Determining the Role of the Human Transient Receptor Potential Subfamily A Member 1 (hTRPA1) Ion Channel in Pain Sensation

Benjamin G Saward

A Thesis Submitted to the Board of the Faculty of Physical Sciences of the University of Oxford in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy

Merton College

Trinity Term 2021

Table of Contents

Abstract

TRPA1 is a thermosensitive, calcium permeable ion channel present in small diameter neurons present in most organs including the skin, lungs and trachea. TRPA1 is activated by small-molecule ligands, notably electrophiles including α,β-unsaturated aldehydes, tear gases, allyl isothiocyanate, iodoacetamide, and formaldehyde (HCHO). Deletion of TRPA1 in mice causes loss of sensitivity to HCHO. Modulation of TRPA1 activity is of interest as a target for pain treatment. **GRC-17536** is in clinical trials, however, the mechanism by which it acts as a TRPA1 antagonist is unknown. The mechanism by which channel opening results from electrophile (*e.g.* HCHO) binding to cysteine residues within TRPA1 is unclear. The work described in this thesis attempted to explore these mechanisms.

The reaction of HCHO with biological nucleophiles is of interest from disease and regulatory perspectives. The reaction of cysteine containing peptides with HCHO was shown to involve selective formation of an *S*-hydroxymethyl adduct. The stability of these adducts was explored through the addition of the ubiquitous cellular nucleophile glutathione (GSH), which was found to remove HCHO from the adducts. A recombinant protein model of the TRPA1 electrophile binding domain was designed, purified and its reactions with electrophiles studied.

To investigate the binding site of **GRC-17536**, a series of photolabile ligands based on **GRC-17536** were designed and synthesised. A structure-activity relationship was constructed from the compounds' ability to inhibit cinnamaldehyde evoked TRPA1 activity. Binding of a compound showing the most potent inhibition (**3-60**) was studied by cryo-electron microscopy (cryo-EM). From a cryo-EM structure of **3-60** bound to TRPA1, it was possible to propose modifications to the ligand which would improve binding and pharmacokinetic properties of theophylline-based ligands such as **GRC-17536**.

TRPA1 shows a "U"-shaped sensitivity towards agonists under changing oxygen concentration. It is controversially reported that this is a result of hydroxylation of the channel by PHD2, a hydroxylase involved in the hypoxic response. Peptides corresponding to the ankyrin loops of TRPA1 were synthesised and screened against FIH, a different HIF-hydroxylase, the results reveal that TRPA1 ankyrin 8 is a substrate of FIH at the peptide level.

The combined results provide new insight into the reaction of HCHO with cysteine residues involved in signalling pathways and open up new avenues for modulating the activity of TRPA1.

Impact of COVID-19

The COVID-19 pandemic has had a profound effect on the work contained in this thesis. The chemistry Research Laboratory (CRL), where the majority of the work contained in this thesis was carried out, was shut between the 17th March 2020 and the 17th July 2020, allowing no access for experimental work. From that date until the present day, the laboratories have been running with reduced occupancy with the number of people working in one laboratory module reduced to 4. From the 17th July the working hours in the CRL were significantly reduced from 24/7 access to between the hours of 8:00 to 22:00. During this time, I shared access to the CRL with one other researcher, allowing me a maximum of half of the amount of time in the CRL as would normally be expected.

From the 19th October, I was placed into a working group of 3 researchers, one of whom was in their final term as a DPhil student. This meant that I was only able to access the CRL for approximately a third to a half of the usual time that I would normally be able to access the CRL for. Additionally, I needed to transfer laboratories during this time from the Schofield group synthetic chemistry lab (LG9) to the chemical biology lab (LG2) in order to complete aspects of the biochemical work contained in this thesis. LG2 has significantly more users and thus I received pushback on this. From the 7th December 2020 the restrictions on working hours were lifted slightly to give access to the CRL between 06:00 to 24:00.

On account of the number of researchers wanting to work in LG2, I struggled with finding time to work on aspects of the biochemical work contained in this thesis. Additionally, the conflict created by a large number of researchers wanting access to a small amount of time was very detrimental to my mental health, especially contributing to a feeling of guilt if an experiment didn't go to plan due to feeling like the time could have been better used by another researcher. Additionally, there were supply issues with consumables including pipette tips, centrifuge tubes, and media and antibiotics for cell culture, further complicating the ability to work. The supply chain issues and problems with laboratory occupancy also affected the small research facilities including the mass spectrometry laboratory whom I wanted to collaborate with on the work concerning the tryptic digests of TRPA1. Additionally, the work could not be outsourced to other laboratories such as the Mohammed group in the biochemistry department due to staffing issues.

On the 27th January 2021, the Schofield group was notified of a positive case of COVID-19 that had come in to contact with several members of the group, forcing them to isolate. The case originated outside of the group but walked into LG2 to use the shared equipment. This profoundly affected my feeling of COVID security and thus I elected to work from home from that point forward.

Acknowledgements

First and foremost, I would like to thank Prof. Chris Schofield for all of his support during the time I have worked as part of his group. Additionally, I would like to thank the Biotechnology and Biological Sciences Research Council (BBSRC) and the Defence, Science & Technology Laboratory (DSTL) for their financial support of this project.

I am extremely grateful to Dr. Richard Hopkinson and Dr. Chris Lindsay for their support and for their expert advice in aspects of the experimental work contained in this thesis and day to day supervision. I would also like to thank Dr. Chris Timperley and Dr. Chris Green for their project management and conceptualisation of the project at DSTL.

The collaborative nature of my work would not have been possible without help from within the university and without. I would especially like to thank Dr. Mariana Grieben and Dr. Ashley Pike for their work on cryo-EM, Dr. Thomas Leissing for his work on FIH crystallography and Dr. Anthony Tumber for his work on peptide hydroxylation. Additionally, I would like to thank Dr. Ya-Min Tian, Dr. Thom Durrant and Dr. Yihua Wang for their kind gifts. Special thanks must be given to Anna Olerinyova for her work on mass photometry, and for her love and support for many years throughout this DPhil.

I would like to thank Dr. Anka Lucic, Dr. Joanna Storr, Tim Suits, Dr. Alen Krajnc, Tika Malla, Hollie Parsons and Dr. Leah Taylor-Kearney for their friendship and making the CJS group a wonderful place to work. Special thanks must go to Dr. Adam Hardy for his support in keeping the lab running.

I would like to thank the members of the House of Delphi for their love and support, especially during the height of the COVID-19 pandemic. Without the support of this community, this thesis would not have been possible. I would also like to thank Dr. Carole Bataille and Aislin Sheldon for their friendship and support.

Importantly, thanks go to my family for their support, encouragement and understanding throughout my life, without which I would not be typing these words.

Author Contributions

All experiments described within the text were carried out by myself unless otherwise stated. All figures obtained from external sources are used with previously obtained permissions. A list of all work carried out by others is given here:

Chapter II:

Mass Photometry of BK3 and subsequent analysis was performed by Ms. Anna Olerinyova.

