GLPG1690
19

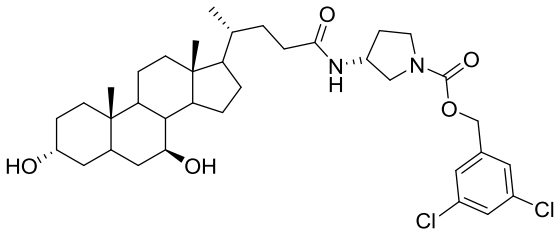
Figure 27 Structure of clinical candidate GLPG1690 **19**.

As a result, between March 2016 and May 2017, GLPG1690 was taken forward to a randomised, double-blind, placebo-controlled Phase IIa clinical trial (FLORA) wherein safety, tolerability, PK and PD were evaluated in patients diagnosed with IPF but not receiving nintedanib or pirfenidone, over the course of 12 weeks. Interestingly, the

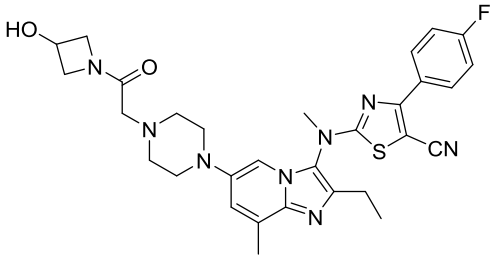
cohort selected for the trial was surprisingly small – only 23 patients were taken forward after eligibility screening with only 20 patients completing the trial.[147]

At the time of writing, Galapagos are recruiting patients for randomised, double-blind, placebo-controlled Phase III clinical trial (ISABELA1) where the main purpose is to establish efficacy and safety of two doses of GLPG1690 in patients who are receiving the current standard of care over a minimum period of 52 weeks. The primary outcome measure will be the rate of decline of forced vital capacity (FVC) in mL with secondary outcome measures also considered. The estimated enrolment is 750 patients, which will overtake bosentan[148] as the largest clinical programme undertaken in IPF, with the anticipated study competition date set for December 2021. If successful, GLPG1690 will be the first FDA-approved autotaxin inhibitor for the treatment of IPF.

A notable observation with Type IV inhibitors is their atypical properties (**Figure 28**). **4** and **19** both possess a high molecular weight and are highly lipophilic, which at first glance, may not be considered as "drug-like" as their values sit outside the parameters stated in Lipinski's "rule of five" for oral drug administration.[149]



4
MW = 664
IC_{50} = 20 nM
cLogP = 8.1



19
MW = 589
IC_{50} = 131 nM
cLogP = 5.8

Figure 28 Physicochemical properties associated with tunnel-hydrophobic pocket binders **4** and **19**.

Pleasingly, however, this works in favour of ATX as both compounds have performed well in *in vivo* studies and this is most apparent in the case of GLPG1690 which has progressed into Phase III clinical trials. From the design of these Type IV inhibitors, and extensive structural investigation, we have started to build a better understanding around the tractability of autotaxin, and the ways in which we can successfully exploit its affinity for higher molecular weight compounds.

Due to the success GLPG1690 and its validation of ATX as a clinical target, we have incentive to further investigate and explore unique and previously unreported binding modes of ATX.

2 Project Aims

Modulation of ATX has been predominantly based around the generation of orthosteric inhibitors which structurally resemble the natural product, LPA, however despite being active, their therapeutic development has been largely limited due to poor physicochemical properties. Investigation into allosteric inhibition of ATX has paved a route towards a new class of compounds which show more promise in the clinic, for example, GLPG1690. Notwithstanding these efforts, there is still space to further explore the structural capacity of ATX in order to investigate new vectors and probe the potential for differentiated downstream *in vitro* pharmacology by exploiting the variety of binding modes available.

In this study, we aimed to demonstrate the amalgamation of key structural features from two design hypotheses shown below in **Figure 29**, based on both competitive orthosteric ATX inhibitors (Type I) and allosteric hybrids (Type IV).

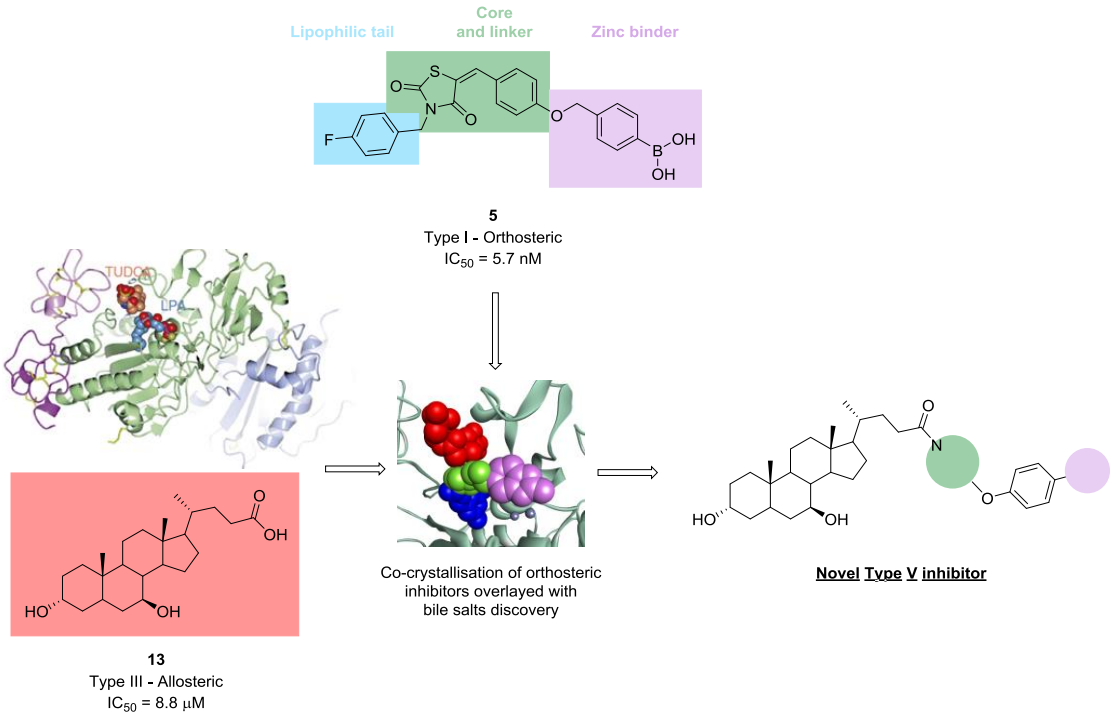


Figure 29 Design hypothesis behind novel steroid-derived Type V ATX inhibitors.

We anticipated this novel design would enable the delivery of a unique class of steroid-derived inhibitors of Type V designation which engage with the tunnel and active site of ATX – a binding mode previously unreported in the literature (**Figure 30**).

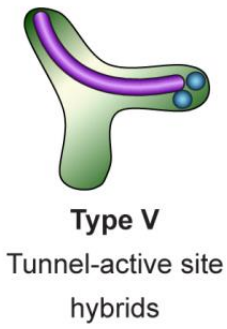


Figure 30 Cartoon depiction of our postulated binding mode for Type V inhibition.

Due to the notorious insolubility of warheads such as the benzoxazolone motif employed in PF-8380,[141,142] we first aimed to employ a fragment-type biochemical screen in order to identify alternative functionalities that could act as suitable ATX warheads, in addition to complimenting the properties of the steroid construct. We then intended to combine these findings with a range of core scaffolds to produce a thorough SAR analysis. We anticipated that this would allow us to establish the optimal compound length and vector required for activity. Despite ATX having an apparent preference for high molecular weight compounds, which can create problems for controlling physicochemical properties, we were interested in tuning clogP values throughout the design process.

We envisaged confirmation of potency for our compound library using a natural substrate assay and aimed to generate full biochemical evaluation and cell-based analysis on our exemplar compounds in order to assess their downstream fate in the signalling pathway. In collaboration with NKI, we intended to utilise a cell migration assay, reflective of a key cellular process in cancer and fibrosis, as a phenotypic means to ascertain the potential of our active compounds as therapeutic tools to understand the pathological role of ATX in proliferative diseases.

Additionally, we aspired to supplement this data with co-crystallography studies which would provide tangible verification that the steroid component is located in the tunnel and the warhead interacts in the active site, therefore demonstrating an entirely novel class of ATX inhibitors of Type V designation.

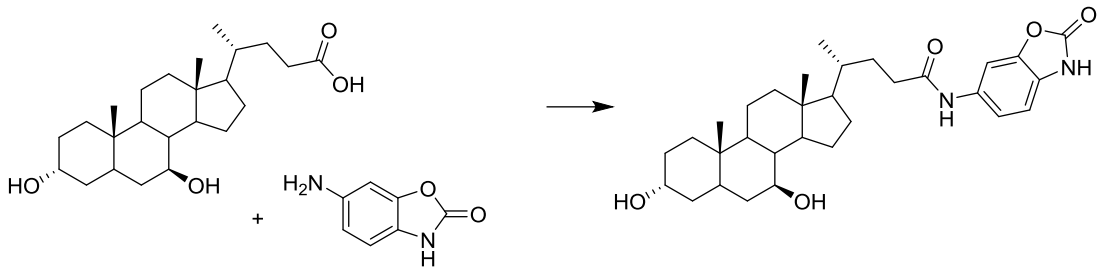
3 Results and Discussion

3.1 Binding Mode Hypothesis

The overarching goal of this project was to design and synthesise a library of novel steroid-derived inhibitors of ATX which engaged with the tunnel and active site in order to probe a unique binding mode in the phosphodiesterase domain of ATX. A major limitation in clinical development for ATX inhibitors is their inherently poor physicochemical properties. We were interested in exploring other avenues of ATX modulation by maximising and building upon the plethora of information available on known inhibitors in order to generate novel structures with improved properties.

PF-8380 has been unquestionably invaluable as a tool compound in ATX inhibitor development. The compound itself has excellent potency and has many structural diversification points, however the extremely modest solubility associated with its structure is problematic, predominantly due to the benzoxazolone functionality. This drawback, in conjunction with the solubility issues associated with the two pendant alcohol groups on the steroid construct, required careful consideration in the design strategy moving forward.

With this in mind, we were aware, from analysis through predictive modelling studies, that simply lifting an exemplar warhead such as the benzoxazolone and initiating a coupling directly with the steroid handle (**Scheme 5**) would not be satisfactory in terms of both discerning the optimal length and trajectory for compound activity and also from a solubility perspective. From rudimentary modelling studies, it could be asserted that direct coupling between the warhead and steroid would generate a compound too short to interact with the active site of ATX (**Figure 31**).



Scheme 5 Synthetic scheme to demonstrate direct coupling between the steroid and warhead.

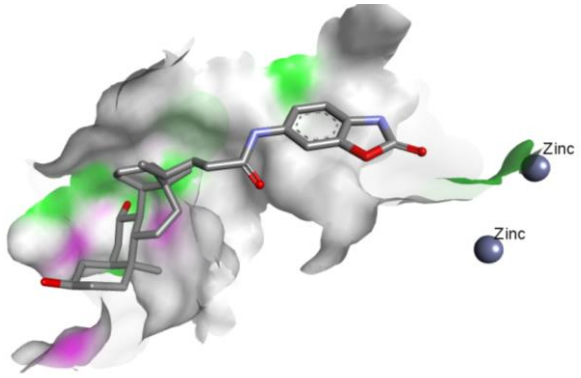


Figure 31 Predicted binding of steroid and warhead coupled in the absence of a spacer unit.

From the project outset, interactions between steroids and the tunnel region had been investigated during the work carried out by NKI, so whilst it would be interesting to explore, we were confident that there was no pressing need to explore structural modification in the steroid region. Therefore, it was evident that novelty could be introduced from the core and warhead regions. These were considered to be the key areas to explore in terms of SAR and in order to establish the limitations of compound design. The core region could be examined through implementation of various linear and cyclic cores, synonymous with most ATX-based SAR projects, with the aim of triaging compound length and trajectory (**Figure 32**).

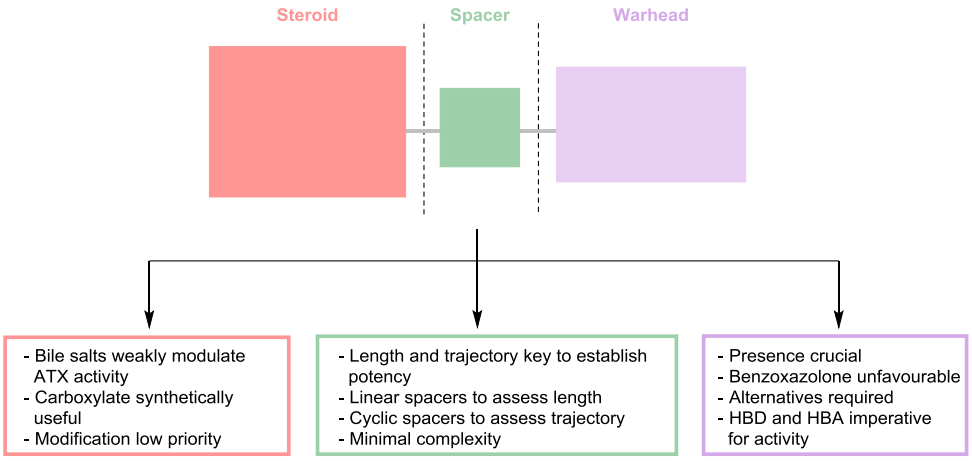
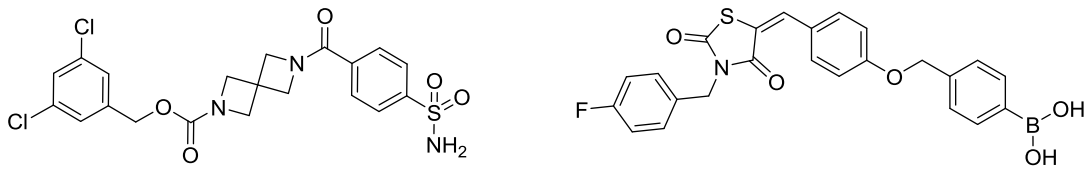


Figure 32 Approach to exploration of novel steroid-derived inhibitors and areas of diversification.

As mentioned previously, the limitations associated with the benzoxazolone and other zinc binding groups related to PF-8380 meant it was necessary to look elsewhere for warhead inspiration. To initiate our SAR campaign, we believed it was prudent to investigate alternative warheads which would complement the steroid architecture rather than work against it, which we aimed to do through a fragment-type biochemical screen in collaboration with NKI.

3.2 Identification of Alternative Warheads

Initially, the main region of interest for assessment was the warhead region as it was crucial to instigate interactions in the active site. A small fragment-type biochemical screen (**Table 11**) identified four potential warheads for attachment to the steroid manifold which were predicted to interact with the active site based on their analogy with potent SAR-derived analogue of PF-8380 **20** and tool compound **5** (**Figure 33**).



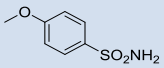
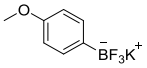
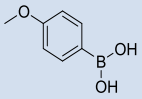
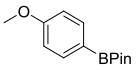
SAR analogue of PF-8380
20
IC_{50} = 56 nM

HA155
5
IC_{50} = 5.7 nM

Figure 33 Structures of existing successful ATX warheads.

With this in mind, activity of the following fragments (**Entries 21-24**, **Table 11**) against ATX was measured (**Figure 34**). We recorded fragment activity as the percentage inhibition calculated at the top concentration (100 μM) with respect to DMSO which acts as the control and constitutes 100% inhibition. Our results indicated that boron-containing warheads were substantially more favourable for progression than the sulfonamide. Despite this however, the success of SAR analogues such as **20** cannot be disregarded therefore we were still keen to pursue both avenues of warhead design.

Table 11 Fragment-type biochemical screen for selected entries **21-24**

Entry	Fragment	% Activity (100 µM)
21		92
22		43
23		63
24		69

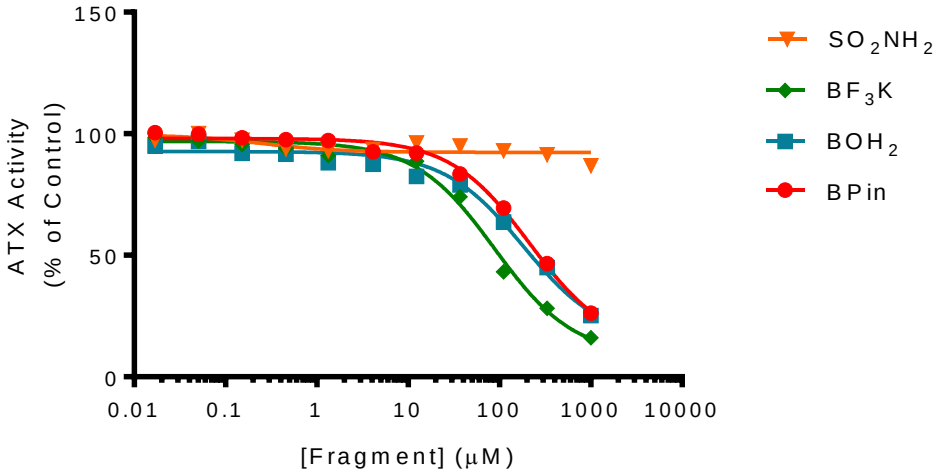


Figure 34 Preliminary screen (n=1) of fragment-type compounds indicate potential warhead candidates – the choline oxidase-coupled activity assay was used to assess inhibition of LPC hydrolysis by selected fragment compounds **21-24.**

Due to the novelty associated with the postulated binding mode, our preliminary SAR investigation was focused on probing the length of compound required for activity. Despite the significantly lower inhibition of the sulfonamide fragment in the screen, sulfonamides have literature precedent as active site binders for ATX[142] and are more synthetically tractable in comparison to the boron-containing warheads. Therefore, the

SAR was initially focused on optimising compound length with the sulfonamide warhead.

3.3 Route Design and Compound Synthesis

3.3.1 Design Strategy

It was reasoned that compounds could be generally constructed via a three-step synthetic route. Due to the inherent solubility issues associated with the steroid and the reactivity of the pendant hydroxyls on the A and B rings of the steroid (**Figure 35**), it was imperative to implement a rational and synthetically logical approach to the step order.

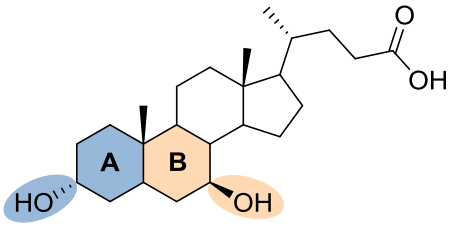
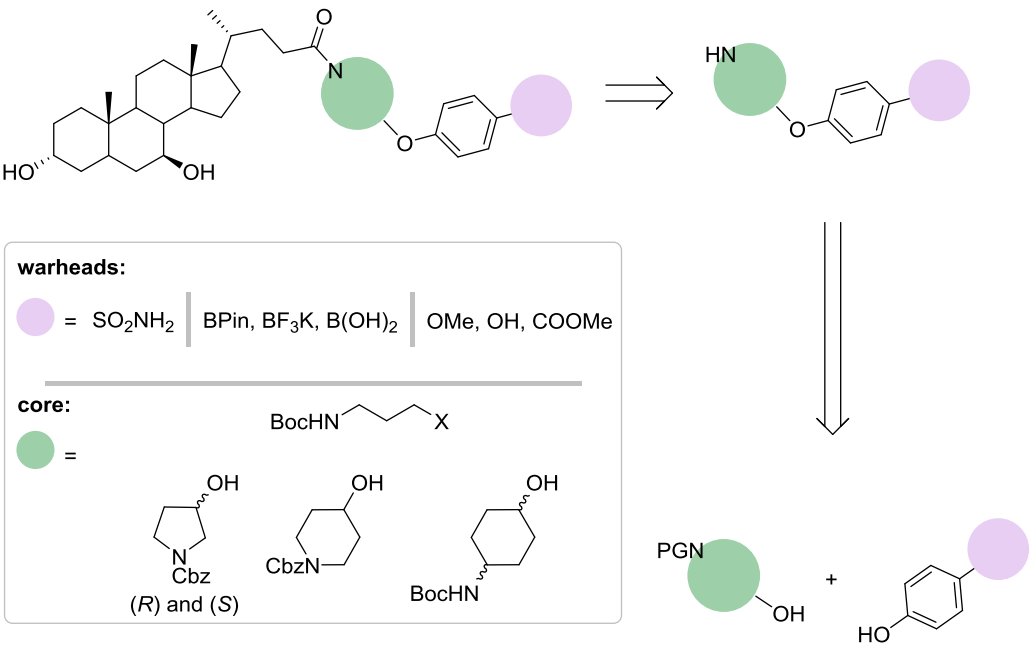


Figure 35 A and B rings of UDCA framework.

It was envisaged that amide coupling of the steroid to the core and warhead component would be the final step of the synthesis to circumvent any potential solubility issues early on in the synthesis as depicted in **Scheme 6**. This would be prefaced with a protecting group removal freeing the amine functionality which could be synthesised from S_N2 nucleophilic substitution between varying *N*-protected amines and substituted phenols.

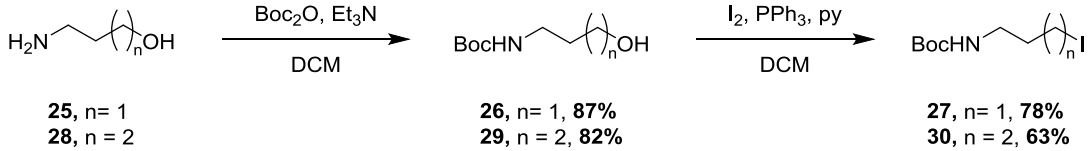


Scheme 6 Proposed retrosynthetic analysis towards inhibitor design with example cores and warheads.

Early in the project, it was evident that this would not be a particularly modular synthesis due to the general inability to implement late-stage diversification after steroid coupling. However, to re-iterate, we were predominantly interested in rapidly identifying a hit candidate through a rational compound design strategy as opposed to carrying out an in-depth SAR optimisation on a known lead candidate.

3.3.2 Synthesis of Linear Analogues

In order to generally assess compound length, synthesis towards the sulfonamide analogues began through the generation of ethyl, propyl and butyl linkers (**Scheme 7**). The precursor to the alkylation reaction was synthesised through a Boc protection of either 3-aminopropanol **25** or 4-aminobutanol **28** with di-*tert*-butyl carbonate (Boc_2O) and triethylamine (Et_3N). This was followed by an Appel reaction with elemental iodine, triphenylphosphine (PPh_3) and pyridine (py) in both cases to produce the respective alkyl halides **27** and **30**. These were then implemented directly as the starting materials for the subsequent alkylation reactions.



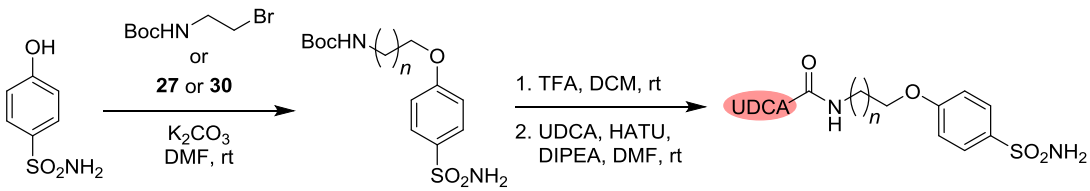
25, n= 1
28, n = 2

26, n= 1, **87%**
29, n = 2, **82%**

27, n= 1, **78%**
30, n = 2, **63%**

Scheme 7 Boc protection and Appel reaction for the generation of alkyl halide starting materials **27** and **30**.

N-Boc protected ether-linked intermediates **32, 34** and **36** were synthesised via an alkylation reaction (**Scheme 8**) using 4-hydroxybenzenesulfonamide **31** and either commercially available *tert*-butyl (2-bromoethyl)carbamate or previously synthesised alkyl halides **27** or **30** in the presence of potassium carbonate (K_2CO_3). Intermediates **32, 34** and **36** were deprotected under acidic conditions to expose the free amine which was used directly in the HATU-mediated amidation with UDCA to generate final compounds **33, 35** and **37** in moderate yield over two steps.



31

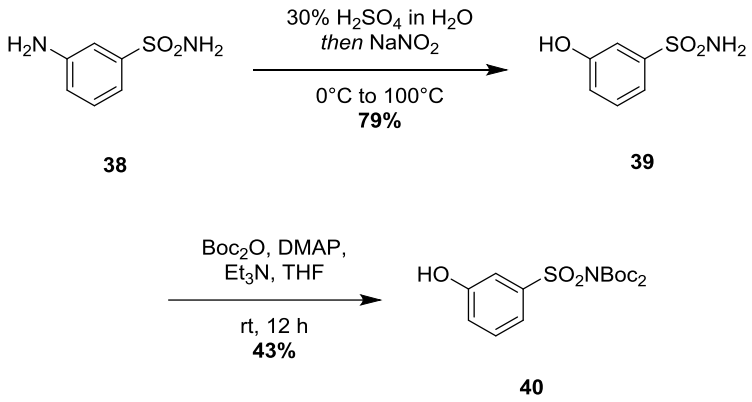
32, n = 1. **35%**
34, n = 2, **31%**
36, n = 3, **39%**

33, n = 1, **30%** (over two steps)
35, n = 2, **42%** (over two steps)
37, n = 3, **36%** (over two steps)

Scheme 8 Synthesis of *para*-substituted linear sulfonamides.

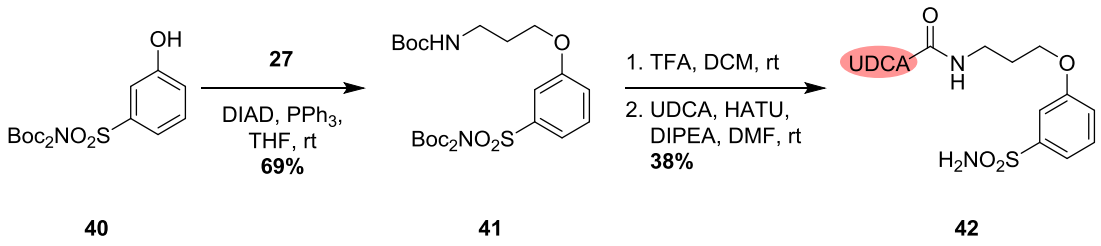
Initial attempts to access the *meta*-substituted sulfonamides intermediates using the same methodology as above failed. Therefore, generation of *meta*-substituted analogues required a more protracted synthesis. 3-hydroxybenzenesulfonamide **39** was synthesised from 3-aminobenzenesulfonamide **38** via a tandem one-pot diazotisation/hydrolysis reaction with 30% sulfuric acid and sodium nitrate in water (**Scheme 9**). Double Boc protection of the sulfonamide functionality was undertaken using standard Boc protection conditions and 4-dimethylaminopyridine (DMAP) employed as an additive to

generate **40**. This strategy was essential to prevent unwanted reaction of the sulfonamide in the ensuing S_N2 reaction.



Scheme 9 Synthesis of *tert*-butyl (*tert*-butoxycarbonyl)((3-hydroxyphenyl)sulfonyl)carbamate.

Synthesis of the *meta*-substituted final compound **42** was facilitated through a Mitsunobu reaction with intermediates **40** and **27** in the presence of PPh_3 and diisopropylazodicarboxylate (DIAD) yielding ether-linked intermediate **41** (**Scheme 10**). Global Boc deprotection released both the sulfonamide and amine functionalities which allowed for amidation with UDCA to proceed generating final compound **42** in moderate yield.



Scheme 10 Synthesis of linear *meta*-substituted sulfonamide analogue **42**.

A sub-series of linear boronate ester analogues were delivered using a similar synthetic strategy. Close-running impurities using the DIAD-mediated Mitsunobu method led to

the application of an alternative Mitsunobu reagent – cyanomethylenetributylphosphorane (CMBP), shown in **Figure 36**.[150] CMBP is a phosphorane ylide pioneered by Tsunoda and colleagues[150] and has been showcased in the literature as an "all-in-one" Mitsunobu catalyst.[151] It serves as both the azo-type reagent and as the phosphine meaning that it can accommodate a wider range of substrates, specifically secondary alcohols, which was beneficial in anticipation for our advancement into cyclic spacer units.

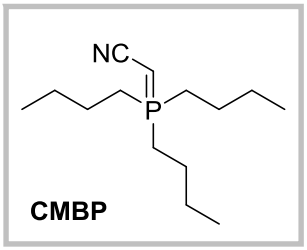
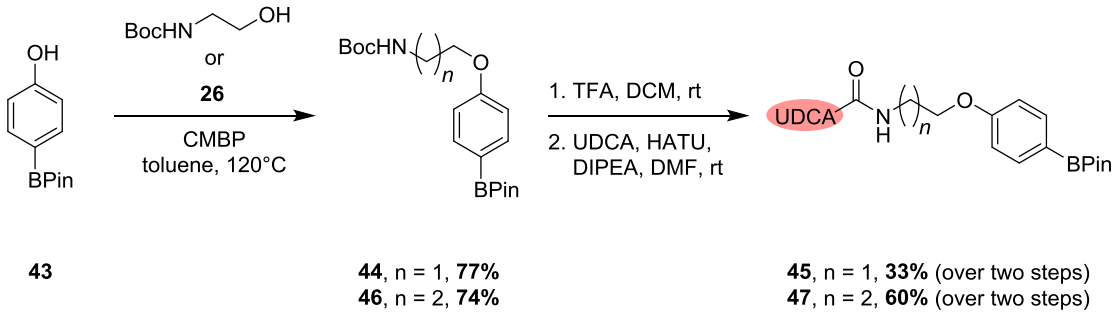


Figure 36 Structure of cyanomethylenetributylphosphorane (CMBP).

Some minor disadvantages in using this catalyst are its expense in addition to the requirement for reaction temperatures greater than the solvent reflux. Therefore, sealed reaction vessels are essential and there are limitations on thermally labile groups. However, from a synthesis perspective, the aforementioned caveats are trivial when comparing the reaction profile and synthetic ease to more traditional methods; DIAD-mediated Mitsunobu reactions tend to be poor yielding, exhibit a limited substrate scope and are plagued with stubborn by-products which limit successful purification as was observed first-hand.

Based on this, linear boronate intermediates were generated using 4-hydroxyphenol boronic acid pinacol ester **43** and either *N*-Boc ethanolamine or intermediate **26** in the presence of CMBP and toluene at 120 °C (**Scheme 11**). Subsequent Boc deprotection and amidation with UDCA provided *para*-substituted final compounds **45** and **47**.



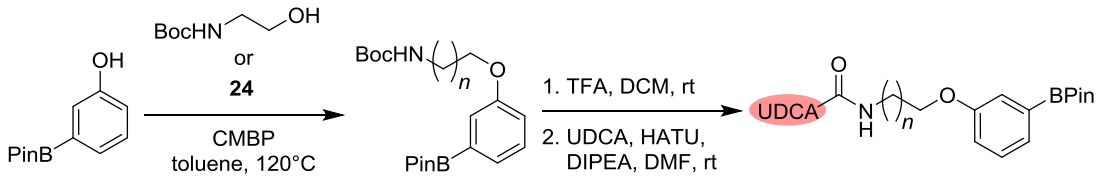
43

44, n = 1, **77%**
46, n = 2, **74%**

45, n = 1, **33%** (over two steps)
47, n = 2, **60%** (over two steps)

Scheme 11 Synthesis of linear 4-substituted boronate ester analogues **45** and **47**.

The same methodology was carried out using 3-hydroxyphenol boronic acid pinacol ester **48** and *N*-Boc ethanolamine or intermediate **26** to generate *meta*-substituted final compounds **50** and **52** (**Scheme 12**).



48

49, n = 1, **56%**
51, n = 2, **88%**

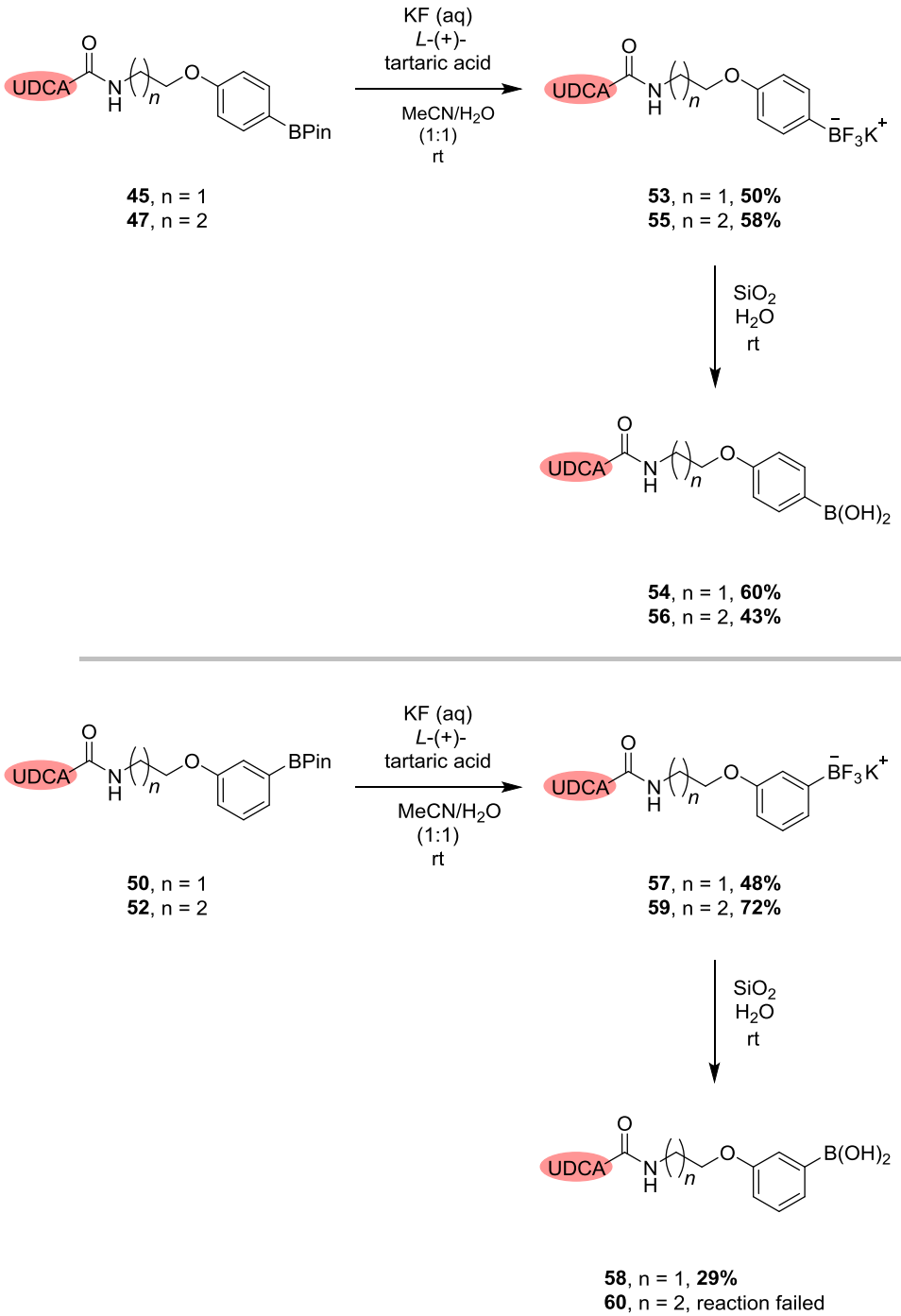
50, n = 1, **44%** (over two steps)
52, n = 2, **33%** (over two steps)

Scheme 12 Synthesis of linear 3-substituted boronate ester analogues **50** and **52**.

Boronate esters **45**, **47**, **50** and **52** were taken forward for further synthetic diversification in order to access the free boronic acid. It was envisaged that, in the interest of maximising yield, the most efficient method of releasing the boronic acid was via the trifluoroborate salt.

This was achieved by subjecting compounds **45**, **47**, **50** and **52** to conditions pioneered by Lloyd-Jones and Lennox[152] which employ a specific ratio of aqueous potassium fluoride, (+)-tartaric acid in a 1:1 mixture of methanol and acetonitrile, successfully forming corresponding trifluoroborates **53, 55, 57** and **59** (**Scheme 13**). Typical

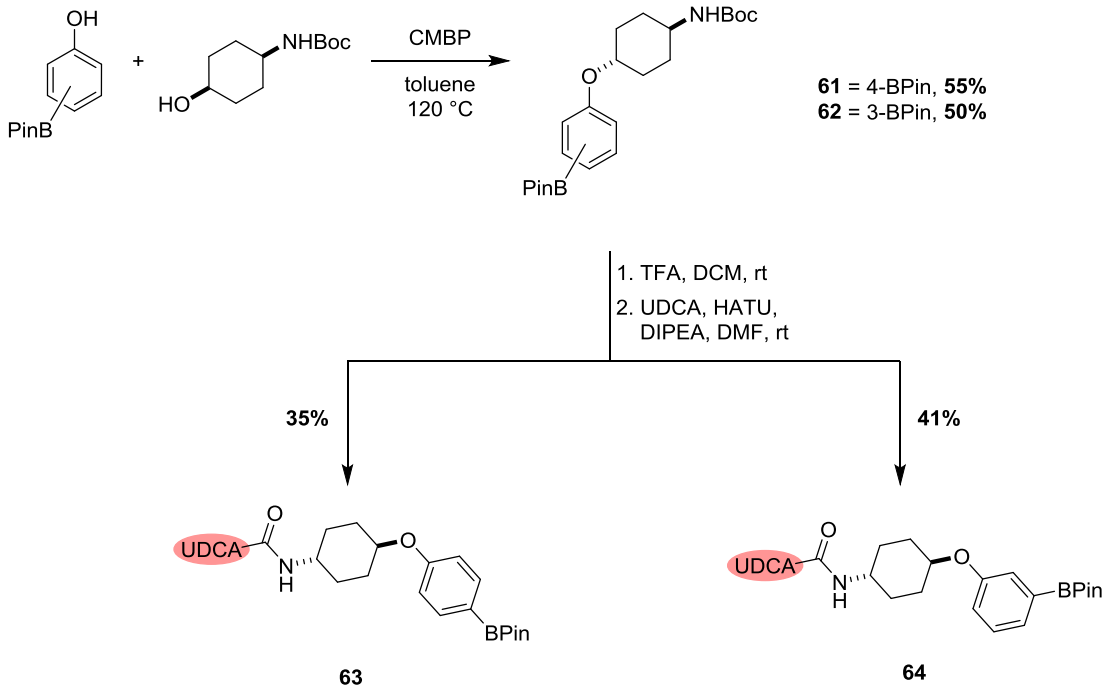
conditions utilising KHF_2 were attempted prior, however did not result in analytically pure material. Subsequent hydrolysis using excess silica in water furnished boronic acids **54, 56** and **58**.



Scheme 13 Synthesis of linear trifluoroborate and boronic acid derivatives using Lloyd-Jones conditions[152] followed by standard $B(OH)_2$ hydrolysis procedure.

3.3.3 Synthesis of Cyclic Analogues

The vector of the warhead was explored through application of various conformationally restrained aminoalcohol cores. Initially, cyclohexyl intermediates **61** and **62** with an exocyclic amino alcohol were synthesised via CMBP-mediated Mitsunobu reaction between *tert*-butyl ((1*s*,4*s*)-4-hydroxycyclohexyl)carbamate and either 3- or 4-hydroxyphenol boronic acid pinacol ester. These building blocks were then used to synthesise final compounds **63** and **64** which could be easily generated through deprotection of **61** and **62** followed by amidation with UDCA (**Scheme 14**).

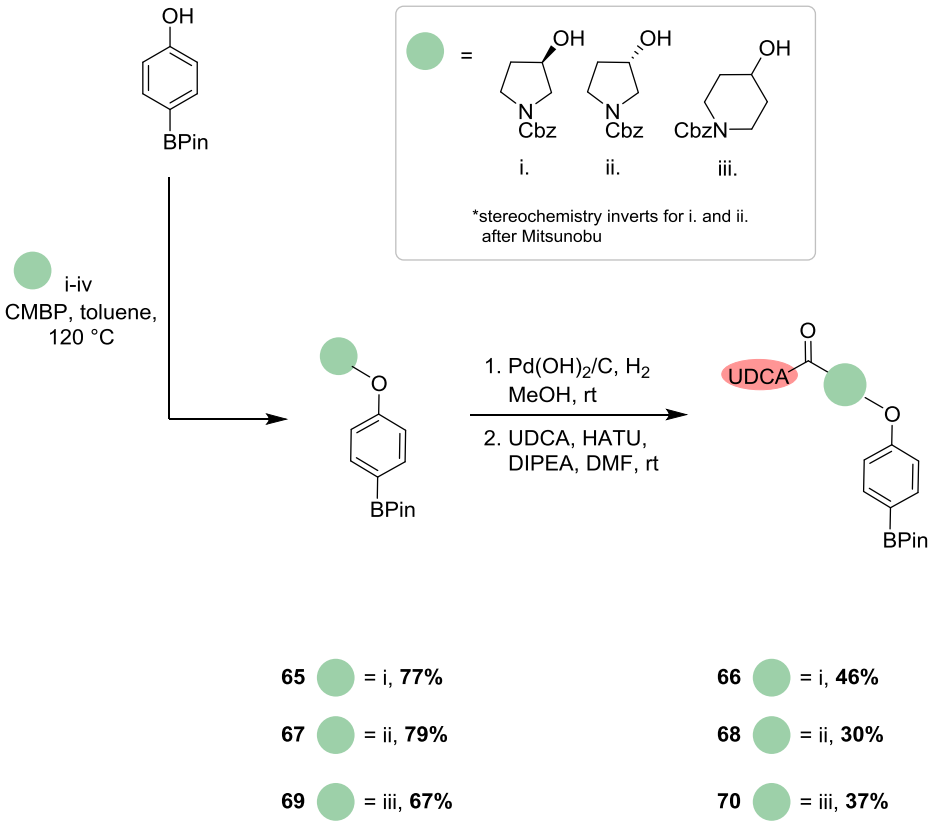


Scheme 14 Synthesis of cyclohexyl-derived inhibitors with both 3- and 4-substituted aryl boronate esters.

A small range of cyclic building blocks were generated using Cbz-protected (piperidinol, (*S*)-pyrrolidinol and (*R*)-pyrrolidinol) with 4-hydroxyphenol boronic acid pinacol ester under CMBP-mediated Mitsunobu conditions (**Scheme 15**).