Chapter III:

Production and purification of hTRPA1 from Sf9 cells was performed by Dr. Mariana Grieben (MG). Cryo-electron microscopy (cryo-EM) studies were performed by MG. Processing cryo-EM and structure building of TRPA1 in complex with 3-60 was performed by MG and Dr. Ashley Pike.

Chapter IV:

Crystalography of FIH in complex with various peptides was performed by Dr. Thomas Leissing. Kinetic analysis of peptide hydroxylation by RapidFire mass spectrometry was performed by Dr. Anthony Tumber.

Abbreviations

$[Ca^{2+}]_i$	Intracellular Ca^{2+} concentration
µw	Microwave irradiation
1CA	Schofield group consensus ankrin
2OG	2-oxogluterate
Ac	Acetyl
AITC	Allyl-isothiocyanate
AKAP	Protein kinase A anchoring protein
AM	Acetoxymethyl
API-ES	Atmospheric pressure chemical ionisation
ARD	Ankyrin repeat domain
ASC	Sodium ascorbate
ATP	Adenosine triphosphate
ATR	Attenuated total reflection
B_2Pin_2	Bis-pinacolato-diboron
BAPTA	1,2-Bis(o-aminophenoxy)-ethane-N,N,N',N'-tetraacetic acid
bHLH	Basic-helix-loop-helix motif
BINAP	2,2′-Bis(diphenylphosphino)-1,1′-binaphthyl
BITC	Benzyl isothiocyanate
BK3	Benkyrin 3 (synthetic ankyrin protein)
Boc	*tert*-Butyloxycarbonyl
BTFA	3-bromo-1,1,1-trifluoroacetone
CaM	Calmodulin
CCC	C-terminal coiled-coil
CDK5	Cyclin dependent kinase 5
CGRP	Calcitonin gene related peptide
CPMG	Carr-Purcell-Meiboom-Gill
CPMG-PROJECT	CPMG with refocusing pulses to supress *J* modulation
cryo-EM	Cryogenic electron microscopy
CTAD	C-Terminal transactivation domain
CTD	C-Terminal cytosolic domain
CV	Column volume
dba	Dibenzylideneacetone
DCC	Dicyclocarbodiimide
DIC	1,3-Diisopropylcarbodiimide
DIPEA	*N*-Diisopropylethylamine
DMA	Dimethylacetamide
DMEM	Dulbecco's modified Eagle's medium
DMF	Dimethyl formamide
DMOG	Dimethyloxalylglycine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dppf	1,1′-Bis(diphenylphosphino)ferrocene

DRG	Dorsal root ganglion/ganglia
DSTL	Defence, Science and Technology Laboratory
DTT	Dithioreitol
EBD	Electrophile binding domain
EGL-9	Hypoxia-inducible factor prolyl hydroxylase
ELSD	Evaporative light scattering detector
EMEM	Eagle's minimal essential medium
ESI	Electrospray ionisation
ESPE	Excitation sculpting with perfect echo
Et	Ethyl
FAS	Ferric ammonium sulphate
FBS	Fetal bovine serum
FIH	Factor inhibiting hypoxia induced factor
FLIPR	Fluorometric imaging plate reader
FPLC	Fast protein liquid chromatography
FrmR	Formaldehyde-induced regulator
GSH	Glutathione
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HB	Hygromycin B
HBSS	Hank's balanced salt solution
HCHO	Formaldehyde
HCTU	*O*-(6-Chlorobenzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate
HEK	Human embryonic kidney cells
HIF	Hypoxia-inducible factor
His_6	Hexa-histidine tag
HMBC	Heteronuclear multiple bond correlation
HMG	*S*-hydroxymethylglutathione
HNE	4-Hydroxynonenal
HPLC	High Performance Liquid Chromatography
HRE	Hypoxia response element
HRMS	High Resolution Mass Spectrometry
HSQC	Heteronuclear single quantum correlation
hTRPA1	Human transient receptor potential subfamily A member 1
IAA	Iodoacetamide
IFH	Interfacial helix
^{i}Pr	Isopropyl
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IVTT	*In-vitro* transcription/translation
J	Coupling constant
KO	Knock-out
LC-MS	Liquid chromatography coupled mass spectrometry
LIC	Ligation-independent cloning

LMNG	Lauryl maltose neopentyl glycol
m/z	Mass to charge ratio
MALDI-MS	Matrix assisted laser desorption ionisation mass spectrometry
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionisation – Time of Flight
Me	Methyl
mRNA	Messenger ribonucleic acid
MS	Mass Spectrometry
Ms	Methanesulfonyl
MS/MS	Mass spectrometry coupled mass spectrometry
MW	Molecular Weight
NG	Nodose ganglion
NMDA	N-methyl-d-aspartate
NMR	Nuclear Magnetic Resonance
NOG	N-oxalylglycine
NTD	N-Terminal cytosolic domain
ODD	Oxygen-dependent degradation domain
ONE	4-Oxynonenal
PAS	per-Arnt-Sim motif
PCR	Polymerase chain reaction
PD	Potential difference
PDB	Protein Data Bank
Ph	Phenyl
PHD	HIF prolyl hydroxylase
PHD2	Prolyl hydroxylase domain-containing protein 2
Pip	Pipecolic acid
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PKA	Protein kinase A
PKC	Protein kinase C
PLC	Phospholipase C
PMF	Peptide mass fingerprint
POPC	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
PTM	Post-translational modification
pVHL	Von Hippel-Lindau factor
rcf	Relative centrifugal force
RFMS	RapidFire mass spectrometry
RP-FCC	Reverse phase flash column chromatography
Rpm	Revolutions per minute
SAR	Structure-activity relationship
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SGC	Structural Genomics Consortium
SP	Substance P
SPE	Solid phase extraction

SPPS	Solid phase peptide synthesis
TFA	Trifluoroacetic acid
TG	Trigeminal ganglion
THC	Δ^9-tetrahydrocannabinol
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TM	Transmembrane
TRIS	2-Amino-2-(hydroxymethyl)propane-1,3-diol
TRP	Transient receptor potential
TRPM8	Transient receptor potential melastatin member 8
TRPV1	Transient receptor potential vanilloid member 1
TSP	3-(Trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt
UV	Ultraviolet
VGPC	voltage-gated potassium channel
VGSC	voltage-gated sodium channel
VWD	Variable wavelength detector
WT	Wild type
δ	Chemical shift
νmax	Ifrared absorption

Chapter I:

Introduction

1. The Molecular Basis of Pain

1.1 Nociception

Pain can be defined as “complex constellation of unpleasant sensory, emotional and cognitive experiences provoked by real or perceived tissue damage and manifested by certain autonomic, psychological, and behavioural reactions”. It is proposed that “pain” should be considered to always be a psychological state, rather than the activation of particular receptors or even as the result of a stimulus.[1] Nociception, therefore, is the neural process by which damaging or potentially damaging (noxious) stimuli are encoded to produce the psychological state of pain. The neurons and proteins which are sensitive towards noxious stimuli are known as nociceptors and nocisensors resepectively.[2]

Nociceptors are part of the somatosensory system, a large group of neurons that are responsible for relaying the sensations of touch, temperature and position from nerve endings in the skin, muscles and joints to the central nervous system.[3] The neurons in this system are split into 3 categories depending on the level of myelination and axon diameter. A thick myelin sheath is present in Aβ and Aδ nerve fibres, and not present in C nerve fibres.[4] Aβ nerve fibres are differentiated from Aδ nerve fibres by their diameter, with Aβ nerve fibers having a thicker diameter. Aβ nerve fibres have high conduction velocities on account of their myelination and thick diameter, and are responsible for relaying information on muscle and joint position and movement.[5] Aδ and C nerve fibres are both nociceptors, with Aδ nerve fibres contributing to the generation of sharp, fast pain and forming the reflex arc, while C fibres generate slow, lasting and spread-out pain.[4]

Nerve cells have a resting potential difference (PD) of approximately -70 mV, although the exact resting potential varies from nerve to nerve on account of ion channel/transporter expression.[6] This arises from differences in concentration between ionic species in the extracellular and intracellular space, and the plasma membrane of a cell being impervious to ions. The concentration gradient is established and maintained by ion transporters which are embedded in the membrane. Of these ion

transporters, the sodium-potassium pump uses ATP to export three Na^+ ions from the cytosol in exchange for two K^+ ions.[7] This leads to a net accumulation of positive charge outside of the cell. In addition to this, the sodium-calcium exchanger (NCX) passively transports three Na^+ ions into the cytosol in exchange for a single Ca^{2+} ion.[8] Ca^{2+} is also actively exported from the cytosol by the plasma membrane Ca^{2+} ATPase (PMCA) which exports a single Ca^{2+} ion in exchange for the hydrolysis of an ATP molecule. The PMCA has a much higher affinity for Ca^{2+} than the NCX but has a much lower turnover, thus the PMCA typically maintains the low intracellular Ca^{2+} concentration in resting cells while the NCX restores the Ca^{2+} gradient after an action potential.[9] This leads to a low concentration of Ca^{2+} in the cytosol relative to the concentration in the extracellular space, leading to a concentration gradient of Ca^{2+} ions across the membrane which can be used for cellular signalling.

Information is relayed by nerves *via* an action potential. This is the cyclic pattern of an input stimulus raising the membrane PD to a threshold voltage, this is followed by the opening of voltage-gated sodium channels (VGSC, *e.g.*$Na_V1.7$) which allows Na^+ ions to flow into the cell, raising the membrane PD to a maximum, positive voltage. At this point the VGSC close and voltage gated potassium channels open (VGPC), causing K^+ ions to flow out of the cell. This leads to the membrane PD becoming smaller and gradually more negative.[10,11] Due to the varied kinetics of open, closing and inactivation of different VGPCs, the membrane PD overshoots the resting PD. The resting PD is restored by the action of the sodium-potassium exchanger (fig. 1-1).

Different nerves respond to different input stimuli. In general, depolarisation of the proceeding portion of a nerve axon serves as the input stimulus for the next portion of the axon, causing the propagation of the action potential.[12–14] At the nerve endings, nocisensors serve as an input stimulus for nociceptors. Nocisensors are typically ligand-gated ion channels, opening in response to a certain ligand, allowing ions to flow down the concentration gradient, depolarising the membrane and initiating an action potential if the threshold PD is reached.[15] Nociceptors are typically polymodal,

meaning that they respond to a variety of different input stimuli. This is the result of the presence of several different classes of nocisensors.[16]

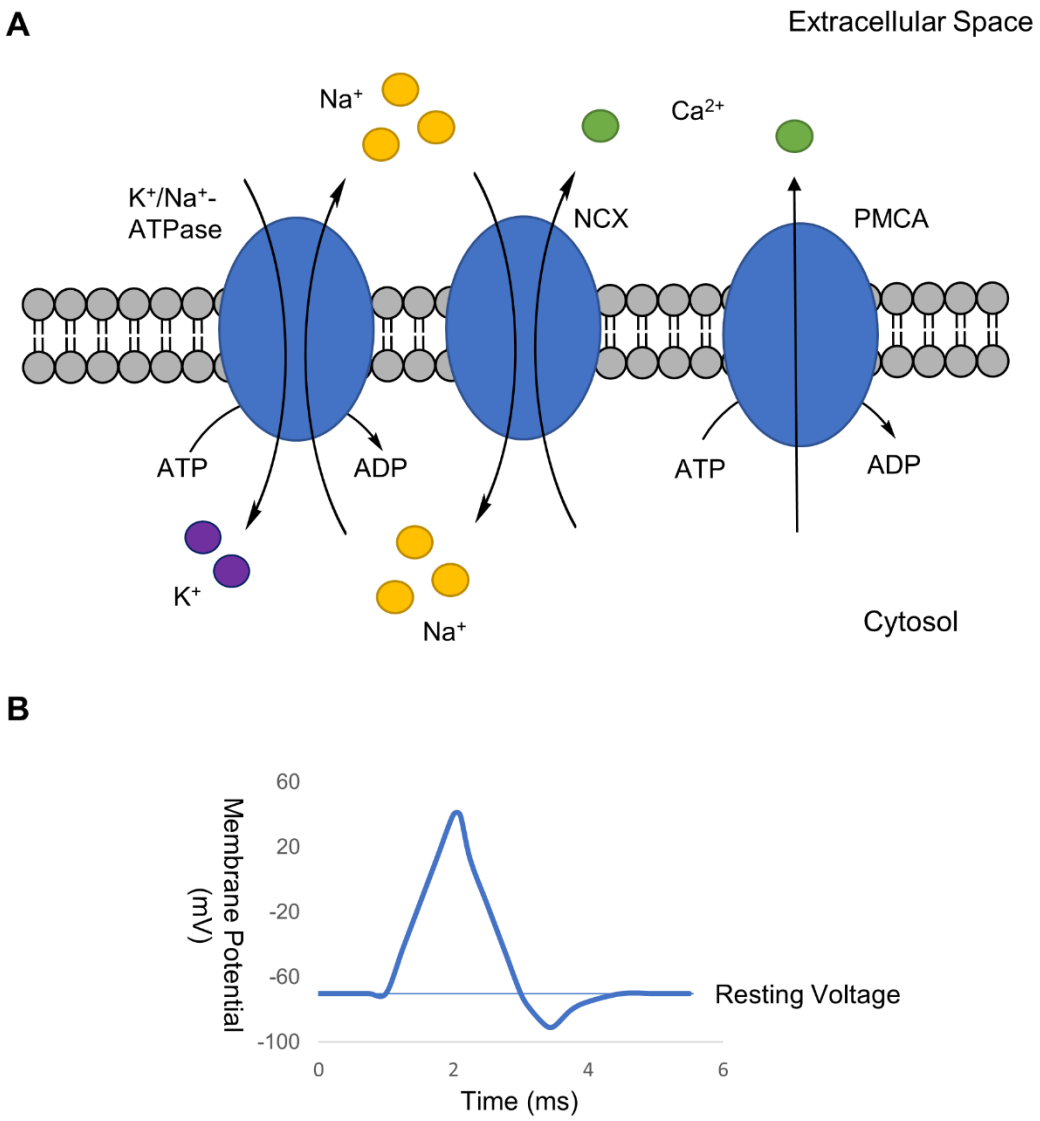


Figure 1-1: Ion transporters in the plasma membrane are responsible for establishing and maintaining a membrane potential difference (PD). A) Overview of the three classes of ion transporter: potassium/sodium transporters (K^+/Na^+-ATPase), sodium/calcium exchangers (NCX) and plasma-membrane calcium-ATPases (PMCA). **B)** Representative graph of change in membrane potential over time in an action potential.[124]