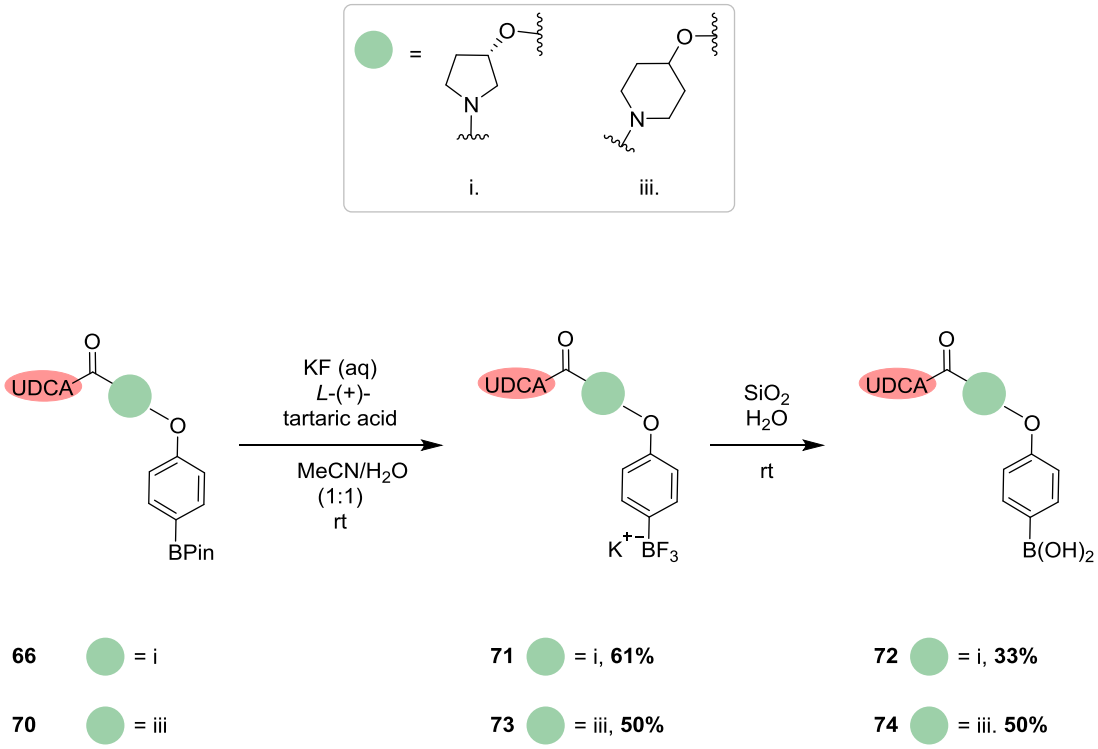
The usefulness of altering the Mitsunobu reagent from DIAD to CMBP was extremely apparent whereby intermediates **65, 67** and **69** were formed in good to excellent yields.

Protecting group cleavage was facilitated through Pd-catalysed hydrogenation using Pearlman's catalyst allowing access to boronate ester final compounds **66, 68** and **70.**



Scheme 15 Synthesis of conformationally restrained steroid-derived inhibitors with boronate ester warhead.

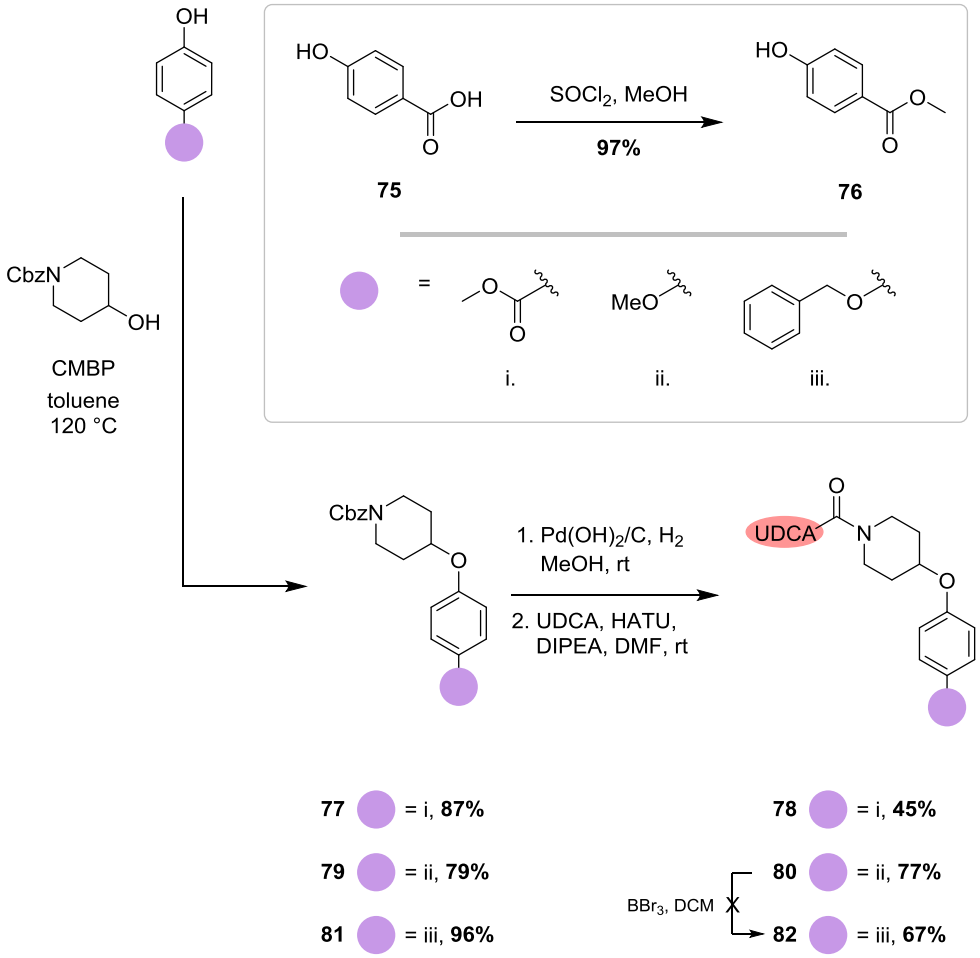
Based on the success of previous boronic acid formation, boronate esters **66** and **70** were further diversified using the same trifluoroborate conditions as described previously[152] (Section 3.3.2) providing compounds **71** and **73** in moderate yield (**Scheme 16**). Subsequent hydrolysis using excess silicon dioxide and water yielded boronic acids **72** and **74.**



Scheme 16 Synthesis of conformationally restrained steroid-derived inhibitors with trifluoroborate and boronic acid warheads, respectively.

3.3.4 Synthesis of Alternative Warhead Analogues

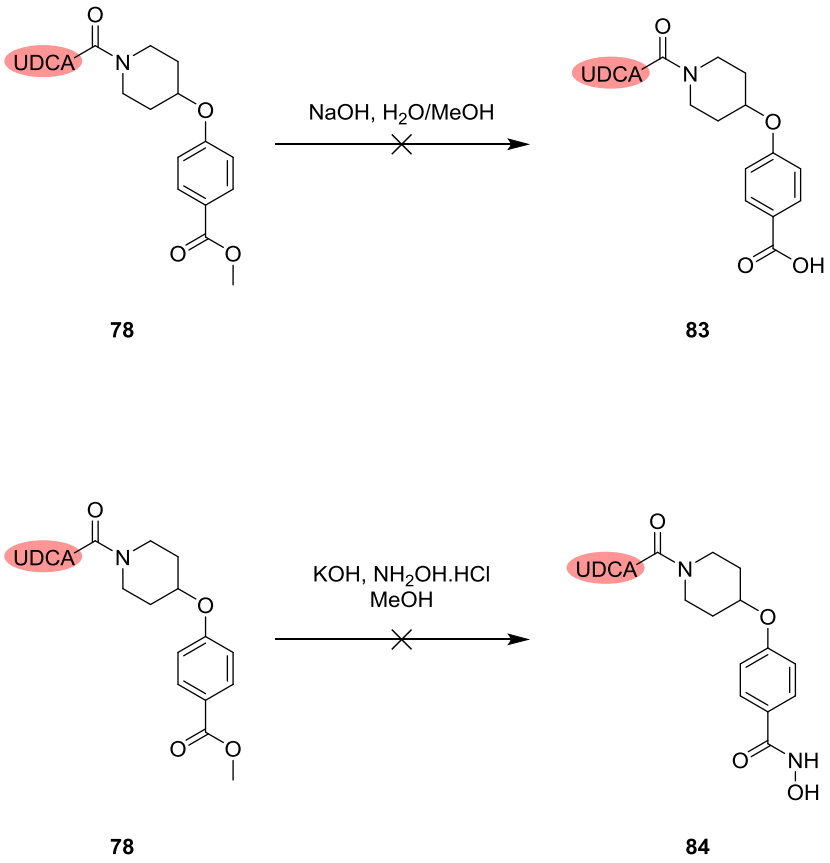
A sub-series of non-boron-containing warheads were also synthesised using the piperidinol core. Esterification of 4-hydroxybenzoic acid **75** with thionyl chloride ($SOCl_2$) provided starting material 4-hydroxybenzoate **76**. CMBP-mediated Mitsunobu conditions were employed to couple methyl 4-hydroxybenzoate, 4-methoxyphenol and 4-(benzyloxy)phenol with benzyl 4-hydroxypiperidine-1-carboxylate thereby generating intermediates **77, 79** and **81** with ester, methoxy and O-benzyl groups respectively. Palladium-catalysed hydrogenation using Pearlman's catalyst facilitated global Cbz deprotection in all three cases and final compounds **78, 80** and **82** were then formed via HATU-mediated amide coupling with UDCA (**Scheme 17**).



Scheme 17 Synthesis of alternative warhead analogues **78, 80** and **82.**

Efforts to transform methoxy **80** into phenol **82** using standard demethylation conditions involving BBr_3 failed; therefore a double protecting group strategy was implemented using **81** which successfully led to the generation of compound **82** whereby the O-benzyl group was cleaved during the Pd-catalysed Cbz deprotection to release the phenol.

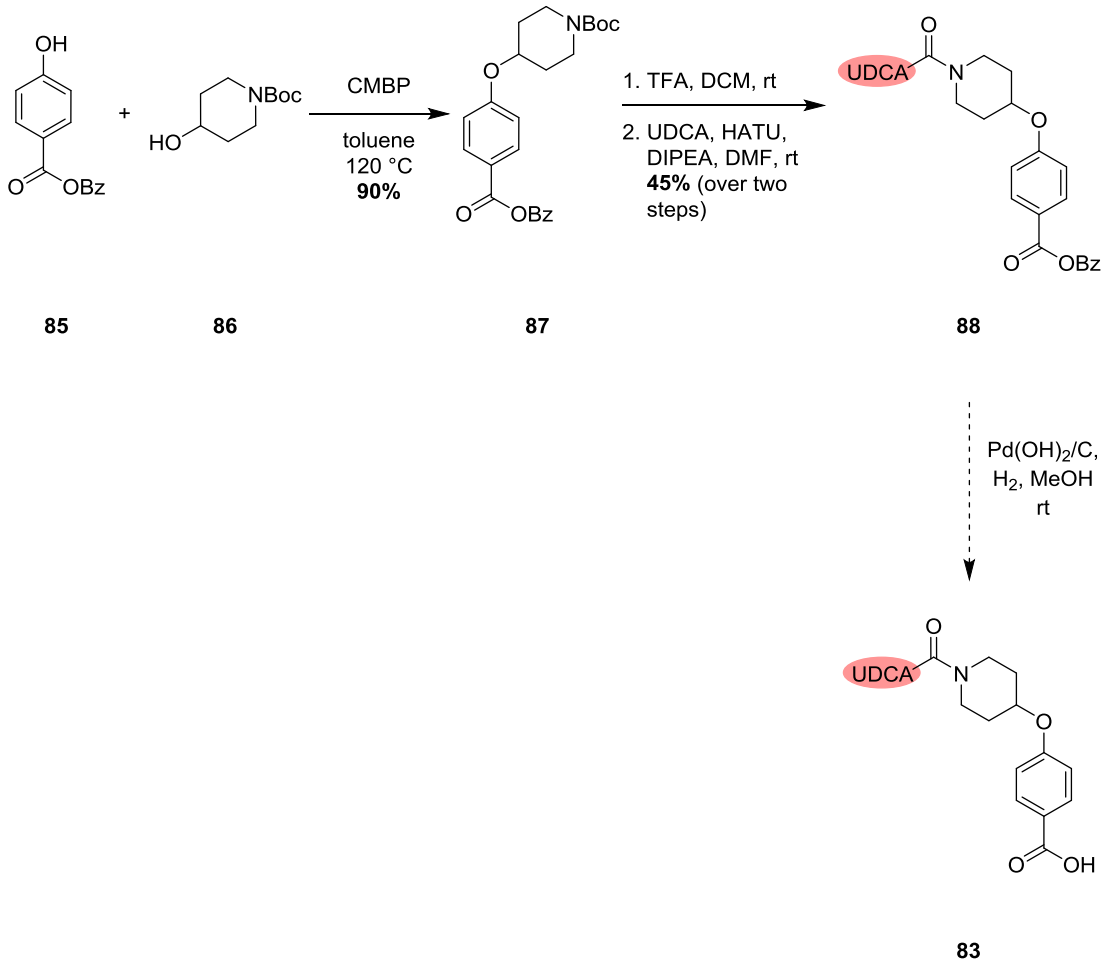
Further late stage diversification attempts to access one-point changes from compound **78** through base-mediated ester hydrolysis to access carboxylic acid **83** or through hydroxamic acid formation using potassium hydroxide and hydroxylamine hydrochloride to access hydroxamic acid **84** also failed (**Scheme 18**).



Scheme 18 Failed late-stage diversification attempts on compound **78**.

As a result, bespoke syntheses were required to access carboxylic acid and hydroxamic acid motifs. Exploiting the utility of protecting groups again, intermediate **87** could be generated through CMBP-mediated Mitsunobu reaction with benzyl 4-hydroxypiperidine-1-carboxylate **85** and piperidinol **86**. Boc deprotection unmasked the amine component which could be successfully coupled with UDCA to form amide **88**. Subsequent Pd-catalysed hydrogenation with Pearlman's catalyst would be the final step of this synthesis towards carboxylic acid **83** however this was halted in favour of other priority compounds (**Scheme 19**).

.



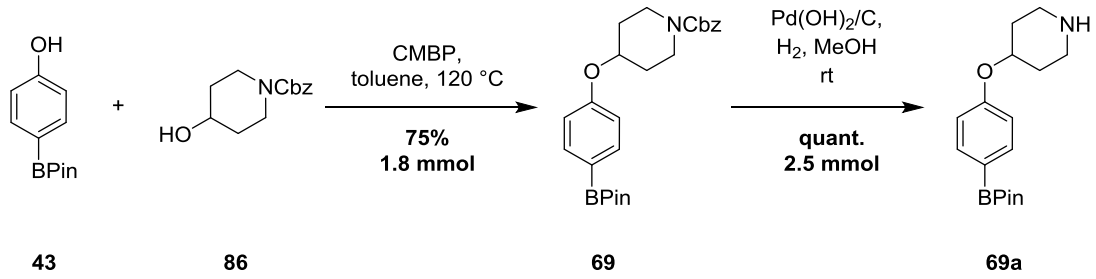
85 **86** **87** **88** **83**

Scheme 19 Bespoke access to carboxylic acid warhead **83.**

Plans to access the hydroxamic acid were also paused to focus on higher priority compounds however it was envisaged that the hydroxamic acid could be accessed through intermediate **77** with hydroxamic acid formation taking place prior to amidation with UDCA.

3.3.5 Synthesis of Alternative Steroid Analogues

A small range of alternative steroids were considered which employed the same conditions as discussed previously. For the first time, intermediate **69a** could be generated on a relatively large scale as a building block in order to facilitate the coupling of different steroid scaffolds. Intermediates **69** and **69a** were successfully synthesised on 1.8 mmol and 2.5 mmol scales respectively via Mitsunobu and subsequent Pearlman's catalyst-mediated hydrogenation which proves the amenability of the first two steps to scale-up conditions (**Scheme 20**).



Scheme 20 Scale-up of compound **69a.**

A range of structurally similar steroid motifs with a carboxylate terminus were selected for coupling with intermediate **69a** (**Figure 37**).

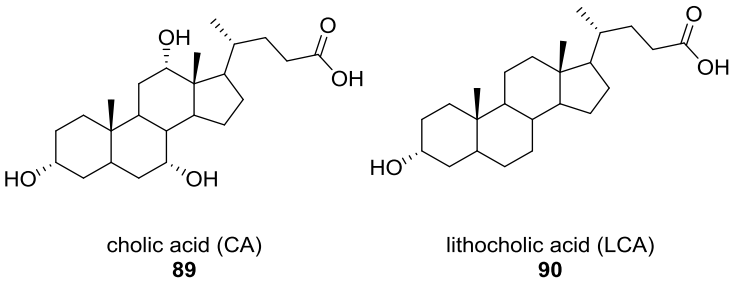
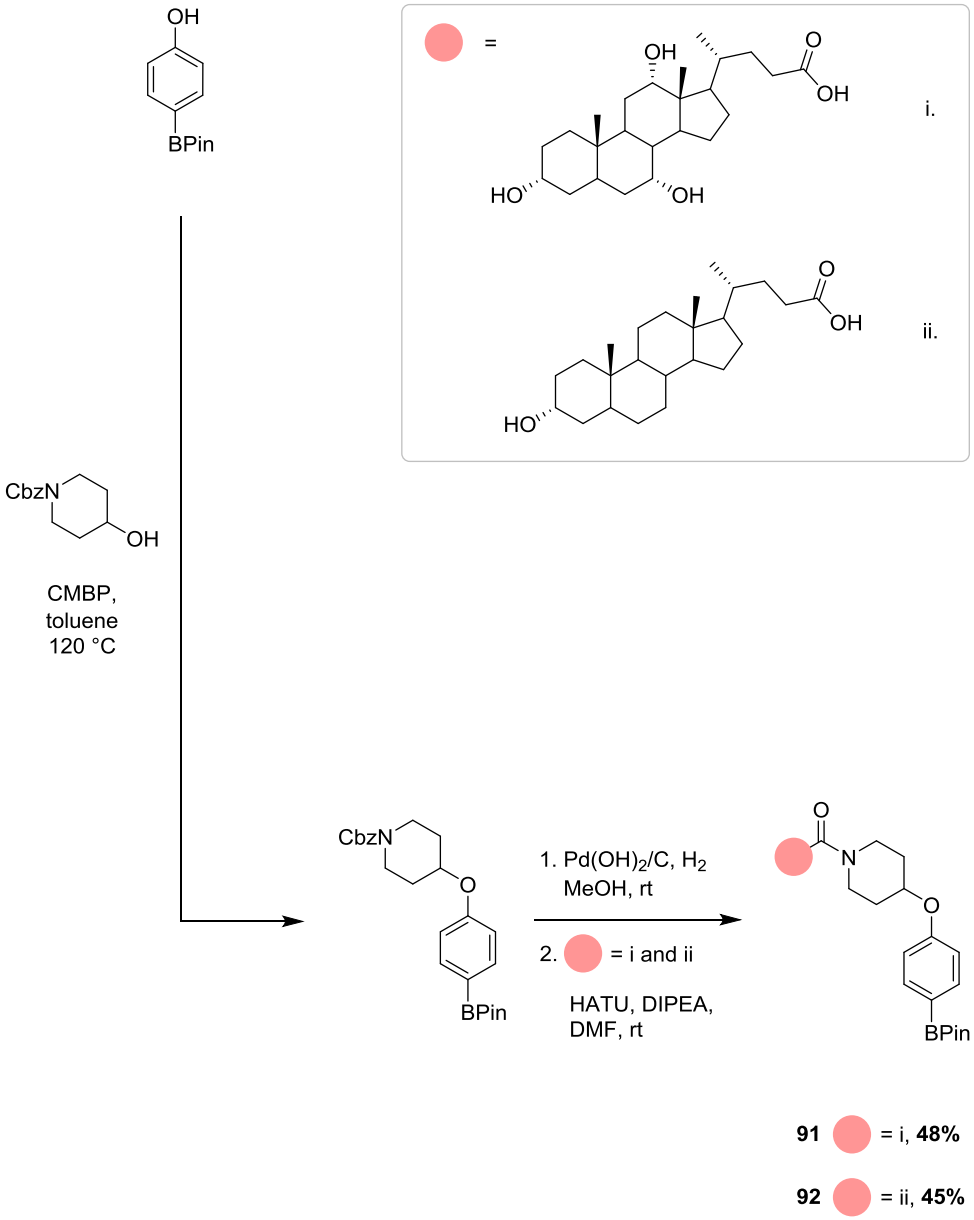


Figure 37 Selection of structurally related steroids for synthesis.

Cholic acid **89** and lithocholic acid **90** were coupled, respectively, using HATU-mediated amide coupling conditions providing boronate ester final compounds **91** and **92** in moderate yields (**Scheme 21**).



Scheme 21 Synthesis of alternative steroid analogues based on direct derivatives of UDCA.

3.4 Initial SAR Analysis

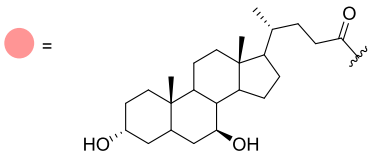
3.4.1 Delineation around Linear Sulfonamide Analogues

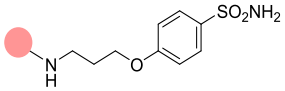
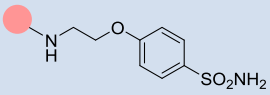
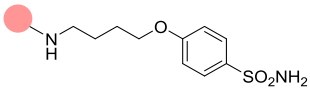
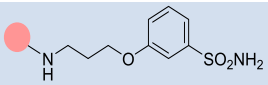
Due to the novelty associated with the proposed binding mode, our preliminary SAR investigation was focused on probing the length of compound required for activity. Despite the lower inhibition of the sulfonamide fragment in the screen, sulfonamides have literature precedent as active site binders for ATX and have been successful potent SAR candidates in other ATX inhibitor libraries, for example, **20**.

Sulfonamides were initially considered to be more synthetically tractable in comparison to the boron-containing warheads. Therefore, the SAR was initially focused on optimising compound length with the sulfonamide warhead.

To our delight, with reference to **Table 12**, propyl-linked compound **35** with *para*-sulfonamide substitution on the aryl ring showed improved potency (**35**, IC_{50} = 0.4 μM) compared to the progenitor steroid (**13**, IC_{50} = 9 μM) when measured using an LPC hydrolysis assay, shown in Table 12. The calculated logP based on ChemDraw algorithms predicted a value of 4.7 which is fairly moderate considering the molecular weight. Activity was lost completely on truncation (**33**, IC_{50} > 10 μM) and homologation (**37**, IC_{50} > 10 μM) demonstrating that there was a narrow window in optimal linker length in order to achieve potency. Negligible change is observed in clogP due to very minimal structural change in adding or detracting one carbon. Additionally, no improvement in activity was observed on switching to the *meta*-substituted derivative (**42**, IC_{50} > 10 μM) and clogP remained the same as **35**. These results indicated that a three-carbon linker and *para*-substitution were most beneficial to potency.

Table 12 Evaluation of sulfonamide analogues



Compound Number	Structure	IC_{50} (μM)	cLogP
13		9	-
35		0.4	4.7
33		> 10	4.4
37		> 10	4.6
42		> 10	4.7

3.4.2 Delineation around Linear Boron Analogues

Early indications from our fragment-type screen anticipated that the introduction of boron-containing warheads could result in significant levels of activity given the success of HA155 early in the development of ATX inhibitors,[153] and its continued importance in the present day.

The application of boron in therapeutics has been somewhat overlooked in favour of more “drug-like” features, however FDA-approval of Bortezomib in 2003[154] has piqued the interest of many medicinal chemists and prompted the inclusion of boronic acids as useful motifs in medicinal chemistry design strategies, particularly when considering them as warheads for enzyme inhibition.

Bortezomib falls in to the category of anti-cancer medication and is prescribed for the treatment of multiple myeloma and mantle cell lymphoma. Structurally, it consists of a modified dipeptide in addition to an alkyl boronic acid which reversibly inhibits the 26S proteasome (**Figure 38**).

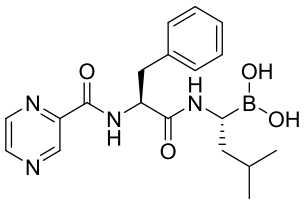


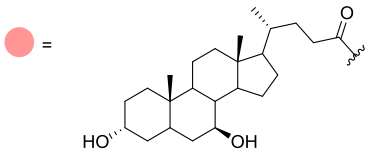
Figure 38 Structure of Bortezomib, an FDA-approved proteasome inhibitor containing a boronic acid functionality.

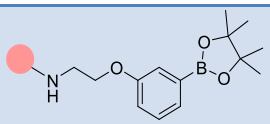
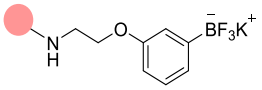
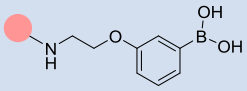
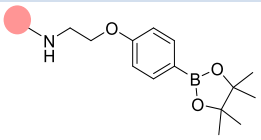
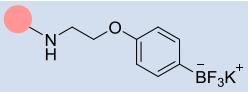
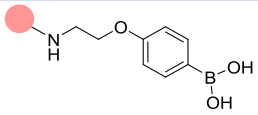
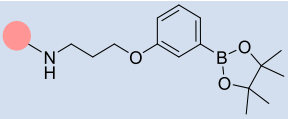
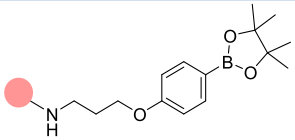
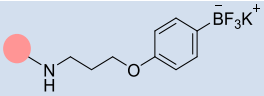
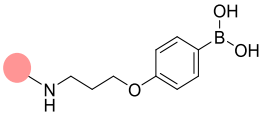
With this in mind, and based on the findings from our screen, we were interested in building on our observations from the sulfonamide sub-series in order to generate matched pairs and explore a series of boron-containing analogues.

The pinacol boronate group was an essential intermediate in the synthesis towards the boronic acid functionality due to the p*K*a similarity between the phenol and boronic acid in the initial Mitsunobu step necessitating protection of the latter functional group. Therefore, testing of the pinacol boronate was to some degree serendipitous, however its demonstrable inhibition in the fragment screen drove our curiosity.

As demonstrated in **Table 13**, a ten-fold increase in potency was observed moving from the ethyl-linked *meta*-substituted pinacol boronate (**50**, IC_{50} = 14 μM) to the *para*-substituted pinacol boronate (**45**, IC_{50} = 1.5 μM) which corroborated our hypothesis that *para*-substitution was preferred. This hypothesis was further substantiated by homologating the linear linker to the propyl derivative which resulted in a further ten-fold increase in ATX inhibition for *meta*- (**52**, IC_{50} = 0.15 μM) and *para*-substituted pinacol boronate (**47**, IC_{50} = 0.07 μM). This indicated that the intrinsic functionality associated with boron-containing motifs plays an important role in inhibitory activity, and was better tolerated in combination with an appropriate length of linker. Naturally, a significant increase in cLogP was associated with this pinacol group in comparison to the sulfonamide functionality due to its lipophilicity.

Table 13 Evaluation of linear boron analogues



Compound Number	Structure	IC_{50} (μM)	cLogP
50		14	7.4
57		N/A	4.7
58		N/A	5.3
45		1.5	7.4
53		N/A	4.7
54		0.101	5.3
52		0.152	7.7
47		0.071	7.7
55		0.07	5.0
56		0.052	5.6

Pleasingly, propyl analogues bearing a pendant trifluoroborate (**55**, IC_{50} = 0.07 μM) or boronic acid (**56**, IC_{50} = 0.05 μM) retained inhibition of ATX when compared to the parent pinacol boronate **47**. Other boronic acids such as **54** were also active despite having a shorter linker which is surprising, however, we believe this demonstrates how well-tolerated the boronic acid is in the active site.

As anticipated, clogP improves considerably on removal of the pinacol protecting group to release the trifluoroborate. Negligible difference is apparent between the trifluoroborate and boronic acid species in each case.

These results corroborated the observations generated from the fragment-type biochemical screen, revealing the inherent versatility of all three boron-containing warheads in practice. We believe that partial hydrolysis of the boronate ester occurs under enzymatic assay conditions which explains the comparable activity between the protected boronic acids (boronates) and the free boronic acid however; this observation lends itself well to exploitation of the pinacol group derivatives as potential prodrug strategies.

Interestingly, returning to Bortezomib, the product itself is provided in a lyophilised form as the mannitol boronic ester[155] which is in equilibrium with its hydrolysis product, the boronic acid **(Figure 39**).

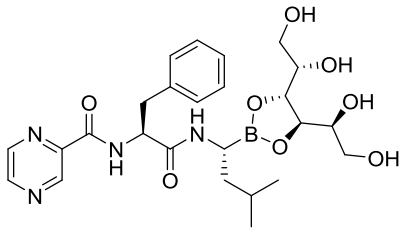


Figure 39 Mannitol boronic ester of Bortezomib for administration.

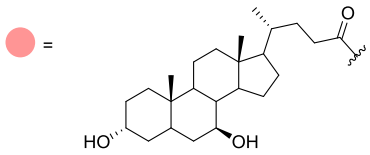
3.4.3 Delineation around Cyclic Analogues

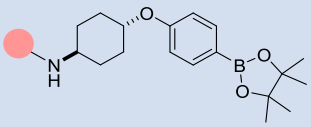
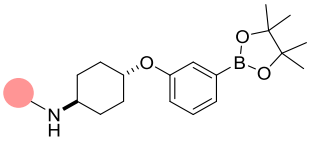
The flexibility associated with linear compounds is largely unfavorable based on thermodynamic considerations; however, we reasoned that this could be easily rectified

by introducing conformational restraint. With suitable warheads in hand, we next examined compound trajectory by restricting rotational freedom in the linker region whilst preserving length.

At first, attempts to constrict the linker by implementing a cyclohexyl-type core with both heteroatoms being exocyclic resulted in inactivity (**Table 14**). This is perhaps unsurprising as the spacer unit equates to four carbons therefore, may still be too long and therefore, not accommodated by ATX. Quite interestingly however, the observed difference in potency between the 3- and 4-substituted analogues remains which again, confirmed our hypothesis that 4-substitution is most optimal.

Table 14 Evaluation of cyclohexyl boronate warheads



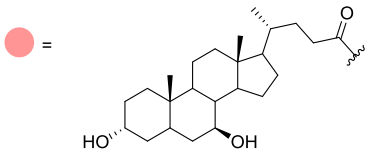
Compound Number	Structure	IC_{50} (μM)	cLogP
63		6	7.9
64		16	7.9

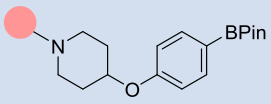
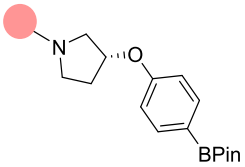
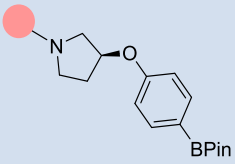
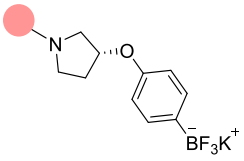
From previous studies, we knew that a secondary amide, for example, in steroid hybrid **4**, did not result in any key interactions. Therefore, the next step was to consider N-containing saturated heterocycles. In order to retain the desired length that we anticipated would be required for activity, we implemented piperidinol and both (*R*)- and (*S*)-pyrrolidine. We anticipated that making small changes from this point onwards would aid us in identifying the exact requirements to correlate trajectory and activity.

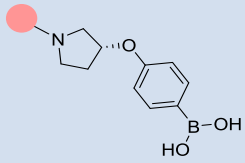
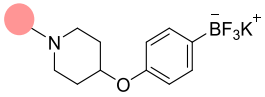
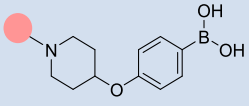
Pleasingly, as depicted in **Table 15**, all 3 cores combined with the pinacol boronate warhead maintained potency with negligible difference noted. Based on this, (*S*)-pyrrolidine was selected for further optimisation and a slight improvement in activity was observed on derivatisation to the respective trifluoroborate and boronic acid functionalities (**71** and **72**).

To our delight, however, the most significant result was obtained by accessing trifluoroborate (**73**, IC_{50} = 0.05 μM) and boronic acid (**74**, IC_{50} = 0.03 μM) derivatives, which contain the piperidine core. Boronic acid **74** represented a 400-fold increase in activity in comparison to the progenitor steroid (**15**, IC_{50} = 11 μM).

Table 15 Evaluation of cyclic boron analogues



Compound Number	Structure	IC_{50} (μM)	cLogP
70		0.08	7.0
66		0.20	7.4
68		0.21	7.4
71		0.11	4.7

72		0.09	5.3
73		0.05	4.3
74		0.03	4.9

3.5 Co-crystallisation of Exemplar Compounds

With potent compounds in hand, and in order to confirm our binding mode hypothesis, we sought to obtain co-crystallisation data of our exemplar compounds with ATX. Crystal structures were generated for two of our most potent compounds, **73** and **74**, to extremely high quality resolution, by our collaborators at NKI.

As anticipated, in all three cases, our intended binding mode was confirmed with the steroid scaffold occupying the hydrophobic tunnel, synonymous with its observed orientation in the crystal structure of steroid hybrid, **4**, and ATX. Gratifyingly, interactions were also observed in the active site with several contacts visible to key neighbouring residues depending on the nature of the warhead and the trajectory of the core motif. Each of the crystal structures will be discussed in turn with respect to the specific binding of each example.

*3.5.1 Co-crystallisation of **73***

For **73**, co-crystallisation with ATX resulted in generation of a crystal structure to an excellent resolution of 2.5 Å. Compound **73** was modelled and subsequently refined using molecular modelling software, which resulted in the assignment of a clear conformation which fits the collected electron density map as depicted in **Figure 40**.

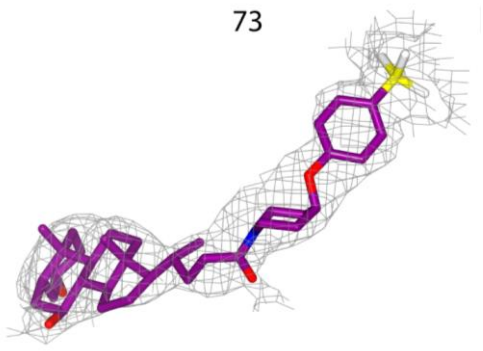


Figure 40 Fit of **73** to electron density.

Analysing each section of the compound in turn, the active site indicated hydrogen bonding interactions which were visible between one of the fluorine atoms and both Asn230 and Thr209 (**Figure 41**), bearing similarity to the observed binding of the phosphate head group of LPA. An additional interaction was observed from the second fluorine which appeared to coordinate to the proximal zinc ion in the active site. Satisfyingly, hydrogen bond interactions between the 3-OH and 7-OH and Trp260 and Tyr82 respectively were apparent which is analogous with the known crystal structures of sterol TUDCA **3** and steroid hybrid, **4**.

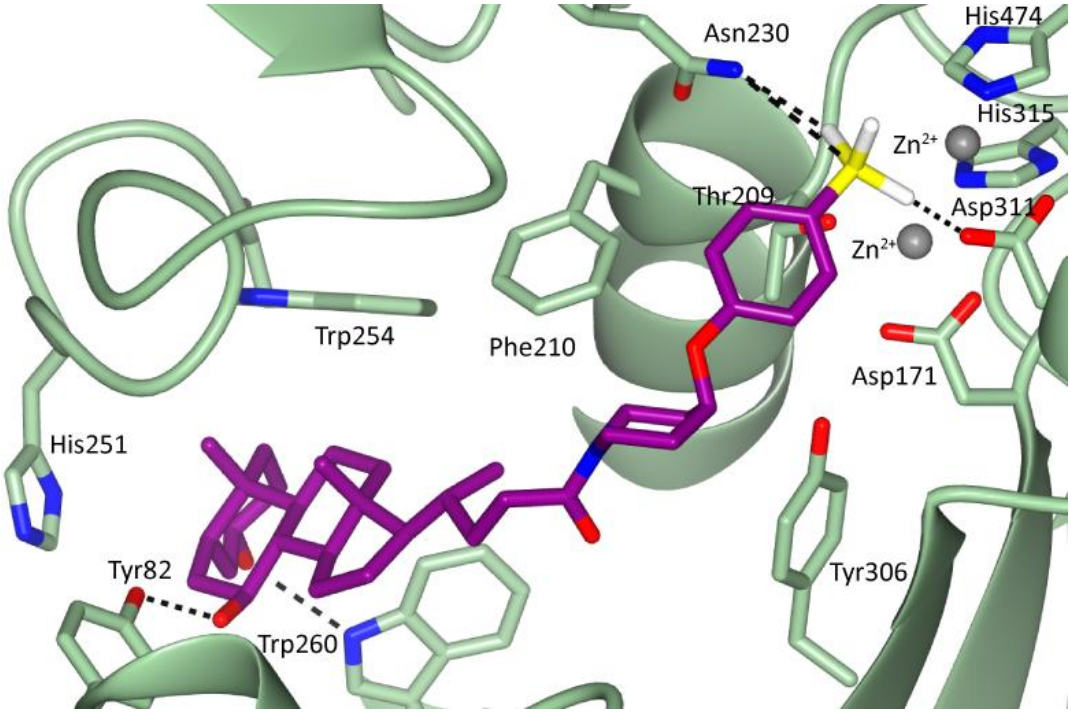


Figure 41 Crystal structure of ATX bound to **73**. Dashed lines indicate hydrogen bonds.

A two-dimensional depiction of the crystal structure is shown in **Figure 42** and summarises the key interactions made between the steroid and the tunnel and also the trifluoroborate warhead and the active site. Side and main chain hydrogen bonds are exemplified in green and blue arrows, respectively. Coordination to a metal atom is indicated in grey.

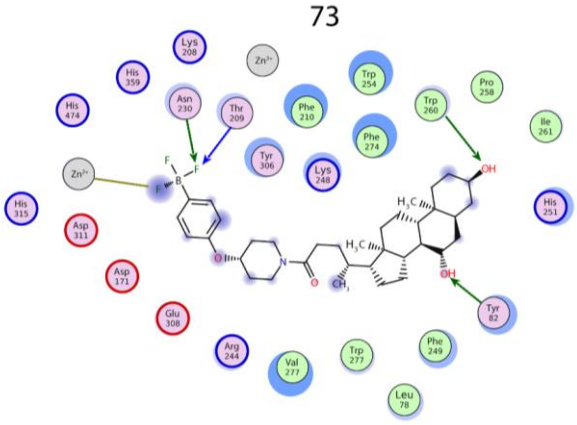


Figure 42 Two-dimensional depiction of the binding mode of **73**.

We believe compound **73** is relevant for two reasons. Firstly, it depicts a novel "Type V" binding mode for ATX inhibition and secondly, we believe that it is the first reported example of a trifluoroborate motif functioning as a zinc binding group, which possibly affords new opportunities for warhead design going forward.

*3.5.2 Co-crystallisation of **74***

Co-crystallisation of **74** with ATX resulted in generation of a crystal structure to 2.1 Å resolution. Modelling and refinement again generated an excellent compound fit for the observed electron density (**Figure 43**).

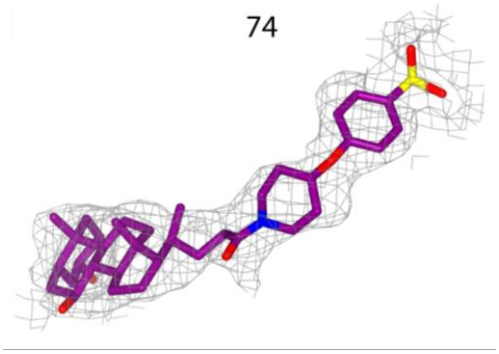


Figure 43 Fit of **74** to electron density.

Boronic acids as covalent modifiers of a nucleophilic residue in a protein have been more widely represented in the literature, e.g. HA155, so certain interactions were more anticipated. Gratifyingly, the crystal structure of **74** bound to ATX mirrored that of HA155,[153] which confirmed our success in identifying a compound which had optimal length and vector (**Figure 44**). The location of the boronic acid has facilitated the formation of a reversible covalent bond with the γ-OH group of Thr209 which is of interest as it is postulated that this reversibility is a key quality for success in ATX inhibition. One hydroxyl present on the boronic acid forms hydrogen bonding interactions with Asp171 and Asp311 in the active site which are supplemented with coordination to the proximal zinc ion. The alternative hydroxyl interacts via hydrogen bonding with the main chain of Thr209 in addition to a neighbouring water molecule. The compact network of bonding present reflects the high potency of **74** and explains its effect in blocking the active site.

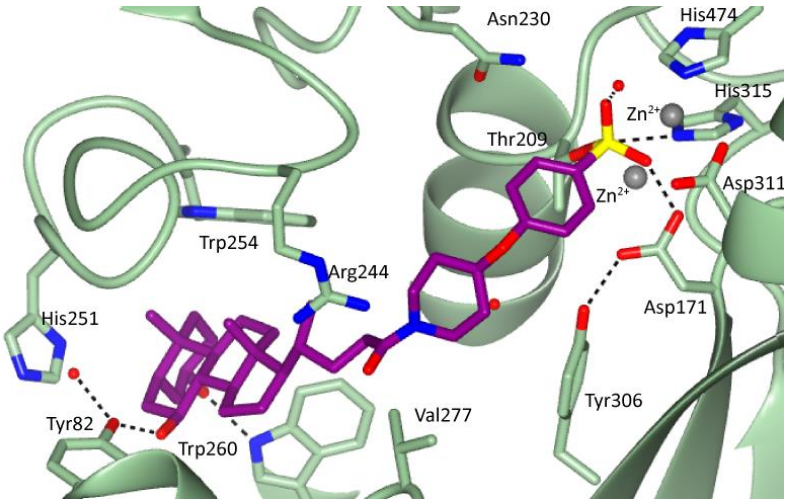


Figure 44 Crystal structure of ATX bound to **74**. Dashed lines indicate hydrogen bonds.

As observed with compound **73**, the steroid portion of compound **74** behaves in an identical fashion with key hydrogen bonds present between the pendant hydroxyls and Trp260 and Tyr82. Interestingly, compound binding in both instances was accommodated by a number of hydrophobic residues in the tunnel, for example, Leu78, Phe210, Leu243, Phe249, Trp254, Trp260, Phe274 and Trp275.

As shown for **73**, a two-dimensional depiction of the crystal structure for **74** has also been generated (**Figure 45**). Key interactions in all regions are specified with side and main chain hydrogen bonds indicated in green and blue arrows, respectively, in addition to hydrogen bonding with water molecules in gold. Coordination to the metal atom is highlighted in grey and the covalent bond between **74** and Thr209 is represented in solid purple. An appreciation can be given to the complex bonding network exhibited between the boronic acid warhead and active site region of ATX.

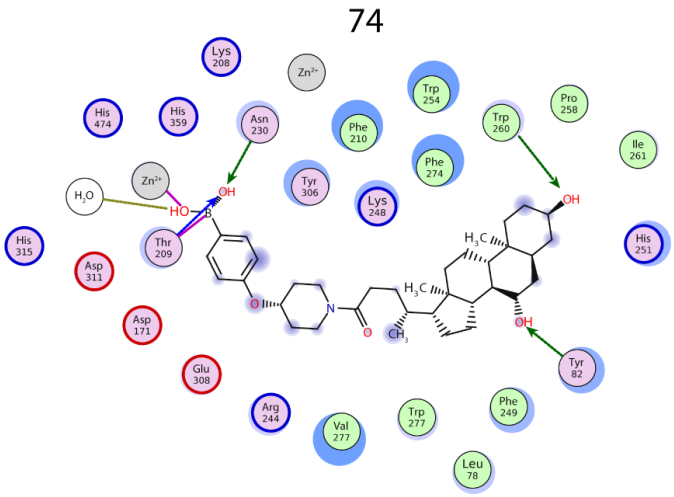


Figure 45 Two-dimensional depiction of the binding mode of **74**.

Taking all of this into account, in addition to the complementarity in binding that we observe between our compounds and their respective parent Type I and Type III/IV binding modes, we consider compounds **73** and **74** as the first examples of Type V inhibitors – hybrid compounds which interact in the tunnel and active site of ATX.

*3.5.3 Co-crystallisation of **72***

A crystal structure of compound **72** was also generated which substantiates our Type V binding hypothesis (**Figure 46**). Compound **72** contains the pyrrolidine core in addition to the boronic acid motif and again, a covalent bond is observed between the boronic acid and the catalytic threonine residue in the active site. A notable observation is the loss of an interaction with one of the hydroxyls in the tunnel region – this could be due to the small difference in trajectory between the pyrrolidine and piperidine spacer units.

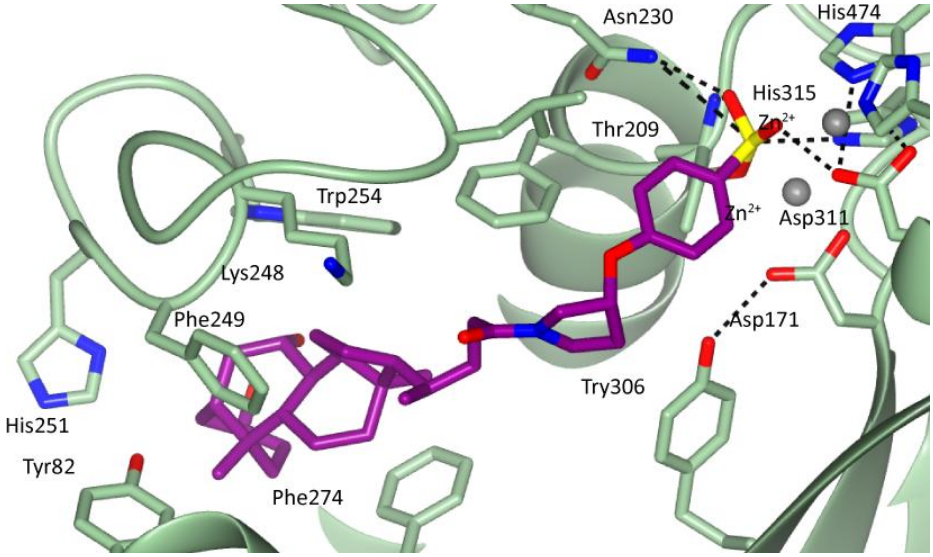


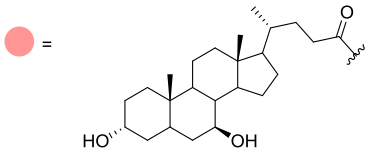
Figure 46 Crystal structure of ATX bound to **72**. Dashed lines indicate hydrogen bonds.

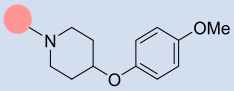
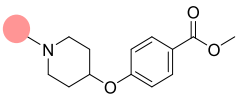
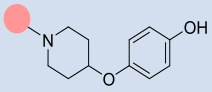
3.6 Extension of SAR Study

3.6.1 Delineation around Alternative Warhead Analogues

Following the success of compound **73** and **74,** we were keen to further assess the parameters of tolerated warheads in this new binding mode. A small assortment of related analogues were assessed in order to establish control compounds without a boron-containing warhead (**Table 16**). Exchanging the boronic acid for a methoxy **80** or ester **78** functionality resulted in negligible activity in the hydrolysis assay suggesting that affinity for the active site predominantly favours hydrogen bond donation from the warhead. Interestingly, unmasking of methoxy **80** to phenol **82** showed no improvement in inhibition of ATX which also suggests that at least two hydrogen bond donors are required for optimal activity.

Table 16 Evaluation of alternative warheads



Compound Number	Structure	IC_{50} (µM)	cLogP
80		> 10	5.5
78		> 10	5.7
82		> 10	4.9

Additionally, the lack of activity of phenol **82** is a significant result in so far as it acted to allay any queries over the possibility of boronic acid oxidation under assay conditions.

We were certain that the activity of the boronic acid analogue **74** was not due to any manipulation of the functionality under assay conditions, as this was not observed with HA155 however, phenol **82** (**Figure 47**) serves as definitive proof that deviation from the boronic acid changes the activity from low nanomolar to high micromolar through oxidation. This is also reflective of zinc binding group SAR analysis observed previously within our laboratories for the PF-8380 analogues with phenol **11** showing complete inactivity when tested against ATX.[142]

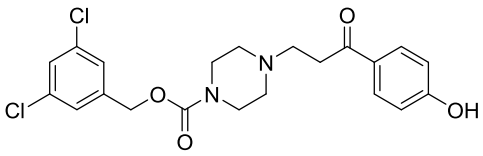


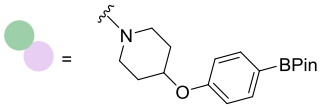
Figure 47 Structure of SAR-derived phenol analogue **11** from PF-8380 series.[142]

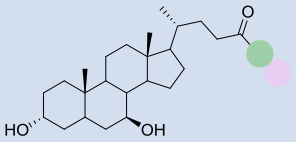
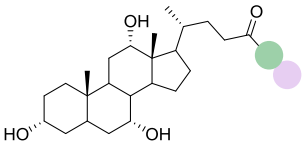
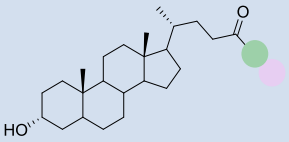
3.6.2 Delineation around Alternative Steroid Analogues

Investigation into the tolerance for different steroids in the tunnel was previously examined by NKI[139] however, based on the success of our compound library so far, and in order to more fully explore all three possible diversification points on the molecule, we were eager to establish if minor changes to the steroid construct would impart any difference in activity.

Therefore, we chose representative steroids, such as cholic acid (CA) **89** and lithocholic acid (LCA) **90,** which are structurally similar to UDCA and available as commodity chemicals. In both cases, no improvement in potency was observed in comparison to compound **70** which contains the UDCA steroid framework as shown in **Table 17**.

Table 17 Evaluation of alternative steroids



Compound Number	Structure	IC_{50} (μM)	cLogP
70		0.08	7.0
91		0.55	4.9
92		0.21	9.0

At this point, it is quite apparent the importance of 7-β hydroxylation on the B ring of the steroid construct. Both examples contain either the epimer (**91**) with an additional hydroxyl at the 12-position or no hydroxyl at the 7-position (**92**) which result in sub micromolar potency with regard to **70**.

Two caveats worth considering here are firstly, that despite having moderate potency, these compounds are more than likely within error of one another. Secondly, the algorithm used to calculate cLogP may not take into account small differences on a scaffold such as the steroid therefore, these values and throughout, must be analysed with caution. Whilst having been tested as the boronate ester, it would be worthwhile to advance these compounds to their respective boronic acid in order to obtain more comparable results.

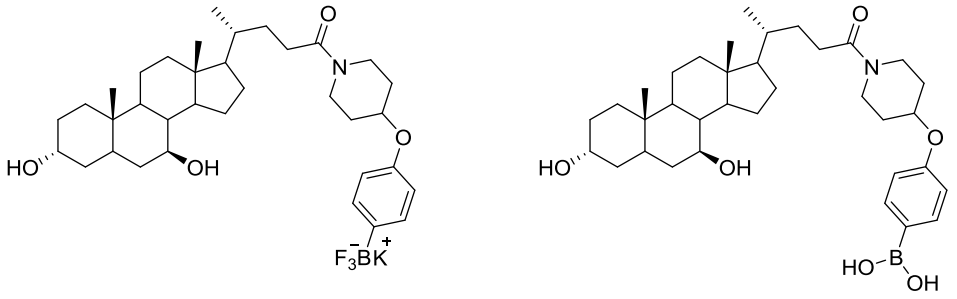
Therefore, it can be preliminarily concluded that UDCA represents the optimal steroid backbone for binding in the tunnel region. However, in terms of tuning physicochemical properties, it would be of interest to consider compounds bearing the LCA scaffold, or a

framework more lipophilic in nature, if interactions between the 7-OH and ATX are established to be non-essential for activity so long as the compound contains a warhead pharmacophore. In this instance, a small decrease in potency may be observed but the revenue in clogP may be beneficial. Compounds **91** and **92** provide latitude to vary the steroid though alternative frameworks for pursuing novelty.

3.7 Biological Evaluation of Exemplar Compounds

As alluded to previously, there is a clear disconnect between potent *in vitro* compounds and their progression through the clinic for diseases such as IPF and NAFLD. Against this background, we were keen to maximise our collaboration with NKI and their extensive experience of ATX in order to better understand the behaviour of our compounds in a biochemical and cellular based context.

With this in mind, we advanced two of our most active compounds, compounds **73** and **74** for further biochemical analysis (**Figure 48**).



73 **74**

Figure 48 Lead compounds **73** and **74** progressed for in-depth biochemical and cell-based analysis.

3.7.1 Hydrolysis Assays

ATX compounds are often triaged through unnatural substrate assays such as *p*NP-TMP and bis-*p*NPP however, in this instance, the LPC hydrolysis assay was carried out first as it does more accurately represent the hydrolysis conditions as it uses an endogenous substrate. Having stated this, we were keen to examine a range of assay formats in order to confirm our binding mode.

An interesting trend that was observed during the SAR screen of Type IV steroid-derived compounds was the appearance of no activity in the bis-*p*NPP assay yet full inhibition in the LPC hydrolysis assay. This was a common theme with such allosteric modulators which questioned the reliability of using this assay to triage compounds from non-orthosteric binding modes, and was a contributing factor in our decision to move directly to the LPC hydrolysis assay as our binding mode was completely without precedent.

Retrospective testing of compound **73** and **74** however, indicated that it may well be the case that these unnatural substrate assays are better suited to compounds that target the active site in some way through either Type I or V inhibition, as we observed full inhibition of ATX with both *p*NP-TMP and bis-*p*NPP substrates (**Figure 49**).

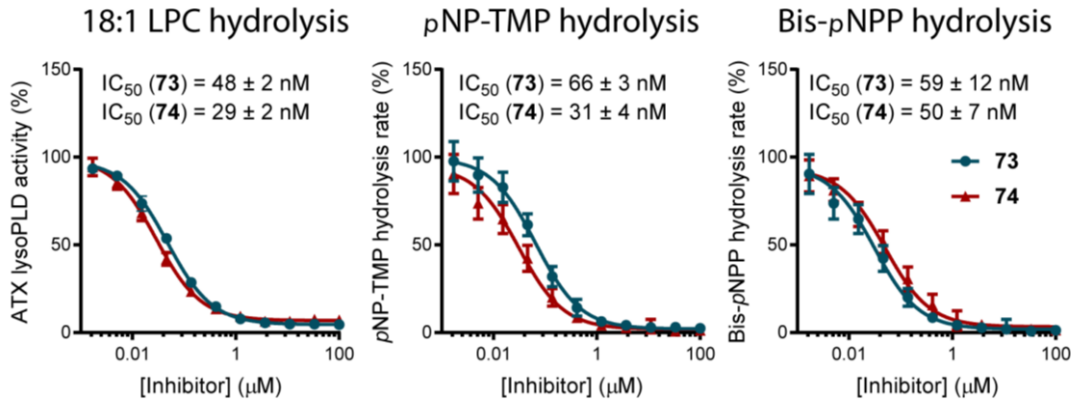


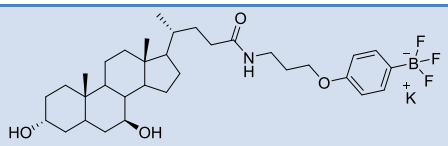
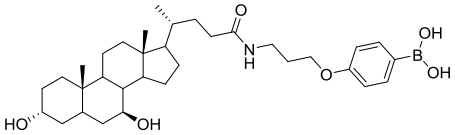
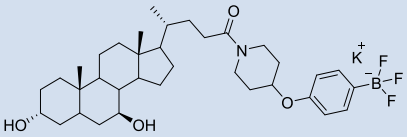
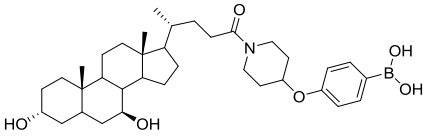
Figure 49 Measure of LPC hydrolysis using natural substrate LPC and two unnatural substrates, *p*NP-TMP and bis-*p*NPP, for compounds **73** and **74,** with respect to DMSO as the control.

3.7.2 Choline Oxidase Assay

An additional screen was implemented to supplement our biochemical armamentarium in order to establish the fate of oxidatively sensitive functionlaities. During the biological evalution of HA155, Albers *et al.* noted the potential for oxidation of aryl boronic acids by hydrogen peroxide,[153] which is generated (but also quickly consumed) *in situ* during the choline release cascade. We felt it was important to recapitulate this assay in order to ensure that our assay conditions were not influencing the integrity of the boronic acid motif.

Compounds with the two important boron-containing warheads were screened in the LPC hydrolysis assay, as shown in **Table 18**, to confirm that even the slightest deviation was not affected.

Table 18 Selection of compounds containing oxidatively sensitive functional groups**.**

Compound Number	Structure
55	
56	
73	
74	

Using DMSO as the control, it was apparent that no aspect of the assay conditions were detrimental to our compounds (**Figure 50**) which is analogous with the experimental observations reported by Albers for HA155 and its relations. This substantiated the lack of activity observed from phenol analogue **82** and alleviates any concerns regarding potential for oxidation of the boronic acid functionality.

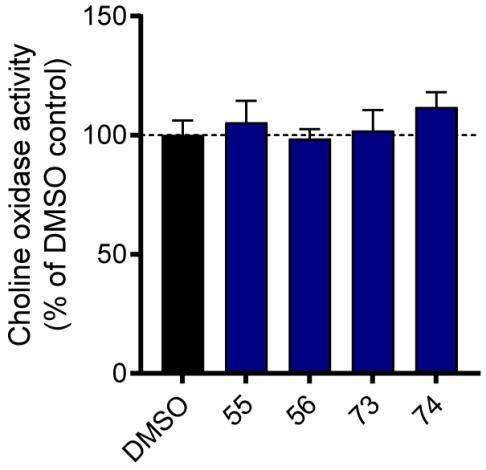


Figure 50 Internal control to assess the integrity of boron-containing warheads under assay conditions. 50 μM choline chloride was incubated with all reaction components in the absence of ATX and LPC.

3.7.3 Mechanism of Inhibition and Selectivity Assays

With our binding mode confirmed, we next sought to determine the mechanism by which our novel compounds were competing with LPA.

Enzyme inhibition is most commonly defined as the diminishment of enzyme activity in the presence of an external inhibitor and natural substrate (**Figure 51**). Compounds classed as enzyme inhibitors can be sub-divided into four distinct categories however for the purpose of this thesis, only competitive and non-competitive inhibition will be discussed.

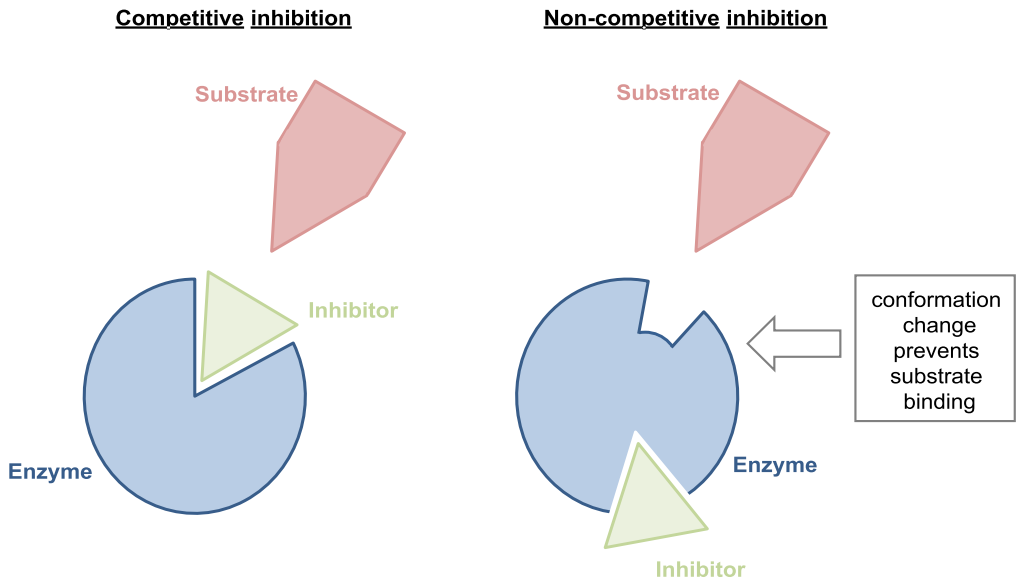


Figure 51 Cartoon representation of competitive vs non-competitive inhibition.

Competitive inhibitors, as expected, act in a competitive manner with the natural substrate and tend to interact with the orthosteric site of the enzyme, and in doing so, out-compete the natural substrate. Non-competitive inhibitors often bind allosterically and tend not to compete with the natural substrate for the active site. Uncompetitive inhibitors bind in a sequential fashion post enzyme –substrate complex formation. For the purpose of this thesis, only competitive and non-competitive inhibition will be discussed further.

To establish mechanism of inhibition, reaction kinetics can be followed over time and plotted using the Michaelis-Menten equation in linear regression form. This can then be manipulated into a Lineweaver-Burk plot where the parameters for values of K_M and V_{max} determine whether inhibition is competitive or non-competitive (**Figure 52**).

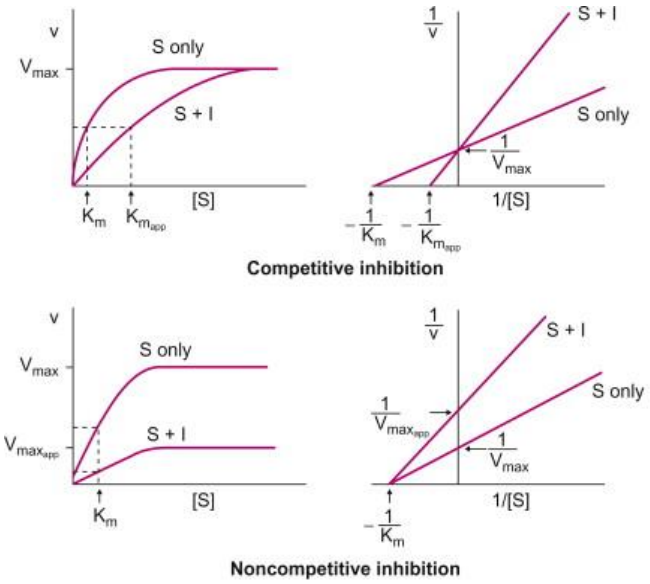


Figure 52 Representative examples of Michaelis-Menten (left) and Lineweaver-Burk (right) plots for competitive and non-competitive enzyme inhibition.

K_M is often referred to as the "Michaelis constant" and can be defined as the substrate concentration at which the enzyme reaches half V_{max}. V_{max} is the maximum rate of reaction, or, the rate of reaction when the enzyme is fully saturated with substrate. A lower K_M value is often characteristic of high affinity for the substrate.

Taking competitive inhibition for example, when inhibitor and compound compete for the active site, V_{max} remains unchanged however the apparent K_M is higher. Conversely, for non-competitive inhibitors, the V_{max} decreases but the K_M remains the same.

Modes of inhibition for orthosteric and allosteric inhibitors have been reported in the literature wherein "Type I" or orthosteric inhibitors, for example, HA155 and PF-8380, are established competitive inhibitors and "Type III" or allosteric modulators, such as TUDCA **3**,[146] are partial non-competitive inhibitors of ATX.

An interesting observation from the generation of Type IV inhibitors, which bind in the tunnel and hydrophobic pocket, was the switch from non-competitive inhibition when analysing the steroid alone to competitive when the steroid framework is tethered to a lipophilic tail (**Figures 53** and **54**).[140]

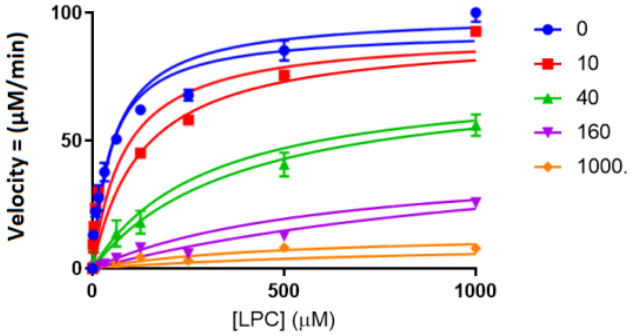


Figure 53 Michaelis-Menten graph for steroid hybrid **4.**

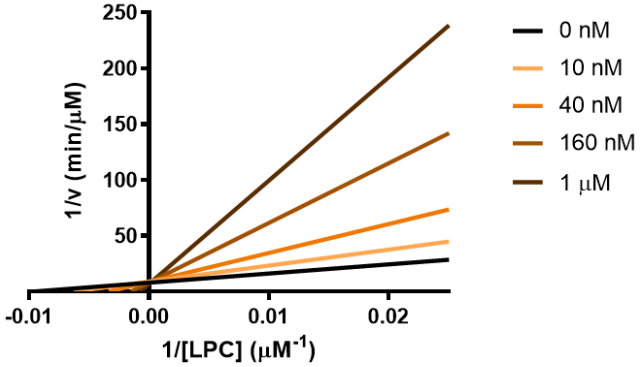


Figure 54 Lineweaver-Burk plot for steroid hybrid **4.**

As anticipated, analysis of inhibition of compound **73** and **74** confirmed competitive inhibition with K*i* values of 24 ± 4 nM and 9 ± 1 nM, respectively (**Figure 55**). We observe the same switch from non-competitive to competitive inhibition through tethering the warhead as was seen with the steroid in combination with the lipophilic tail for Type IV compounds such as steroid hybrid **4**.

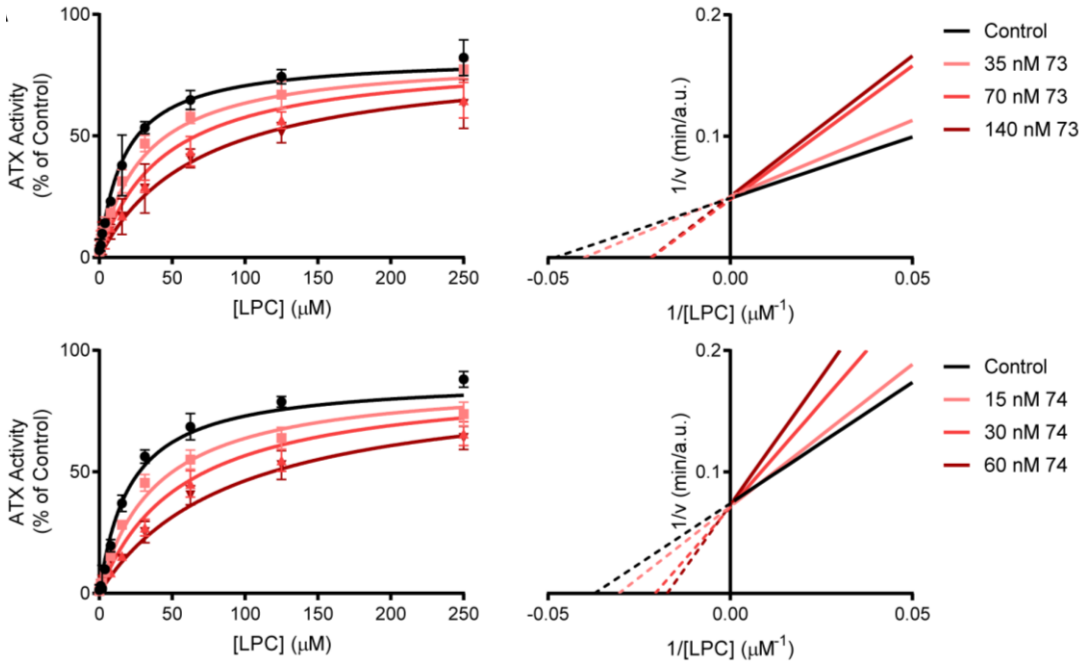


Figure 55 Michaelis-Menten and Lineweaver-Burke graphs for **73** and **74.**

Due to its complexity, many aspects of the LPC hydrolysis pathway remain poorly understood and although the variety in LPC species in circulation is well-known, the effect that the production of co-existing LPA species has on the multiple receptors remains unclear. Recent studies undertaken at NKI suggest that the ATX tunnel acts an allosteric site of LPA binding, and on extension, it was discovered that activation by LPA in the tunnel stabilises incoming LPC in the orthosteric site which encourages hydrolysis.[145]

This model also suggests that there could be interplay between shorter and longer LPC/LPA species at a physiological level and subsequent LPA activation of ATX. Certain pairings of LPC/LPA species resulted in a more "activated" ATX and it is hypothesised that acceleration of LPC hydrolysis of longer species could be accessed through allosteric binding of shorter LPA species (and *vice versa*).

From this, our collaborators have intimated that it could be possible that targeting different binding modes may have differentiated *in vitro* pharmacology in terms of receptor signalling by different variations of LPC/LPA, which in turn may have a direct affect the therapeutic conclusion.

Based on this, it was pertinent to examine the relevance of the inhibitory effect of our compounds on representative examples of LPC with differing acyl chain lengths, specifically, 14:0, 16:0, 18:1 and 22:0.

Both compounds **73** and **74** exerted a similar effect on 14:0, 16:0 and 18:1 LPC hydrolysis however a surprising result was obtained from 22:0 LPC hydrolysis (**Figure 56**). Inhibition of 22:0 only reached ~50% which potentially indicates that both compounds have an ability to act as partial inhibitors of longer acyl chains.

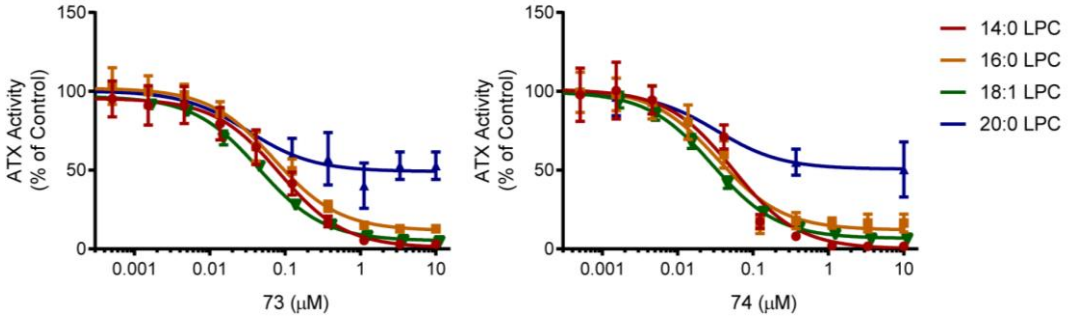


Figure 56 LPC hydrolysis of varying acyl chain lengths with 14:0, 16:0, 18:1 and 22:0 LPC.

3.7.4 LPA Activation Assay

Armed with the knowledge that our compounds were competitively inhibiting ATX, we were interested in augmenting this with some additional studies which were specifically aimed at confirming binding in tunnel region.

As mentioned in Section 3.6.5, NKI reported, through a series of biochemical studies involving both orthosteric and allosteric inhibitors that the ATX tunnel functions as an allosteric regulatory site.

When ATX is inhibited by an allosteric modulator such as TUDCA **3**, an increase in LPA activation is observed which implies that TUDCA can be displaced by LPA (**Figure 57**). As the concentrations of **3** increase, LPA-mediated activation is also more significant which would suggest that there is competition between the two for the tunnel region.

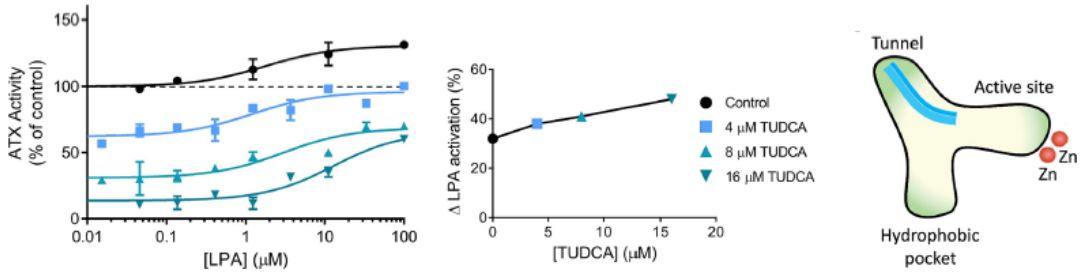


Figure 57 LPA-mediated activation of ATX in the presence of allosteric inhibitor, TUDCA, **3**.

These observations are contrasted in the case of orthosteric inhibitors, wherein partial inhibition of ATX with PF-8380 **1** results in similar LPA-mediated activation regardless of the presence or absence of **1** (**Figure 58**). From this, it could be deduced that binding of orthosteric inhibitors does not interfere with LPA-mediated activation.

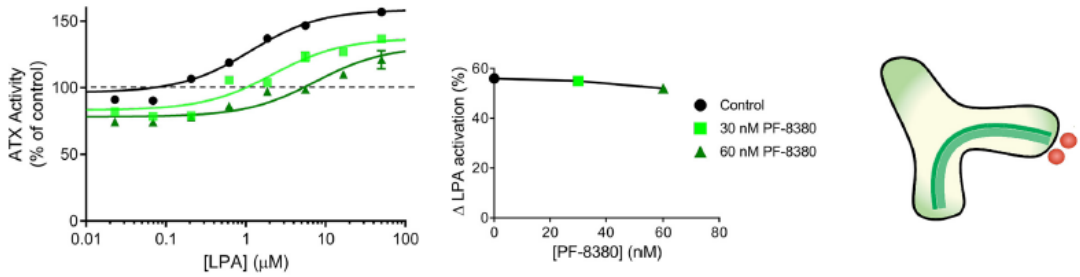


Figure 58 LPA-mediated activation of ATX in the presence of orthosteric inhibitor, PF-8380, **1**.

This effect is even more striking when considering mixed allosteric-orthosteric inhibitors such as steroid hybrid, **4**. At higher concentrations of inhibitor, LPA-mediated activation is completely shut down which implies that neither displacement nor activation by LPA is possible when ATX is inhibited by tunnel-hydrophobic pocket compounds (**Figure 59**).

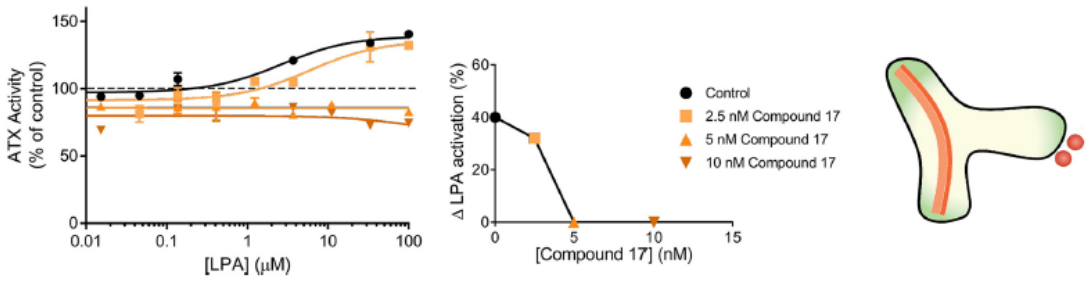


Figure 59 LPA-mediated activation of ATX in the presence of mixed allosteric-orthosteric inhibitor, **4**.

On considering these results, we were intrigued to assess the behaviour of our novel "Type V" binding mode in LPA-mediated activation. Drawing parallels between allosteric and orthosteric inhibitors, TUDCA **3** and PF-8380 **1**, respectively, we were anticipating results similar to that of steroid hybrid **4** for both compounds **73** and **74.**

As expected in both instances, we observe a significant decrease in LPA-mediated activation which indicates that LPA has negligible effect on ATX when it is inhibited by a tunnel-active site binder (**Figure 60**).

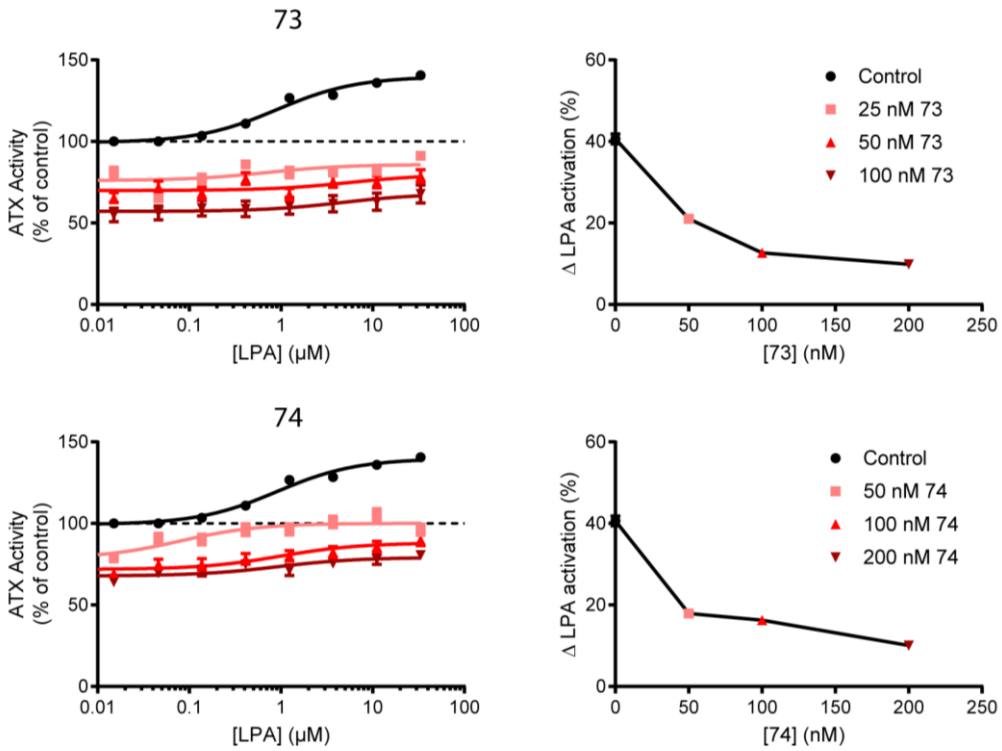


Figure 60 LPA-mediated activation of ATX in the presence of novel mixed allosteric-orthosteric inhibitors, **73** and **74.**

The multi-faceted binding capacity of Type IV and now, Type V, to prevent both orthosteric and allosteric binding of LPA is remarkable. Based on the clinical success of Type IV inhibitors, combined with striking results in emerging kinetic models such as LPA activation, a two-pronged allosteric-orthosteric inhibitor approach cannot be ruled out as a more successful method for ATX inhibition where the end point is clinical development.

3.7.5 Cell-based Assays

After establishing activity of these new compounds at an enzymatic level, we were next interested in obtaining a clearer view of the behaviour of this binding mode from a downstream signalling perspective. Due to the novelty associated with this Type V binding mode, it was of interest to develop a comparison between robust tool compounds with known cellular activity.

The success of these compounds is based upon their ability to decrease LPA receptor activation therefore, in collaboration with Netherlands Cancer Institute, we investigated the phosphorylation of Protein kinase B (AKT) and Extracellular Signal-Regulated Kinase (ERK) proteins, a fundamental ATX-mediated process initiated by LPA_1. The cascade is instigated by activation of LPA_1 in turn driving $G_{\alpha i}$ and $G_{\alpha 12/13}$ to activate Phosphatidylinositol-3-kinase (PI3K) and RhoA pathways, which are direct contributors to cellular proliferation, migration and apoptosis (**Figure 61**).[156]

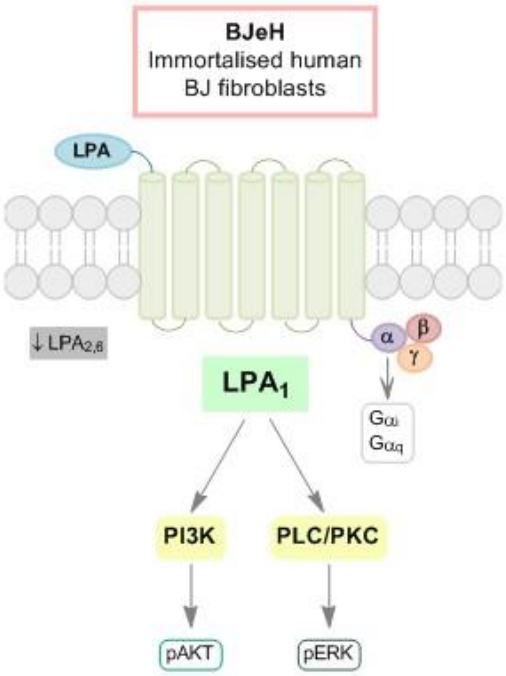


Figure 61 LPA_1 mediated activation of phosphorylated AKT and ERK proteins**.**

Through Western blot analysis using human transformed skin fibroblasts (BJeH), we observed the ability of our compounds to block phosphorylation of AKT through ATX inhibition. Through cell stimulation with either uninhibited ATX (control) or compound-bound ATX, in both cases, we observe a distinct decrease in phosphorylation of AKT on increasing inhibitor concentration of **73** and **74** compared with the control (**Figure 62**).

Dose-response curves were prepared in triplicate generating EC_{50} values of ~100 nM and ~60 nM for compounds **73** and **74**, respectively. This correlates to a two-fold increase with respect to their *in vitro* biochemical potencies for LPC hydrolysis. The consistency between these values provides confidence that the enzymatic effect is translated downstream, which suggest that compounds from this novel binding mode can significantly impact on LPA-mediated downstream signalling.

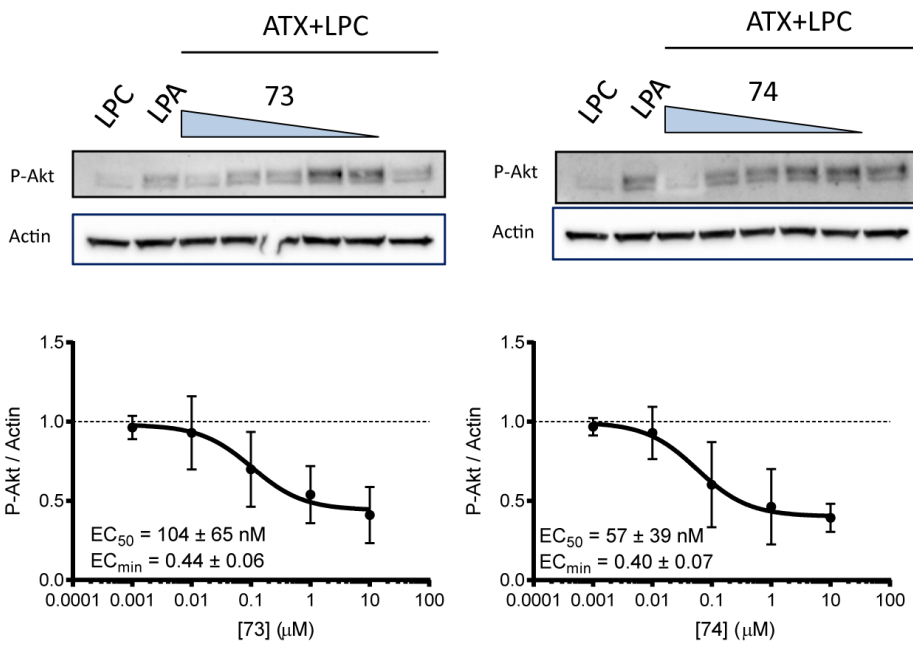


Figure 62 Western blot analysis for compounds **73** and **74** with respective dose-response curves plotted as a mean of independent triplicate experiments.

Both orthosteric SAR-derived analogues of PF-8380 and tunnel-pocket hybrids significantly attenuate the phosphorylation of AKT, therefore we were pleased to observe similar levels of activity with our novel compounds.

A caveat in the use of fibroblasts is their co-expression of LPA_1 and LPA_6 in almost equal quantities therefore it can be particularly difficult to differentiate which receptor is being activated by a novel compound. There is somewhat contradictory evidence in the literature as to which G_α subunits couple to LPA_6, therefore, we exploited a known active $LPA_{1/2/3}$ antagonist Ki16425[157] (**Figure 63**) in order to confirm LPA_1 activation and exclude any effect of LPA_6.

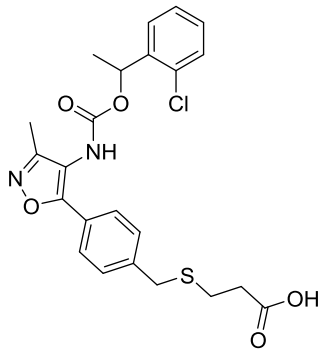


Figure 63 Structure of LPA_1 antagonist, Ki16425.

As expected, employing Ki16425 resulted in the deactivation of LPA_1 and subsequent downstream responses, which are comparable to the effects observed by compounds **73** and **74** (**Figure 64**).

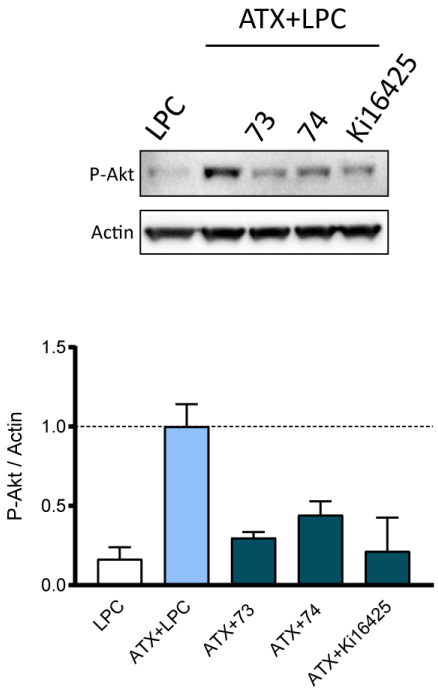


Figure 64 Western blot analysis for $LPA_{1/2/3}$ antagonist Ki16425 compared to compounds **73** and **74** with respective dose-response curves plotted as a mean of independent triplicate experiments.

3.7.6 Phenotypic Assay

Finally, we were keen to use a robust phenotypic study in order to bridge the gap between standard cell-based analysis and *in vivo* testing. After several discussions, we believed the most effective method to pursue this was through a cell migration approach.

Although our therapeutic targets of interest are fibroproliferative diseases such as NAFLD and IPF, there are significant similarities in their underlying pathophysiology to that of cancer. In addition, cellular studies involving immortalised cancer cell lines are better established in comparison to those involving fibroblasts. Due to their fundamental structure, fibroblasts can be particularly difficult to culture and maintain therefore; we believed that using an immortalised cancer cell-line with a known connection to the ATX-LPA signalling pathway, such as triple negative human adenocarcinoma MDA-MB-231 cells, would make a substantial starting point for our investigation into disease-relevant phenotypic studies.