1.2 Transient Receptor Potential (TRP) channels

The transient receptor potential (TRP) channels form a family of ion channels which are responsible for sensing environmental stimuli. The TRP channels were first identified in *Drosophila melanogaster* where mutants showing abnormal photoreception were characterised according to their

behavioural phenotype.[17] In the case of *transient receptor potential* (*trp*) mutants, the flies behaved as if blind in the presence of bright light but exhibited behaviour that was indistinguishable from the wild-type (WT) under low light conditions.[18] In addition, *D. melanogaster trp* mutants require a long period in the dark to be able to respond to bright light, unlike the WT flies which respond to light irrespective of a period in the dark.[18] The name TRP originates from the transient ability of a receptor in the optical nerves of *D. melanogaster trp* mutants to potentially generate action potentials (*i.e.* the transient receptor potential). Identification of the *trp* locus in *D. melanogaster* and subsequent sequencing revealed a 1275 amino acid membrane protein which serves as a calcium channel critical to the phototransduction process in *D. melanogaster*.[19] Further studies identified an N-terminal cytosolic domain (NTD), 6 trans-membrane (TM) helices and a C-terminal cytosolic domain, features which all members of the TRP family share.[20]

The TRP channel superfamily is further divided into 11 families based on sequence homology, which are named after a defining feature of one or all the sub-family members. The 9 families present in animals are: canonical (TRPC), vanilliod (TRPV), vanilliod-like (TRPVL), melastatin (TRPM), mucolipin (TRPML), ankyrin (TRPA), polycystic kidney disease (TRPP), no mechanoreceptor potential (TRPN) and soromelastatin (TRPS). TRPY and TRPF are fungus specific and form a phylogenetic outgroup from the other TRP families.[21] TRPC channels share the highest similarity (30-35% amino acid sequence identity) with the original *D. melanogaster* TRP protein and are typically activated downstream of receptors that use the phospholipase C signalling cascade.[22] The TRP channels most relevant to human pain are TRPA1 (*vide infra*), TRPV1[23] and TRPM8.[24] TRPA1, TRPV1 and TRPM8 are present in nociceptive neurons, typically small diameter afferent neurons.

TRPA1 is present in a subset of TRPV1 expressing neurons, while TRPM8 is typically present in a distinct population of small diameter afferent neurons from TRPV1 positive neurons.[25]

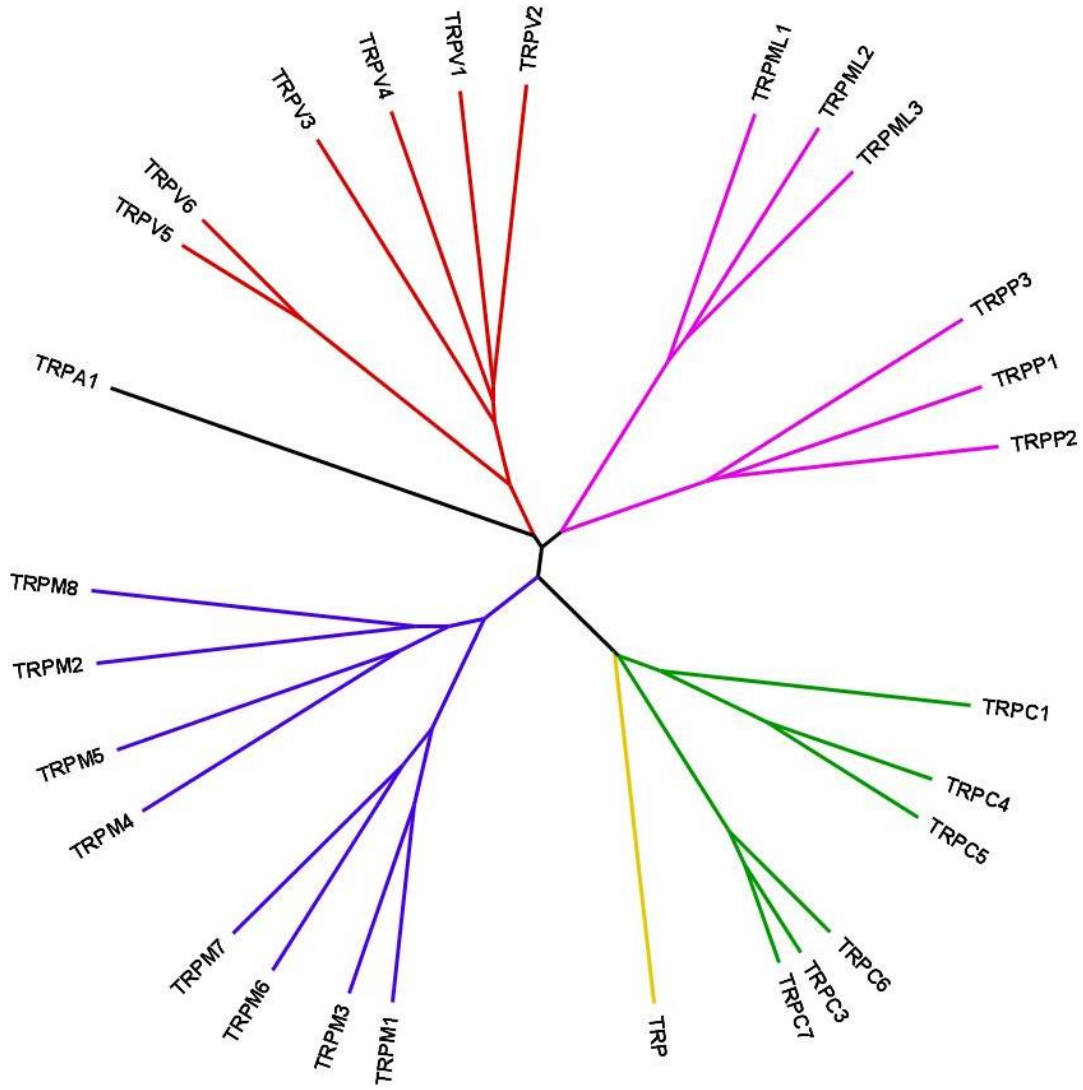


Figure 1-2: Phylogenetic tree of TRP channels in humans. The tree is rooted to the TRP channel found in *Drosophila Melangaster* (TRP, yellow). Tree built using Clustal Omega[125] and visualised using FigTree v1.4.4.

TRPM8 is a thermosensitive TRP channel present in animals that responds to innocuous cold temperatures and is activated by agonists such as menthol and icilin.[24] Activation of TRPM8 typically decreases nociceptive response or generates pain relief (analgesia) in several animal models of pain.[26] Potent anti-inflammatory effects for TRPM8 WT mice are also seen a murine model of ulcerative colitis compared to TRPM8 knockout mice.[27] Interestingly TRPM8 antagonists also provide relief

from certain types of pain, especially cold allodynia.[28,29] Taken together, TRPM8 modulation provides an avenue for targeted pain relief. TRPV1 is another thermosensitive TRP channel that responds to noxious heat and capsaicin, the active ingredient in Capsicum peppers (*i.e.* chilli peppers), which gives a warming/burning sensation when ingested.[30] TRPV1 is implicated in pro-inflammatory pathways and inflammatory pain. TRPV1 antagonists and TRPV1 knockout decrease thermal hyperalgesia and increase the threshold for which heat is perceived to be noxious.[31,32]

2. TRPA1

2.1 Expression of TRPA1

TRP subfamily A member 1 (TRPA1), previously known as ANKTM1, is a calcium-permeable ion channel present in animals that belongs to the TRP family. The TRPA1 gene is expressed in a subset of TRPV1-expressing small diameter neurons, typically C-fiber neurons.[25] Most neurons where TRPA1 is present are peptigerdic nociceptors, using either calcitonin gene related peptide (CGRP) or substance P (SP) as a neurotransmitter and relaying a sense of pain to the brain.[33] The TRPA1 gene is strongly expressed in neurons originating in the dorsal root ganglia (DRG). The DRG are nerve root bundles in the vertebral column (*i.e.* spine). DRG sections can be grouped into the areas of the body where the nerves end. In mice, TRPA1 messenger RNA (mRNA) is found in all DRG sections,[34] suggesting the presence of functional TRPA1 in a subsection of all neurons originating in from DRG. Additionally, TRPA1 mRNA has also been detected in the trigeminal ganglia (TG) and the nodose ganglion (NG) in mice.[35] The TG have nerve endings in the face and are a key part of relaying orofacial pain to the brain.[36] The NG is part of the vagal nerve, relaying information from the upper thorax, including the lungs and heart, where the free nerve endings reside.[36] The expression of TRPA1 in ganglia that relay information from large sections of the body, including those that innervate organs such as the colon,[37] heart, lungs and gastro-intestinal tract,[38] means TRPA1 activation contributes to visceral pain which is described as non-localised, dull pain.[39] TRPA1 is also expressed in the mouse inner ear in the organ of Corti within the cochlea which contains hair cells which are responsible for

hearing.[40] TRPA1 mRNA has also been detected in the murine central nervous system (CNS) with a low level of expression in the cerebellum, cerebrum, forebrain and hippocampus.[41]

Similar patterns of TRPA1 gene expression observed in rats and mice are also seen in humans, with direct evidence of protein expression in the TG,[42] lung fibroblasts and alveolar epithelium.[43] The TRPA1 gene is also expressed in a variety of non-neuronal cells in humans, in keratinocytes, fibroblasts and melanocytes in the skin[44] and in odontoblasts isolated from the dental pulp.[45]

2.2 Function of TRPA1

TRPA1 is highly promiscuous towards activation by electrophiles as a result of its electrophile binding domain (EBD) on its NTD which contains the highly reactive C621.[46] TRPA1 shows a preference for lipophilic electrophiles including α,β-unsaturated aldehydes[35] such as 4-hydroxynonenal (**HNE**, **1**) which is considered its endogenous ligand, 4-oxynonenal (**ONE**, **2**), cinnamaldehyde (**3**) and the prostaglandins 15dPGJ$_2$ (**4**) and Δ^{12}-PGJ$_2$ (**5**) (fig. 1-3).[47] Both HNE and ONE are products of lipid peroxidation which result from oxidative cellular stress.[48] Other lipophilic electrophiles that activate TRPA1 are the tear gases **CS** and **CR**, along with other benzylidinemalonitriles and dibenzoxazepines with varying potencies.[49,50] Additionally, **JT010** is possibly one of the most potent and selective agonists of TRPA1, which undergoes reaction on the chloroacetyl group (fig. 1-3).[51] This mode of reaction is mirrored for other agonists such as iodoacetamide (**IAA**).[46,52]

TRPA1 is also activated by some hydrophilic electrophiles such at formaldehyde (HCHO)[53] and allyl isothiocyanate (**AITC**) which is the principal component in mustard oil.[54,55] Neurons from the DRG of TRPA1 knockout mice are completely insensitive to HCHO while still responding to the TRPV1 agonist capsaicin,[53] suggesting that TRPA1 is the molecular target for HCHO (fig. 1-3).

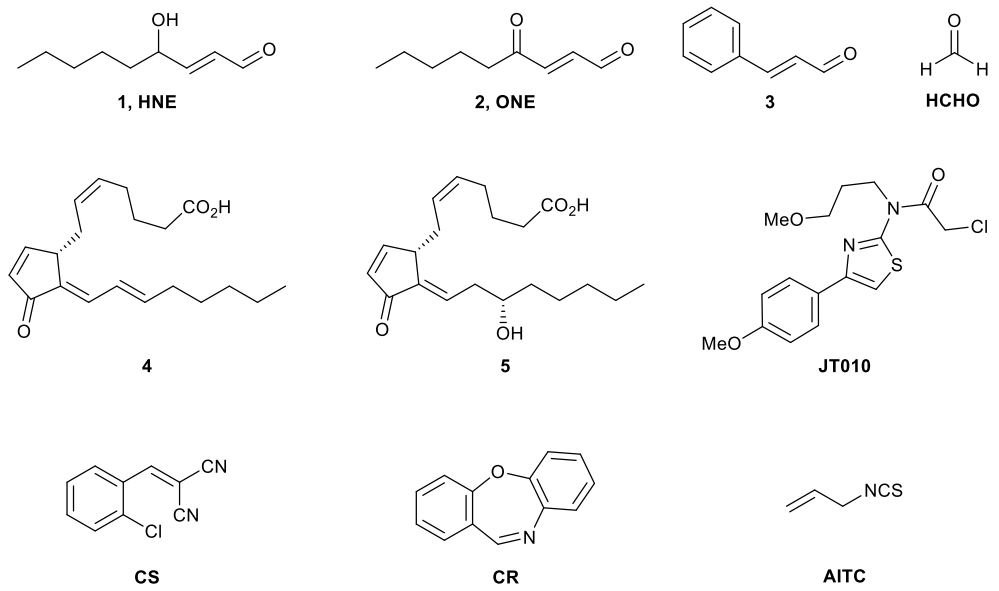


Figure 1-3: Electrophilic agonists of TRPA1. A common feature of agonists is the α,β-unsaturated aldehyde moiety. **AITC** = allyl isothiocyanate, **HCHO** = formaldehyde.