Interruption of the cell migration process is one of the fundamental pathways associated with fibrosis and cancer. In a healthy individual, cell migration is managed and kept under control; however in patients suffering from proliferative diseases, such as fibrosis or cancer, this process becomes dysregulated leading to the uncontrolled migration and subsequent accumulation of cells which forms the basis for tumour formation/fibrous scar build-up. Therefore, in the knowledge that these compounds can block downstream phosphorylation, we were eager to further characterise the mechanism of action and determine if our compounds were capable of preventing cell migration.

As described previously, activation of LPA_1 initiates signal transduction through the PI3K pathway which results in cell migration and proliferation (**Figure 65**). Studies have reported that the PI3K/ERK pathway is relevant to breast cancer cell migration[156] therefore we elected to use these cells to demonstrate the effect of our compounds and to align with our Western blot analysis in Section 3.6.7.

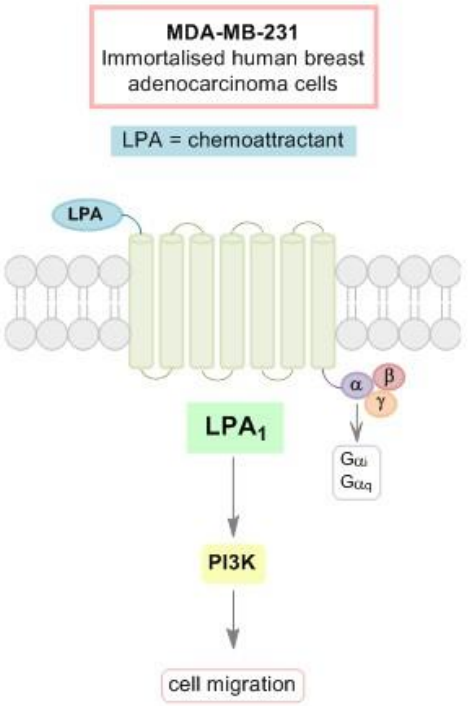


Figure 65 Representation of signalling cascade in cell migration process.

Again through collaboration with NKI, we intended to demonstrate the anti-migratory properties of this emerging series of compounds using immortalised cancer cells (MDA-MB-231 – triple negative breast cancer cells). Most tumour cells have a magnified migratory response when subjected to LPA which is described as positive chemotaxis or chemoattraction.

We employed a Boyden chamber technique which allowed for quantification of migratory cells which traversed through a fibronectin-coated filter. The purpose of the fibronectin coating is to represent the extracellular matrix – a key driver of fibrosis. To summarise, two million cells are loaded onto the upper chamber and allowed to migrate through the fibronectin-coated filter towards the chemotactic solutions present in the lower chambers over the course of several hours. Any non-migrated cells are removed from one side of the filter and the remaining cells on the underside of the filter are stained. A basic representation is exemplified in **Figure 66** with an example of the visual output on the filter post-staining.

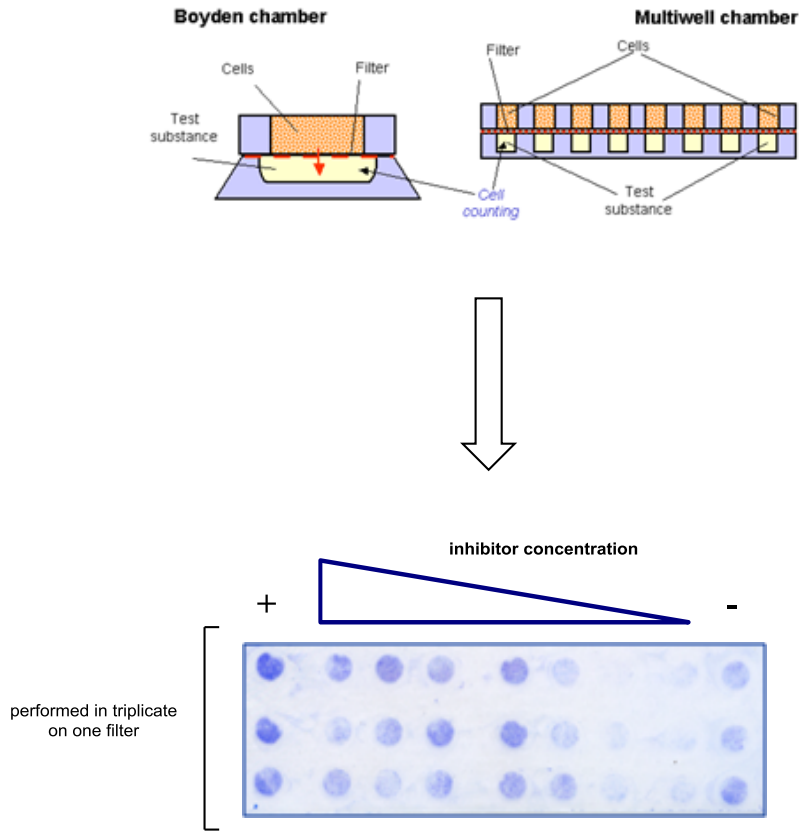


Figure 66 Basic representation of Boyden chamber technique and example densitometry output.

Due to the significant complexity associated with the technical aspects of this assay, we decided to nominate **74** alone as the representative compound. As expected, **74** efficiently reduces cell migration by ~50% in comparison to uninhibited ATX which acts as the control (**Figure 67**).

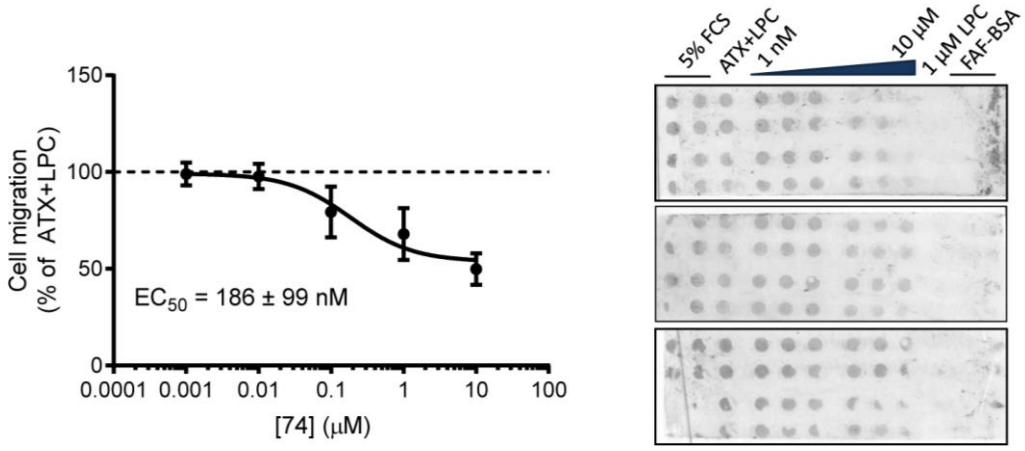


Figure 67 Cell migration studies using breast cancer cell-line MDA-MB-231 with **74** in a concentration range of 10-0.001 μM.

A dose-response curve was generated which provided an EC_{50} in the range of 200 nM and we believe this provides robust evidence that our compounds can efficiently attenuate LPA activation and chemotaxis. This cell line predominantly expresses LPA_1, in addition to lower levels of $LPA_{2/3}$, therefore it can be hypothesised that Type V inhibitors have the capacity to effectively negate LPA production and signalling through these receptors.

With the evident success of this preliminary phenotypic study, we anticipate that the results place compound **74** in a favourable position to benefit from further studies such as DMPK and further cell-based analysis to obtain a more rounded perspective of the future trajectory of these compounds.

4 Conclusions

The central aim of this research project was to investigate the development of a novel Type V inhibitor of ATX which binds in the tunnel and active site of the catalytic domain. A rational design approach was constructed through careful consideration and consolidation of previous potent SAR libraries from both orthosteric and allosteric inhibitors, which led to identification of our lead analogue.

In an attempt to move away from the benzoxazolone motif featured in orthosteric tool compound PF-8380, we took inspiration from HA155 and developed a series of boron-containing motifs as alternative warheads. Throughout our SAR campaign (**Figure 68**), we identified the optimal length and trajectory for activity with regard to ATX which led us to a piperidinol-based core in conjunction with a *para*-substituted aryl boronic acid **74**.

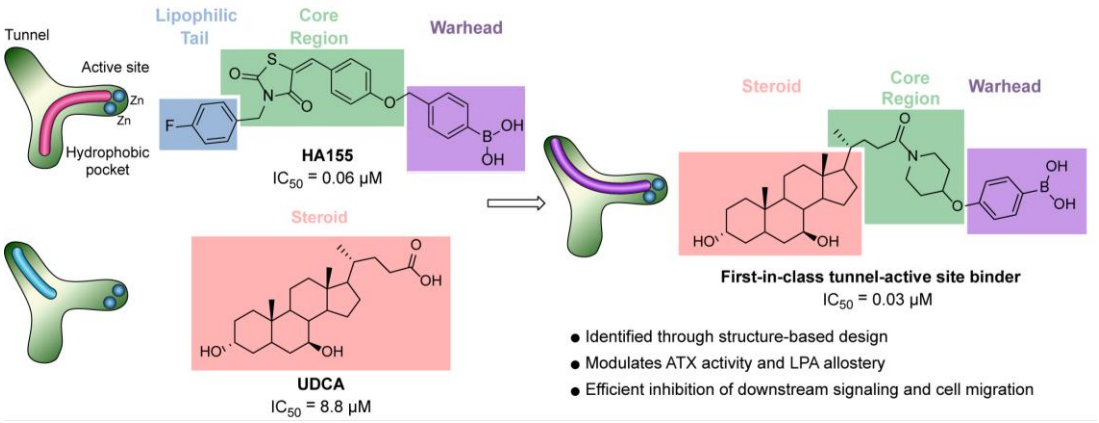


Figure 68 Overview of design and synthesis campaign towards Type V ATX inhibitors.

Co-crystallisation of our lead compounds in conjunction with ATX provided excellent quality structures which unequivocally confirm that these compounds are first-in-class "Type V" tunnel-active site binders of ATX. *In vitro* biochemical data has endorsed the competitive mechanism of inhibition anticipated for this binding mode, and comprehensive cellular studies have revealed the intrinsic ability of these novel compounds to attenuate LPA-mediated downstream signalling in a disease-relevant

context. Finally, we established the significant anti-migratory property of **74** in cancer cells, suggesting it could have utility in *in vivo* models if appropriate pharmacokinetics can be demonstrated.

Compound **74** exhibits excellent affinity for ATX in addition to improved physicochemical properties in comparison to progenitor ATX inhibitors currently in the clinic (**Figure 69**). We believe these data place **74** in a unique and favourable position to benefit from progression into more advanced studies.

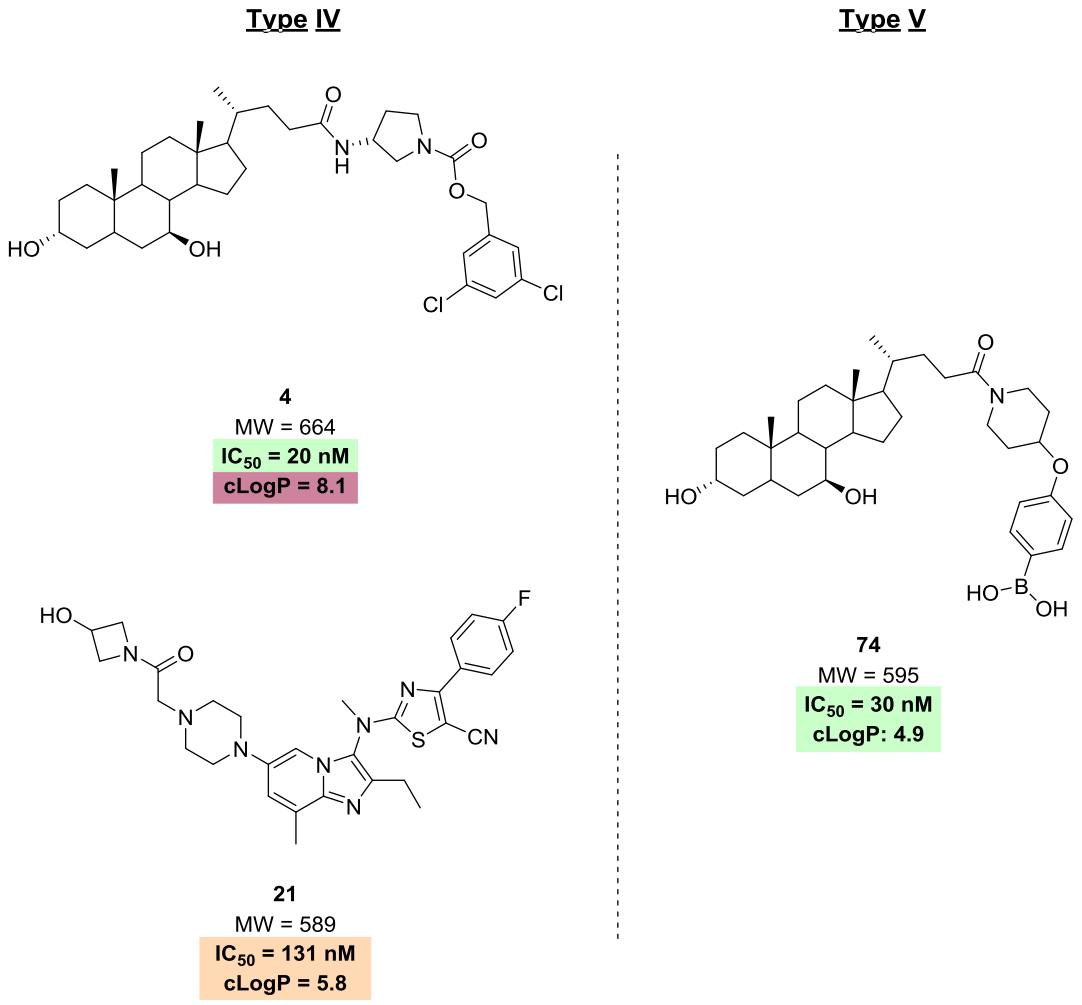


Figure 69 Comparison of physicochemical properties between two clinical "Type IV" inhibitors and novel "Type V" inhibitor **74.**

5 Future Work

Ultimately, the goal of this project was to establish a tool compound capable of exploring a new binding mode of ATX. Following the success of **74** in biochemical and cell-based studies, and in the knowledge that boronic acids are tolerated physiologically, the next immediate step would be to consider DMPK studies to determine physicochemical parameters, followed by testing in *in vivo* models to correlate with LPA signalling and more advanced phenotypic studies for fibrosis and cancer (**Figure 70**).

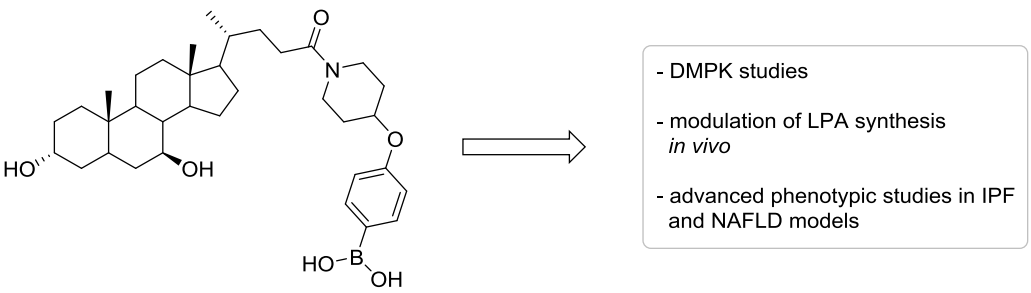


Figure 70 Future investigation for lead compound **74.**

In addition, despite the therapeutic use of boron gaining more traction,[158,159] it may be worthwhile to investigate other, more "drug-like" warheads, for example, a hydroxamic acid or tetrazole functionality **(Figure 71**), which maintain acceptable cLogP. Hydroxamic acids are reported zinc binders in both ATX[160] and histone deacetylase (HDAC)[161] targets, therefore this would be an interesting comparison. However, queries have been raised over their metabolic stability and potential toxicity issues for diseases other than cancer, for example, in the case of vorinostat and belinostat, which despite being approved as pharmaceutical agents, display poor pharmacokinetic and selectivity profiles.[162] Therefore the synthesis would be purely out of curiosity rather than to replace the boronic acid motif. In addition to boronic acids, tetrazoles are known bioisosteres of carboxylic acids therefore it would be of interest to compare their activity *in vitro*.

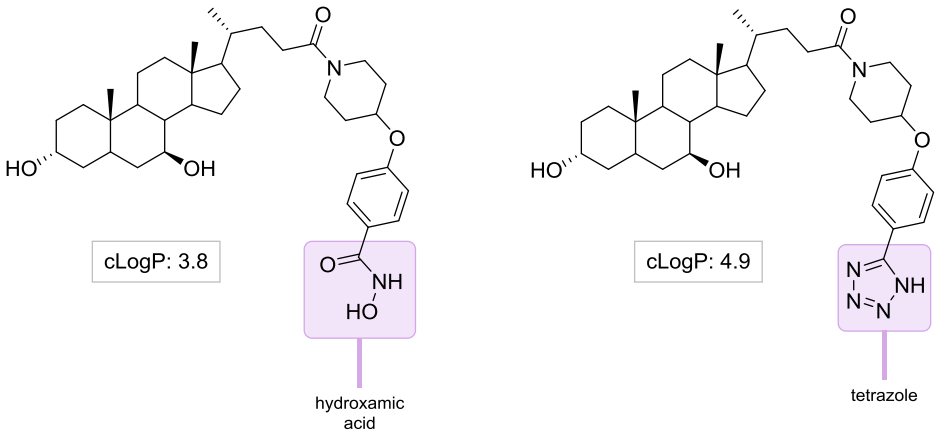
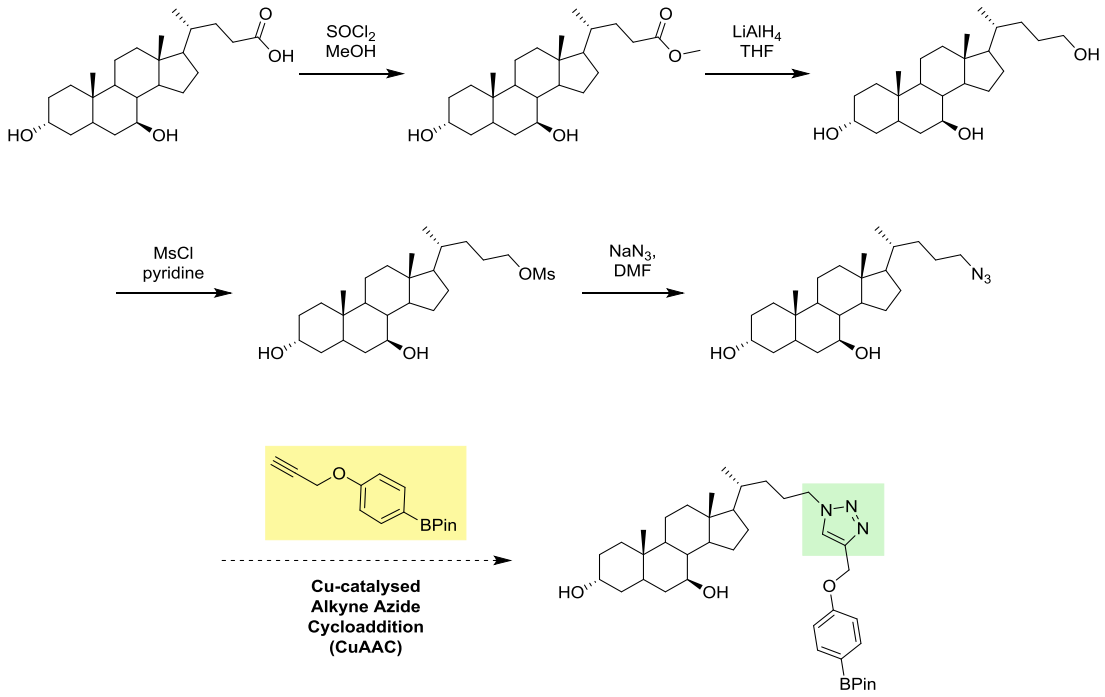


Figure 71 Potential diversification to non-boron containing warheads with steroidal scaffold based on literature precedent.

Furthermore, we have also been interested in introducing the concept of biorthogonality into the autotaxin projects. Biorthogonality has recently emerged as a powerful means to generate biomolecules in their native environment[163–165] under physiologically relevant conditions quickly with both high selectivity and functional group tolerance.

With this in mind, a related project is currently underway to investigate the possibility of accessing these compounds through the implementation of copper-catalysed alkyne-azide cycloaddition (CuAAC) chemistry (**Scheme 22**).[166] On the steroid component, functional group interconversion from the pendant carboxylic acid to an azide moiety could provide a building block which, in turn, could be decorated with a range of different alkyne handles in order to access a raft of 1,2,4-triazole-linked compounds with either a warhead or a lipophilic tail.



Scheme 22 Proposed route towards a biorthogonal synthesis of steroid-derived ATX inhibitors.

6 Experimental

6.1 General Techniques

All reagents and solvents were obtained from commercial suppliers and were used without further purification unless otherwise stated. Purification was carried out according to standard laboratory methods.

6.1.1 Purification of Solvents

All solvents used for dry reactions (PhH, THF, Et_2O) were obtained from a PureSolv SPS-400-5 solvent purification system. These solvents were transferred to and stored in a septum-sealed oven-dried flask over previously activated 3 Å molecular sieves and purged with and stored under nitrogen. CH_2Cl_2, Et_2O, EtOAc, MeOH, and petroleum ether 40-60^o for purification purposes were used as obtained from suppliers without further purification.

6.1.2 Experimental Details

Air-sensitive reactions were carried out using conventional glassware. The glassware was oven-dried at 150 °C and purged with N_2 before use. Purging refers to a vacuum/nitrogen-refilling procedure. Reactions were carried out at 0 °C using ice/water baths. Room temperature was generally *ca.* 18 °C. Reactions were carried out at elevated temperatures using a temperature-regulated hotplate/stirrer.

6.1.3 Purification of Products

Thin layer chromatography was carried out using Merck silica plates coated with fluorescent indicator UV254. These were analysed under 254 nm UV light or developed using potassium permanganate or vanillin solution. Flash chromatography was carried out using ZEOprep 60 HYD 40-63 μm silica gel or IST Isolute Flash silica cartridges.

6.1.4 Analysis of Products

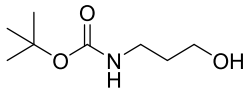
Fourier Transformed Infra-Red (FTIR) spectra were obtained on a Shimadzu IRAffinity-1 machine. ^1H, ^{13}C, ^{11}B and ^{19}F NMR spectra were obtained on a Bruker DPX 400 spectrometer at 400 and 101 MHz, respectively and a Bruker AVIIIHD500 at 500, 126,

160 and 471 MHz, respectively. Chemical shifts are reported in ppm and coupling constants are reported in Hz with $CDCl_3$ referenced at 7.26 (^1H) and 77.0 ppm (^{13}C), DMSO-d_6 referenced at 2.50 (^1H) and 39.5 (^{13}C) and MeOD-d_4 referenced at 3.30 (^1H) and 49.0 (^{13}C). Multiplicities are reported as observed and coupling constants have been corrected where appropriate. High resolution mass spectrometry (HRMS) data were obtained through analysis at the EPSRC UK National Mass Spectrometry Facility at Swansea University.

6.2 General Experimental Procedures

General Procedure A: *N*-Boc Protection

For example, for the preparation of *tert*-butyl (3-hydroxypropyl)carbamate **26**



To a round-bottomed flask was added 3-aminopropanol (525 μL, 6.87 mmol, 1.0 equiv), Boc_2O (1.5 g, 6.87 mmol, 1.0 equiv) and Et_3N (1.43 mL, 10.3 mmol, 1.5 equiv) in DCM (10 mL) and the reaction mixture was then stirred at room temperature for 3.5 hours. The reaction was quenched with 1M HCl (2 x 50 mL), and the organics were washed with DCM (40 mL) and washed with brine (80 mL). The organics were dried over Na_2SO_4, filtered, concentrated *in vacuo* to a residue that was purified by column chromatography on silica (50% EtOAc in petroleum ether 40-60) to yield **26** as a clear oil (976 mg, 81%).

^{1}H NMR ($CDCl_3$, 500 MHz): 4.87 (br. s, 1H), 3.63 (br. s, 2H), 3.25 (q, J = 5.0 Hz, 2H), 3.14 (br. s, 1H), 1.64 (quint, J = 5.0 Hz, 2H), 1.42 (s, 9H).

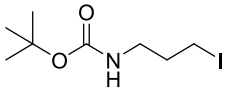
^{13}C NMR ($CDCl_3$, 126 MHz): 157.3, 79.7, 59.4, 37.1, 33.0, 28.5.

Consistent with previously reported spectral analysis.[167]

ν_{max} (neat): 3324, 2932, 2974, 2876, 1682, 1519 cm^{-1}.

General Procedure B: Appel

For example, for the preparation of *tert*-butyl (3-iodopropyl)carbamate **27**



To a round-bottomed was added *tert*-butyl (3-hydroxypropyl)carbamate **26** (976 mg, 5.57 mmol, 1.0 equiv) and triphenylphosphine (2.56 g, 9.75 mmol, 1.75 equiv) in DCM (20 mL) and cooled to 0 °C using an ice-bath. Pyridine (90 μL, 11.14 mmol, 2.0 equiv) was added dropwise followed by the dropwise addition of iodine (2.47 g, 9.75 mmol, 1.75 equiv) in THF (2 mL). The reaction mixture was allowed to warm to room temperature and left to stir for 4 hours. The reaction mixture was quenched with 1M HCl (2 x 50 mL), and the organics were washed with diethyl ether (40 mL) and washed with brine (80 mL). The organics were combined, dried with a hydrophobic frit and concentrated *in vacuo* to a residue that was purified by column chromatography on silica (60% EtOAc in petroleum ether 40-60) to afford title compound **27** as brown oil (1.24 g, 78%).

^{1}H NMR ($CDCl_3$, 500 MHz): 4.63 (br. s, 1H), 3.23-3.17 (m, 4H), 2.00 (quint, J = 5.0 Hz, 2H), 1.44 (s, 9H).

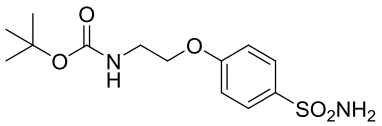
^{13}C NMR ($CDCl_3$, 126 MHz): 156.1, 79.6, 41.2, 33.5, 28.5, 3.23.

Consistent with previously reported spectral analysis.[168]

ν_{max} (neat): 3255, 2952, 1666, 1489, 1446 cm^{-1}.

General Procedure C: Alkylation Reaction

For example, for the preparation of *tert*-butyl (2-(4-sulfamoylphenoxy)ethyl)carbamate **32**



To a round-bottomed flask equipped with a reflux condenser was added 4-hydroxybenzenesulfonamide **31** (200 mg, 1.15 mmol, 1.1 equiv), K_2CO_3 (158 mg, 1.15 mmol, 1.1 equiv) and DMF (5 mL). The reaction mixture was heated to 80 °C and stirred for 20 minutes before addition of *tert*-butyl (2-bromoethyl)carbamate (235 mg, 1.04 mmol, 1.0 equiv). Reaction was stirred overnight at 80 °C. The reaction mixture was diluted with H_2O (80 mL) and the organics were extracted with EtOAc (2 x 30 mL) and washed with brine (80 mL). The organics were dried with a hydrophobic frit and concentrated *in vacuo* and purified by column chromatography on silica (2% MeOH/DCM) to afford the desired product **32** as a white solid (117 mg, 35%).

^{1}H NMR (DMSO-d_6, 500 MHz): 7.73 (d, J = 8.9 Hz, 2H), 7.19 (s, 2H), 7.07 (d, J = 8.9 Hz, 2H), 7.02 (t, J = 5.5 Hz, 1H), 4.03 (t, J = 5.5 Hz, 2H), 3.28-3.33 (m, coincident with solvent), 1.38 (s, 9H).

^{13}C NMR (DMSO-d_6, 126 MHz): 160.8, 155.7, 136.3, 127.6, 114.5, 77.8, 66.8, 28.2. 1 x C not observed.

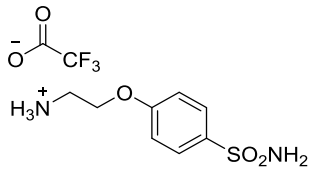
Consistent with previously reported spectral analysis.[167]

ν_{max} (neat): 3398, 3317, 3218, 3112, 2969, 1667, 1597, 1521, 1597, 1521, 1502 cm^{-1}.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_{13}H_{20}N_2O_5SNa$) requires 339.0985 *m/z*, found 339.0989 *m/z*.

General Procedure D: *N*-Boc Deprotection

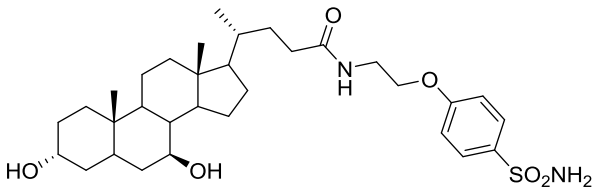
For example, for the preparation of 2-(4-sulfamoylphenoxy)ethan-1-aminium 2,2,2-trifluoroacetate **32a**



To a round-bottomed flask was added *tert*-butyl (2-(4-sulfamoylphenoxy)ethyl)carbamate (100 mg, 0.316 mmol) and a mixture of trifluoroacetic acid (370 μL, 2.21 mmol) and DCM (5 mL). The reaction was then stirred at room temperature for 4 hours. The reaction was concentrated *in vacuo* yielding a yellow gum which was telescoped through to the subsequent step as the trifluoroacetate salt without further purification.

General Procedure E: HATU Amidation

For example, for the preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(4-sulfamoylphenoxy)ethyl)pentanamide **33**



To a round-bottomed flask was added ursodeoxycholic acid (UDCA, 106 mg, 0.27 mmol, 1.0 equiv) in dimethylformamide (DMF, 2 mL), *N*,*N*-diisopropylethylamine (DIPEA, 237 µL, 1.36 mmol, 5.0 equiv) and then 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU, 114 mg, 0.30 mmol, 1.1 equiv). After stirring at room temperature for 15 minutes, crude intermediate 2-(4-sulfamoylphenoxy)ethan-1-aminium 2,2,2-trifluoroacetate **32a** (90 mg, 0.27 mmol, 1.0 equiv.) was added and stirred for 16 hours at room temperature. The reaction mixture was diluted with H_2O (30 mL) and the organics were extracted with EtOAc (2 x 20 mL) and washed with brine (80 mL). The organics were dried with a hydrophobic frit and concentrated *in vacuo* to a residue that was purified by column chromatography on silica (10% MeOH/DCM) to afford desired product **33** as an off-white solid (56 mg, 30% over two steps).

^{1}H NMR (DMSO-d_6, 500 MHz): 8.04 (t, *J* = 5.6 Hz, 1H), 7.73 (d, *J* = 8.9 Hz, 2H), 7.19 (s, 2H), 7.07 (d, *J* = 8.9 Hz, 2H), 4.42 (d, *J* = 4.6 Hz, 1H), 4.04 (t, *J* = 5.6 Hz, 2H), 3.86 (d, *J* = 6.8 Hz, 1H), 3.41 (q, *J* = 5.5 Hz, 2H), 3.30-3.24 (m, 1H, coincident with solvent peak), 3.17 (d, *J* = 5.3 Hz, 2H), 2.14-2.09 (m, 1H), 2.03-1.97 (m, 1H), 1.92-1.90 (m, 1H), 1.85-1.79 (m, 1H), 1.77-1.72 (m, 1H), 1.68-1.63 (m, 3H), 1.49-1.43 (m, 3H), 1.38-1.27 (m, 6H), 1.21-1.07 (m, 6H), 1.02-0.91 (m, 2H), 0.88-0.87 (m, 6H), 0.58 (s, 3H).

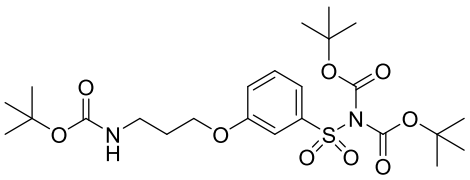
^{13}C NMR (DMSO-d_6, 126 MHz): 172.9, 160.8, 136.3, 127.7, 114.5, 69.7, 69.5, 66.8, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 38.0, 37.7, 37.3, 34.9, 34.8, 33.7, 32.3, 31.6, 30.2, 28.2, 26.7, 23.3, 20.8, 18.4, 12.0. 1 x C not observed.

ν_{max} (neat): 3634, 3566, 3396, 3313, 2928, 2863, 1649, 1630, 1500, 1455 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{32}H_{51}N_2O_6S$) requires 591.3462 *m/z*, found 591.3462 *m/z*.

General Procedure F: DIAD-Mediated Mitsunobu

For the preparation of *tert*-butyl (*tert*-butoxycarbonyl)((3-(3-((*tert*-butoxycarbonyl)amino)propoxy)phenyl)sulfonyl)carbamate **41**



To an oven-dried, round-bottomed flask purged with N_2 was added *tert*-butyl (3-iodopropyl)carbamate **27** (122 mg, 0.70 mmol, 1.3 equiv), *tert*-butyl (*tert*-butoxycarbonyl)((3-hydroxyphenyl)sulfonyl)carbamate **40** (200 mg, 0.53 mmol, 1.0 equiv) and triphenylphosphine (184 mg, 0.70 mmol, 1.3 equiv) in anhydrous THF (3 mL) and cooled to 0 °C using an ice bath. Diisopropyl azodicarboxylate (DIAD, 140 μL, 0.70 mmol, 1.3 equiv.) was added dropwise and the reaction was left to warm to room temperature and stirred for 16 hours. Reaction mixture was concentrated *in vacuo* to a residue that was purified by column chromatography on silica (10% EtOAc in petroleum ether 40-60) to yield title compound **41** as an off-white solid (118 mg, 42%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.76-7.74 (m, 1H), 7.72 (t, *J* = 1.9 Hz, 1H), 7.52 (t, *J* = 8.0 Hz, 1H), 7.44-7.41 (m, 1H), 4.99 (br. s, 1H), 3.89 (t, *J* = 6.7 Hz, 2H), 3.22-3.19 (m, 2H), 1.95-1.88 (m, 2H), 1.56 (s, 9H), 1.44 (s, 9H), 1.34 (s, 9H).

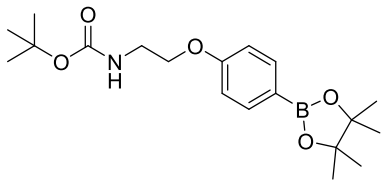
^{13}C NMR ($CDCl_3$, 126 MHz): 156.2, 151.2, 151.10, 151.05, 141.5, 129.9, 126.5, 125.0, 121.0, 85.2, 84.7, 79.3, 45.0, 37.4, 30.6, 28.6, 28.0, 27.8.

ν_{max} (liquid film): 3318, 2978, 2934, 1762, 1719, 1509, 1478 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{24}H_{39}N_2O_9S$) requires 531.2371 *m/z*, found 531.2365 *m/z*.

General Procedure G: CMBP-Mediated Mitsunobu

For example, for the preparation of *tert*-butyl (2-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)carbamate **44**



To a microwave vial was added 4-hydroxyphenyl boronic acid pinacol ester (200 mg, 0.90 mmol, 1.0 equiv.) and *N*-Boc ethanolamine (210 μL, 1.36 mmol, 1.3 equiv.). The vial was sealed with a microwave cap, purged under N_2 followed by addition of anhydrous toluene (4 mL). To the vial was added CMBP (471 μL, 1.09 mmol, 2.0 equiv.) and the reaction heated at 120 °C overnight. Reaction mixture was concentrated *in vacuo* to a residue that was purified by column chromatography on silica (10% EtOAc in hexane) then washed with saturated Na_2CO_3 (3 x 20 mL) to afford the desired product **44** as a clear oil (250 mg, 77%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.74 (d, *J* = 8.4 Hz, 2H), 6.88 (d, *J* = 8.4 Hz, 2H), 4.97 (br. s, 1H), 4.04 (t, *J* = 5.0 Hz, 2H), 3.54-3.53 (m, 2H), 1.45 (s, 9H), 1.33 (s, 12H).

^{13}C NMR ($CDCl_3$, 126 MHz): 161.3, 156.0, 136.7, 114.0, 83.7, 79.7, 67.2, 40.3, 28.5, 25.0. 1 x C bearing B not observed.

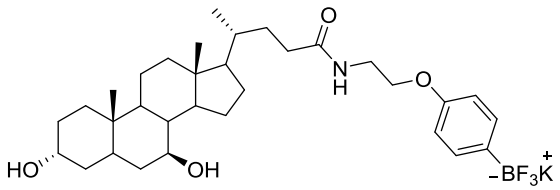
^{11}B NMR ($CDCl_3$, 160 MHz): 30.9.

ν_{max} (neat): 3360, 2977, 2932, 1700, 1606, 1513 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{19}H_{31}BNO_5$) requires 364.2290 *m/z*, found 364.2285 *m/z*.

General Procedure H: Trifluoroborate Formation

For example, for the preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(4-(trifluoro-l4-boraneyl)phenoxy)ethyl)pentanamide, potassium salt **53**



To a round-bottomed flask was added (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)pentanamide **45** (82 mg, 0.13 mmol, 1.0 equiv.) and MeOH (250 μL) followed by MeCN (250 μL). Potassium fluoride (30 mg, 0.51 mmol, 4.0 equiv) in H_2O (500 μL) was then added and the reaction mixture was stirred until complete dissolution of the boronic ester. *L*-(+)-tartaric acid (39 mg, 0.26 mmol, 2.05 equiv) was dissolved in THF (1 mL) and added dropwise to the rapidly (≈ 1000 RPM) stirring biphasic mixture over a period of 5 minutes, as a white precipitate formed. The reaction was stirred for 2 minutes, diluted with MeCN (3 mL) and stirred for a further 2 minutes before being diluted again with MeCN (1 mL) and filtered. The flask and filter were rinsed with further portions of acetonitrile (3 x 5 mL) and the combined filtrates were concentrated *in vacuo* to provide a mixture of the corresponding potassium organotrifluoroborate salt and pinacol. The filter cake was then washed with Et_2O (3 x 5 mL) to remove the excess pinacol leaving behind the product as a white solid which was telescoped through to the next step as the trifluoroborate salt without further purification.

^{1}H NMR (DMSO-d_6, 500 MHz): 7.98 (t, *J* = 5.5 Hz, 1H), 7.20 (d, *J* = 8.2 Hz, 2H), 6.65 (d, *J* = 8.2 Hz, 2H), 4.40 (d, *J* = 4.7 Hz, 1H), 3.88-3.84 (m, 3H), 3.37-3.34 (m, 2H, coincident with solvent), 2.15-2.08 (m, 1H), 2.03-1.97 (m, 1H), 1.97-1.94 (m, 1H), 1.87-

1.81 (m, 1H), 1.78-1.73 (m, 1H), 1.68-1.63 (m, 3H), 1.49-1.44 (m, 3H), 1.40-1.30 (m, 7H), 1.22-1.07 (m, 7H), 1.03-0.93 (m ,2H), 0.89-0.87 (m, 6H), 0.61 (s, 3H).

^{11}B NMR (DMSO-d_6, 160 MHz): 31.6.

^{13}C NMR (DMSO-d_6, 126 MHz): 172.8, 156.3, 132.3, 112.5, 69.7, 69.4, 66.0, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 38.3, 37.7, 37.3, 34.9, 34.8, 33.7, 32.3, 31.6, 30.2, 28.1, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C not observed, 1 x C bearing B not observed.

^{19}F NMR (DMSO-d_6, 471 MHz): -138.3.

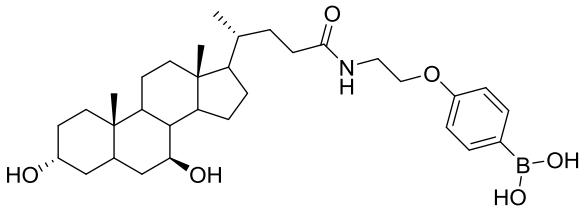
ν_{max} (neat): 3371, 2930, 2867, 1646, 1653, 1456 cm^{-1}.

HRMS: exact mass calculated for $[M\text{-}K]^-$ ($C_{32}H_{48}BF_3NO_4$) requires 578.3640 *m/z*, found 578.3639 *m/z.*

Experimental

General Procedure I: $B(OH)_2$ Formation

For example, for the preparation of (4-(2-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)ethoxy)phenyl)boronic acid **54**



To a round-bottomed flask containing (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(4-(trifluoro-l4-boraneyl)phenoxy)ethyl)pentanamide, potassium salt **53** (60 mg, 0.10 mmol, 1.0 equiv.) and excess SiO_2 under argon was added H_2O (3 mL). The reaction mixture was stirred at rt for 1 h before being was then filtered, and the filter cake was thoroughly washed with EtOAc. The organic phases were separated, and the aqueous phase was extracted with EtOAc (2x 10 mL). The combined organic phases were combined, washed with brine (10 mL), dried over $MgSO_4$, and concentrated *in vacuo* to afford title compound **54** as a white solid (33.6 mg, 60 %).

^{1}H NMR (DMSO-d_6, 400 MHz): 8.01 (app. br. s., 1H), 7.82-7.71 (m, 4H), 6.88-6.87 (m, 2H), 4.47-4.43 (m, 1H), 4.02-3.89 (m, 2H), 3.89-3.80 (m, 1H), 3.48-3.35 (m, coincident with solvent), 2.11-2.09 (m, 1H), 2.01-1.96 (m, 1H), 1.92-1.80 (m, 2H), 1.73-1.56 (m, 5H), 1.55-1.41 (m, 4H), 1.40-1.33 (m, 6H), 1.24-1.10 (m, 7H), 1.03-0.95 (m, 2H), 0.91-0.80 (m, 7H), 0.58 (s, 3H).

^{13}C NMR (DMSO-d_6, 101 MHz): 172.9, 160.1, 135.8, 113.4, 69.7, 69.5, 66.1, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 38.2, 37.7, 37.3, 34.9, 34.8, 33.8, 32.3, 31.6, 30.2, 28.1, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C not observed, 1 x C bearing B not observed.

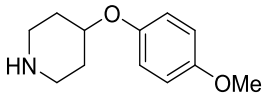
^{11}B NMR (MeOD, 160 MHz): 28.4.

ν_{max} (neat): 3332, 2932, 2857, 1636, 1605 cm^{-1}.

HRMS: exact mass calculated for $[M-H]^-$ ($C_{32}H_{49}BNO_6$) requires 554.3664 *m/z*, found 554.3662 *m/z*.

General Procedure J: Cbz Deprotection

For example, for the preparation of 4-(4-methoxyphenoxy)piperidine **79a**

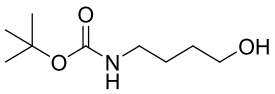


To a round-bottomed flask charged with benzyl 4-(4-methoxyphenoxy)piperidine-1-carboxylate (328 mg, 0.96 mmol, 1.0 equiv.) was added $Pd(OH)_2/C$ (136 mg, 20 mol %) and methanol (8 mL). The reaction was sparged with H_2 (balloon) for 1 minute, and stirred under an atmosphere of H_2 (balloon) for 5 hours. The reaction was filtered through celite, eluting MeOH. The organics were concentrated *in vacuo* and the crude material was carried through to the next step without further purification as an off-white solid. (191 mg, 96%).