TRPA1 also has several ligand binding sites in its transmembrane domain.[56–58] These binding sites are neither agonistic nor antagonistic by nature, unlike the agonistic EBD, with different ligands showing agonism or antagonism of the channel. Potent non-electrophilic activators of TRPA1 include cannabinoids (*e.g.* **THC**),[59–61] the anaesthetic propofol (**6**)[57] and **PF-4840154** which is now used as a reference agonist for TRPA1 in place of mustard oil by Pfizer.[62] The antagonist **A-967079** shares the same binding site as the above agonists in a pocket between the S5, S6 and P1 helices.[63] A key interaction between **A-967079** and TRPA1 involves a π-stacking interaction with F909.[56] The theophylline class of antagonists including **HC-030031** bind to a distinct but poorly characterised site near N855 (fig. 1-4).[58]

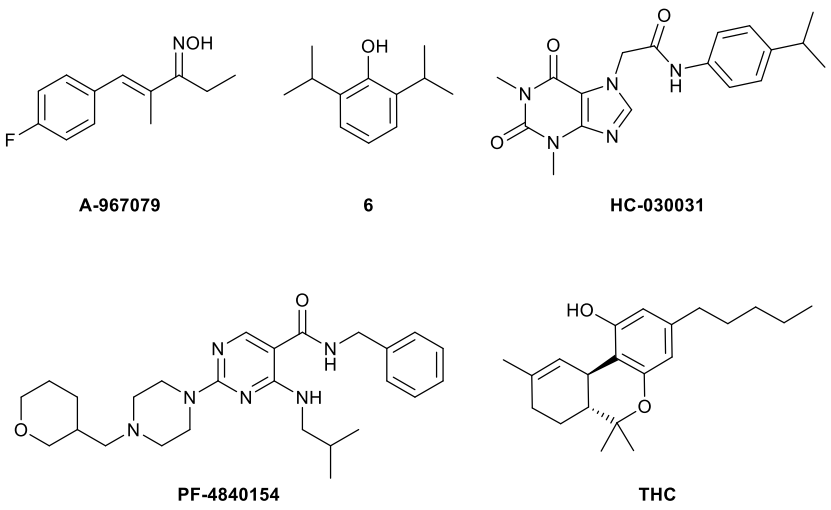


Figure 1-4: Ligands that bind in the transmembrane domain of TRPA1 and stabilise the pore region in the closed or open state. THC = Δ^9-tetrahydrocannabinol.

2.3 Structure of TRPA1

TRPA1 shares many similarities with other TRP channels such as assembling as a tetramer, having 6-TM spanning helices, pore forming helices, and N- and C-terminal cytosolic domains (NTD and CTD respectively). The NTD of TRPA1 contains 16 putative ankyrin repeats, a Ca^{2+} binding EF-hand,[64] and the electrophile binding domain (EBD).[65] Ankyrin repeats are a 30-34 residue secondary structure characterised by a helix-turn-helix motif and a flexible loop. Ankyrins are typically responsible for aiding protein folding and mediating protein-protein interactions such as with factor inhibiting hypoxia induced factor (FIH, *vide-infra*).[66,67] The CTD forms a coiled-coil which helps stabilise the quaternary structure of the protein.[56]

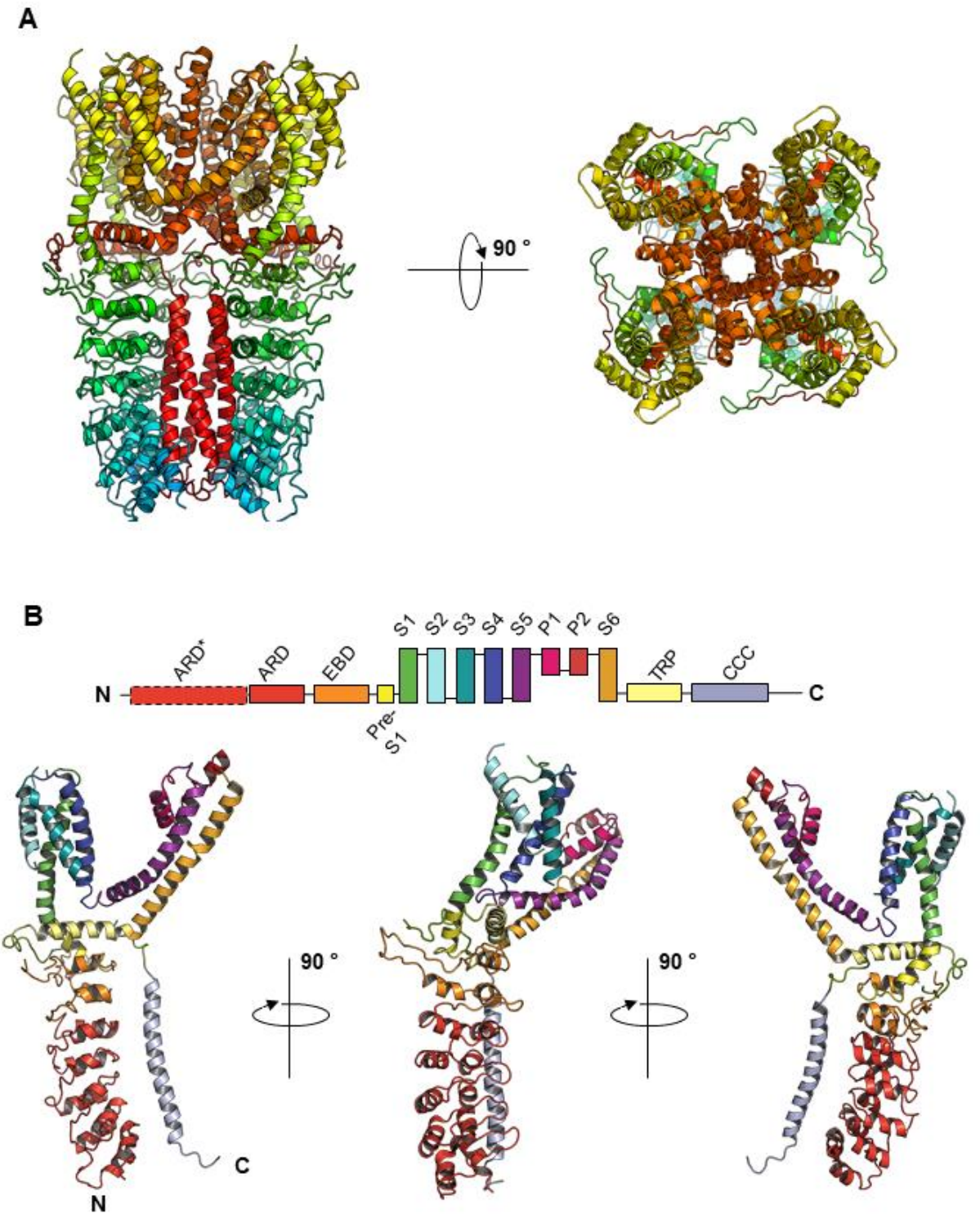


Figure 1-5: Views from a cryo-electron microscopy (cryo-EM) structure of TRPA1 (PDB accession code: 3J9P): A) TRPA1 forms a tetramer with a clear pore, allowing ions to flow through the membrane. **B)** Individual monomer with structural features highlighted from N- to C- terminus: ankyrin repeat domain (ARD), electrophile binding domain (EBD), voltage-sensor domain (S1-S4), pore forming helices (P1-P2), TRP-like box (TRP) and C-terminal coiled-coil (CCC).

The EBD of TRPA1 is lined with hydrophobic and aromatic residues, such as F612, H614, F669, Y680, Y662 and W605, which allow for the entry of electrophiles into the EBD from the membrane.[68] The presence of these residues explains the preference of TRPA1 for lipophilic electrophiles such as **JT010**, PGJ_2 and cinnamaldehyde (fig. 1-6).

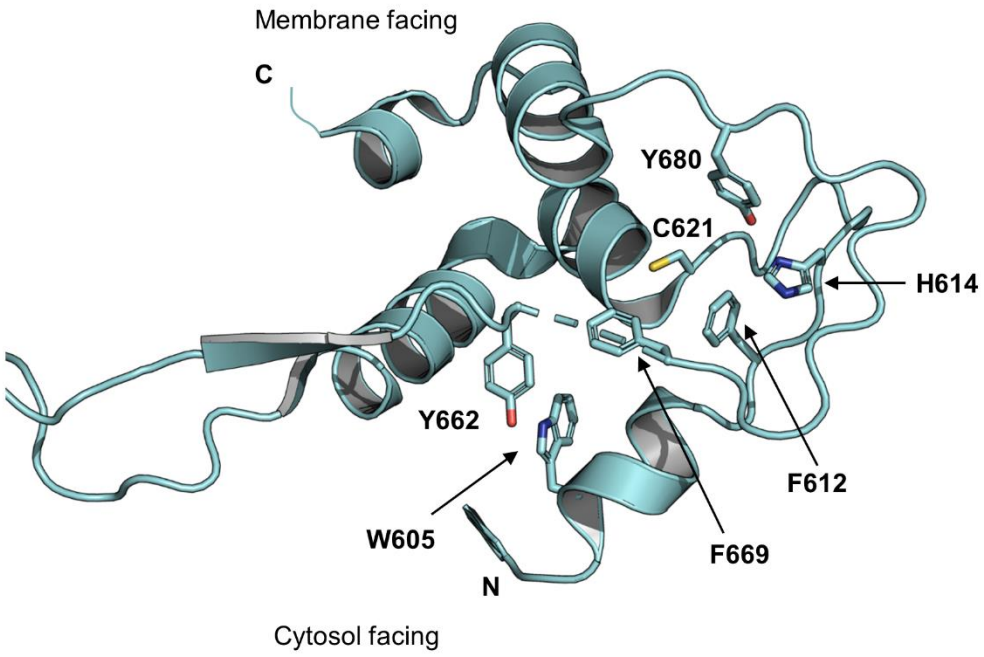


Figure 1-6: Views of the electrophile binding domain of TRPA1 by cryo-EM (PDB accession code: 6V9W). NB: residues 663-668 omitted for clarity.

How TRPA1 transitions from a closed to open state in response to covalent modification on C621 and why C621 has such high nucleophilicity remained unclear until 2020. Suo *et al.*[68] hypothesise that the nucleophilicity of C621 is a result of the interactions of C621 with the dipole of helix H2, stabilising the thiolate anion. H2 is held in place by hydrogen bonding interactions between E628 and E625 of H2 with K620. C621 also interacts with F612 via a thiol-π interaction. This interaction lowers the pKa of C621 and holds C621 in the correct orientation to react with electrophiles (fig. 1-7).

Suo *et al.*[68] and Zhao *et al.*[52] published complimentary mechanisms for TRPA1 opening in response to modification on C621 by a bulky electrophile. Although Suo *et al.*[68] captured TRPA1 in the deactivated/closed state by cryo-electron microscopy (cryo-EM), it was observed that the binding of

JT010 caused a displacement of P666 and a large rearrangement of the region containing C665. Suo *et al.*[68] therefore hypothesised that the top half of the binding site (containing C665, P666 and Y680) acts as a switch that is activated on reaction with electrophiles. The non-specific, hydrophobic interaction formed between Y680 and the ligand allows for TRPA1 to be activated by a wide variety of electrophiles. Zhao *et al.*[52] reported that the rearrangement of the top half of the binding site upon channel activation by a bulky electrophile allows K671 to form hydrogen bonding interactions with the C-terminus of the TRP-like helix. This interaction enhances the positive electrostatic surface potential at the N-terminus of the TRP-like helix via strengthening the helix dipole (fig. 1-7). Zhao *et al.*[52] propose this causes repulsion of any nearby subunits from the TRP-like helix, including the S6 helix which contains the lower portion of the gate, biasing the channel to move to the open state.

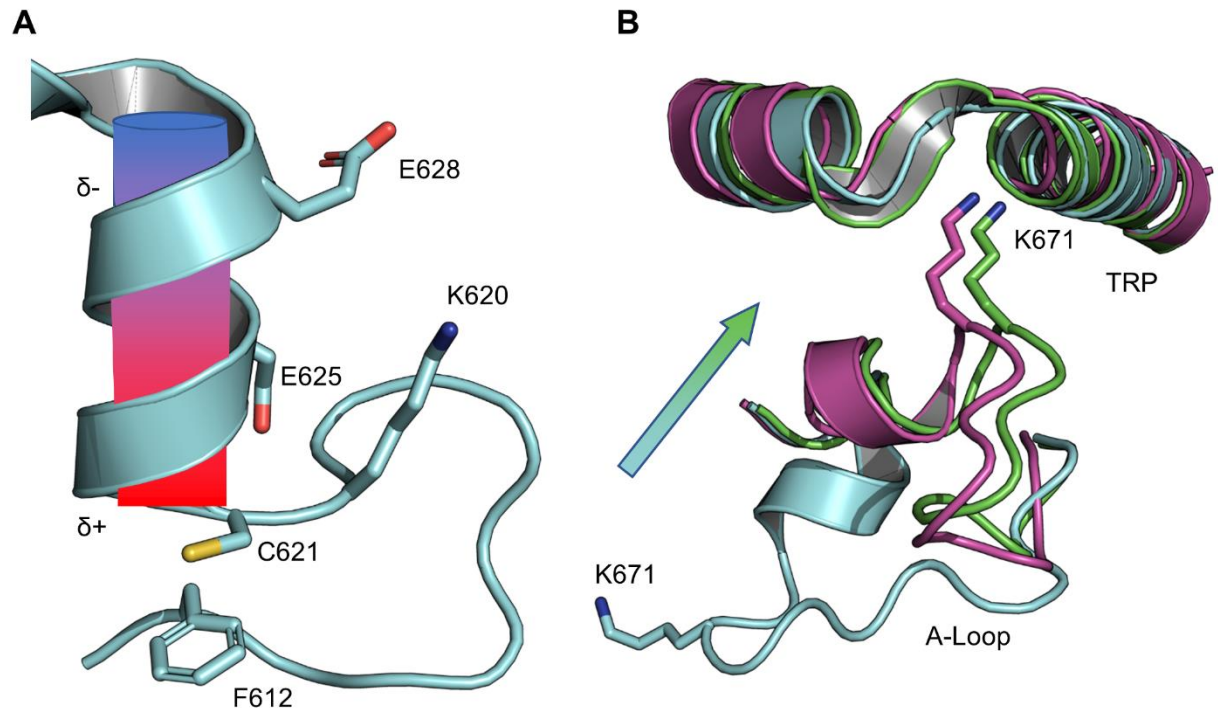


Figure 1-7: Views of the electrophile binding domain of TRPA1 by cryo-EM. A) The highly nucleophilic C621 interacts with the dipole of helix H2 and F612. **B)** Electrophile binding causes a large shift in the orientation of the A-loop, allowing K671 to interact with the TRP-like helix. PDB accession codes: 6V9W (blue), 6V9V (green), 6PQO (pink).