6.3 Characterisation Data

6.3.1 Linker Synthesis

Preparation of *tert*-butyl (4-hydroxybutyl)carbamate **29**



Prepared according to General Procedure A, using 4-aminobutanol (0.63 mL, 6.87 mool, 1.0 equiv.), Boc_2O (1.50 g, 6.87 mmol, 1.0 equiv.) and Et_3N (1.44 mL, 10.31 mmol, 1.5 equiv.) in DCM (10 mL). The crude material was subjected to the purification outlined in General Procedure A (silica gel, 50 % EtOAc in petroleum ether 40-60) to afford desired product **29** as a clear oil (1.06 g, 82%).

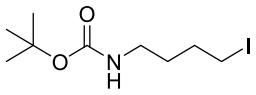
^{1}H NMR ($CDCl_3$, 500 MHz): 4.62 (br. s, 1H), 3.67-3.65 (m, 2H), 3.17-3.14 (m, 2H), 1.60-1.53 (m, 4H), 1.44 (s, 9H).

^{13}C NMR ($CDCl_3$, 126 MHz): 156.5, 78.2, 61.4, 40.4, 30.8, 29.2, 27.2.

Consistent with previously reported spectral analysis.[169]

ν_{max} (liquid film): 3342, 2936, 2871, 1685, 1526 cm^{-1}.

Preparation of *tert*-butyl (3-iodobutyl)carbamate **30**



Prepared according to General Procedure B, *tert*-butyl (3-hydroxypropyl)carbamate **29** (1.06 g, 5.66 mmol, 1.0 equiv.) and triphenylphosphine (2.59 g, 9.88 mmol, 1.75 equiv.) in THF (10 mL). Pyridine (0.90 mL, 11.22 mmol, 2.0 equiv.) was added dropwise followed by the dropwise addition of iodine (2.51 g, 9.88 mmol, 1.75 equiv.) in THF (2 mL). The crude material was subjected to the purification outlined in General Procedure B (silica gel, 10 % EtOAc in petroleum ether 40-60) to afford the desired product **30** as a brown oil (1.05 g, 63%).

^{1}H NMR ($CDCl_3$, 500 MHz): 4.52 (br. s, 1H), 3.20 (t, J = 6.9 Hz, 2H), 3.13-3.15 (m, 2H), 1.88-1.82 (m, 2H), 1.62-1.56 (m, 2H), 1.44 (s, 9H).

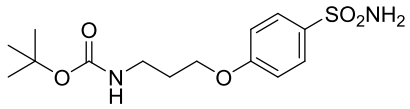
^{13}C NMR ($CDCl_3$, 126 MHz): 156.1, 79.4, 39.6, 31.2, 30.8, 28.6, 6.3.

Consistent with previously reported spectral analysis.[170]

ν_{max} (neat): 3416, 2971, 1750, 1653, 1405, 875 cm^{-1}.

6.3.2 Sulfonamide Analogues

Preparation of *tert*-butyl (3-(4-sulfamoylphenoxy)propyl)carbamate **34**



Prepared according to General Procedure A, using 4-hydroxybenzenesulfonamide (200 mg, 1.15 mmol, 1.1 equiv.), *tert*-butyl (3-iodopropyl)carbamate **27** (300 mg, 1.05 mmol, 1.0 equiv.) K_2CO_3 (158 mg, 1.15 mmol, 1.1 equiv.) and DMF (5 mL). The crude material was subjected to the purification outlined in General Procedure A (silica gel, % EtOAc in petroleum ether 40-60) to afford desired product **34** as an off-white solid (117 mg, 31%).

^{1}H NMR (DMSO-d_6, 500 MHz): 7.73 (d, J = 8.9 Hz, 2H), 7.18 (s, 2H), 7.06 (d, J = 8.9 Hz, 2H), 6.90-6.89 (m, 1H), 4.04 (t, J = 6.5 Hz, 2H), 3.08 (q, J = 6.5 Hz, 2H), 1.84 (quint., J = 6.5 Hz, 2H), 1.37 (s, 9H).

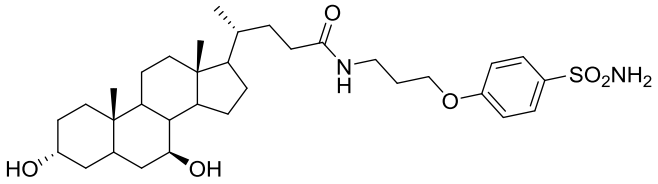
^{13}C NMR (DMSO-d_6, 101 MHz): 161.0, 155.6, 136.1, 127.6, 114.4, 77.5, 65.7, 36.8, 29.0, 28.2.

Consistent with previously reported spectral analysis.[167]

ν_{max} (neat): 3381, 3332, 3239, 2930, 1660, 1601, 1630, 1500 cm^{-1}.

HRMS: exact mass calculated for $[M+NH_4]^+$ ($C_{14}H_{26}N_3O_5S$) requires 348.1588 *m/z*, found 348.1590 *m/z*.

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(3-(4-sulfamoylphenoxy)propyl)pentanamide **35**



tert-butyl (3-(4-sulfamoylphenoxy)propyl)carbamate **34** (100 mg, 0.44 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using trifluoroacetic acid (230 µL, 3.05 mmol, 7.0 equiv.) in DCM (2 mL) to generate crude intermediate 3-(4-sulfamoylphenoxy)propan-1-aminium 2,2,2-trifluoroacetate. Compound **35** was prepared according to General Procedure E, using UDCA (85 mg, 0.22 mmol, 1.0 equiv.), 3-(4-sulfamoylphenoxy)propan-1-aminium 2,2,2-trifluoroacetate (75 mg, 0.22 mmol, 1.0 equiv.), DMF (2 mL), DIPEA (190 µL, 1.09 mmol, 5.0 equiv.) and HATU (90 mg, 0.24 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-10% MeOH in CH_2Cl_2) to afford desired product **35** as a clear solid (77 mg, 42% over two steps).

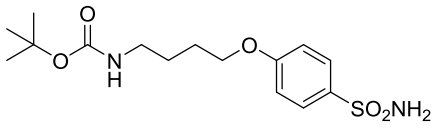
^{1}H NMR (DMSO-d_6, 500 MHz): 7.84 (t, J = 5.6 Hz, 1H), 7.74 (d, J = 8.9 Hz, 2H), 7.18 (s, 1H), 7.06 (d, J = 8.9 Hz, 2H), 4.42 (d, J = 4.6 Hz, 1H), 4.05 (t, J = 6.3 Hz, 2H), 3.85 (d, J = 6.8 Hz, 1H), 3.33-3.24 (m, 2H, coincident with solvent peak), 3.18 (app. q, J = 6.6 Hz, 2H), 2.11-2.06 (m, 1H), 1.99-1.91 (m, 2H), 1.88-1.81 (m, 3H), 1.78-1.72 (m, 1H), 1.69-1.63 (m, 3H), 1.49-1.26 (m, 11H), 1.20-0.91 (m, 9H), 0.89-0.87 (m, 6H), 0.60 (s, 3H).

^{13}C NMR (DMSO-d_6, 101 MHz): 172.6, 160.9, 136.1, 127.6, 114.4, 69.7, 69.5, 65.7, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 37.7, 37.3, 35.3, 34.9, 34.8, 33.7, 32.4, 31.7, 30.2, 28.8, 28.2, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C not observed.

ν_{max} (neat): 3381, 2929, 2864, 1633, 1597, 1547 cm^{-1}.

HRMS: exact mass calculated for $[M-H]^-$ ($C_{33}H_{51}N_2O_6S$) requires 603.3473 *m/z*, found 603.3473 *m/z*.

Preparation of *tert*-butyl (4-(4-sulfamoylphenoxy)butyl)carbamate **36**



Prepared according to General Procedure A, using 4-hydroxybenzenesulfonamide (191 mg, 1.10 mmol, 1.1 equiv.), *tert*-butyl (4-iodobutyl)carbamate **30** (300 mg, 1.00 mmol, 1.0 equiv.), K_2CO_3 (152 mg, 1.10 mmol 1.1 equiv.) and DMF (5 mL). The crude material was subjected to the purification outlined in General Procedure C (silica gel, 60% EtOAc in petroleum ether 40-60) to afford the desired product **36** as an off-white solid (148 mg, 39%).

^{1}H NMR (DMSO-d_6, 500 MHz): 7.73 (d, J = 8.9 Hz, 2H), 7.18 (s, 2H), 7.06 (d, J = 8.9 Hz, 2H), 6.85-6.83 (m, 1H), 4.04 (t, J = 6.6 Hz, 2H), 2.97 (app. q, J = 6.6 Hz, 2H), 1.74-1.68 (m, 2H), 1.55-1.49 (m, 2H), 1.37 (s, 9H).

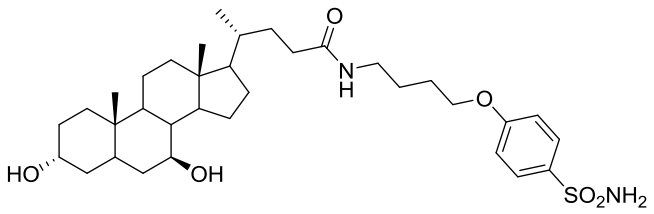
^{13}C NMR (DMSO-d_6, 101 MHz): 161.0, 155.6, 136.0, 127.6, 114.4, 77.4, 67.6, 28.2, 26.0, 25.9. 1 x C not observed.

Consistent with previously reported spectral analysis.[167]

ν_{max} (neat): 3363, 3281, 2952, 2926, 2865, 1682, 1599, 1526, 1498 cm^{-1}.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_{15}H_{24}N_2O_5SNa$) requires 367.1298 *m/z*, found 367.1299 *m/z*.

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(4-(4-sulfamoylphenoxy)butyl)pentanamide **37**



tert-Butyl (4-(4-sulfamoylphenoxy)butyl)carbamate **36** (150 mg, 0.44 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using trifluoroacetic acid (230 μL, 3.05 mmol, 7.0 equiv.) in DCM (2 mL) to generate crude intermediate 4-(4-sulfamoylphenoxy)butan-1-aminium 2,2,2-trifluoroacetate. Compound **37** was prepared according to General Procedure E, using UDCA (114 mg, 0.22 mmol, 1.0 equiv.), 4-(4-sulfamoylphenoxy)butan-1-aminium 2,2,2-trifluoroacetate (100 mg, 0.29 mmol, 1.0 equiv.), DMF (3 mL), DIPEA (250 μL, 1.46 mmol, 5.0 equiv.) and HATU (122 mg, 0.32 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-10% MeOH in DCM) to afford the desired product **37** as a clear solid (96 mg, 36% over two steps).

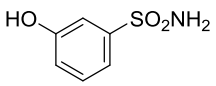
^{1}H NMR (DMSO-d_6, 500 MHz): 7.77 (t, J = 5.6 Hz, 1H), 7.73 (d, J = 8.9 Hz, 2H), 7.17 (s, 1H), 7.06 (d, J = 8.9 Hz, 2H), 4.42 (d, J = 4.6 Hz, 1H), 4.04 (t, J = 6.5 Hz, 2H), 3.85 (d, J = 6.8 Hz, 1H), 3.34-3.24 (m, coincident with solvent peak), 3.08 (app. q, J = 6.5 Hz, 2H), 2.11-2.05 (m, 1H), 1.99-1.92 (m, 2H), 1.86-1.81 (m, 1H), 1.78-1.63 (m, 6H), 1.55-1.44 (m, 5H), 1.41-1.26 (m, 7H), 1.21-1.03 (m, 6H), 1.03-0.97 (m, 1H), 0.89-0.87 (m, 6H), 0.60 (s, 3H).

^{13}C NMR (DMSO-d_6, 101 MHz): 172.4, 161.0, 136.0, 127.6, 114.4, 69.7, 69.4, 67.6, 55.9, 54.9, 43.1, 43.0, 42.1, 38.7, 37.9, 37.7, 37.3, 34.9, 34.8, 33.7, 32.4, 31.7, 30.2, 28.1, 26.7, 26.0, 25.7, 23.3, 20.8, 18.5, 12.0. 1 x C not observed.

ν_{max} (neat): 3293, 3078, 2927, 2862, 1644, 1549, 1501, 1452 cm^{-1}.

HRMS: exact mass calculated for $[M-H]^+$ ($C_{34}H_{53}N_2O_6S$) requires 617.3630 *m/z*, found 617.3627 *m/z*.

Preparation of 3-hydroxybenzenesulfonamide **39**



To a solution of 3-aminobenzenesulfonamide (300 mg, 1.73 mmol, 1.0 equiv.) in 30% sulfuric acid (3 mL H_2SO_4 in 7 mL H_2O), stirred at 0 °C in an ice-water bath was added dropwise a solution of sodium nitrite (155 mg) in water (2 mL). Reaction solution as stirred at 0 °C for 30 minutes before addition of water (2 mL). Solution was then stirred at 100 °C for 30 minutes. Reaction mixture was cooled to room temperature, partitioned between EtOAc and brine. Aqueous phase was washed with EtOAc and organics were combined and dried over sodium sulfate, filtered and concentrated *in vacuo* to yield **39** as a pale yellow solid (199 mg, 66%). The crude material was taken forward without further purification.

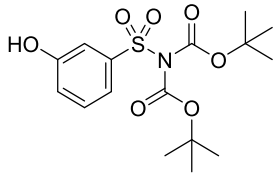
^{1}H NMR (DMSO-d_6, 500 MHz): 10.0 (br. s, 1H), 7.36 (t, J = 7.9 Hz, 1H), 7.27-7.22 (m, 4H), 6.97-6.95 (m, 1H).

^{13}C NMR (DMSO-d_6, 101 MHz): 157.5, 145.2, 130.0, 118.7, 116.1, 112.3.

Consistent with previously reported spectral analysis.[171]

ν_{max} (neat): 3316, 3241, 2815, 2693, 1594, 1547, 1484 cm^{-1}.

Preparation of *tert*-butyl (*tert*-butoxycarbonyl)((3-hydroxyphenyl)sulfonyl)carbamate **40**



A flask was charged with 3-hydroxybenzenesulfonamide **39** (1000 mg, 5.77 mmol, 1.0 equiv.), 4-dimethylaminopyridine (70 mg, 0.14 mmol, 0.1 equiv.), THF (20 mL), and triethylamine (1.6 mL, 11.54 mmol, 2.0 equiv.), then Boc_2O (3.8 g, 17.32 mmol, 3.0 equiv.) was added in one portion.[172] The resulting mixture was stirred at room temperature under ambient atmosphere for 24 hours, then washed with 2M HCl (2 x 20 mL) and brine (20 mL), dried over sodium sulfate and concentrated *in vacuo*. The crude material was subjected to purification (silica gel, 30% EtOAc in petroleum ether 40-60) and then washed with 4M NaOH (2 x 20 mL) to afford the desired product **40** as an off-white solid (275 mg, 13%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.90-7.88 (m, 2H), 7.85 (t, *J* = 6.9 Hz, 1H), 7.55 (t, *J* = 6.9 Hz, 1H), 7.47-7.45 (m, 1H), 7.25 (br. s. 1H), 1.56 (s, 9H), 1.40 (s, 9H).

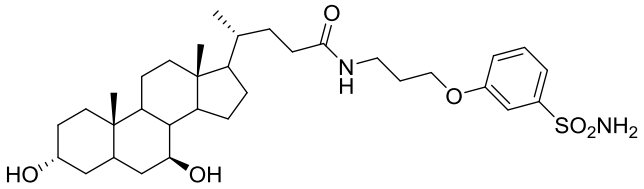
^{13}C NMR ($CDCl_3$, 101 MHz): 151.2, 151.2, 148.9, 140.3, 130.1, 127.0, 125.5, 121.4, 84.7, 84.6, 28.0, 27.8.

ν_{max} (liquid film): 3250, 2984, 1748 cm^{-1}.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_{16}H_{23}NO_7SNa$) requires 396.1087 *m/z*, found 396.1085 *m/z.*

Experimental

Preparation of **(**4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(3-(3-sulfamoylphenoxy)propyl)pentanamide **42**



tert-butyl (*tert*-butoxycarbonyl)((3-(3-((*tert*-butoxycarbonyl)amino)propoxy)phenyl)sulfonyl)carbamate **41** (400 mg, 0.75 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using trifluoroacetic acid (2 mL, 21.0 mmol, 28.0 equiv.) in DCM (2 mL) to generate crude intermediate 3-(3-sulfamoylphenoxy)propan-1-aminium 2,2,2-trifluoroacetate. Compound **42** was prepared according to General Procedure E, using UDCA (184 mg, 0.47 mmol, 1.0 equiv.), 3-(3-sulfamoylphenoxy)propan-1-aminium 2,2,2-trifluoroacetate (162 mg, 0.47 mmol, 1.0 equiv.), DMF (3 mL), DIPEA (246 μL, 1.41 mmol, 6.0 equiv.) and HATU (196 mg, 0.52 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure C (silica gel, 0-8% MeOH in DCM) to afford the desired product **42** as a white solid (107 mg, 24% over two steps).

^{1}H NMR (DMSO-d_6, 500 MHz): 7.71 (t, *J* = 5.6 Hz, 1H), 7.48 (t, *J* = 5.9 Hz, 1H), 7.37 (t, *J* = 5.6 Hz, 1H) 7.19-7.15 (m, 2H), 6.99 (dd, *J* = 8.1, 2.1 Hz, 1H), 4.42 (d, *J* = 4.6 Hz, 1H), 4.08 (q, *J* = 5.3 Hz, 1H), 3.85 (d, *J* = 6.8 Hz, 1H), 3.34-3.25 (m, 2H, coincident with solvent peak), 3.17 (d, *J* = 5.3 Hz, 2H), 2.99 (q, *J* = 6.6 Hz, 2H), 2.70 (q, *J* = 6.6 Hz, 2H), 2.06-2.01 (m, 1H), 1.94-1.81 (m, 3H), 1.77-1.59 (m, 4H), 1.52-1.26 (m, 13H), 1.19-1.07 (m, 6H), 1.02-0.91 (m, 2H), 0.87-0.86 (m, 6H), 0.60 (s, 3H).

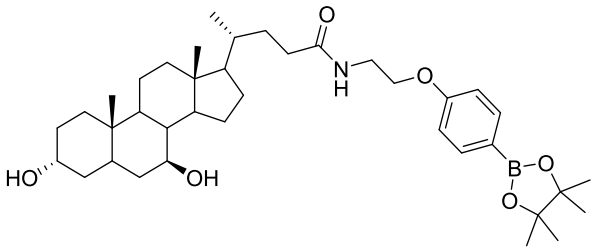
^{13}C NMR (DMSO-d_6, 101 MHz): 172.6, 157.7, 141.4, 130.2, 119.3, 116.9, 113.0, 69.7, 69.5, 55.9, 54.7, 48.6, 43.1, 42.2, 40.5, 38.7. 37.7. 37.3, 36.0, 34.9, 34.8, 33.8, 32.4, 31.7, 30.2, 29.4, 28.1, 26.7, 23.3, 20.8, 18.4, 12.0.

ν_{max} (neat): 2928, 2863, 1640, 1632, 1586, 1547, 1450 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{33}H_{53}N_2O_6S$) requires 605.3619 *m/z*, found 605.3611 *m/z*.

6.3.3 Linear Boronate Analogues

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10, 13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)pentanamide **45**



tert-Butyl (2-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)carbamate **44** (207 mg, 0.57 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using a 1:1 mixture of trifluoroacetic acid (1 mL) in DCM (1 mL) to generate crude intermediate **44a**. Compound **45** was prepared according to General Procedure E, using ursodeoxycholic acid (247 mg, 0.63 mmol, 1.0 equiv.), DIPEA (436 μL, 2.50 mmol, 4.0 equiv.) and HATU (262 mg, 0.69 mmol, 1.1 equiv) in DMF (5 mL). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-3% MeOH in DCM) to afford the desired product **45** as a clear solid (131 mg, 33% over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.75 (d, *J* = 8.3 Hz, 2H), 6.88 (d, *J* = 8.3 Hz, 2H), 5.88 (br.s, 1H), 4.09-4.06 (m, 2H), 3.70-3.65 (m, 2H), 3.60-3.55 (m, 2H), 2.31-2.22 (m, 1H), 2.15-2.05 (m, 1H), 2.01-1.95 (m, 1H), 1.93-1.85 (m, 1H), 1.81-1.76 (m, 4H), 1.68-1.66 (m, 2H), 1.61-1.56 (m, 6H), 1.51-1.47 (m, 2H), 1.44-1.39 (m, 4H), 1.33 (s, 12H), 1.28-1.23 (m, 4H), 1.21-1.14 (m, 1H), 1.11-0.99 (m, 3H), 0.94-0.92 (m, 6H), 0.94 (s, 3H).

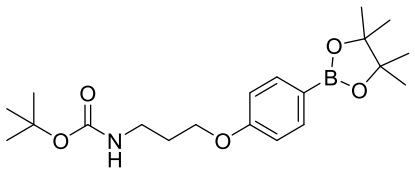
^{13}C NMR ($CDCl_3$, 126 MHz): 173.8, 161.2, 136.8, 113.9, 83.8, 71.6, 71.5, 66.9, 55.9, 55.1, 43.9, 42.6, 40.3, 39.3, 39.0, 37.5, 37.0, 35.5, 35.1, 34.2, 33.7, 31.9, 30.5, 28.8, 27.0, 25.0, 23.5, 21.3, 18.6, 12.3. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 31.4.

ν_{max} (neat): 3421, 2928, 2861, 1651, 1604, 1547, 1454 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{38}H_{61}BNO_6$) requires 638.4593 *m/z*, found 638.4582 *m/z.*

Preparation of *tert*-butyl (3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)carbamate **46**



Prepared according to General Procedure G, using 4-hydroxyphenyl boronic acid pinacol ester (250 mg, 1.13 mmol, 1.0 equiv.), *tert*-butyl (3-hydroxypropyl)carbamate **26** (257 mg, 1.47 mmol, 1.3 equiv.), CMBP (471 μL, 1.09 mmol, 2.0 equiv.) in anhydrous toluene (4 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-7% EtOAc in hexane) then washed with saturated Na_2CO_3 (3 x 20 mL) to afford the desired product **46** as a clear oil (250 mg, 74%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.74 (d, *J* = 8.5 Hz, 2H), 6.88 (d, *J* = 8.5 Hz, 2H), 4.74 (br. s, 1H), 4.04 (t, *J* = 6.0 Hz, 2H), 3.33-3.31 (m, 2H), 2.00-1.95 (m, 2H) 1.44 (s, 9H), 1.33 (s, 12H).

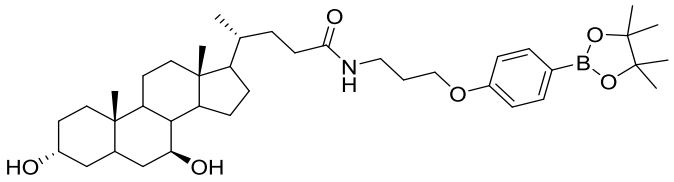
^{13}C NMR ($CDCl_3$, 101 MHz): 161.5, 156.1, 136.7, 114.0, 83.7, 79.4, 65.7, 38.1, 29.6, 28.5, 25.0. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 31.7.

ν_{max} (neat): 3357, 2977, 2932, 1694, 1606, 1517 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{20}H_{33}BNO_5$) requires 378.2446 *m/z*, found 378.2441 *m/z*.

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)pentanamide **47**



tert-Butyl (3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)carbamate **46** (231 mg, 0.61 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using a 1:1 mixture of trifluoroacetic acid (1 mL) in DCM (1 mL) to generate crude intermediate 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propan-1-aminium 2,2,2-trifluoroacetate **46a**. Compound **47** was prepared according to General Procedure E, using ursodeoxycholic acid (239 mg, 0.61 mmol, 1.0 equiv.), 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propan-1-aminium 2,2,2-trifluoroacetate, DIPEA (424 μL, 2.43 mmol, 4.0 equiv.) and HATU (254 mg, 0.67 mmol, 1.1 equiv.) in DMF (3 mL). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-3% MeOH in DCM) to afford the desired product **47** as a clear solid (247 mg, 60% over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.75 (d, *J* = 8.4 Hz, 2H), 6.88 (d, *J* = 8.4 Hz, 2H), 5.81 (t, *J* = 5.3 Hz, 1H), 4.07 (t, *J* = 5.8 Hz, 2H), 3.63-3.54 (m, 2H), 3.45 (app. q, *J* = 6.2 Hz, 2H), 2.20-2.26 (m, 1H), 1.98-2.09 (m, 4H), 1.85-1.92 (m, 1H), 1.76-1.82 (m, 5H), 1.66-1.68 (m, 2H), 1.57-1.68 (m, 2H), 1.57-1.61 (m, 3H), 1.40-1.51 (m, 9H), 1.33 (s, 12 H), 0.99-1.17 (m, 4H), 0.94 (s, 3H), 0.92-0.93 (m, 3H), 0.67 (s, 3H).

^{13}C NMR ($CDCl_3$, 101 MHz): 173.7, 161.3, 136.8, 113.9, 83.8, 71.6, 71.5, 66.3, 55.9, 55.1, 43.9, 43.9, 42.6, 40.3, 39.3, 37.5, 37.5, 37.0, 35.5, 35.1, 34.2, 33.8, 32.0, 30.5, 29.1, 28.8, 27.0, 25.0, 23.5, 21.3, 18.7, 12.3. 1 x C bearing B not observed.

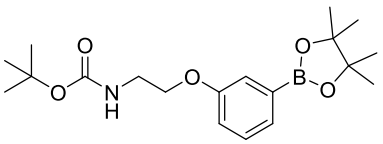
^{11}B NMR ($CDCl_3$, 160 MHz): 31.7.

Experimental

ν_{max} (neat): 3324, 2926, 2863, 1651, 1604, 1547, 1450 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{39}H_{63}BNO_6$) requires 652.4755 *m/z*, found 652.4756 *m/z.*

Preparation of *tert*-butyl (2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)carbamate **49**



Prepared according to General Procedure G, using 3-hydroxyphenyl boronic acid pinacol ester **48** (200 mg, 0.90 mmol, 1.0 equiv.), *N*-Boc ethanolamine (210 μL, 1.36 mmol, 1.5 equiv.), CMBP (471 μL, 1.09 mmol , 2.0 equiv.) in anhydrous toluene (4 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-7% EtOAc in hexane) then washed with saturated Na_2CO_3 (3 x 20 mL) to afford the desired product **49** as a clear oil (183 mg, 56%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.42 (d, *J* = 7.2 Hz, 1H), 7.32-7.28 (m, 2H), 7.01-6.98 (m, 1H), 4.98 (br. s, 1H), 4.05 (t, *J* = 5.0 Hz, 2H), 3.54-3.53 (m, 2H), 1.45 (s, 9H), 1.34 (s, 12H).

^{13}C NMR ($CDCl_3$, 126 MHz): 159.3, 158.2, 129.2, 127.7, 120.0, 118.2, 84.0, 79.4, 67.3, 40.3, 28.6, 25.0. 1 x C bearing B not observed.

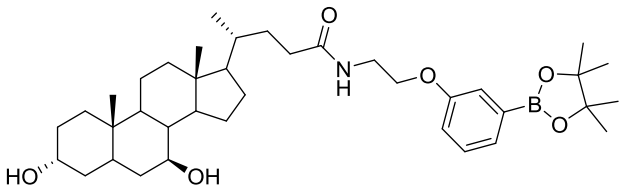
^{11}B NMR ($CDCl_3$, 160 MHz): 30.9.

ν_{max} (neat): 3370, 2975, 2930, 1687, 1536 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{19}H_{31}BNO_5$) requires 364.2290 *m/z*, found 364.2289 *m/z*.

Experimental

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)pentanamide **50**



tert-Butyl (2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)carbamate **49** (183 mg, 0.50 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using a 1:1 mixture of trifluoroacetic acid (1 mL) in DCM (1 mL) to generate crude intermediate 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethan-1-aminium 2,2,2-trifluoroacetate. Compound **50** was prepared according to General Procedure E, using UDCA (196 mg, 0.50 mmol, 1.0 equiv.), 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethan-1-aminium 2,2,2-trifluoroacetate, DIPEA (350 μL, 2.00 mmol, 4.0 equiv.) and HATU (209 mg, 0.55 mmol, 1.1 equiv) in DMF (5 mL). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-4% MeOH in DCM) to afford the desired product **50** as a clear solid (130 mg, 44% over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.41 (d, J = 7.2 Hz, 1H), 7.32-7.27 (m, 2H), 7.00-6.97 (m, 1H), 5.99 (br. s, 1H), 4.06 (t, J = 5.0 Hz, 2H), 3.66-3.63 (m, 2H), 3.59-3.54 (m, 2H), 2.28-2.22 (m, 1H), 2.11-2.05 (m, 2H), 1.98-1.96 (m, 1H), 1.91-1.85 (m, 1H), 1.81-1.75 (m, 8H), 1.67-1.64 (m, 2H), 1.60-1.56 (m, 2H), 1.49-1.38 (m, 7H), 1.33 (s, 12H), 1.30-1.19 (m, 5H), 1.15-0.97 (m, 4H), 0.93-0.91 (m, 6H), 0.64 (s, 3H).

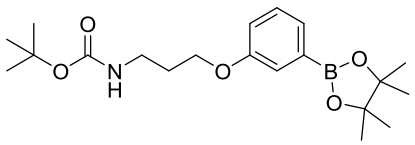
^{13}C NMR ($CDCl_3$, 126 MHz): 173.7, 158.1, 129.2, 127.8, 120.0, 118.1, 84.1, 71.6, 71.5, 67.0, 55.9, 55.1, 44.0, 43.9, 42.6, 40.3, 39.3, 39.1, 37.5, 37.0, 35.5, 35.1, 34.2, 33.7, 31.9, 30.5, 28.8, 27.0, 25.0, 23.5, 21.3, 18.6, 12.3. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 30.8.

ν_{max} (neat): 3330, 2926, 2863, 1651, 1547, 1428 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{38}H_{61}BNO_6$) requires 638.4593 *m/z*, found 638.4580 *m/z*.

Preparation of *tert*-butyl (3-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)carbamate **51**



Prepared according to General Procedure G, using 3-hydroxyphenyl boronic acid pinacol ester (200 mg, 0.90 mmol, 1.0 equiv.), *tert*-butyl (3-hydroxypropyl)carbamate **24** (238 mg, 1.36 mmol, 1.5 equiv.), CMBP (471 μL, 1.09 mmol , 2.0 equiv.) in anhydrous toluene (4 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-7% EtOAc in hexane) then washed with saturated Na_2CO_3 (3 x 20 mL) to afford the desired product **51** as a clear oil (298 mg, 88%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.40-7.39 (m, 1H), 7.32-7.31 (m, 1H), 7.28 (m, 1H, coincident with solvent peak), 7.00-6.98 (m, 1H), 4.77 (br. s., 1H), 4.05 (t, 2H, J = 5.9 Hz), 3.33-3.32 (m, 2H), 1.99-1.94 (m, 2H), 1.44 (s, 9H), 1.34 (s, 12H).

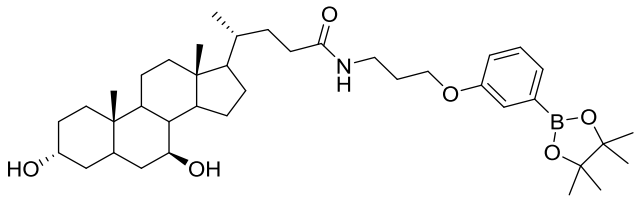
^{13}C NMR ($CDCl_3$, 126 MHz): 158.4, 156.1, 129.1, 127.5, 119.8, 118.4, 84.0, 66.0, 38.3, 29.6, 28.6, 25.0. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 31.2.

ν_{max} (neat): 3375, 2980, 2932, 1688, 1534, 1433 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{20}H_{33}BNO_5$) requires 378.2446 *m/z*, found 378.2441 *m/z*.

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)pentanamide **52**



tert-Butyl (3-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)carbamate **51** (300 mg, 0.80 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using a 1:1 mixture of trifluoroacetic acid (1 mL) in DCM (1 mL) to generate crude intermediate 3-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propan-1-aminium 2,2,2-trifluoroacetate. Compound **52** was prepared according to General Procedure E, using UDCA (345 mg, 0.88 mmol, 1.0 equiv.), 3-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propan-1-aminium 2,2,2-trifluoroacetate, DIPEA (614 µL, 3.52 mmol, 4.0 equiv.) and HATU (368 mg, 0.97 mmol, 1.1 equiv) in DMF (4 mL). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-4% MeOH in DCM) to afford the desired product **52** as a clear solid (190 mg, 33% over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.41 (d, *J* = 7.2 Hz, 1H), 7.28-7.32 (m, 2H), 6.99 (dd, *J* = 8.1, 2.4 Hz, 1H), 5.86 (br. s, 1H), 4.08 (t, *J* = 5.6 Hz, 2H), 3.56-3.61 (m, 2H), 3.45-3.48 (m, 2H), 2.22-2.27 (m, 1H), 2.04-2.10 (m, 2H), 1.97-2.01 (m, 3H), 1.85-1.92 (m, 1H), 1.76-1.81 (m, 6H), 1.54-1.68 (m, 8H), 1.40-1.51 (m, 8H), 1.34 (s, 14H), 1.19-1.30 (m, 6H), 0.99-1.16 (m, 4H), 0.93-0.94 (m, 6H), 0.66 (s, 3H).

^{13}C NMR ($CDCl_3$,126 MHz): 173.7, 158.2, 129.2, 127.7, 119.7, 118.3, 84.0, 71.6, 71.5, 66.6, 55.9, 55.1, 44.0, 43.9, 42.6, 40.3, 39.3, 37.7, 37.5, 37.0, 35.6, 35.1, 34.2, 33.9, 32.0, 30.5, 29.1, 28.8, 27.0, 25.0, 23.5, 21.3, 18.7, 12.3. 1 x C bearing B not observed.

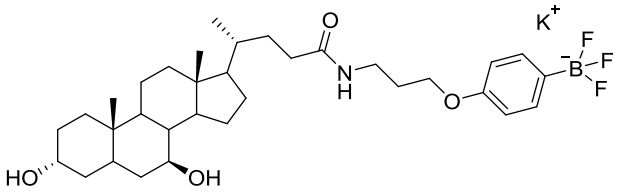
^{11}B NMR ($CDCl_3$, 160 MHz): 31.2.

ν_{max} (neat): 3811, 2928, 2863, 1573, 1548, 1649, 1428 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{39}H_{63}BNO_6$) requires 652.4755 *m/z*, found 652.4753 *m/z*.

6.3.4 Linear Trifluoroborate and Boronic Acid Analogues

Preparation of potassium (4-(3-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)propoxy)phenyl)trifluoroborate **55**



Prepared according to General Procedure H, using (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)pentanamide **47** (146 mg, 0.21 mmol, 1.0 equiv.), potassium fluoride (52 mg, 0.89 mmol, 1.0 equiv) in H_2O (0.4 mL), L-(+)-tartaric acid (69 mg, 0.46 mmol, 2.05 equiv.) and MeOH/MeCN (1:1). The crude material carried through to the next step as a white solid (80.9 mg, 58%) without further purification.

^{1}H NMR (DMSO-d_6, 500 MHz): 7.83 (t, J = 5.3 Hz, 1H), 7.20 (d, J = 8.1 Hz, 2H), 6.64 (d, J = 8.1 Hz, 2H), 4.41 (br. s, 1H), 3.90-3.84 (m, 3H), 3.16 (m, 2H), 2.11-2.05 (m, 1H), 1.99-1.92 (m, 2H), 1.87-1.73 (m, 4H), 1.68-1.63 (m, 3H), 1.51-1.46 (m, 3H), 1.41-1.27 (m, 8H), 1.21-1.07 (m, 8H), 1.03-0.98 (m, 1H), 0.89-0.87 (m, 6H), 0.61 (s, 3H).

^{13}C NMR (DMSO-d_6, 126 MHz): 172.5, 156.4, 132.2, 112.5, 69.7, 69.5, 64.7, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 37.7, 37.3, 35.5, 34.9, 34.8, 33.8, 32.4, 31.7, 30.2, 29.1, 28.2, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C bearing B not observed, 1 x C not observed.

^{11}B NMR (DMSO-d_6, 160 MHz): 3.63.

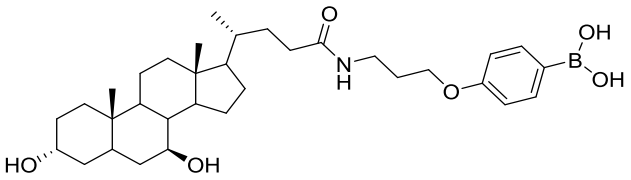
^{19}F NMR (DMSO-d_6, 471 MHz): -138.3.

ν_{max} (neat): 3364, 2930, 2865, 1646, 1605 cm^{-1}.

HRMS: exact mass calculated for $[M-K]^-$ ($C_{33}H_{50}BF_3NO_4$) requires 592.3796 *m/z*, found 592.3796 *m/z*.

Experimental

Preparation of (4-(3-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)propoxy)phenyl)boronic acid **56**



Prepared according to General Procedure I, using potassium (4-(3-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)propoxy)phenyl)trifluoroborate **55** (75 mg, 0.13 mmol, 1.0 equiv.) excess SiO_2 and H_2O (3 mL). The crude material was subjected to the purification outlined in General Procedure I (silica gel, 0-8% MeOH in DCM) to afford the desired product **56** as a white solid (32.1 mg, 43 %).

^{1}H NMR (DMSO-d_6, 400 MHz): 7.82 (app. br. s., 2H), 7.72 (d, 2H, *J* = 8.3 Hz), 6.92 (d, 2H, *J* = 8.3 Hz), 4.69-4.62 (m, 1H), 4.48-4.42 (m, 1H), 3.87-3.85 (m, 2H), 3.75-3.68 (m, 1H), 3.16-3.32 (m, coincident with solvent), 2.36-2.32 (m, 1H), 2.23-2.18 (m, 1H), 1.95-1.82 (m, 5H), 1.75-1.57 (m, 6H), 1.48-1.42 (m, 4H), 1.41-1.29 (m, 6H), 1.23-1.03 (m, 8H), 0.91-0.87 (m, 6H), 0.62 (s, 3H).

^{1}H NMR (DMSO-d_6 with D_2O shake, 500 MHz): 7.63 (d, 2H, *J* = 7.2 Hz), 6.83 (d, 2H, *J* = 7.2 Hz), 3.96-3.87 (m, 2H), 3.37-3.24 (m, 2H), 3.22-3.09 (m, 2H), 2.12-2.03 (m, 1H), 2.00-1.91 (m, 1H), 1.86-1.74 (m, 3H), 1.70-1.53 (m, 5H), 1.48-1.34 (m, 4H), 1.32-1.16 (m, 6H), 1.12-0.96 (m, 7H), 0.92-0.83 (m, 2H), 0.82-0.78 (m, 6H), 0.44 (s, 3H).

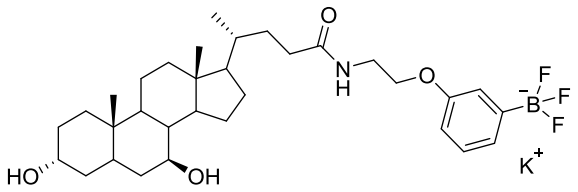
^{13}C NMR (DMSO-d_6, 101 MHz): 172.6, 160.3, 135.8, 113.4, 69.7, 69.5, 64.9, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 37.7, 37.3, 35.4, 34.9, 34.8, 33.7, 32.4, 31.7, 30.2, 28.9, 28.2, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C bearing B not observed, 1 x C not observed.

^{11}B NMR (MeOD, 160 MHz): 28.4.

ν_{max} (neat): 3322, 2924, 2854, 1651, 1602, 1513, 1454, 1411 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{33}H_{53}BNO_6$) requires 570.3966 *m/z*, found 570.3959 *m/z*.

Preparation of potassium (3-(2-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)ethoxy)phenyl)trifluoroborate **57**



Prepared according to General Procedure H, using (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)ethyl)pentanamide (130 mg, 0.20 mmol, 1.0 equiv.), potassium fluoride (47 mg, 0.81 mmol, 1.0 equiv) in H_2O (0.3 mL), L-(+)-tartaric acid (62 mg, 0.42 mmol, 2.05 equiv.) and MeOH/MeCN (1:1). The crude material carried through to the next step as a white solid (58.4 mg, 48%) without further purification.

^{1}H NMR (DMSO-d_6, 400 MHz): 7.98 (app. br. s., 1H), 6.98 (t, 1H, J = 7.3 Hz), 6.90-6.86 (m, 2H), 6.57 (d, 1H, J = 7.9 Hz), 4.46-4.44 (m, 1H), 3.91-3.86 (m, 3H), 2.16-2.07 (m, 1H), 2.02-1.97 (m, 1H), 1.93-1.89 (m, 1H), 1.87-1.80 (m, 1H), 1.78-1.73 (m, 1H), 1.68-1.61 (m, 3H), 1.52-1.43 (m, 4H), 1.41-1.28 (m, 8H), 1.19-1.10 (m, 7H), 1.01-0.93 (m, 2H), 0.90-0.87 (m, 7H), 0.60 (s, 3H).

^{13}C NMR (DMSO-d_6, 126 MHz): 173.0, 157.2, 127.2, 124.1, 117.1, 111.3, 69.8, 69.5, 65.8, 55.9, 54.7, 43.1, 43.0, 42.2, 38.8, 38.5, 37.7, 37.3, 35.0, 34.9, 33.8, 32.4, 31.6, 30.3, 28.2, 26.7, 23.3, 20.9, 18.5, 12.1. 1 x C bearing B not observed, 1 x C not observed.

^{11}B NMR (DMSO-d_6, 101 MHz): 3.08.

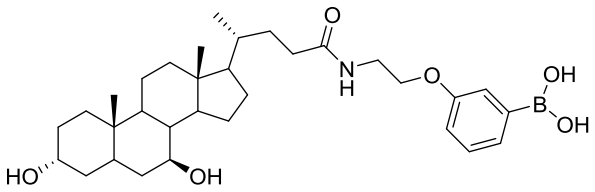
^{19}F NMR (DMSO-d_6, 471 MHz): -139.2.

ν_{max} (neat): 3381, 2930, 2867, 1646 cm^{-1}.

HRMS: exact mass calculated for $[M\text{-}K]^-$ ($C_{32}H_{48}BF_3NO_4$) requires 578.3640 *m/z*, found 578.3637 *m/z*.

Experimental

Preparation of (3-(2-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)ethoxy)phenyl)boronic acid **58**



Prepared according to General Procedure I, using potassium (3-(2-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)ethoxy)phenyl)trifluoroborate **57** (50 mg, 0.09 mmol, 1.0 equiv.), excess SiO_2 and H_2O (2 mL). The crude material was subjected to the purification outlined in General Procedure I (silica gel, 0-8% MeOH in DCM) to afford the desired product **58** as a white solid (14.3 mg, 29%).

^{1}H NMR (MeOD, 500 MHz): 8.23 (br. s., 1H), 7.37-7.25 (m, 2H), 7.21-7.18 (m, 1H), 7.01-6.99 (m, 1H), 4.08 (t, 2H, J = 5.4 Hz), 3.60-3.58 (m, 2H), 3.55-3.47 (m, 2H), 2.32-2.26 (m, 1H), 2.20-2.14 (m, 1H), 2.04-2.02 (m, 1H), 1.92-1.89 (m, 2H), 1.87-1.83 (m, 3H), 1.67-1.63 (m, 3H), 1.60-1.56 (m, 2H), 1.53-1.42 (m, 7H), 1.36-1.28 (m, 5H), 1.26-1.17 (m, 2H), 1.13-1.03 (m, 3H), 0.99-0.98 (m, 6H), 0.69 (s, 3H).

^{13}C NMR (MeOD, 126 MHz): 177.1, 159.7, 129.9, 129.7, 127.6, 127.0, 120.5, 117.9, 117.0, 72.1, 72.0, 67.5, 57.5, 56.6, 49.9, 44.8, 44.5, 44.0, 41.5, 40.7, 40.2, 38.6, 38.0, 36.7, 36.1, 35.2, 34.1, 33.3, 31.0, 29.6, 27.9, 23.9, 22.4, 19.0, 12.6. 1 x C bearing B not observed.

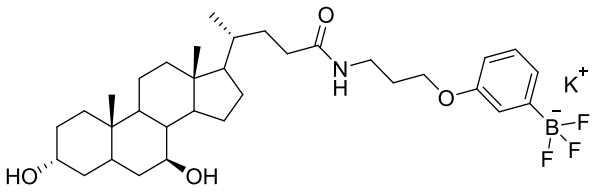
^{11}B NMR (MeOD, 160 MHz): 28.5.

Required use of MeOD resulted in partial OH for D exchange; however this can be rectified by utilising DMSO as the NMR solvent.

ν_{max} (neat): 3325 2932, 2867, 1646, 1636, 1605 cm^{-1}.

HRMS: exact mass calculated for $[M-H]^-$ ($C_{32}H_{49}BNO_6$) requires 554.3664 m/z, found 554.3661 m/z.

Preparation of potassium (3-(3-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanamido)propoxy)phenyl)trifluoroborate **59**



Prepared according to General Procedure H, using (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-(3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)propyl)pentanamide (160 mg, 0.24 mmol, 1.0 equiv.), potassium fluoride (57 mg, 0.98 mmol, 1.0 equiv) in H_2O (0.5 mL), L-(+)-tartaric acid (74 mg, 0.49 mmol, 2.05 equiv.) and MeOH/MeCN (1:1). The crude material carried through to the next step as a white solid (109.6 mg, 72%) without further purification.

^{1}H NMR (DMSO-d_6, 500 MHz): 7.85 (t, *J* = 5.5 Hz, 1H), 6.98 (app. t, *J* = 7.5 Hz, 1H), 6.89-6.86 (m, 2H), 6.55 (dd, *J* = 7.9, 1.9 Hz, 1H), 4.41 (d, *J* = 4.4 Hz, 1H), 3.92-3.88 (m, 2H), 3.85 (d, *J* = 6.8 Hz, 1H), 3.18-3.15 (m, 2H), 2.11-2.06 (m, 1H), 1.98-1.92 (m, 1H), 1.84-1.73 (m, 4H), 1.68-1.73 (m, 3H), 1.49-1.46 (m, 3H), 1.40-1.27 (m, 7H), 1.19-1.10 (m, 10H), 1.01-0.91 (m, 2H), 0.89-0.87 (m, 6H), 0.61 (s, 3H).

^{13}C NMR (DMSO-d_6, 126 MHz): 172.5, 157.3, 127.1, 123.8, 117.0, 111.2, 69.7, 69.5, 64.6, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 37.7, 37.3, 35.5, 34.9, 34.8, 33.8, 32.4, 31.7, 30.2, 29.2, 28.2, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C bearing B not observed, 1 x C not observed.

^{11}B NMR (DMSO-d_6, 160 MHz): 3.19.

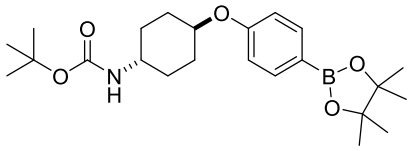
^{19}F NMR (DMSO-d_6, 471 MHz): 139.2.

ν_{max} (neat): 3314, 2930, 2865 cm^{-1}.

HRMS: exact mass calculated for $[M-K]^-$ ($C_{33}H_{50}BF_3NO_4$) requires 592.3796 *m/z*, found 592.3795 *m/z*.

6.3.5 Cyclohexyl Boronate Analogues

For example, for the preparation of *tert*-butyl ((1*R*,4*R*)-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexyl)carbamate **61**



To a microwave vial was added *cis*-4-Boc-aminocyclohexanol (174 mg, 0.81 mmol, 1.5 equiv.) and 4-hydroxyphenolboronic acid pinacol ester (120 mg, 0.54 mmol, 1.0 equiv.). The vial was sealed with a microwave cap, purged under N_2 followed by addition of anhydrous toluene (6 mL). To the vial was added CMBP (0.29 mL, 1.09 mmol, 2.0 equiv.) and the reaction heated at 120 °C for 16 hours. Reaction mixture was concentrated *in vacuo* to a residue that was purified by column chromatography on silica (10% EtOAc in petroleum either 40-60) to yield desired product **61** as a brown solid (125 mg, 55%).

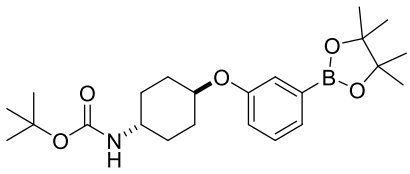
^{1}H NMR ($CDCl_3$, 500 MHz): 7.72 (d, *J* = 8.2 Hz, 2H), 6.86 (d, *J* = 8.2 Hz, 2H), 4.41 (br. s, 1H), 4.26-4.21 (m, 1H), 3.51 (br. s, 1H), 2.13-2.07 (m, 4H), 1.59-1.53 (m, 2H), 1.45 (s, 9H), 1.33 (s, 12H), 1.28-1.26 (m, 2H).

^{13}C NMR ($CDCl_3$, 126 Hz): 160.4, 155.4, 136.7, 115.2, 83.7, 74.7, 48.9, 30.9, 30.2, 28.6, 25.0. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 Hz): 31.6.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{23}H_{37}BNO_5$) requires 418.2769 *m/z*, found 418.2763 *m/z.*

Preparation of *tert*-butyl ((1*R*,4*R*)-4-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexyl)carbamate **62**



Prepared according to General Procedure G, using *cis*-4-Boc-aminocyclohexanol (174 mg, 0.81 mmol, 1.5 equiv.) 3-hydroxyphenolboronic acid pinacol ester (120 mg, 0.54 mmol, 1.0 equiv.), CMBP (0.29 mL, 1.09 mmol, 2.0 equiv.) in anhydrous toluene (6 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-10% EtOAc in petroleum ether 40-60) to afford the desired product **62** as a white solid (113 mg, 50 %).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.38 (d, J = 7.2 Hz, 1H), 7.32 (d, J = 2.1 Hz, 1H), 7.25-7.28 (m, coincident with solvent peak), 6.98 (dd, J = 8.2, 2.3 Hz, 1H), 4.40 (br. s, 1H), 4.25-4.20 (m, 1H), 3.51 (br. s, 1H), 2.12-2.06 (m, 4H), 1.59-1.52 (m, 2H), 1.45 (s, 9H), 1.34 (s, 12H), 1.30-1.23 (m, 2H).

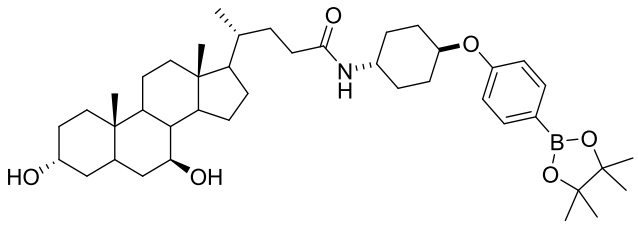
^{13}C NMR (DMSO-d_6, 101 MHz): 157.3, 129.1, 127.6, 121.9, 119.9, 84.0, 75.1, 30.9, 30.4, 28.6, 25.0. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 30.6.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_{23}H_{36}BNO_5Na$) requires 440.2584 *m/z*, found 440.2585 *m/z*.

Experimental

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-((1*r*,4*R*)-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexyl)pentanamide **63**



tert-butyl ((1*R*,4*R*)-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexyl)carbamate **61** (77 mg, 0.18 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using a 1:1 mixture of trifluoroacetic acid (1 mL) in DCM (1 mL) to generate crude intermediate (1*R*,4*R*)-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexan-1-aminium 2,2,2-trifluoroacetate. Compound **63** was prepared according to General Procedure E, using UDCA (69 mg, 0.18 mmol, 1.0 equiv.), (1*R*,4*R*)-4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexan-1-aminium 2,2,2-trifluoroacetate (76 mg, 0.18 mmol, 1.0 equiv), DMF (3 mL), DIPEA (123 μL, 0.71 mmol, 4.0 equiv.) and HATU (74 mg, 0.19 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-3% MeOH in DCM) to afford the desired product **63** as a brown solid (42.3 mg, 33% over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.72 (d, *J* = 8.3 Hz, 2H), 6.87 (d, *J* = 8.3 Hz, 2H), 7.41 (d, *J* = 7.2 Hz, 1H), 5.26 (br. s, 1H), 4.26-4.22 (m, 1H), 3.87-3.80 (m, 1H), 3.61-3.57 (m, 2H), 2.25-2.19 (m, 1H), 2.14-2.07 (m, 5H), 2.07-2.01 (m, 1H), 1.95-1.87 (m, 1H), 1.84-1.76 (m, 4H), 1.68-1.66 (m, 3H), 1.62-1.56 (m, 9H), 1.51-1.41 (m, 8H), 1.33 (s, 12H), 1.29-1.22 (m, 6H), 1.17-1.02 (m, 4H), 0.95-0.93 (m, 6H), 0.68 (s, 3H).

^{13}C NMR ($CDCl_3$,101 MHz): 173.1, 160.4, 136.7, 115.3, 83.7, 74.8, 71.6, 71.5, 55.9, 55.2, 47.6, 43.9, 42.6, 40.3, 39.4, 37.5, 37.0, 35.6, 35.1, 34.2, 33.9, 32.1, 30.7, 30.5, 30.3, 28.9, 27.1, 25.0, 23.5, 21.3, 18.7, 12.3. 1 x C bearing B not observed.

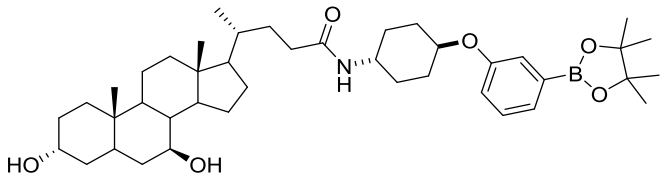
^{11}B NMR ($CDCl_3$, 128 MHz): 30.2 (background borosilicate glass signal observed).

ν_{max} (neat): 3398, 2928, 2861, 1649, 1602, 1537, 1513, 1452 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{42}H_{67}BNO_6$) requires 692.5063 *m/z*, found 692.5055 *m/z.*

Experimental

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-*N*-((1*R*,4*R*)-4-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexyl)pentanamide **64**



tert-Butyl ((1*R*,4*R*)-4-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexyl)carbamate **62** (100 mg, 0.24 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using using a 1:1 mixture of trifluoroacetic acid (1 mL) in DCM (1 mL) to generate crude intermediate (1*r*,4*r*)-4-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexan-1-aminium 2,2,2-trifluoroacetate. Compound **64** was prepared according to General Procedure E, using UDCA (71 mg, 0.18 mmol, 1.0 equiv.), (1*R*,4*R*)-4-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)cyclohexan-1-aminium 2,2,2-trifluoroacetate, DMF (3 mL), DIPEA (125 μL, 0.72 mmol, 4.0 equiv.) and HATU (75 mg, 0.20 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-3% MeOH in DCM) to afford the desired product **64** as an off-white solid (51.2 mg, 31% over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.39-7.38 (m, 1H), 7.32-7.31 (m, 1H), 7.29-7.26 (m, coincident with solvent), 5.25 (br. s., 1H), 4.25-4.21 (m 1H), 3.90-3.81 (m, 1H), 3.62-3.58 (m, 2H), 2.27-2.19 (m, 1H), 2.13-2.06 (m, 4H), 2.01-1.99 (m, 1H), 1.95-1.88 (m, 1H), 1.84-1.76 (m, 3H), 1.69-1.66 (m, 3H), 1.62-1.55 (m, 9H), 1.52-1.41 (m, 8H), 1.34 (s, 12H), 1.29-1.23 (m, 7H), 1.17-1.02 (m, 3H), 0.95-0.93 (m, 6H), 0.88 (s, 3H).

^{13}C NMR ($CDCl_3$, 101 MHz): 173.1, 157.3, 129.2, 127.6, 121.9, 120.0, 84.0, 75.2, 71.6, 71.5, 55.9, 55.2, 47.7, 43.9, 42.6, 40.3, 39.4, 37.5, 37.0, 35.6, 35.1, 34.2, 33.9, 32.1, 30.7, 30.5, 30.4, 28.9, 28.5, 28.5, 27.1, 25.0, 23.5, 21.3, 18.7, 12.3. 1 x C bearing B not observed.

Experimental

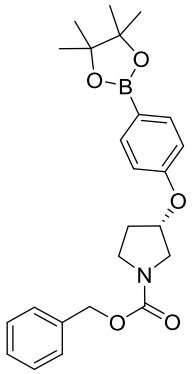
^{11}B NMR ($CDCl_3$, 128 MHz): 31.2 (background borosilicate glass signal observed).

ν_{max} (neat): 3313, 2928, 2861, 1645, 1537, 1426 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{42}H_{67}BNO_6$) requires 692.5063 *m/z*, found 692.5039 *m/z*.

6.3.6 Cyclic Boronate Analogues

Preparation of (*S*)-benzyl 3-(4-(4, 4, 5, 5-tetramethyl-1, 3, 2-dioxaborolan-2-yl)phenoxy)pyrrolidine-1-carboxylate **65**



Prepared according to General Procedure G, using (*R*)-(-)-1-Cbz-3-hydroxypyrrolidine (301 mg, 1.36 mmol, 1.5 equiv.) 4-hydroxyphenol boronic acid pinacol ester (200 mg, 0.91 mmol, 1.0 equiv.), CMBP (477 μL, 1.82 mmol, 2.0 equiv.) in anhydrous toluene (5 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-12% EtOAc in hexane) to afford the desired product **65** as a pale yellow gum (307 mg, 77%).

^{1}H NMR ($CDCl_3$, 500 MHz, 300 K): 7.75 (d, 2H, *J* = 8.1 Hz), 7.39-7.30 (m, 5H), 6.85 (d, 2H, *J* = 8.1 Hz), 5.19-5.09 (m, 2H), 4.95 (br. s, 1H), 3.72-3.54 (m, 4H), 2.24-2.20 (m, 1H), 2.15-2.06 (m, 1H), 1.33 (s, 12H). Rotameric mixture observed.

^{1}H NMR Variable Temperature (DMSO-d_6, 400 MHz, 373 K): 7.63 (d, 2H, *J* = 8.3 Hz), 7.37-7.30 (m, 5H), 6.94 (d, 2H, *J* = 8.3 Hz), 5.10 (br. s, 2H), 5.07 (br. s, 1H), 3.69-3.66 (m, 1H), 3.56-3.45 (m, 3H), 2.25-2.17 (m, 1H), 2.10-2.07 (m, 1H), 1.30 (s, 12H).

^{13}C NMR ($CDCl_3$, 126 MHz, 300 K): 159.8, 159.7, 154.99, 154.96, 137.0, 136.9, 136.8, 128.6, 128.08, 128.06, 114.9, 83.7, 76.1, 75.3, 67.0, 52.1, 51.6, 44.5, 44.1, 31.7, 30.9, 25.0. 1 x C bearing B not observed. Rotameric mixture observed.

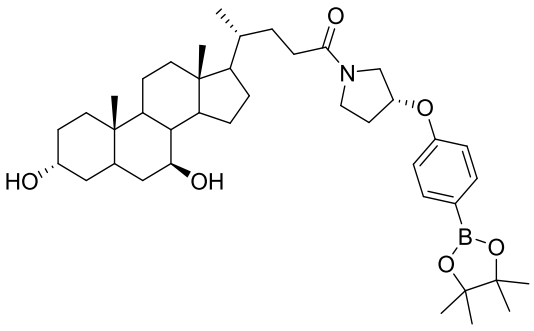
^{13}C NMR Variable Temperature (DMSO-d_6, 101 MHz, 373 K): 159.2, 153.6, 136.7, 135.7, 127.8, 127.1, 126.8, 114.7, 82.9, 75.4, 65.4, 50.9, 43.4, 30.1, 24.1. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 30.9.

ν_{max} (neat) 2977, 2866, 1705, 1606, 1415 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{24}H_{31}BNO_5$) requires 424.2290 *m/z*, found 424.2284 *m/z*.

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-((*S*)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)pyrrolidin-1-yl)pentan-1-one **66**



(*S*)-benzyl 3-(4-(4, 4, 5, 5-tetramethyl-1, 3, 2-dioxaborolan-2-yl)phenoxy)pyrrolidine-1-carboxylate **65** (269 mg, 0.63 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure J to generate crude intermediate **65a**. Compound **66** was prepared according to General Procedure E, using UDCA (255 mg, 0.65 mmol, 1.0 equiv.), **65a** (194 mg, 0.65 mmol, 1.0 equiv), DMF (5 mL), DIPEA (340 μL, 1.95 mmol, 3.0 equiv.) and HATU (273 mg, 0.72 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure C (silica gel, 0-6% MeOH in DCM) to afford the desired product **66** as a white solid (194 mg, 46 %).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.76-7.72 (m, 2H), 6.87-6.84 (m, 2H), 5.06-4.93 (m, 1H), 3.83-3.56 (m, 6H), 2.37-2.07 (m, 4H), 2.02-1.97 (m, 1H), 1.94-1.88 (m, 1H), 1.83-1.76 (m, 4H), 1.68-1.57 (m, 5H), 1.51-1.40 (m, 9H), 1.33 (s, 12H), 1.28-1.20 (m, 4H), 0.99-0.91 (m, 6H), 0.68-0.67 (m, 3H). Rotameric mixture observed.

^{13}C NMR ($CDCl_3$, 126 MHz): 172.6, 172.5, 159.7, 159.6, 136.83, 136.78, 114.94, 114.88, 83.83, 83.77, 76.3, 74.8, 71.6, 71.5, 55.9, 55.2, 52.3, 51.6, 44.8, 44.0, 43.94, 43.86, 42.6, 40.3, 39.3, 37.5, 36.9, 35.6, 35.5, 35.1, 34.2, 32.0, 31.9, 31.7, 31.01, 30.99, 30.5, 30.2, 28.81, 28.77, 27.1, 25.0, 23.5, 21.3, 18.8, 12.3. 1 x C bearing B not observed. Rotameric mixture observed.

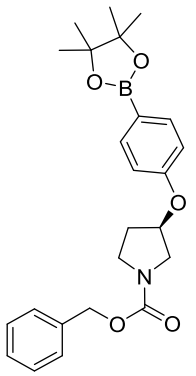
^{11}B NMR ($CDCl_3$, 160 Hz): 31.1.

Experimental

ν_{max} (neat): 3418, 2930, 2867, 1606 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{40}H_{62}BNO_6Na$) requires 686.4569 *m/z*, found 686.4562 *m/z*.

Preparation of benzyl (*R*)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)pyrrolidine-1-carboxylate **67**



Prepared according to General Procedure G, using (*S*)-benzyl 3-hydroxypyrrolidine-1-carboxylate (301 mg, 1.36 mmol, 1.5 equiv.) 4-hydroxyphenol boronic acid pinacol ester (200 mg, 0.91 mmol, 1.0 equiv.), CMBP (477 µL, 1.82 mmol, 2.0 equiv.) in anhydrous toluene (5 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-12% EtOAc in hexane) to afford the desired product **67** as a yellow gum (298 mg, 79 %).

^{1}H NMR ($CDCl_3$, 500 MHz, 300 K): 7.75 (d, 2H, *J* = 8.0 Hz), 7.39-7.30 (m, 5H), 6.85 (d, 2H, *J* = 8.0 Hz), 5.19-5.09 (m, 2H), 4.95 (br. s, 1H), 3.72-3.56 (m, 4H), 2.23-2.20 (m, 1H), 2.15-2.06 (m, 1H), 1.33 (s, 12H). Rotameric mixture observed.

^{1}H NMR Variable Temperature (DMSO-d_6, 400 MHz, 373 K): 7.62 (d, 2H, *J* = 8.5 Hz), 7.35-7.28 (m, 5H), 6.93 (d, 2H, *J* = 8.5 Hz), 5.09-5.04 (m, 3H), 3.69-3.65 (m, 1H), 3.57-3.47 (m, 3H), 2.25-2.16 (m, 1H), 2.11-2.05 (m, 1H), 1.30 (s, 12H).

^{13}C NMR ($CDCl_3$, 126 MHz, 300 K): 159.8, 159.7, 154.98, 154.96, 137.0, 136.9, 136.8, 128.6, 128.08, 128.06, 114.9, 83.7, 76.1, 75.3, 67.0, 52.1, 51.6, 44.5, 44.1, 31.6, 30.8, 25.0. 1 x C bearing B not observed. Rotameric mixture observed.

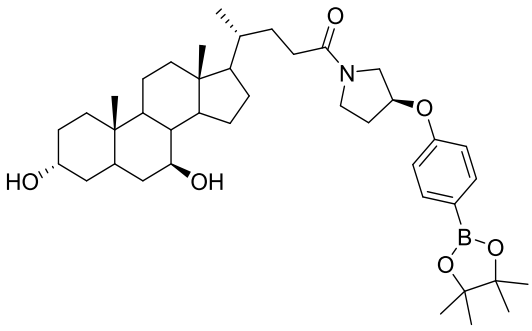
^{13}C NMR Variable Temperature (DMSO-d_6, 101 MHz, 373 K): 159.2, 153.6, 136.7, 135.7, 127.8, 127.1, 126.8, 114.7, 82.9, 75.5, 65.4, 50.9, 43.4, 30.1, 24.1. 1 x C bearing B not observed.

^{11}B NMR ($CDCl_3$, 128 MHz): 30.6.

ν_{max} (neat): 2979, 2887, 1700, 1605, 1415, 1359 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{24}H_{31}BNO_5$) requires 424.2290 *m/z*, found 424.2278 *m/z*.

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-((*R*)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)pyrrolidin-1-yl)pentan-1-one **68**



Benzyl (*R*)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)pyrrolidine-1-carboxylate **67** (269 mg, 0.63 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure J to generate crude intermediate **67a**. Compound **68** was prepared according to General Procedure E using UDCA (302 mg, 0.77 mmol, 1.0 equiv.), **67a** (231 mg, 0.77 mmol, 1.0 equiv), DMF (5 mL), DIPEA (405 μL, 2.32 mmol, 3.0 equiv.) and HATU (323 mg, 0.85 mmol, 1.1 equiv.). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-6% MeOH in DCM) to afford the desired product **68** as a white solid (151 mg, 30%).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.76-7.72 (m, 2H), 6.87-6.84 (m, 2H), 5.03-4.95 (m, 1H), 3.82-3.56 (m, 6H), 2.37-2.19 (m, 3H), 2.16-2.07 (m, 1H), 2.01-1.90 (m, 2H), 1.83-1.76 (m, 5H), 1.68-1.54 (m, 6H), 1.50-1.41 (m, 7H), 1.33 (s, 12H), 1.25-1.15 (m, 4H), 1.10-1.01 (m, 2H), 0.94-0.91 (m, 6H), 0.68-0.67 (m, 3H). Rotameric mixture observed.

^{13}C NMR ($CDCl_3$, 126 MHz): 172.6, 172.5, 159.7, 159.6, 136.82, 136.77, 114.94, 114.87, 83.82, 83.76, 76.3, 74.8, 71.6, 71.51, 71.48, 55.9, 55.8, 55.2, 55.1, 52.2, 51.6, 44.8, 43.9, 43.9, 42.6, 40.30, 40.27, 39.3, 37.5, 37.0, 35.6, 35.1, 34.2, 32.0, 31.8, 31.7, 31.1, 31.0, 30.5, 30.2, 28.8, 27.1, 25.0, 23.5, 21.3, 18.80, 18.75, 12.3. 1 x C bearing B not observed. Rotameric mixture observed.

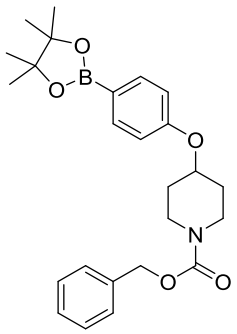
^{11}B NMR ($CDCl_3$, 128 MHz): 31.1.

ν_{max} (neat): 3407, 2930, 2867, 1629, 1606 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{40}H_{62}BNO_6Na$) requires 686.4569 *m/z*, found 686.4558 *m/z*.

Experimental

Preparation of benzyl 4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidine-1-carboxylate **69**



Prepared according to General Procedure G, using benzyl-4-hydroxypiperidine-1-carboxylate (479 mg, 2.04 mmol, 1.5 equiv.) 4-hydroxyphenol boronic acid pinacol ester (300 mg, 1.36 mmol, 1.0 equiv.), CMBP (715 µL, 2.73 mmol, 2.0 equiv.) in anhydrous toluene (6 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-12% EtOAc in hexane) to afford the desired product **69** as a white solid (400 mg, 67 %) (75% on 1.8 mmol scale).

^{1}H NMR ($CDCl_3$, 400 MHz): 7.74 (d, 2H, *J* = 8.5 Hz), 7.31-7.37 (m, 5H), 6.89 (d, 2H, *J* = 8.5 Hz), 5.14 (s, 2H), 4.54-4.58 (m, 1H), 3.70-3.75 (m, 2H), 3.46-3.51 (m, 2H), 1.92 (br. s, 2H), 1.80 (br. s, 2H), 1.33 (s, 12H).

^{13}C NMR ($CDCl_3$, 101 MHz): 159.9, 155.5, 137.0, 136.8, 128.7, 128.2, 128.0, 115.3, 83.8, 71.4, 67.3, 40.8, 30.2, 25.0.

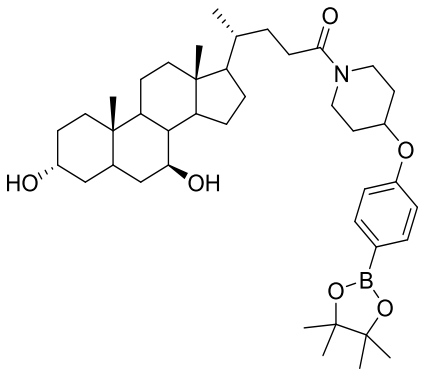
^{11}B NMR ($CDCl_3$, 126 MHz): 31.9.

ν_{max} (neat): 2977, 2941, 2873, 1690, 1603, 1454 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{25}H_{33}BNO_5$) requires 438.2456 *m/z*, found 438.2464 *m/z*.

Experimental

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-(4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidin-1-yl)pentan-1-one **70**



Benzyl 4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidine-1-carboxylate **69** (300 mg, 0.80 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure J to generate crude intermediate 4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidine **69a**. Compound **70** was prepared according to General Procedure E, using UDCA (263 mg, 0.67 mmol, 1.0 equiv.), 4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidine **69a** (204 mg, 0.67 mmol, 1.0 equiv), DMF (5 mL), DIPEA (350 µL, 3.0 mmol, 3.0 equiv.) and HATU (281 mg, 0.74 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-5% MeOH in DCM) to afford the desired product **70** as a white solid (167 mg, 37 %).

^{1}H NMR ($CDCl_3$, 400 MHz): 7.75 (d, 2H, *J* = 8.2 Hz), 6.90 (d, 2H, *J* = 8.2 Hz), 4.61-4.60 (m, 1H), 3.74-3.66 (m, 3H), 3.61-3.55 (m, 2H), 3.42-3.40 (m, 1H), 2.42-2.37 (m, 1H), 2.26-2.22 (m, 1H), 2.01-1.99 (m, 1H), 1.96-1.88 (m, 3H), 1.83-1.76 (m 5H), 1.68-1.66 (m, 2H), 1.62-1.51 (m, 5H), 1.48-1.41 (m, 5H), 1.33 (s, 12H), 1.28-1.21 (m, 4H), 1.18-0.99 (m, 4H), 0.96-0.95 (m, 6H), 0.89-0.80 (m, 2H), 0.68 (s, 3H). Rotameric mixture observed.

^{13}C NMR ($CDCl_3$, 126 MHz): 172.2, 159.8, 136.8, 115.3, 83.8, 71.6, 71.5, 71.3, 55.9, 55.2, 44.0, 42.6, 42.5, 40.3, 39.3, 38.4, 37.5, 37.0, 35.8, 35.1, 34.2, 31.7, 31.3, 30.6,

30.5, 30.3, 28.9, 27.1, 25.0, 23.5, 21.3, 18.8, 12.3. 1 x C bearing B not observed, 1 x C not observed. Rotameric mixture observed.

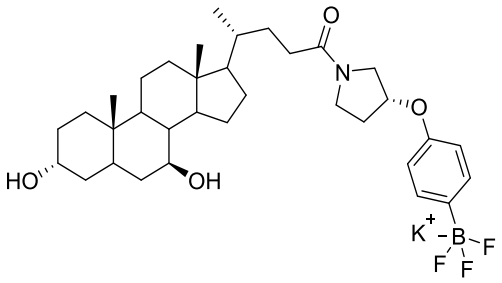
^{11}B NMR ($CDCl_3$, 160 MHz): 31.4.

ν_{max} (neat): 3446, 2928, 2865, 1629, 1606 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{41}H_{65}BNO_6$) requires 678.4912 *m/z,* found 678.4907 *m/z.*

6.3.7 Cyclic Trifluoroborate and Boronic Acid Analogues

Preparation of potassium (4-(((3*R*)-1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)pyrrolidin-3-yl)oxy)phenyl)trifluoroborate **71**



Prepared according to General Procedure H, using (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-((*S*)-3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)pyrrolidin-1-yl)pentan-1-one **66** (122 mg, 0.18 mmol, 1.0 equiv.), potassium fluoride (43 mg, 0.73 mmol, 1.0 equiv) in H_2O (0.75 mL), *L*-(+)-tartaric acid (56 mg, 0.37 mmol, 2.05 equiv.) and MeOH/MeCN (0.75 mL:0.75 mL, 1:1). The crude material carried through to the next step as a white solid (70.1 mg, 61 %) without further purification.

^{1}H NMR (DMSO-d_6, 126 MHz): 7.23-7.20 (m, 2H), 6.68-6.65 (m, 2H), 4.97-4.89 (m, 1H), 4.41 (app. br. s., 1H), 3.92-3.72 (m, 2H), 3.61-3.48 (m, 3H), 2.32-2.09 (m, 3H), 2.02-1.83 (m, 3H), 1.76-1.66 (m, 4H), 1.46-1.29 (m, 10H), 1.27-1.09 (m, 8H), 0.91-0.87 (m, 7H), 0.62-0.61 (m, 3H). Rotameric mixture observed.

^{13}C NMR (DMSO-d_6, 126 MHz): 171.1, 171.0, 154.6, 154.6, 132.4, 132.4, 113.82, 113.75, 76.0, 74.3, 73.5, 69.7, 69.4, 55.8, 54.7, 51.6, 51.0, 43.1, 43.0, 42.2, 38.7, 37.7, 37.3, 35.0, 34.9, 34.8, 33.7, 31.1, 30.8, 30.7, 30.6, 30.5, 30.2, 29.5, 28.2, 26.7, 23.3, 20.8, 18.6, 12.0. 1 x C bearing B not observed. Rotameric mixture observed.

^{11}B NMR (DMSO-d_6, 160 MHz): 3.4.

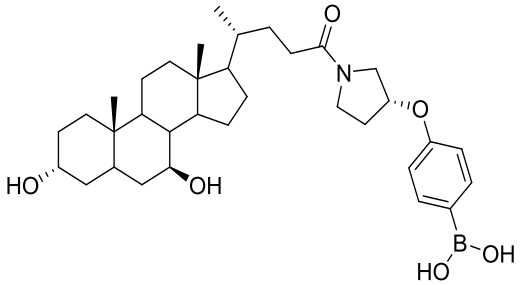
^{19}F NMR (DMSO-d_6, 376 MHz): -138.4.

ν_{max} (neat): 3319, 2928, 2865, 1603, 1448 cm^{-1}.

HRMS: exact mass calculated for $[M-K]^-$ ($C_{34}H_{50}BF_3NO_4$) requires 604.3796 *m/z*, found 604.3796 *m/z.*

Experimental

Preparation of (4-(((3*R*)-1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)pyrrolidin-3-yl)oxy)phenyl)boronic acid **72**



Prepared according to General Procedure I, using potassium (4-(((3*R*)-1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)pyrrolidin-3-yl)oxy)phenyl)trifluoroborate **71** (31 mg, 0.05 mmol, 1.0 equiv.), excess SiO_2 and H_2O (3 mL). The crude material was subjected to the purification outlined in General Procedure I (silica gel, 0-12% MeOH in DCM) to afford the desired product **72** as a white solid (10.1 mg, 33 %).

^{1}H NMR (MeOD, 400 MHz): 7.75-7.62 (m, 2H), 6.98-6.90 (m, 2H), 5.16-5.09 (m, 1H), 3.87-3.64 (m, 4H), 3.56-3.47 (m, 2H), 2.50-2.21 (m, 2H), 2.09-2.02 (m, 1H), 1.95-1.74 (m, 5H), 1.70-1.57 (m, 4H), 1.55-1.43 (m, 6H), 1.39-1.06 (m, 10H), 1.04-0.96 (m, 6H), 0.75-0.71 (m, 3H). Rotameric mixture observed.

^{13}C NMR (MeOD, 126 MHz): 175.22, 175.19, 136.8, 136.7, 115.9, 115.8, 77.5, 76.0, 72.1, 72.0, 57.5, 57.5, 56.6, 56.5, 53.5, 52.6, 46.1, 45.1, 44.8, 44.8, 44.5, 44.1, 41.6, 41.5, 40.7, 38.6, 38.6, 38.0, 36.89, 36.86, 36.1, 35.2, 32.7, 32.5, 32.3, 31.0, 30.9, 30.7, 29.7, 29.6, 27.98, 27.96, 23.9, 22.4, 19.2, 12.7. 1 x C bearing B not observed. Rotameric mixture observed.

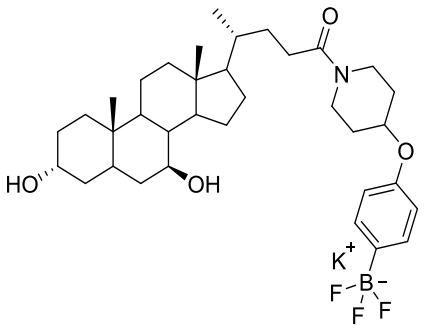
Required use of MeOD resulted in partial OH for D exchange; however this can be rectified by utilising DMSO as the NMR solvent.

^{11}B NMR (MeOD, 160 MHz): 28.6.

HRMS: exact mass calculated for $[M+Ethylene\ Glycol+H]^+$ ($C_{35}H_{55}BNO_6$) requires 608.4129 *m/z*, found 608.4130 *m/z.*

Experimental

Preparation of potassium (4-((1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)piperidin-4-yl)oxy)phenyl)trifluoroborate **73**



Prepared according to General Procedure H, using (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-(4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidin-1-yl)pentan-1-one **70** (120 mg, 0.18 mmol, 1.0 equiv.), potassium fluoride (42 mg, 0.71 mmol, 1.0 equiv) in H_2O (1 mL), *L*-(+)-tartaric acid (55 mg, 0.36 mmol, 2.05 equiv.) in THF (0.5 mL) and MeOH/MeCN (0.75 mL:0.75 mL, 1:1). The crude material carried through to the next step as a white solid (108 mg, 50 %) without further purification.

^{1}H NMR (DMSO-d_6, 400 MHz): 7.21 (d, 2H, J = 8.0 Hz), 6.90 (d, 2H, J = 8.0 Hz), 4.49-4.46 (m, 1H), 4.42-4.41 (m, 1H), 3.86-3.81 (m, 2H), 3.72-3.63 (m, 1H), 3.24-3.16 (m, 1H), 2.36-2.30 (m, 1H), 2.26-2.16 (m, 1H), 1.95-1.83 (m, 4H), 1.79-1.72 (m, 1H), 1.68-1.64 (m, 3H), 1.58-1.54 (m, 1H), 1.51-1.45 (m, 4H), 1.42-1.27 (m, 7H), 1.24-1.12 (m, 5H), 1.11-1.02 (m, 6H), 0.91-0.90 (m, 3H), 0.87 (s, 3H), 0.62 (s, 3H).

^{13}C NMR (DMSO-d_6, 126 MHz): 170.9, 154.7, 132.3, 114.3, 71.5, 69.7, 69.4, 55.8, 54.7, 43.1, 43.0, 42.2, 38.7, 38.2, 37.7, 37.3, 35.0, 34.8, 33.7, 31.3, 31.2, 30.5, 30.2, 29.5, 28.2, 26.7, 23.3, 20.8, 18.5, 12.0. 1 x C bearing B not observed.

ν_{max} (neat) 3399, 2930, 2865, 1605, 1605, 1454 cm^{-1}.

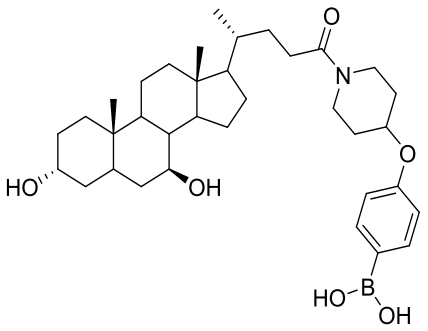
^{11}B NMR (DMSO-d_6, 101 MHz): 3.38.

^{19}F NMR (DMSO-d_6, 471 MHz): -138.4.

HRMS: exact mass calculated for $[M-K]^-$ ($C_{35}H_{52}BF_3NO_4$) requires 618.3953 *m/z*, found 618.3953 *m/z*.

Experimental

Preparation of (4-((1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)piperidin-4-yl)oxy)phenyl)boronic acid **74**



Prepared according to General Procedure I, using potassium (4-((1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)piperidin-4-yl)oxy)phenyl)trifluoroborate **73** (100 mg, 0.11 mmol, 1.0 equiv.), excess SiO_2 and H_2O (5 mL). The crude material was subjected to the purification outlined in General Procedure I (silica gel, 0-12% MeOH in DCM) to afford the desired product **74** as a white solid (45 mg, 50 %).

^{1}H NMR (DMSO-d_6, 500 MHz): 7.82 (app. br. s., 2H), 7.72 (d, 2H, *J* = 8.3 Hz), 6.92 (d, 2H, *J* = 8.3 Hz), 4.69-4.62 (m, 1H), 4.48-4.42 (m, 1H), 3.87-3.85 (m, 2H), 3.75-3.68 (m, 1H), 3.16-3.32 (m, coincident with solvent), 2.36-2.32 (m, 1H), 2.23-2.18 (m, 1H), 1.95-1.82 (m, 5H), 1.75-1.57 (m, 6H), 1.48-1.42 (m, 4H), 1.41-1.29 (m, 6H), 1.23-1.03 (m, 8H), 0.91-0.87 (m, 6H), 0.62 (s, 3H).

^{13}C NMR (DMSO-d_6, 101 MHz): 170.9, 158.7, 135.9, 114.7, 71.5, 69.7, 69.5, 55.9, 54.7, 43.1, 43.0, 42.2, 38.7, 38.2, 37.7, 37.3, 35.1, 34.8, 33.8, 31.2, 31.1, 30.3, 30.2, 29.5, 28.2, 26.8, 23.3, 20.8, 18.6, 12.0. 1 x C bearing B not observed.

^{11}B NMR (MeOD, 128 MHz): 29.7.

ν_{max} (neat): 3360, 2928, 2863, 1601 cm^{-1}.

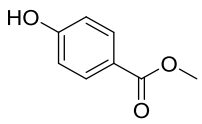
HRMS: exact mass calculated for [M+Ethylene Glycol+H]$^+$ ($C_{37}H_{57}BNO_6$) requires 622.4285 *m/z*, found 622.4288 *m/z*.

Experimental

Purity 99.6% by analytical HPLC.

6.3.8 Alternative Warhead Analogues

Preparation of methyl 4-hydroxybenzoate **76**



To a round-bottomed flask was added 4-hydroxybenzoic acid (1.58 g, 11.4 mmol, 1.0 equiv.) in methanol (15 mL). Thionyl chloride (1 mL, 13.7 mmol, 1.2 equiv.) was then added dropwise over 15 minutes. The reaction mixture was stirred at reflux (~ 65 °C) for 4 hours. The solvent was removed *in vacuo* and the crude material was carried through to the next step without further purification as an off-white solid (1.68 g, 97%).

^{1}H NMR (DMSO-d_6, 400 MHz): 10.3 (s, 1H), 7.81 (d, J = 8.7 Hz, 2H), 6.84 (d, J = 8.7 Hz, 2H), 3.78 (s, 3H).

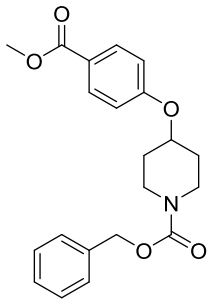
^{13}C NMR (DMSO-d_6, 101 MHz): 166.0, 161.9, 131.4, 120.2, 115.3, 51.6.

Consistent with previously reported spectral analysis.[173]

ν_{max} (neat): 3299, 1681, 1608, 1590, 1515, 1435 cm^{-1}.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_8H_8O_3Na$) requires 175.0366 *m/z*, found 175.0364 *m/z*.

Preparation of benzyl 4-(4-(methoxycarbonyl)phenoxy)piperidine-1-carboxylate **77**



Prepared according to General Procedure G, using benzyl-4-hydroxypiperidine-1-carboxylate (231 mg, 0.98 mmol, 1.5 equiv.) methyl 4-hydroxybenzoate **76** (100 mg, 0.66 mmol, 1.0 equiv.), CMBP (344 μL, 1.32 mmol, 2.0 equiv.) in anhydrous toluene (5 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-20% EtOAc in hexane) to afford the desired product **77** as a yellow oil (188 mg, 77 %).

^{1}H NMR ($CDCl_3$, 500 MHz): 7.97 (d, *J* = 8.8 Hz, 2H), 7.37-7.30 (m, 5H), 6.91 (d, *J* = 8.8 Hz, 2H), 5.15 (s, 2H), 4.59 (m, 1H), 3.88 (s, 3H), 3.77-3.72 (m, 2H), 3.52-3.47 (m, 2H), 1.94 (app. br. s, 2H), 1.81 (app. br. s, 2H).

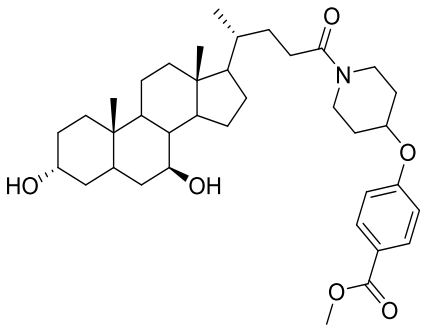
^{13}C NMR ($CDCl_3$, 126 MHz): 166.9, 161.2, 155.4, 136.9, 131.8, 128.7, 128.2, 128.1, 123.0, 115.3, 71.9, 67.4, 52.0, 40.8, 30.4.

ν_{max} (neat): 2953, 2871, 1696, 1606, 1435 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{21}H_{24}NO_5$) requires 370.1649 *m/z*, found 370.1648 *m/z*.

Experimental

Preparation of methyl 4-((1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)piperidin-4-yl)oxy)benzoate **78**



Benzyl 4-(4-(methoxycarbonyl)phenoxy)piperidine-1-carboxylate **77** (269 mg, 0.63 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure J to generate crude intermediate **77a.** Compound **78** was prepared according to General Procedure E, using UDCA (169 mg, 0.43 mmol, 1.0 equiv.), **77a** (100 mg, 0.43 mmol, 1.0 equiv), DMF (3 mL), DIPEA (223 μL, 1.28 mmol, 3.0 equiv.) and HATU (178 mg, 0.47 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-5% MeOH in DCM) to afford the desired product **78** as a white solid (141 mg, 54 %).

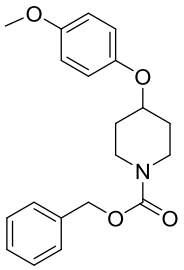
^{1}H NMR (DMSO-d_6, 500 MHz): 7.90 (d, J = 8.9 Hz, 2H), 7.09 (d, J = 8.9 Hz, 2H), 4.77-4.72 (m, 1H), 4.42 (d, J = 4.6 Hz, 1H), 3.89-3.85 (m, 2H), 3.81 (s, 3H), 3.72-3.70 (m, 1H), 3.37-3.20 (m, 3H, solvent overlap), 2.37-2.32 (m, 1H), 2.25-2.19 (m, 1H), 2.01-1.90 (m, 3H), 1.82-1.82 (m, 1H), 1.79-1.72 (m, 1H), 1.69-1.60 (m, 4H), 1.53-1.44 (m, 4H), 1.39-1.31 (m, 6H), 1.22-1.08 (m, 7H), 1.06-1.00 (m, 1H), 0.92-0.87 (m, 6H), 0.62 (s, 3H).

^{13}C NMR (DMSO-d_6, 126 MHz): 170.9, 165.8, 161.0, 131.3, 121.8, 115.4, 72.2, 69.7, 69.4, 55.8, 54.7, 51.8, 43.1, 43.0, 42.2, 38.7, 38.1, 37.7, 37.3, 35.0, 34.8, 33.7, 31.2, 31.0, 30.2, 30.1, 29.4, 28.2, 26.7, 23.3, 20.8, 18.5, 12.0.

ν_{max} (neat) 3440, 3429, 2930, 2865, 1716, 1606, 1512, 1437 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{37}H_{56}NO_6$) requires 610.4108 *m/z*, found 610.4110 *m/z*.

Preparation of benzyl 4-(4-methoxyphenoxy)piperidine-1-carboxylate **79**



Prepared according to General Procedure G, using benzyl-4-hydroxypiperidine-1-carboxylate (425 mg, 1.81 mmol, 1.5 equiv.) methyl 4-hydroxybenzoate (150 mg, 1.21 mmol, 1.0 equiv.), CMBP (659 μL, 2.42 mmol, 2.0 equiv.) in anhydrous toluene (5 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-30% EtOAc in hexane) to afford the desired product **79** as a yellow oil (328 mg, 79 %).

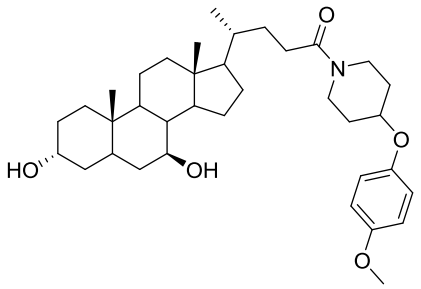
^{1}H NMR ($CDCl_3$, 500 MHz): 7.37-7.30 (m, 5H), 6.87-6.82 (m, 4H), 5.14 (s, 2H), 4.35 (tt, *J* = 6.9, 3.4 Hz, 1H), 3.79-3.74 (m, 5H, multiplet coincident with OMe), 3.45-3.40 (m, 2H), 1.89 (app. br. s, 2H), 1.77 (app. br. s, 2H).

^{13}C NMR ($CDCl_3$, 126 MHz): 155.5, 154.5, 151.2, 137.0, 128.7, 128.1, 128.0, 118.0, 114.9, 73.2, 67.3, 55.8, 41.0, 30.6.

ν_{max} (neat): 2951, 2871, 2835, 1696, 1508, 1430 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{20}H_{24}NO_4$) requires 342.1700 *m/z*, found 342.1694 *m/z.*

Preparation of (4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-(4-(4-methoxyphenoxy)piperidin-1-yl)pentan-1-one **80**



Benzyl 4-(4-methoxyphenoxy)piperidine-1-carboxylate **79** (269 mg, 0.63 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure J to generate crude intermediate **79a**. Prepared according to General Procedure E, using UDCA (361 mg, 0.92 mmol, 1.0 equiv.), **79a** (191 mg, 0.92 mmol, 1.0 equiv), DMF (5 mL), DIPEA (482 μL, 2.75 mmol, 3.0 equiv.) and HATU (384 mg, 1.01 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-5% MeOH in DCM) to afford the desired product **80** as a white solid (183 mg, 34 %).

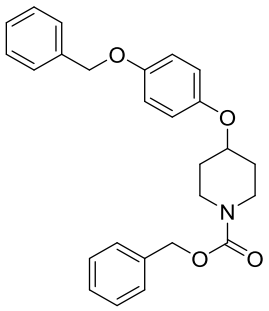
^{1}H NMR (DMSO-d_6, 500 MHz): 6.92-6.84 (m, 4H), 4.47-4.41 (m, 2H), 3.86-3.80 (m, 2H), 3.69-3.64 (m, 4H), 3.35-3.18 (m, solvent overlap, 3H), 2.36-2.30 (m, 1H), 2.24-2.18 (m, 1H), 1.95-1.82 (m, 4H), 1.78-1.63 (m, 4H), 1.59-1.52 (m, 1H), 1.50-1.39 (m, 5H), 1.38-1.27 (m, 6H), 1.24-1.05 (m, 8H), 0.91-0.87 (m, 6H), 0.62 (s, 3H). Rotameric mixture observed.

^{13}C NMR ($CDCl_3$, 126 MHz): 172.2, 159.5, 151.3, 117.9, 114.9, 73.1, 71.6, 71.5, 55.9, 55.8, 55.2, 44.0, 42.6, 40.3, 39.3, 38.5, 37.5, 37.0, 35.7, 35.1, 34.2, 31.7, 31.5, 30.6, 30.5, 30.4, 28.9, 27.1, 23.5, 21.3, 18.8, 12.3. 2 x C not observed, rotameric mixture observed.

ν_{max} (neat): 3386, 2928, 2863, 1623, 1508, 1444 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{36}H_{56}NO_5$) requires 582.4158 *m/z*, found 582.4163 *m/z*.

Preparation of benzyl 4-(4-(benzyloxy)phenoxy)piperidine-1-carboxylate **81**



Prepared according to General Procedure G, using benzyl-4-hydroxypiperidine-1-carboxylate (300 mg, 1.27 mmol, 1.5 equiv.) 4-benzyloxyphenol (200 mg, 0.85 mmol, 1.0 equiv.), CMBP (463 μL, 1.70 mmol, 2.0 equiv.) in anhydrous toluene (5 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-6% EtOAc in hexane) to afford the desired product **81** as a clear oil (339 mg, 96 %).

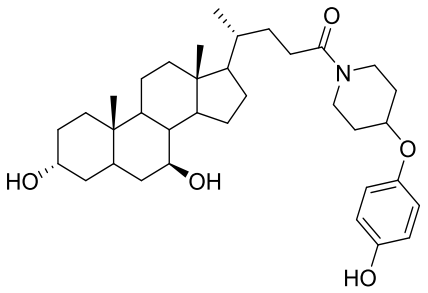
^{1}H NMR ($CDCl_3$, 500 MHz): 7.43-7.31 (m, 10H), 6.88 (m, 4H), 5.14 (s, 2H), 5.02 (s, 2H), 4.35 (m, 1H), 3.78-3.73 (m, 2H), 3.45-3.40 (m, 2H), 1.95-1.84 (m, 2H), 1.81-1.70 (m, 2H).

^{13}C NMR ($CDCl_3$, 126 MHz): 155.5, 153.6, 151.4, 137.4, 137.0, 128.7, 128.7, 128.1, 128.06, 128.0, 127.6, 117.9, 116.0, 73.1, 70.8, 67.3, 41.0, 30.7.

ν_{max} (neat): 2949, 2865, 1696, 1506, 1430 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{26}H_{28}NO_4$) requires 418.2018 *m/z*, found 418.2014 *m/z*.

Preparation of (4R)-4-((3R,7S,10S,13R)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1H-cyclopenta[a]phenanthren-17-yl)-1-(4-(4-hydroxyphenoxy)piperidin-1-yl)pentan-1-one **82**



Benzyl 4-(4-(benzyloxy)phenoxy)piperidine-1-carboxylate **81** (310 mg, 0.74 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure J to generate crude intermediate **81a.** Prepared according to General Procedure E, using UDCA (282 mg, 0.72 mmol, 1.0 equiv.), 4-(piperidin-4-yloxy)phenol **81a** (141 mg, 0.72 mmol, 1.0 equiv), DMF (6 mL), DIPEA (377 μL, 2.16 mmol, 3.0 equiv.) and HATU (301 mg, 0.79 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-4% MeOH in DCM) to afford the desired product **82** as a white solid (284 mg, 69%).

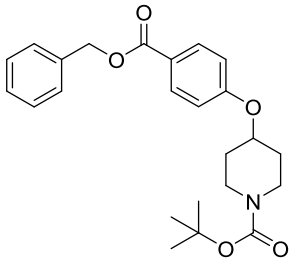
^{1}H NMR (MeOD, 500 MHz): 6.81 (d, 2H, *J* = 8.8 Hz), 6.70 (d, 2H, *J* = 8.8 Hz), 4.43-4.41 (m, 1H), 3.85-3.74 (m, 2H), 3.52-3.45 (m, 4H), 2.49-2.43 (m, 1H), 2.35-2.29 (m, 1H), 1.06-2.03 (m, 1H), 1.98-1.80 (m, 7H), 1.76-1.67 (m, 2H), 1.67-1.54 (m, 5H), 1.49-1.44 (m, 6H), 1.34-1.19 (m, 6H), 1.19-1.03 (m, 2H), 1.00-0.99 (m, 3H), 0.97 (s, 3H), 0.72 (s, 3H). Rotameric mixture observed.

^{13}C NMR (MeOD, 126 MHz): 174.7, 153.1, 151.7, 119.3, 117.1, 74.2, 74.6, 72.2, 72.0, 57.7, 56.6, 44.9, 44.7, 44.0, 41.6, 40.8, 39.9, 38.6, 38.1, 37.0, 36.2, 35.2, 33.1, 32.6, 31.7, 31.4, 31.1, 29.7, 27.9 23.9, 22.4, 19.2, 12.7. Rotameric mixture observed.

ν_{max} (neat): 3358, 2930, 2865, 1713, 1512, 1448 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{35}H_{54}NO_5$) requires 568.4002 *m/z*, found 568.4009 *m/z.*

Preparation of *tert*-butyl 4-(4-((benzyloxy)carbonyl)phenoxy)piperidine-1-carboxylate **87**



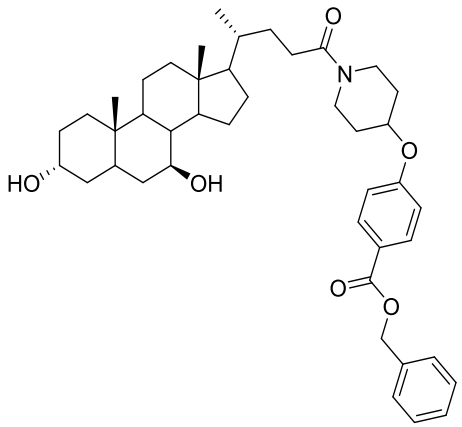
Prepared according to General Procedure G, using *tert*-butyl 4-hydroxypiperidine-1-carboxylate (237 mg, 1.18 mmol, 1.5 equiv.) benzyl 4-hydroxybenzoate (180 mg, 0.79 mmol, 1.0 equiv.), CMBP (429 μL, 1.58 mmol, 2.0 equiv.) in anhydrous toluene (6 mL). The crude material was subjected to the purification outlined in General Procedure G (silica gel, 0-5% EtOAc in hexane) to afford the desired product **87** as a white solid (294 mg, 91 %).

^{1}H NMR ($CDCl_3$, 500 MHz): 8.02 (d, 2H, *J* = 8.8 Hz), 7.44-7.32 (m, 5H), 6.91 (d, 2H, *J* = 8.8 Hz), 5.34 (s, 9H), 4.58-4.54 (m, 1H), 3.71-3.66 (m, 2H), 3.39-3.34 (m, 2H), 1.95-1.91 (m, 2H), 1.79-1.73 (m, 2H), 1.47 (s, 9H).

^{13}C NMR ($CDCl_3$, 126 MHz): 166.2, 161.4, 155.0, 136.5, 132.0, 128.7, 128.3, 128.3, 122.8, 115.4, 79.9, 72.4, 66.6, 30.5, 28.6.

HRMS: exact mass calculated for $[M+NH_4]^+$ ($C_{24}H_{33}N_2O_5$) requires 429.2384 *m/z*, found 429.2385 *m/z*.

Preparation of benzyl 4-((1-((4*R*)-4-((3*R*,7*S*,10*S*,13*R*)-3,7-dihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentanoyl)piperidin-4-yl)oxy)benzoate **88**



tert-butyl 4-(4-((benzyloxy)carbonyl)phenoxy)piperidine-1-carboxylate **87** (290 mg, 0.70 mmol, 1.0 equiv.) was subjected to deprotection conditions in accordance with General Procedure D using a 1:1 mixture of TFA and DCM (2 mL: 2 mL) to generate crude intermediate 4-(4-((benzyloxy)carbonyl)phenoxy)piperidin-1-ium 2,2,2-trifluoroacetate **87a**. Compound **88** was prepared according to General Procedure E, using UDCA (282 mg, 0.72 mmol, 1.0 equiv.), 4-(4-((benzyloxy)carbonyl)phenoxy)piperidin-1-ium 2,2,2-trifluoroacetate **87a** (218 mg, 0.70 mmol, 1.0 equiv), DMF (6 mL), DIPEA (489 μL, 2.80 mmol, 3.0 equiv.) and HATU (292 mg, 0.7 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-3% MeOH in DCM) to afford the desired product **88** as a white solid (215 mg, 45 % over two steps).

^{1}H NMR ($CDCl_3$, 500 MHz): 8.03 (d, 2H, *J* = 8.8 Hz), 7.44-7.32 (m, 5H), 6.92 (d, 2H, *J* = 8.8 Hz), 5.34 (s, 2H), 4.65-4.61 (m, 1H), 3.79-3.75 (m, 1H), 3.71-3.64 (m, 2H), 3.61-3.56 (m, 2H), 3.45-3.40 (m, 1H), 2.43-2.36 (m, 1H), 2.28-2.21 (m, 1H), 2.02-1.90 (m, 4H), 1.86-1.76 (m, 6H), 1.68-1.66 (m, 2H), 1.62-1.51 (m, 3H), 1.49-1.40 (m, 8H), 1.39-1.23 (m, 6H), 1.18-1.02 (m, 3H), 0.96-0.95 (m, 6H), 0.68 (s, 3H). Rotameric mixture observed.

Experimental

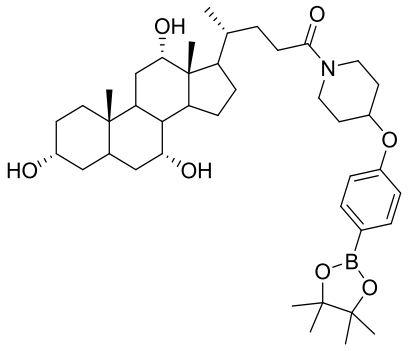
^{13}C NMR ($CDCl_3$, 126 MHz): 172.2, 166.2, 161.2, 136.4, 132.0, 128.7, 128.32, 128.26, 123.0, 115.4, 71.9, 71.6, 71.5, 66.6, 55.9, 55.2, 44.0, 42.6, 42.4, 40.3, 39.3, 38.3, 37.5, 37.0, 35.7, 35.1, 34.2, 31.7, 31.2, 30.6, 30.5, 30.2, 28.9, 27.1, 23.5, 21.3, 18.8, 12.3. 1 x C not observed, rotameric mixture observed.

ν_{max} (neat): 3392, 2930, 2863, 1714, 1605 cm^{-1}.

HRMS: exact mass calculated for $[M+H]^+$ ($C_{43}H_{60}NO_6$) requires 686.4415 *m/z*, found 686.4410 *m/z.*

6.3.9 Alternative Steroid Analogues

Preparation of (4*R*)-1-(4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidin-1-yl)-4-((3*R*,7*R*,10*S*,12*S*,13*R*)-3,7,12-trihydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)pentan-1-one **91**



Prepared according to General Procedure E, using cholic acid (204 mg, 0.50 mmol, 1.0 equiv.), **69a** (150 mg, 0.50 mmol, 1.0 equiv), DMF (3 mL), DIPEA (350 μL, 4.0 mmol, 3.0 equiv.) and HATU (209 mg, 0.55 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-7% MeOH in DCM) to afford the desired product **91** as a white solid (167 mg, 46 %).

^{1}H NMR ($CDCl_3$, 400 MHz): 7.75 (d, 2H, *J* = 8.6 Hz), 6.90 (d, 2H, *J* = 8.6 Hz), 4.63-4.58 (m, 1H), 3.97 (br.s, 1H), 3.84-3.82 (m, 1H), 3.77-3.63 (m, 3H), 3.48-3.38 (m, 2H), 2.46-2.37 (m, 1H), 2.30-2.17 (m, 4H), 1.97-1.87 (m, 7H), 1.82-1.75 (m, 5H), 1.71-1.62 (m, 3H), 1.59-1.48 (m, 4H), 1.44-1.38 (m, 4H), 1.33 (s, 12H), 1.30-1.24 (m, 2H), 1.20-1.06 (m, 2H), 1.02-1.00 (m, 3H), 0.89 (s, 3H), 0.69 (s, 3H). Rotameric mixture observed.

^{13}C NMR ($CDCl_3$, 101 MHz): 172.3, 159.8, 136.8, 115.3, 83.8, 73.1, 72.1, 71.3, 68.6, 47.1, 46.7, 42.5, 42.0, 41.6, 39.8, 39.7, 38.4, 35.6, 35.4, 34.9, 34.8, 31.5, 31.3, 30.7, 30.4, 30.3, 28.4, 27.7, 26.7, 25.0, 23.4, 22.7, 17.7, 12.7. Rotameric mixture observed.

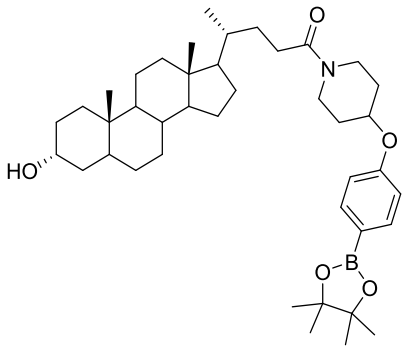
^{11}B NMR ($CDCl_3$, 160 MHz): 31.6.

ν_{max} (neat): 3396, 2930, 2867, 1605 cm^{-1}.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_{41}H_{64}BNO_7Na$) requires 716.4669 *m/z*, found 716.4666 *m/z*.

Experimental

Preparation of (4*R*)-4-((3*R*,10*S*,13*R*)-3-hydroxy-10,13-dimethylhexadecahydro-1*H*-cyclopenta[*a*]phenanthren-17-yl)-1-(4-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)piperidin-1-yl)pentan-1-one **92**



Prepared according to General Procedure E, using lithocholic acid (186 mg, 0.50 mmol, 1.0 equiv.), **69a** (150 mg, 0.50 mmol, 1.0 equiv), DMF (3 mL), DIPEA (350 μL, 4.0 mmol, 3.0 equiv.) and HATU (209 mg, 0.55 mmol, 1.1 equiv). The crude material was subjected to the purification outlined in General Procedure E (silica gel, 0-2% MeOH in DCM) to afford the desired product **92** as a pale pink solid (181 mg, 45 %).

^{1}H NMR ($CDCl_3$, 400 MHz): 7.75 (d, 2H, *J* = 8.1 Hz), 6.90 (d, 2H, *J* = 8.1 Hz), 4.61-4.59 (m, 1H), 3.78-3.73 (m, 1H), 3.70-3.60 (m, 3H), 3.43-3.40 (m, 1H), 2.43-2.36 (m, 1H), 2.27-2.20 (m, 1H), 1.98-1.75 (m, 10H), 1.67-1.65 (m, 1H), 1.58-1.57 (m,1H), 1.46-1.38 (m, 8H), 1.33 (s, 12H), 1.27-1.20 (m, 4H), 1.17-1.05 (m, 5H), 0.95-0.93 (m, 3H), 0.92 (s, 3H), 0.65 (s, 3H). Rotameric mixture observed.

^{13}C NMR ($CDCl_3$, 126 MHz): 172.3, 159.9, 136.8, 115.3, 83.8, 72.0, 71.4, 56.7, 56.2, 42.9, 42.5, 42.3, 40.6, 40.3, 38.4, 36.7, 36.0, 35.9, 35.5, 34.7, 31.7, 31.3, 30.7, 30.6, 30.3, 28.4, 27.4, 26.6, 25.0, 24.4, 23.5, 21.0, 18.7, 12.2. Rotameric mixture observed.

^{11}B NMR ($CDCl_3$, 160 MHz): 31.8.

ν_{max} (neat): 3418, 2932, 2865, 1623, 1606 cm^{-1}.

HRMS: exact mass calculated for $[M+Na]^+$ ($C_{41}H_{64}BNO_5Na$) requires 684.4777 *m/z*, found 684.4769 *m/z*.

7 Appendices

7.1 Appendix 1 – General Biological Protocols

7.1.1 Protein Production and Crystallisation

Rat ATX was over-expressed and purified as described previously. For the crystallisation studies, ATX was incubated with each screened compound at a 1:10 (protein:compound) ratio for at least 30 minutes. Crystals were grown for at least 7 days in a 24-well optimization screen (18−20% PEG 3350, 0.1−0.4 M NaSCN, and 0.1−0.4 M NH_4I). In all cases, the best diffracting crystals were obtained at room temperature (293 K) by mixing 1 μL of protein:compound solution and 1 μL of reservoir solution. All crystals were vitrified in cryoprotectant, which consisted of reservoir solution with the addition of 20% (v/v) glycerol. Other solvent:protein ratios tested per condition were 1:2, 2:1.

7.1.2 Crystallographic Data and Methods

The X-ray diffraction data for the ATX−inhibitor complexes with **22** were collected at ESRF on beamline MASSIF1[33] and complexes with **23** were collected at SLS on beamline PXIII[34] at 100 K, and were recorded on a PILATUS 2M-F detector to a resolution of 2.00 and 2.48 Å, respectively. All data were processed and integrated with XDS.[34] All compounds were processed on site using the SLS automated processing pipeline and scaled with AIMLESS.[35] The structures were determined by molecular replacement using MOLREP[22] with the structure of ATX (PDB 2XR9) as the search model. Model building and subsequent refinement were performed iteratively with COOT,[23] REFMAC5,[24] and PDB_REDO.[25] Structure validation was carried out by MolProbity.[26] The structure models and experimental diffraction data were deposited at the PDB for compounds **22** and **23**, respectively. Crystallographic data and refinement details are available in **Table 19**.

Table 19 Crystallographic details

	73	**74**	**72**
Data collection			
wavelength (Å)	1.0000	1.0000	1.278295
resolution (Å)	2.00	2.48	1.57
space group	$P12_11$	$P12_11$	$P12_11$
unit cell a, b, c (Å), α, β, γ (deg)	62.3 89.0 76.5 90.0 102.7 90.0	62.8 89.6 77.6 90.0 102.8 90.0	62.5 88.7 77.6 90.0 103.1 90.0
$CC_{1/2}$	0.998 (0.834)	0.982 (0.760)	0.968 (0.236)
R_{merge}	0.043 (0.484)	0.099 (0.621)	0.214 (0.000)
I/σI	8.3 (1.2)	7.8 (1.7)	1.7 (0.0)
completeness (%)	99 (98)	99.8 (100)	83.4 (7.9)
redundancy	2.9 (2.9)	3.5 (3.4)	2.3 (1.1)
Refinement			
no. atoms	6735	3794	
protein	6299	6300	
ligand/metal/glycan	127	134	
water/iodine	309	360	
TLS groups	1	1	1
R_{work}/R_{free} (%)	19.2/26.8	18.5/24.4	0.391/0.421
Validation			
rsms/rmsZ bond lengths (Å)	0.0099/0.386	0.0087/0.401	0.0067
rsms/rmsZ bond angles (deg)	1.775/0.674	1.684/0.703	1.553
Ramachandran preferred/outliers	94.35/0.13	100/0	56/11
Rotamers preferred (%)	94	97	
Ramachandran Z score	-1859	-1639	
MolProbity/Clash score (%ile)	100/100	100/100	100

[a]High resolution shell in parentheses.

7.1.3 Biochemical Assays and Modelling of Kinetic Data

The biochemical studies of ATX lysoPLD activity were performed with ATX. Activity was measured by a coupled reaction with 1 U ml^{-1} choline oxidase and 2U ml^{-1} horseradish peroxidase (HRP) and 2 mM homovanillic acid (HVA) (all from Sigma-Aldrich). For the assays, 14:0, 16:0, 18:1 20:0 LPC (Avanti polar Lipids Inc.) was incubated with 20 nM ATX, reaching a final volume of 100 µl buffer, which contained 50 mM Tris, 0.01%, 50 mM $CaCl_2$, Triton X-100, pH 7.4. Steady-state choline release was measured at 37 °C by HVA fluorescence at λ_{ex} / λ_{em} = 320/460 nm in Corning® 96- or 384-well OptiPlate (Sigma-Aldrich) and with a Pherastar plate reader (BMG Labtech). To determine the IC_{50} for the different inhibitors on ATX activity, the velocity of the reaction was monitored for each compound as a function of time and the linear phase of the kinetics was taken from 60 min after the addition of ATX to the reaction buffer. The resulting fluorescence intensity signal over time was used to model all inhibitor concentrations simultaneously using the following formula, where v_{max} and v_{min} were fitted for the minimum and maximum relative velocity, and c_i corresponds to the inhibitor concentration for each assay:

$$V = \frac{v_{max} - v_{min}}{(1 + c_i / IC_{50})} + v_{min} \qquad \text{(Equation } \mathbf{1}\text{)}$$

7.1.4 Competition with LPA Allostery

The activation assays using LPA were performed in a similar way as those done for the inhibitors. In this case, LPA was dissolved in Ethanol: H_20 (1:1), 0.01% TX-100 and was added to the reaction buffer. The presence of ethanol was taken into account and controls in the absence of ATX and/or LPC were employed to correct the kinetic data. ATX was incubated for 30 min with different concentrations of inhibitors and subsequently added to the reaction buffer containing 150 µM 18:1 LPC and different starting concentrations of 18:1 LPA. The slopes were calculated from at least 60 min after the addition of ATX. The percentage of LPA-driven activation was normalised to ATX in the absence of LPA and inhibitors, which represented 100% activity. Lastly, the activation constant or AC_{50} was obtained from the following equation:

$$V = \frac{v_{max} - v_{min}}{(1 + c_i/AC_{50})} + v_{min} \qquad \text{(Equation 2)}$$

7.1.5 Mechanistic Studies with ATX Inhibitors

For initial comparison between competitive and non-competitive inhibition, we performed assays measuring LPC hydrolysis in the presence of three inhibitor concentrations (0.5 IC_{50}, IC_{50} and 2 IC_{50}) or with uninhibited ATX, from which slopes taken from 60 min after the start of the reaction were fit in two non-linear equations[19]:

$$V = \frac{V_{max} c_{LPC}}{K_M(1 + c_i/K_i) + c_{LPC}} \text{ competitive inhibition} \qquad \text{(Equation 3)}$$

$$V = \frac{(\frac{V_{max}}{1 + c_i/K_i}) c_{LPC}}{K_M + c_{LPC}} \text{ non-competitive inhibition} \qquad \text{(Equation 4)}$$

where V is the observed velocity and c_{LPC} is the corresponding LPC concentration for each data point, and c_i is the inhibitor concentration for each curve; and K_i is the inhibition constant. To statistically determine the chance of each type of inhibition, we calculated the α value in the partial mixed inhibition model (Equation 5), where *Part* defines the partiality of the inhibition, and $\alpha > 1$ or $\alpha = 1$ correspond to competitive and non-competitive inhibition modes, respectively. Lastly, we utilised the Akaike's Informative Criteria to assess the significance of the analysis.

$$V = \frac{V_{max} c_{LPA}(1 + P_{art})}{K_M(1 + c_i/K_i) + c_{LPA}(1 + c_i/\alpha K_i)} \text{ partial mixed inhibition} \qquad \text{(Equation 5)}$$

7.1.6 AKT and ERK Phosphorylation by Western-blotting

100,000 BJeH cells were seeded in 12-well tissue culture plates and allowed to grow for 24 h in DMEM (GIBCO, Life Technologies) containing 10% FCS (Thermo Scientific) and 100 U mL−1 streptomycin/penicillin (GIBCO, Life Technologies). Next, they were washed twice with PBS and serum starved O/N. ATX 20 nM was incubated with inhibitors for 30 min in serum free medium containing 0.05% (w/v) fatty acid-free BSA (total volume 1 mL). Medium from the 12-well plates was removed and replaced with 1 mL of ATX-inhibitor mixture. Cells were stimulated for 15 min, medium was removed, and plates were immediately frozen on dry ice and stored at −80 °C. For Western

blotting, cells were washed with cold PBS, lysed in RIPA buffer, supplemented with protease inhibitors (Pierce), 20 mM NaF and 1 mM Orthovanadate, and spun down. Protein concentration was measured using a BCA protein assay kit (Pierce), and LDS sample buffer (NuPAGE, Invitrogen) 1 mM dithiothreitol. (DTT) were added to the lysate. 20 μg total protein were loaded on SDS-PAGE pre-cast gradient gels (4–12% Nu-Page Bis-Tris, Invitrogen), followed by transfer to nitrocellulose membrane. Non-specific protein binding was blocked by 5% BSA in PBS-Tween (0.1%); primary antibodies were: phospho-Akt (Ser473), catalogue number: D93, dilution: 1:1,000; phospho-ERK1/2 (Thr202/Tyr204), catalogue number: D13.14.4E, dilution: 1:2,000; from Cell Signaling Technology. They were incubated O/N at 4°C in PBS-tween with 5% BSA, containing 0.1% NaN_3. Blots were then incubated for 1 hour at RT with monoclonal anti-β-actin antibody, clone AC-15, dilution: 1:10,000, from Sigma, which was dissolved in PBS-tween with 5% skimmed milk containing 0.1% NaN_3. Horseradish peroxidase-conjugated secondary antibodies (goat anti-mouse (Bio-Rad), catalogue number: 1721011; goat anti-rabbit (Pierce), catalogue number: 1858415) were incubated for 1h at room temperature in PBS-tween with 2.5% BSA, and developed using ECL Western blot reagent.

7.1.7 Production of LPA_1-HA-expressing HeLa-Flp-In Cells

Human LPA_1 cDNA was amplified by PCR to remove its stop codon and add the restriction sites for *BamH*I and *Xho*I, after which it was subcloned in an in-house produced pDNA5.1-HA vector. 0.2 μg of the resulting vector and 1.8 μg pOG44 Flp-Recombinase Expression Vector (Invitrogen) were incubated with 6 μl Fugene6 (Invitrogen) in 200 μl OptiMEM (Gibco) for 30 minutes, after which the mix was added to previously produced HeLa-Flp-In cells. Their medium was refreshed 24 hours later, and selection with 1 μg/mL puromycin was started and maintained with resistant cells.

7.1.8 Cell Migration Assay

Migration of MDA-MB-231 cells was performed using 48-well chemotaxis chambers (Neuro Probe, Inc.) equipped with 8 mm-pore polycarbonate membranes, which were coated with either fibronectin (10 μg/mL) (F1141, Sigma-Aldrich). Cells ($2x10^6$ cells/mL) were added to the upper chamber. 0.05% fatty acid-free BSA was

used as a lysophospholipid carrier. Cells were allowed to migrate for 4 h at 37 °C in humidified air containing 5% CO_2. Migrated cells were fixed in Diff-Quik Fix and stained using Diff-Quik II. Migration was quantified by colour intensity measurements using Photoshop software.

7.2 Appendix 2 – Extended Biological Protocols

7.2.1 Bis-pNPP hydrolysis by ATX

Final volume: 100 µL/well

- 50 µL Z buffer containing 0.2 mg/ml FAF-BSA.
- 25 µL 4 mM bis-*p*NPP (4x), diluted from 20 mM stock with Z buffer. [] in assay = 1 mM.
- 12.5 µL inhibitor/DMSO/LPA (8x concentrated).
- 12.5 µL 160 nM ATX (8x, diluted from 33.5 mg/mL stock with Z buffer. [] in assay = 20 nM

Z buffer composition:

Concentration in Buffer Solution	Stock	Volume
140 mM NaCl	1 M	7 mL
5 mM mM KCl	1 M	250 µL
1 mM $MgCl_2$	1 M	50 µL
1 mM $CaCl_2$	2.5 M	20 µL
100 mM Tris, pH 7.4	1 M	2.5 mL
MQ water		40.18 mL

1. Divide Z buffer (with 0.2 mg/mL FAF-BSA) over wells: 50 µL/well.
2. Add 25 µL substrate/well.
3. Add 12.5 µL inhibitor/LPA.
4. Warm to 37 ^{o}C in the plate reader.
5. Add 12.5 µL ATX and measure at absorbance at 405 nm.

7.2.2 Choline release assay

Final volume: 100 µL/well

- 50 µL reaction buffer containing choline oxidase and HRP.
- 25 µL 600 µM LPC (4x concentrated) from 47.9 mM stock in chloroform. [] in assay = 150 µM
- 12.5 µL inhibitor/DMSO (8x concentrated).
- 12.5 µL 160 nM ATX (8x concentrated) from 33.5 mg/ml stock. [] in assay = 20 µM

Reaction buffer:

Concentration in Buffer Solution	**Stock**	**Volume**
50 mM Tris pH 7.4	1 M	300 µL
0.01% TX100	1%	60 µL
25 mM $CaCl_2$	2.5 M	60 µL
2.5 mM HVA	0.05 M	300 µL
2 U/mL HRP	20 U/mL	600 µL
2 U/mL choline oxidase	10 U/mL	1,200 µL
MQ water		3,480 µL

1. Divide mix of reaction buffer over wells: 50 µL/well.
2. Add 25 µL LPC/well.
3. Warm to 37 ^{o}C in the plate reader.
4. Add ATX and measure at ex/em 355/460 nm.

7.2.3 Mode of Inhibition Analysis

Final volume: 100 µL/well

- 50 µL reaction buffer containing choline oxidase and HRP.

- 25 µL 600 µM LPC (4x concentrated) from 47.9 mM stock in chloroform. [] in assay = 150 µM
- 12.5 µL inhibitor/DMSO (8x concentrated).
- 12.5 µL 160 nM ATX (8x concentrated) from 33.5 mg/mL stock. [] in assay = 20 µM

Reaction buffer:

Concentration in Buffer Solution	Stock	Volume
50 mM Tris pH 7.4	1 M	300 µL
0.01% TX100	1%	60 µL
25 mM $CaCl_2$	2.5 M	60 µL
2.5 mM HVA	0.05 M	300 µL
2 U/mL HRP	20 U/mL	600 µL
2 U/mL choline oxidase	10 U/mL	1,200 µL
MQ water		3,480 µL

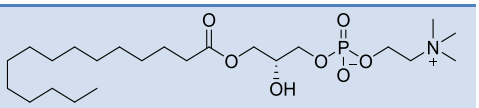
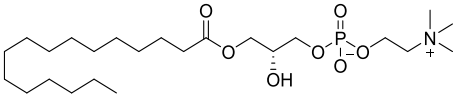
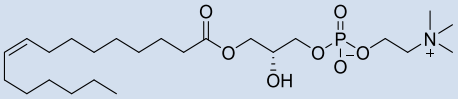
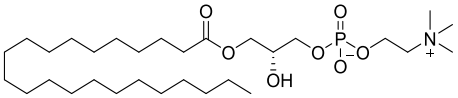
- Titrate 18:1 LPC starting from 4000 µM (= 4x stock) in 1/2 steps across the plate + "no LPC" control in the last column (+ 25 µL 50 mM Tris/TX100 instead of LPC).
  - Add 25 µL LPC to the 50 µL reaction buffer already in the wells.
- Prepare 1 mL 160 nM, 600 nM and 1200 nM inhibitor stocks (all 8x stocks).
  - Add 12.5 µL inhibitor to the wells containing 50 µL reaction buffer + 25 µL LPC.
- Add 12.5 µL 8x rATX (cf = 20 nM) and start measuring.

↓ [Inhibitor] → [LPC] (µM)	(#1) Initial []: 4000 µM LPC (4x)	(#2-10) ½ dilutions of LPC (50 µL 50 mM Tris/TX100 + 50 µL LPC)	(#12) No LPC Ctr.
A-B (0 µM inhibitor)			+ 25 µL 50 mM Tris/TX100
C-D (20 nM inhibitor)			+ 25 µL 50 mM Tris/TX100
E-F (75 nM inhibitor)			+ 25 µL 50 mM Tris/TX100
G-H (150 nM inhibitor)			+ 25 µL 50 mM Tris/TX100

7.2.4 LPC Specificity for Inhibitory Activity

Final volume: 100 µL/well

- 50 µL reaction buffer containing choline oxidase and HRP.
- 25 µL 600 µM LPC species (4x concentrated).

LPC Species	Structure
14:0	
16:0	
18:1	
22:0	

- 12.5 µL inhibitor/DMSO (8x concentrated).
- 12.5 µL 160 nM ATX (8x concentrated) from 33.5 mg/mL stock. [] in assay = 20 µM

Reaction buffer:

Concentration in Buffer Solution	Stock	Volume
50 mM Tris pH 7.4	1 M	300 µL
0.01% TX100	1%	60 µL
25 mM $CaCl_2$	2.5 M	60 µL
2.5 mM HVA	0.05 M	300 µL
2 U/mL HRP	20 U/mL	600 µL
2 U/mL choline oxidase	10 U/mL	1,200 µL
MQ water		3,480 µL

- Prepare 1 ml 600 µM (=4x stock) LPC stocks for 14:0, 16:0, 18:1 and 22:0 species.
  - Add 25 µL LPC to the 50 µL reaction buffer already in the wells.
- Titrate the inhibitor in 1/3 steps across the plate + DMSO control in the absence of the compound.
  - Pipette 12.5 µL form the dilution plate to the wells containing reaction buffer and LPC.
- Add 12.5 µL 8x rATX (Cf = 20 nM) and start measuring at ex/em 355/460 nm.

↓ [LPC] → [Inhibitor] (µM)	**(#1) Initial []: 800 µM inhibitor (8x)**	**(#2-10) 1/3 dilutions of inhibitor (60 µL DMSO + 30 µL inhibitor)**	**(#12) DMSO Ctr.**
A-B (150 µM 14:0 LPC)			+ 12.5 µL DMSO
C-D (150 µM 16:0 LPC)			+ 12.5 µL DMSO
E-F (150 µM 18:1 LPC)			+ 12.5 µL DMSO
G-H (150 µM 22:0			+ 12.5 µL

LPC)		DMSO

7.2.5 Reactivity with Reaction Buffer

Final volume: 100 µL/well

- 50 µL reaction buffer containing choline oxidase and HRP.
- 25 µL 160 µM choline chloride (4x concentrated) in H_2O. [] in assay = 40 µM
- 25 µL 120 µM (4x concentrated). [] in assay = 30 µM

Reaction buffer:

Concentration in Buffer Solution	**Stock**	**Volume**
50 mM Tris pH 7.4	1 M	300 µL
0.01% TX100	1%	60 µL
25 mM $CaCl_2$	2.5 M	60 µL
2.5 mM HVA	0.05 M	300 µL
2 U/mL HRP	20 U/mL	600 µL
2 U/mL choline oxidase	10 U/mL	1,200 µL
MQ water		3,480 µL

1. Divide mix of reaction buffer over wells: 50 µL/well.
2. Add 25 µL inhibitor/DMSO per well.
3. Add 25 µL choline chloride/well.
4. Warm to 37 ^{o}C in the plate reader.
5. Measure at ex/em 355/460 nm.

7.2.6 Western Blotting Analysis

<u>Day 1:</u>

1. Trypsinise BJeH cells and plate 100,000 cells/well on 12-well plates. Typically one confluent (90%) 15-cm plate should be enough to plate 30 wells and keep a stock on a 10-cm plate.

Day 2:

1. Wash cells twice with PBS and add serum-free DMEM. Incubate overnight.

Day 3:

3. In the morning, wash cells twice with PBS and add serum-free DMEM.
4. Prepare solutions for cell stimulation:
    - Take sterile serum-free DMEM into a falcon tube. Add BSA to have a 1 mg/ml solution, also under the hood.
    - Pipette desired stimulation solutions in 1mL Eppendorf tubes and allow to incubate at room temperature for 30 minutes.
5. Wash cells again twice with PBS and then add the stimulation solutions. Let incubate at 37 °C for 20 minutes.
6. After the incubation time is complete, wash cells twice with PBS (using glass Pasteur pipette to aspirate all remaining PBS) and freeze using dry ice.

Western-blotting protocol: **(everything on ice until step #5)**

7. Thaw plates with cells and allow to chill at 4 °C for 5-10 minutes.
8. Add lysis buffer:
    - Add 80 µL/well for 12-well plates, 120 µL/well for 6-well plates and 500 µL/well for 10-cm plates.
    - RIPA buffer composition: **50 mM Tris**, **150 mM NaCl**, **1% NP40, 0.5% DOC, 0.1% SDS, pH 7.4** supplemented with protease inhibitors (from 20x stock), **NaF** (1 mM from 50x stock) and **sodium orthovanadate** (1 mM from 200x stock). The latter two are Ser/Thr

phosphatase and Tyr/alkaline phosphatase inhibitors, respectively, necessary to blot for phosphorylated proteins.

9. Allow cells to lyse for 5-10 minutes. Use yellow pipette tip to scrape cells off the bottom of the wells and collect in Eppendorf tubes.
10. Spin samples down at **14,000 RPM for 15 minutes.**
11. Determine protein concentration by BCA assay. 200 µL/well + 20 µL H_2O + 5 µL sample. Adjust to 20-30 µg total protein.
12. Prepare samples with loading LDS dye and DTT:
    - **Per tube for gel: 50 µL sample + 17.5 µL 4x loading dye + 2.5 µLDTT.**
13. Run gels at 160-180 V.
14. Transfer onto nitrocellulose membranes (0.45 µm pore size) at 90-100 V for 1 hour and 10 minutes. Adjust voltage so that amperage remains constant at **0.15-0.20 Å**.
15. Stain with Ponceau S staining solution to establish if protein transfer was successful. Destain with **PBS/Tween (0.05%)**.
16. Block membranes with **5% BSA** for phosphorylated proteins or **5% milk** for 30 minutes.
17. Wash blots once with PBS/Tween for 15 minutes.
18. Start incubation with primary antibodies:
    - Rabbit anti-**P-Akt**: Overnight incubation at 4 ℃ - 1:2,000 in 5% BSA, 0.01% NaAz. Do not freeze and thaw.
    - Rabbit anti-**P-ERK**: 2 hours incubation at rt - 1:2,000 in 5% BSA, 0.01% NaAz. Do not freeze and thaw.
    - Rabbit anti-**Total Akt**: 2 hours incubation at rt - 1:2,000 in 5% BSA, 0.01% NaAz. Do not freeze and thaw.

- Rabbit anti-**Total ERK**: Overnight incubation at 4 °C - 1:2,000 in 5% BSA, 0.01% NaAz. Do not freeze and thaw.

19. Between incubation with primary and secondary antibodies, wash 2-3 times with PBS/Tween (15 minutes/wash).
20. Incubation with secondary antibodies. All at 1:5,000 in 2.5% Milk (or 2.5% BSA for phosphorylated protein).
21. Wash 2-3 times with PBS/Tween (15 minutes/wash) after incubation with secondary antibody,
22. Develop using ECL (1:1 mix).

7.2.7 Cell Migration

1 hour before starting the assay

1. Float filter for 1 hour upside up in fibronectin (note that filter should submerge, else take new fibronectin) (fibronectin F1141 from Sigma Aldrich 1 mg/mL, working concentration 10 µg/mL) dissolve 100µL in 10 mL PBS.
2. Wash the Boyden chamber **very well** and dry at 37 °C.
3. Start serum starvation of cells.

After 1 hour

4. In the lower chamber: **28 µL** reagents (make in FAF-BSA solution SFM (1 mg/mL in RPMI for B16F10 cells).
5. Air-dry the filter and carefully place it on top of the lower chamber, place middle part in top part and screw together.
6. Upper chamber: **55µL** cells,$1*10^6$ cells/ml (When studying receptors, preferably use EDTA to detach cells instead of trypsin)
7. Trypsinise your cells using 7 mL of serum free medium for the three 10 cm plates, count cells and dilute them to get $1*10^6$ cells/mL.

8. Incubate 3 hours at 37 °C.
9. Open chambers and grab filter with wide clip.
10. Float upside in PBS, attach heavy clip and wipe filter. Repeat 3 times.
11. Dry filter.
12. Use lightblue (fix), red and dark blue colouring solution subsequently, 6 times each.
13. Wash colour of in H_2O and stick filter to glass slide.

7.2.8 ATX Crystallisation Screens

ATX was incubated with each screened compound at a 1:10 (protein:compound) ratio for at least 30 minutes. Crystals were grown for a minimum of 7 days in a 24-well optimisation screen (18−20% PEG 3350, 0.1−0.4 M NaSCN, and 0.1−0.4 M NH_4I). In all cases, the best diffracting crystals were obtained at room temperature (293 K) by mixing 1 μL of protein:compound solution and 1 μL of reservoir solution. All crystals were vitrified in cryoprotectant, which consisted of reservoir solution with the addition of 20% (v/v) glycerol. Other solvent:protein ratios tested per condition were 1:2, 2:1.

Tables with all conditions tested adjusted for preparation of different volume sizes:

	18%	**20%**	**22%**	**18%**	**20%**	**22%**	
	360	400	440	360	400	440	
PEG 3350	100	100	100	200	200	200	PEG 3350
NH_4I (0.1M)	37.5	37.5	37.5	12.5	12.5	12.5	NH_4I (0.2M)
NaSCN (0.3M)	502.5	462.5	422.5	427.5	387.5	347.5	NaSCN (0.1M)
MQ water	1000	1000	1000	1000	1000	1000	MQ water
	360	400	440	360	400	440	
PEG 3350	200	200	200	200	200	200	PEG 3350
NH_4I (0.2M)	37.5	37.5	37.5	25	25	25	NH_4I (0.2M)
NaSCN (0.3M)	402.5	362.5	322.5	415	375	335	NaSCN (0.2M)
MQ water	1000	1000	1000	1000	1000	1000	MQ water
	360	400	440	360	400	440	
PEG 3350	300	300	300	200	200	200	PEG 3350
NH_4I (0.3M)	37.5	37.5	37.5	37.5	37.5	37.5	NH_4I (0.2M)
NaSCN (0.3M)	302.5	262.5	222.5	402.5	362.5	322.5	NaSCN (0.3M)
MQ water	1000	1000	1000	1000	1000	1000	MQ water
	360	400	440	360	400	440	
PEG 3350	400	400	400	200	200	200	PEG 3350
NH_4I (0.4M)	37.5	37.5	37.5	50	50	50	NH_4I (0.2M)
NaSCN (0.3M)	202.5	162.5	122.5	390	350	310	NaSCN (0.4M)
MQ water	1000	1000	1000	1000	1000	1000	MQ water

	18%	**20%**	**22%**	**18%**	**20%**	**22%**	
	720	800	880	720	800	880	
PEG 3350	200	200	200	400	400	400	PEG 3350
NH_4I (0.1M)	75	75	75	25	25	25	NH_4I (0.2M)
NaSCN (0.3M)	1005	925	845	855	775	695	NaSCN (0.1M)
MQ water	2000	2000	2000	2000	2000	2000	MQ water
	720	800	880	720	800	880	
PEG 3350	400	400	400	400	400	400	PEG 3350
NH_4I (0.2M)	75	75	75	50	50	50	NH_4I (0.2M)
NaSCN (0.3M)	805	725	645	830	750	670	NaSCN (0.2M)
MQ water	2000	2000	2000	2000	2000	2000	MQ water
	720	800	880	720	800	880	
PEG 3350	600	600	600	600	600	600	PEG 3350
NH_4I (0.3M)	75	75	75	75	75	75	NH_4I (0.2M)
NaSCN (0.3M)	605	525	445	805	725	645	NaSCN (0.3M)
MQ water	2000	2000	2000	2000	2000	2000	MQ water
	720	800	880	720	800	880	
PEG 3350	800	800	800	800	800	800	PEG 3350
NH_4I (0.4M)	75	75	75	100	100	100	NH_4I (0.2M)
NaSCN (0.3M)	405	325	245	780	700	620	NaSCN (0.4M)
MQ water	2000	2000	2000	2000	2000	2000	MQ water

	18%	**20%**	**22%**	**18%**	**20%**	**22%**	
PEG 3350	1800	2000	2200	1800	2000	2200	PEG 3350
NH_4I (0.1M)	500	500	500	1000	1000	1000	NH_4I (0.2M)
NaSCN (0.3M)	187.5	187.5	187.5	62.5	62.5	62.5	NaSCN (0.1M)
MQ water	2512.5	2312.5	2112.5	2137.5	1937.5	1737.5	MQ water
	5000	5000	5000	5000	5000	5000	
PEG 3350	1800	2000	2200	1800	2000	2200	PEG 3350
NH_4I (0.2M)	1000	1000	1000	1000	1000	1000	NH_4I (0.2M)
NaSCN (0.3M)	187.5	187.5	187.5	125	125	125	NaSCN (0.2M)
MQ water	2012.5	1812.5	1612.5	2075	1875	1675	MQ water
	5000	5000	5000	5000	5000	5000	
PEG 3350	1800	2000	2200	1800	2000	2200	PEG 3350
NH_4I (0.3M)	1500	1500	1500	1000	1000	1000	NH_4I (0.2M)
NaSCN (0.3M)	187.5	187.5	187.5	187.5	187.5	187.5	NaSCN (0.3M)
MQ water	1512.5	1312.5	1112.5	2012.5	1812.5	1612.5	MQ water
	5000	5000	5000	5000	5000	5000	
PEG 3350	1800	2000	2200	1800	2000	2200	PEG 3350
NH_4I (0.4M)	2000	2000	2000	1000	1000	1000	NH_4I (0.2M)
NaSCN (0.3M)	187.5	187.5	187.5	250	250	250	NaSCN (0.4M)
MQ water	1012.5	812.5	612.5	1950	1750	1550	MQ water
	5000	5000	5000	5000	5000	5000	

7.2.9 Crystallisation Data Collection, Data Processing and Structure Refinement

Collection:

Data for co-crystals with JMC4.6-2 and JMC4.29 were collected at the Swiss Light Source (SLS) at the PXIII beamline with the detector PILATUS 2M-F with a maximum resolution of 2.00 Å.

Processing:

- All data from ESRF were processed and integrated with XDS.
- Data from SLS was processed on site using the SLS automated processing pipeline and scaled with AIMLESS.
- The structures were determined by molecular replacement using PHASER Maximum Likelihood Molecular Replacement with the structure of 7-α-hydroxycholesterol-bound rat ATX (PDB 5DLT) as the search model.
- The space groups for all crystal structures was $P2_12_12_1$. With angles ~ 60, 90, 150.

Refinement:

- Model building and subsequent refinement were performed iteratively with:
  - COOT
  - REFMAC6
  - PDB REDO

8 References

1 C. Pavoine and F. Pecker, *Cardiovasc. Res.*, 2009, **82**, 175–183.

2 S. H. Francis and J. D. Corbin, in *Encyclopedia of Biological Chemistry: Second Edition*, Elsevier Inc., 2013, pp. 567–573.

3 K. Massé, S. Bhamra, G. Allsop, N. Dale and E. A. Jones, *Int. J. Dev. Biol.*, 2010, **54**, 181–93.

4 J. W. Goding, B. Grobben and H. Slegers, *Biochim. Biophys. Acta - Mol. Basis Dis.*, 2003, **1638**, 1–19.

5 L. M. Yuelling and B. Fuss, *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids*, 2008, **1781**, 525–530.

6 C. Stefan, S. Jansen and M. Bollen, *Trends Biochem. Sci.*, 2005, **30**, 542–550.

7 K. Kato, H. Nishimasu, S. Okudaira, E. Mihara, R. Ishitani, J. Takagi, J. Aoki and O. Nureki, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 16876–16881.

8 C. Döhler, M. Zebisch and N. Sträter, *Sci. Rep.*, 2018, **8**, 10874.

9 H. Sakagami, J. Aoki, Y. Natori, K. Nishikawa, Y. Kakehi, Y. Natori and H. Arai, *J. Biol. Chem.*, 2005, **280**, 23084–23093.

10 S. Jansen, C. Stefan, J. W. M. Creemers, E. Waelkens, A. Van Eynde, W. Stalmans and M. Bollen, *J. Cell Sci.*, 2005, **118**, 3081–3089.

11 H. Zimmermann, M. Zebisch and N. Sträter, *Purinergic Signal.*, 2012, **8**, 437–502.

12 M. A. Fox, J. K. Alexander, F. S. Afshari, R. J. Colello and B. Fuss, *Mol. Cell. Neurosci.*, 2004, **27**, 140–150.

13 R. Gijsbers, H. Ceulemans and M. Bollen, *Biochem. J.*, 2003, **371**, 321–330.

14 S. Jansen, A. Perrakis, C. Ulens, C. Winkler, M. Andries, R. P. Joosten, M. Van Acker, F. P. Luyten, W. H. Moolenaar and M. Bollen, *Structure*, 2012,

20, 1948–1959.

15 A. Zhou, J. A. Huntington, N. S. Pannu, R. W. Carrell and R. J. Read, *Nat. Struct. Biol.*, 2003, **10**, 541–544.

16 R. Gijsbers, H. Ceulemans, W. Stalmans and M. Bollen, *J. Biol. Chem.*, 2001, **276**, 1361–1368.

17 J. Hausmann, W. J. Keune, A. L. Hipgrave Ederveen, L. van Zeijl, R. P. Joosten and A. Perrakis, *J. Struct. Biol.*, 2016, **195**, 199–206.

18 J. Hausmann, S. Kamtekar, E. Christodoulou, J. E. Day, T. Wu, Z. Fulkerson, H. M. H. G. Albers, L. A. van Meeteren, A. J. S. Houben, L. van Zeijl, S. Jansen, M. Andries, T. Hall, L. E. Pegg, T. E. Benson, M. Kasiem, K. Harlos, C. W. Vander Kooi, S. S. Smyth, H. Ovaa, M. Bollen, A. J. Morris, W. H. Moolenaar and A. Perrakis, *Nat. Struct. Mol. Biol.*, 2011, **18**, 198–204.

19 R. A. Albright, W. C. Chang, D. Robert, D. L. Ornstein, W. Cao, L. Liu, M. E. Redick, J. I. Young, E. M. De La Cruz and D. T. Braddock, *Blood*, 2012, **120**, 4432–4440.

20 A. Gorelik, A. Randriamihaja, K. Illes and B. Nagar, *FEBS J.*, 2017, **284**, 3718–3726.

21 M. L. Stracke, H. C. Krutzsch, E. J. Unsworth, A. Arestad, V. Cioce, E. Schiffmann and L. A. Liotta, *J. Biol. Chem.*, 1992, **267**, 2524–2529.

22 A. Tokumura, E. Majima, Y. Kariya, K. Tominaga, K. Kogure, K. Yasuda and K. Fukuzawa, *J. Biol. Chem.*, 2002, **277**, 39436–39442.

23 M. Umezu-Goto, Y. Kishi, A. Taira, K. Hama, N. Dohmae, K. Takio, T. Yamori, G. B. Mills, K. Inoue, J. Aoki and H. Arai, *J. Cell Biol.*, 2002, **158**, 227–233.

24 A. Perrakis and W. H. Moolenaar, *J. Lipid Res.*, 2014, **55**, 1010–1018.

25 Z. Fulkerson, T. Wu, M. Sunkara, C. Vander Kooi, A. J. Morris and S. S. Smyth, *J. Biol. Chem.*, 2011, **286**, 34654–34663.

26 J. Hausmann, S. Kamtekar, E. Christodoulou, J. E. Day, T. Wu, Z. Fulkerson, H. M. H. G. Albers, L. A. Van Meeteren, A. J. S. Houben, L. Van Zeijl, S. Jansen, M. Andries, T. Hall, L. E. Pegg, T. E. Benson, M. Kasiem, K. Harlos, C. W. V. Kooi, S. S. Smyth, H. Ovaa, M. Bollen, A. J. Morris, W. H. Moolenaar and A. Perrakis, *Nat. Struct. Mol. Biol.*, 2011, **18**, 198–205.

27 J. Lee, I. D. Jung, S. W. Nam, T. Clair, E. M. Jeong, S. Y. Hong, J. W. Han, H. W. Lee, M. L. Stracke and H. Y. Lee, *Biochem. Pharmacol.*, 2001, **62**, 219–224.

28 H. Nishimasu, S. Okudaira, K. Hama, E. Mihara, N. Dohmae, A. Inoue, R. Ishitani, J. Takagi, J. Aoki and O. Nureki, *Nat. Struct. Mol. Biol.*, 2011, **18**, 205–212.

29 H. Nishimasu, S. Okudaira, K. Hama, E. Mihara, N. Dohmae, A. Inoue, R. Ishitani, J. Takagi, J. Aoki and O. Nureki, *Nat. Struct. Mol. Biol.*, 2011, **18**, 205–213.

30 H. Y. Lee, T. Clair, P. T. Mulvaney, E. C. Woodhouse, S. Aznavoorian, L. A. Liotta and M. L. Stracke, *J. Biol. Chem.*, 1996, **271**, 24408–24412.

31 G. B. Mills and W. H. Moolenaar, *Nat. Rev. Cancer*, 2003, **3**, 582–591.

32 K. Nakanaga, K. Hama and J. Aoki, *J. Biochem.*, 2010, **148**, 13–24.

33 T. Eichholtz, K. Jalink, I. Fahrenfort and W. H. Moolenaar, *Biochem. J.*, 1993, **291**, 677–680.

34 Y. Kihara, H. Mizuno and J. Chun, *Exp. Cell Res.*, 2015, **333**, 171–177.

35 K. Yanagida and S. Ishii, *J. Biochem.*, 2011, **150**, 223–232.

36 Y. Shimomura, M. Wajid, Y. Ishii, L. Shapiro, L. Petukhova, D. Gordon and A. M. Christiano, *Nat. Genet.*, 2008, **40**, 335–339.

37 C. W. Lee, R. Rivera, S. Gardell, A. E. Dubin and J. Chun, *J. Biol. Chem.*, 2006, **281**, 23589–23597.

38 K. Yanagida, K. Masago, H. Nakanishi, Y. Kihara, F. Hamano, Y. Tajima, R. Taguchi, T. Shimizu and S. Ishii, *J. Biol. Chem.*, 2009, **284**, 17731–17741.

39 Y. C. Yung, N. C. Stoddard and J. Chun, *J. Lipid Res.*, 2014, **55**, 1192–1214.

40 R. Rivera and J. Chun, *Future Lipidol.*, 2007, **2**, 535–545.

41 M.-E. Lin, D. R. Herr and J. Chun, *Prostaglandins*, 2010, **91**, 130–138.

42 S. Willier, E. Butt and T. G. P. Grunewald, *Biol. Cell*, 2013, **105**, 317–333.

43 Y. C. Yung, T. Mutoh, M. E. Lin, K. Noguchi, R. R. Rivera, J. W. Choi, M. A. Kingsbury and J. Chun, *Sci. Transl. Med.*, 2011, **3**, 99ra87.

44 S. M. Pasternack, I. Von Kügelgen, K. Al Aboud, Y. A. Lee, F. Rüschendorf, K. Voss, A. M. Hillmer, G. J. Molderings, T. Franz, A. Ramirez, P. Nürnberg, M. M. Nöthen and R. C. Betz, *Nat. Genet.*, 2008, **40**, 329–334.

45 M. Gotoh, Y. Fujiwara, J. Yue, J. Liu, S. C. Lee, J. Fells, A. Uchiyama, K. Murakami-Murofushi, S. Kennel, J. Wall, R. Patil, R. Gupte, L. Balazs, D. D. Miller and G. J. Tigyi, *Biochem. Soc. Trans.*, 2012, **40**, 31–36.

46 I. Ninou, C. Magkrioti and V. Aidinis, *Front. Med.*, 2018, **5**, 180.

47 G. Bain, K. E. Shannon, F. Huang, J. Darlington, L. Goulet, P. Prodanovich, G. L. Ma, A. M. Santini, A. J. Stein, D. Lonergan, C. D. King, I. Calderon, A. Lai, J. H. Hutchinson and J. F. Evans, *J. Pharmacol. Exp. Ther.*, 2017, **360**, 1–13.

48 S. G. Bourgoin and C. Zhao, *Curr. Opin. Investig. Drugs*, 2010, **11**, 515–526.

49 S. Liu, M. Murph, N. Panupinthu and G. B. Mills, *Cell Cycle*, 2009, 8, 3695–3701.

50 Y. Xu, *Cancers*, *2019*, **11**, 1523.

51 S. Mulholland, H. Adamali and S. L. Barratt, *Clin. Pharmacol. Adv. Appl.*, 2020, **12**, 97–108.

52 T. Wynn, *J. Pathol.*, 2008, **214**, 199–210.

53 J. H. W. Distler, A. H. Györfi, M. Ramanujam, M. L. Whitfield, M. Königshoff and R. Lafyatis, *Nat. Rev. Rheumatol.*, 2019, **15**, 705–730.

54 D. J. Lederer and F. J. Martinez, *N. Engl. J. Med.*, 2018, **378**, 1811–1823.

55 C. P. Denton and D. Khanna, *Lancet*, 2017, **390**, 1685–1699.

56 C. Trautwein, S. L. Friedman, D. Schuppan and M. Pinzani, *J. Hepatol.*, 2015, **62**, S15–S24.

57 M. Ruiz-Ortega, S. Rayego-Mateos, S. Lamas, A. Ortiz and R. R. Rodrigues-Diez, *Nat. Rev. Nephrol.*, 2020, **16**, 269–288.

58 T. A. Wynn, *J. Clin. Invest.*, 2007, **117,** 524–529.

59 C. B. Nanthakumar, R. J. D. Hatley, S. Lemma, J. Gauldie, R. P. Marshall and S. J. F. Macdonald, *Nat. Rev. Drug Discov.*, 2015, **14**, 693–720.

60 J. A. Kropski, T. S. Blackwell and J. E. Loyd, *Eur. Respir. J.*, 2015, **45**, 1717–1727.

61 C. López-Ramírez, L. Suarez Valdivia and J. Rodríguez Portal, *Med. Sci.*, 2018, **6**, 58.

62 S. Bellentani and M. Marino, *Ann. Hepatol.*, 2009, **8**, S4–S8.

63 T. M. Maher, A. U. Wells and G. J. Laurent, *Eur. Respir. J.*, 2007, **30**, 835–839.

64 A. J. Sanyal, *Nat. Rev. Gastroenterol. Hepatol.*, 2019, **16**, 1.

65 M. Benedict and X. Zhang, *World J. Hepatol.*, 2017, **9**, 715–732.

66 J. K. Dyson, Q. M. Anstee and S. McPherson, *Frontline Gastroenterol.*, 2014, **5**, 211–218.

67 M. Romero-Gómez, S. Zelber-Sagi and M. Trenell, *J. Hepatol.*, 2017, **67**, 829–846.

68 Z. Younossi, Q. M. Anstee, M. Marietti, T. Hardy, L. Henry, M. Eslam, J.

George and E. Bugianesi, *Nat. Rev. Gastroenterol. Hepatol.*, 2018, **15**, 11–20.

69 *Diabetes Care*, 2016, **39**, 632–638.

70 M. D. Beaton, *Can. J. Gastroenterol.*, 2012, **26**, 353–357.

71 R. Loomba, E. Lawitz, P. S. Mantry, S. Jayakumar, S. H. Caldwell, H. Arnold, A. M. Diehl, C. S. Djedjos, L. Han, R. P. Myers, G. M. Subramanian, J. G. McHutchison, Z. D. Goodman, N. H. Afdhal, M. R. Charlton and for the G. Investigators, *Hepatology*, 2018, **67**, 549.

72 S. A. Harrison, V. W. S. Wong, T. Okanoue, N. Bzowej, R. Vuppalanchi, Z. Younes, A. Kohli, S. Sarin, S. H. Caldwell, N. Alkhouri, M. L. Shiffman, M. Camargo, G. Li, K. Kersey, C. Jia, Y. Zhu, C. S. Djedjos, G. M. Subramanian, R. P. Myers, N. Gunn, A. Sheikh, Q. M. Anstee, M. Romero-Gomez, M. Trauner, Z. Goodman, E. J. Lawitz and Z. Younossi, *J. Hepatol.*, 2020, **73**, 26–39.

73 I. Dickson, *Nat. Rev. Gastroenterol. Hepatol.*, 2020, **17**, 260.

74 R. Safadi, F. M. Konikoff, M. Mahamid, S. Zelber-Sagi, M. Halpern, T. Gilat and R. Oren, *Clin. Gastroenterol. Hepatol.*, 2014, **12**, 2085-2091.

75 M. Iruarrizaga-Lejarreta, M. Varela-Rey, D. Fernández-Ramos, I. Martínez-Arranz, T. C. Delgado, J. Simon, V. G. Juan, L. delaCruz-Villar, M. Azkargorta, J. L. Lavin, R. Mayo, S. M. Van Liempd, I. Aurrekoetxea, X. Buqué, D. D. Cave, A. Peña, J. Rodríguez-Cuesta, A. M. Aransay, F. Elortza, J. M. Falcón-Pérez, P. Aspichueta, L. Hayardeny, M. Noureddin, A. J. Sanyal, C. Alonso, J. Anguita, M. L. Martínez-Chantar, S. C. Lu and J. M. Mato, *Hepatol. Commun.*, 2017, **1**, 911–927.

76 R. Pellicciari, M. Pruzanski and A. Gioiello, in *Successful Drug Discovery*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2018, pp. 197–244.

77 M. Eslam, R. Alvani and G. Shiha, *Lancet*, 2019, **394**, 2131–2133.

78 A. Mullard, *Nat. Rev. Drug Discov. 2020,* **19**, 501.

79 S. Deweerdt, *Nature*, 2017, **551**, S92–S93.

80 D. A. Schwartz, *Trans. Am. Clin. Climatol. Assoc.*, 2016, **127**, 34–45.

81 D. L. Clarke, L. A. Murray, B. Crestani and M. A. Sleeman, *Pharmacol. Ther.*, 2017, **169**, 35–46.

82 H. J. Kim, D. Perlman and R. Tomic, *Respir. Med.*, 2015, **109**, 661–670.

83 *Am. J. Respir. Crit. Care Med.*, 2002, **165**, 277–304.

84 L. Plantier, A. Cazes, A.-T. Dinh-Xuan, C. Bancal, S. Marchand-Adam and B. Crestani, *Eur. Respir. Rev.*, 2018, **27**, 170062.

85 D. J. Lederer and F. J. Martinez, *N. Engl. J. Med.*, 2018, **378**, 1811–1823.

86 V. Cottin, B. Crestani, D. Valeyre, B. Wallaert, J. Cadranel, J.-C. Dalphin, P. Delaval, D. Israel-Biet, R. Kessler, M. Reynaud-Gaubert, B. Aguilaniu, B. Bouquillon, P. Carré, C. Danel, J.-B. Faivre, G. Ferretti, N. Just, S. Kouzan, F. Lebargy, S. Marchand-Adam, B. Philippe, G. Prévot, B. Stach, F. Thivolet-Béjui, J.-F. Cordier, the F. N. R. C. and the N. of C. C. for R. L. French National Reference Centre and Network of Competence Centres for Rare Lung Diseases, *Eur. Respir. Rev.*, 2014, **23**, 193–214.

87 A. Devaraj, *Eur. Respir. Rev.*, 2014, **23**, 215–219.

88 C. E. Daniels, E. S. Yi and J. H. Ryu, *Eur. Respir. J.*, 2008, **32**, 170–174.

89 E. Fernández Fabrellas, R. Peris Sánchez, C. Sabater Abad and G. Juan Samper, *Med. Sci.* 2018, **6**, 51.

90 I. V. Yang and D. A. Schwartz, *Transl. Res.*, 2015, **165**, 48–60.

91 R. M. du Bois, *Eur. Respir. Rev.*, 2012, **21**, 141–6.

92 K. Macneal and D. A. Schwartz, *Proc. Am. Thorac. Soc.*, 2012, **9**, 120–125.

93 F. V Castelino and J. Varga, *Arthritis Res. Ther.*, 2010, **12**, 213.

94 M. Kolb, F. Bonella and L. Wollin, *Respir. Med.*, 2017, **131**, 49–57.

95 S. Rangarajan, M. L. Locy, T. R. Luckhardt and V. J. Thannickal, *Drugs*, 2016, **76**, 291–300.

96 *Am. J. Respir. Crit. Care Med.*, 2000, **161**, 646–664.

97 P. J. Wolters, H. R. Collard and K. D. Jones, *Annu. Rev. Pathol. Mech. Dis.*, 2014, **9**, 157–179.

98 T. M. Maher, A. U. Wells and G. J. Laurent, *Eur. Respir. J.*, 2007, **30**, 835–839.

99 C. K. Probst, S. B. Montesi, B. D. Medoff, B. S. Shea and R. S. Knipe, *Eur. Respir. J.* 2007, **29**, 976-985.

100 N. I. Chaudhary, G. J. Roth, F. Hilberg, J. Müller-Quernheim, A. Prasse, G. Zissel, A. Schnapp and J. E. Park, *Eur. Respir. J.*, 2007, **29**, 976–985.

101 S. K. Huang and J. C. Horowitz, *Austin J. Pulm. Respir. Med.*, 2014, **1**, 3.

102 J. Baum and H. S. Duffy, *J. Cardiovasc. Pharmacol.*, 2011, 57, 376–379.

103 A. Caminati, R. Cassandro, O. Torre and S. Harari, *Eur. Respir. Rev.*, 2017, **26**, 170047.

104 T. Kuerner, *Respir. Investig.*, 2015, **53**, 2–6.

105 R. J. Homer, J. A. Elias, ; Chun, G. Lee and E. Herzog, *Arch. Pathol. Lab. Med.*, 2011, **135**, 780–788.

106 L. Fregonese and I. Eichler, *BMC Med.*, 2015, **13**, 239.

107 R. M. Baron, A. J. S. Choi, C. A. Owen and A. M. K. Choi, *Am. J. Physiol. - Lung Cell. Mol. Physiol.*, 2012, **302**, L485.

108 G. Raghu and M. Selman, *Am. J. Respir. Crit. Care Med.*, 2015, **191**, 252–254.

109 E. S. Kim and G. M. Keating, *Drugs*, 2015, **75**, 219–230.

110 M. E. Cho and J. B. Kopp, *Expert Opin. Investig. Drugs*, 2010, **19**, 275–83.

111 G. A. Margaritopoulos, E. Vasarmidi and K. M. Antoniou, *Core Evid.*, 2016, **11**, 11–22.

112 H. Oku, T. Shimizu, T. Kawabata, M. Nagira, I. Hikita, A. Ueyama, S. Matsushima, M. Torii and A. Arimura, *Eur. J. Pharmacol.*, 2008, **590**, 400–408.

113 B.-S. Lee, S. B. Margolin and R. A. Nowak, *Pirfenidone: A Novel Pharmacological Agent That Inhibits Leiomyoma Cell Proliferation and Collagen Production*, 1998, vol. 83.

114 L. Wollin, E. Wex, A. Pautsch, G. Schnapp, K. E. Hostettler, S. Stowasser and M. Kolb, *Eur. Respir. J.*, 2015, **45**, 1434–1445.

115 E. S. Kim and G. M. Keating, *Drugs*, 2015, **75**, 219–230.

116 S. E. Torrisi, M. Pavone, A. Vancheri and C. Vancheri, *Eur. Respir. Rev.*, 2017, 26.

117 G. Raghu, B. van den Blink, M. J. Hamblin, A. W. Brown, J. A. Golden, L. A. Ho, M. S. Wijsenbeek, M. Vasakova, A. Pesci, D. E. Antin-Ozerkis, K. C. Meyer, M. Kreuter, D. Moran, H. Santin-Janin, F. Aubin, G. J. Mulder, R. Gupta and L. Richeldi, *Lancet Respir. Med.*, 2019, **7**, 657–664.

118 L. Richeldi, E. R. Fernández Pérez, U. Costabel, C. Albera, D. J. Lederer, K. R. Flaherty, N. Ettinger, R. Perez, M. B. Scholand, J. Goldin, K. H. Peony Yu, T. Neff, S. Porter, M. Zhong, E. Gorina, E. Kouchakji and G. Raghu, *Lancet Respir. Med.*, 2020, **8**, 25–33.

119 G. Lee, S.-U. Kang, J.-H. Ryou, J.-J. Lim, D.-Y. Lee, H.-J. Kwon, G.-H. Ha and Y.-H. Lee, in *American Thoracic Society International Conference Meetings Abstracts*, American Thoracic Society, 2019, pp. A2577–A2577.

120 N. Euer, M. Schwirzke, V. Evtimova, H. Burtscher, M. Jarsch, D. Tarin and U. H. Weidle, *Anticancer Res.*, 2002, **22**, 733–740.

121 D. Hanahan and R. A. Weinberg, *Cell*, 2000, **100**, 57–70.

122 D. Hanahan and R. A. Weinberg, *Cell*, 2011, **144**, 646–674.

123 N. Oikonomou, M.-A. Mouratis, A. Tzouvelekis, E. Kaffe, C. Valavanis, G. Vilaras, A. Karameris, G. D. Prestwich, D. Bouros and V. Aidinis, *Am. J. Respir. Cell Mol. Biol.*, 2012, **47**, 566–574.

124 B. Ley, H. R. Collard and T. E. King, *Am. J. Respir. Crit. Care Med.*, 2011, **183**, 431–440.

125 E. Kaffe, A. Katsifa, N. Xylourgidis, I. Ninou, M. Zannikou, V. Harokopos, P. Foka, A. Dimitriadis, K. Evangelou, A. N. Moulas, U. Georgopoulou, V. G. Gorgoulis, G. N. Dalekos and V. Aidinis, *Hepatology*, 2017, **65**, 1369–1383.

126 G. Bain, K. E. Shannon, F. Huang, J. Darlington, L. Goulet, P. Prodanovich, G. L. Ma, A. M. Santini, A. J. Stein, D. Lonergan, C. D. King, I. Calderon, A. Lai, J. H. Hutchinson and J. F. Evans, *J. Pharmacol. Exp. Ther.* 2017, **360**, 1-13.

127 N. Desroy, C. Housseman, X. Bock, A. Joncour, N. Bienvenu, L. Cherel, V. Labeguere, E. Rondet, C. Peixoto, J.-M. Grassot, O. Picolet, D. Annoot, N. Triballeau, A. Monjardet, E. Wakselman, V. Roncoroni, S. Le Tallec, R. Blanque, C. Cottereaux, N. Vandervoort, T. Christophe, P. Mollat, M. Lamers, M. Auberval, B. Hrvacic, J. Ralic, L. Oste, E. van der Aar, R. Brys and B. Heckmann, *J. Med. Chem.*, 2017, **60**, 3580–3590.

128 C. G. Thomson, D. Le Grand, M. Dowling, C. E. Brocklehurst, C. Chinn, L. Elphick, M. Faller, M. Freeman, V. Furminger, C. Gasser, A. Hamadi, E. Hardaker, V. Head, J. C. Hill, D. I. Janus, D. Pearce, A.-S. Poulaud, E. Stanley and L. Sviridenko, *Bioorg. Med. Chem. Lett.*, 2018, **28**, 2279–2284.

129 J. Gierse, A. Thorarensen, K. Beltey, E. Bradshaw-Pierce, L. Cortes-Burgos, T. Hall, A. Johnston, M. Murphy, O. Nemirovskiy, S. Ogawa, L. Pegg, M. Pelc, M. Prinsen, M. Schnute, J. Wendling, S. Wene, R. Weinberg, A. Wittwer, B. Zweifel and J. Masferrer, J. Pharm. Exp. Ther. 2010, **331**, 310-

317.

130 D. Castagna, E. L. Duffy, D. Semaan, L. C. Young, J. M. Pritchard, S. J. F. Macdonald, D. C. Budd, C. Jamieson, A. J. B. Watson, J. Aoki, Y. Kakehi, A. M. Michon, D. J. Mitchell, R. K. Prinjha, A. D. Roberts, R. J. Sheppard, R. J. Watson, H. Nochi, K. Tamoto, H. Tomura and F. Okajima, *Med. Chem. Commun.*, 2015, **6**, 1149–1155.

131 H. M. H. G. Albers, A. Dong, L. A. van Meeteren, D. A. Egan, M. Sunkara, E. W. van Tilburg, K. Schuurman, O. van Tellingen, A. J. Morris, S. S. Smyth, W. H. Moolenaar and H. Ovaa, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 7257–62.

132 D. Castagna, D. C. Budd, S. J. F. Macdonald, C. Jamieson and A. J. B. Watson, *J. Med. Chem.*, 2016, **59**, 5604–5621.

133 M. G. K. Benesch, Y. Y. Zhao, J. M. Curtis, T. P. W. McMullen and D. N. Brindley, *J. Lipid Res.*, 2015, **56**, 1134–44.

134 J. Kazius, R. McGuire and R. Bursi, *J. Med. Chem.*, 2005, **48**, 312–320.

135 A. Nikolaou, M. G. Kokotou, D. Limnios, A. Psarra and G. Kokotos, *Expert Opin. Ther. Pat.*, 2017, 1–15.

136 A. N. Matralis, A. Afantitis and V. Aidinis, *Med. Res. Rev.*, 2019, **39**, 976–1013.

137 F. Salgado-Polo and A. Perrakis, *Cancers*, 2019, **11**, 1577.

138 A. J. Stein, G. Bain, P. Prodanovich, A. M. Santini, J. Darlington, N. M. P. Stelzer, R. S. Sidhu, J. Schaub, L. Goulet, D. Lonergan, I. Calderon, J. F. Evans and J. H. Hutchinson, *Mol. Pharmacol.*, 2015, **88**, 982–992.

139 W.-J. Keune, J. Hausmann, R. Bolier, D. Tolenaars, A. Kremer, T. Heidebrecht, R. P. Joosten, M. Sunkara, A. J. Morris, E. Matas-Rico, W. H. Moolenaar, R. P. Oude Elferink and A. Perrakis, *Nat. Commun.*, 2016, **7**, 11248.

140 W.-J. Keune, F. Potjewyd, T. Heidebrecht, F. Salgado-Polo, S. J. F. Macdonald, L. Chelvarajan, A. Abdel Latif, S. Soman, A. J. Morris, A. J. B. Watson, C. Jamieson and A. Perrakis, *J. Med. Chem.*, 2017, **60**, 2006–2017.

141 E. Duffy, University of Strathclyde, Approaches to Lead Generation for Idiopathic Pulmonary Fibrosis Targets, *PhD Thesis*, 2015.

142 F. Potjewyd, University of Strathclyde, Structure-Guided Approaches to the Design of Novel Autotaxin Inhibitors, *PhD Thesis*, 2016.

143 H. M. H. G. Albers and H. Ovaa, *Chem. Rev.*, 2012, **112**, 2593–603.

144 H. Nishimasu, R. Ishitani, J. Aoki and O. Nureki, *Trends Pharmacol. Sci.*, 2012, 33, 138–145.

145 F. Salgado-Polo, A. Fish, M.-T. Matsoukas, T. Heidebrecht, W.-J. Keune and A. Perrakis, *J. Biol. Chem.*, 2018, **293**, 14312-14327.

146 W. Keune, J. Hausmann, R. Bolier, D. Tolenaars, A. Kremer, T. Heidebrecht, R. Joosten, M. Sunkara, A. Morris, E. Matas-Rico, W. Moolenaar, R. Oude Elferink and A. Perrakis, *Nat. Commun.*, 2016, **7**, 11248.

147 T. M. Maher, E. M. van der Aar, O. Van de Steen, L. Allamassey, J. Desrivot, S. Dupont, L. Fagard, P. Ford, A. Fieuw and W. Wuyts, *Lancet Respir. Med.*, 2018, **6**, 627–635.

148 T. E. King, K. K. Brown, G. Raghu, R. M. du Bois, D. A. Lynch, F. Martinez, D. Valeyre, I. Leconte, A. Morganti, S. Roux and J. Behr, *Am. J. Respir. Crit. Care Med.*, 2011, **184**, 92–99.

149 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development q settings*, 2001, vol. 46.

150 I. Sakamoto, H. Kaku and T. Tsunoda, *Pharm. Bull.* 203, **62**, 474-476.

151 S. Fletcher, H. Firouzabadi, B. Akhlaghinia, R. Azadi, Y. Li, H. Hioki, S. Itô, T. Lu, M. A. Young, S. Yu, S. Albeck, K. N. Houk, D. S. Tawfik and D.

Baker, *Org. Chem. Front.*, 2015, **2**, 739–752.

152 A. J. J. Lennox and G. C. Lloyd-Jones, *Angew. Chemie Int. Ed.*, 2012, **51**, 9385–9388.

153 H. M. H. G. Albers, L. A. van Meeteren, D. A. Egan, E. W. van Tilburg, W. H. Moolenaar and H. Ovaa, *J. Med. Chem.*, 2010, **53**, 4958–4967.

154 A. Field-Smith, G. J. Morgan and F. E. Davies, *Ther. Clin. Risk Manag.*, 2006, 2, 271–279.

156 J. Du, C. Sun, Z. Hu, Y. Yang, Y. Zhu, D. Zheng, L. Gu and X. Lu, *PLoS One*, 2010, **5**, e15940.

157 H. Ohta, K. Sato, N. Murata, A. Damirin, E. Malchinkhuu, J. Kon, T. Kimura, M. Tobo, Y. Yamazaki, T. Watanabe, M. Yagi, M. Sato, R. Suzuki, H. Murooka, T. Sakai, T. Nishitoba, D. S. Im, H. Nochi, K. Tamoto, H. Tomura and F. Okajima, *Mol. Pharmacol.*, 2003, **64**, 994–1005.

158 P. C. Trippier and C. McGuigan, *MedChemComm*, 2010, **1**, 183.

159 A. Draganov, D. Wang and B. Wang, *Top. Med. Chem.*, 2016, **17**, 1–28.

160 A. Nikolaou, I. Ninou, M. G. Kokotou, E. Kaffe, A. Afantitis, V. Aidinis and G. Kokotos, *J. Med. Chem.*, 2018, **61**, 3697–3711.

161 S. F. Sleiman, D. E. Olson, M. W. Bourassa, S. S. Karuppagounder, Y. L. Zhang, J. Gale, F. F. Wagner, M. Basso, G. Coppola, J. T. Pinto, E. B. Holson and R. R. Ratan, *J. Neurosci.*, 2015, **35**, 438.

162 S. Shen and A. P. Kozikowski, *ChemMedChem*, 2016, **11**, 15–21.

163 H. C. Hang, C. Yu, D. L. Kato and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 14846–14851.

164 C. R. Bertozzi, *Acc. Chem. Res.*, 2011, **44**, 651–653.

165 E. M. Sletten and C. R. Bertozzi, *Angew. Chemie Int. Ed.*, 2009, **48**, 6974–6998.

166 X. Jiang, X. Hao, L. Jing, G. Wu, D. Kang, X. Liu and P. Zhan, *Expert Opin. Drug Discov.*, 2019, **14**, 779–789.

167 M. Bozdag, M. Pinard, F. Carta, E. Masini, A. Scozzafava, R. McKenna and C. T. Supuran, *J. Med. Chem.*, 2014, **57**, 9673–9686.

168 C. Ensch and M. Hesse, *Helv. Chim. Acta*, 2002, **85**, 1659–1677.

169 R. Taj and J. L. Sorensen, *Tetrahedron Lett.*, 2015, **56**, 7108–7111.

170 A. Niida, H. Tanigaki, E. Inokuchi, Y. Sasaki, S. Oishi, H. Ohno, H. Tamamura, Z. Wang, S. C. Peiper, K. Kitaura, A. Otaka and N. Fujii, *J. Org. Chem.*, 2006, **71**, 3942–3951.

171 S. P. Sahoo, H. Koyama, D. J. Miller, Merck & Co. Inc, WO/2004/066963, 2004.

172 B. M. Trost and C. A. Kalnmals, *Org. Lett.*, 2017, **19**, 2346–2349.

173 A. M. Wawro, Y. Aoki, T. Muraoka, K. Tsumoto and K. Kinbara, *Chem. Commun.*, 2018, **54**, 1913–1916.