For smaller electrophiles such as **IAA**, Zhao *et al.*[52] propose a two-step mechanism for channel opening. C621 is the primary site of covalent reaction and primes the flexible loop as described above,

however, there is evidence that C665 is also implicated in electrophile binding.[52] Thus, the C665S mutant of TRPA1 showed a decreased current flow as measured by patch-clamp electrophysiology on opening compared to the WT channel in response to **IAA**. Zhao *et al.*[52] propose that modification of C665 in addition to C621 by small electrophiles stabilises the loop configuration described above, more efficiently coupling the loop orientation to the channel gate and stabilising the open conformation.

Zhao *et al.*[52] were also the first to capture a cryo-EM structure of TRPA1 in its open state, allowing for changes in the TM region to be discerned on channel opening. The pore of TRPA1, which allows or blocks the passage of ions through the membrane, is comprised of the S5 and S6 helices, and the P1 and P2 helices. There are two narrow sections of the pore: the upper and lower gates. The upper gate is also known as the selectivity filter and is comprised of D915 and the amide backbone of G914. Mutations of D915 alter the ion selectivity of the channel,[69] confirming the importance of this residue. The lower gate is comprised of I957 and V961, which are part of the S6 helix, which provide a hydrophobic region of the pore which undergoes dewetting (*i.e.* the formation of a section of the pore with little to no water between the monomers). This provides a free energy barrier to ion conduction through the pore which is lowered upon the lower gate dilating (fig. 1-8).[70,71]

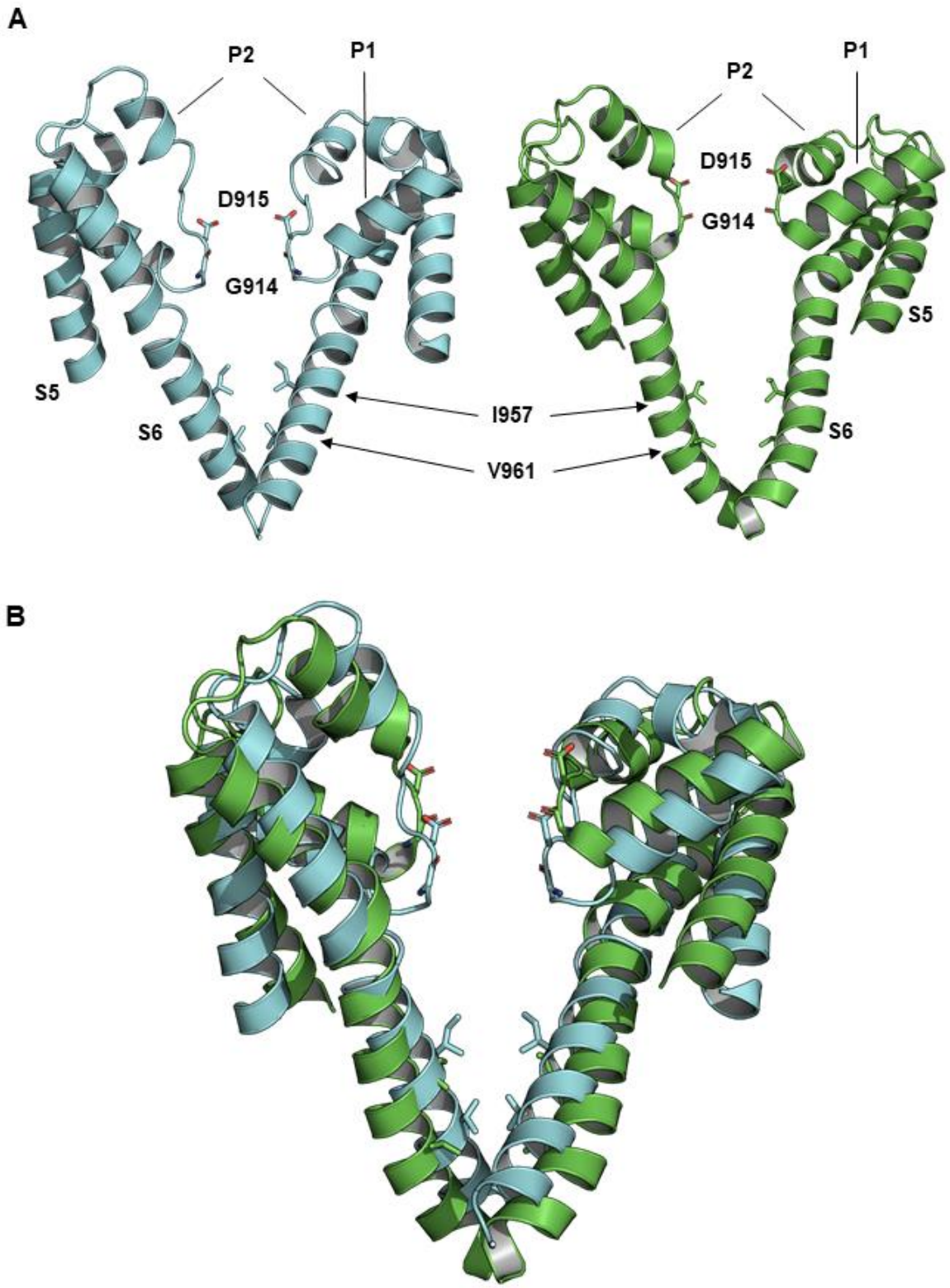


Figure 1-8: Views of the pore region of TRPA1 in the open (green, PDB accession code: 6V9X) and closed states (blue, PDB accession code: 6V9W). A) Residues forming the upper and lower gates of TRPA1 are annotated. **B)** Overlay of the two structures shows the pore helices moving upwards and outwards, and the lower gate dilating upon channel opening.

Upon ligand binding to C621 and transmission of the signal *via* the mechanisms described above, the S5 helix straightens and moves the pore helices upwards and outwards. This exposes acidic residues on the extracellular side of the TM region, attracting cations towards the channel and repelling anions. The lower gate also dilates from a radius of 5.3 Å to 7.8 Å (a hydrated radius of 0.9 and 2.1 Å) allowing the passage of hydrated Ca^{2+} cations (fig. 1-9).[52]

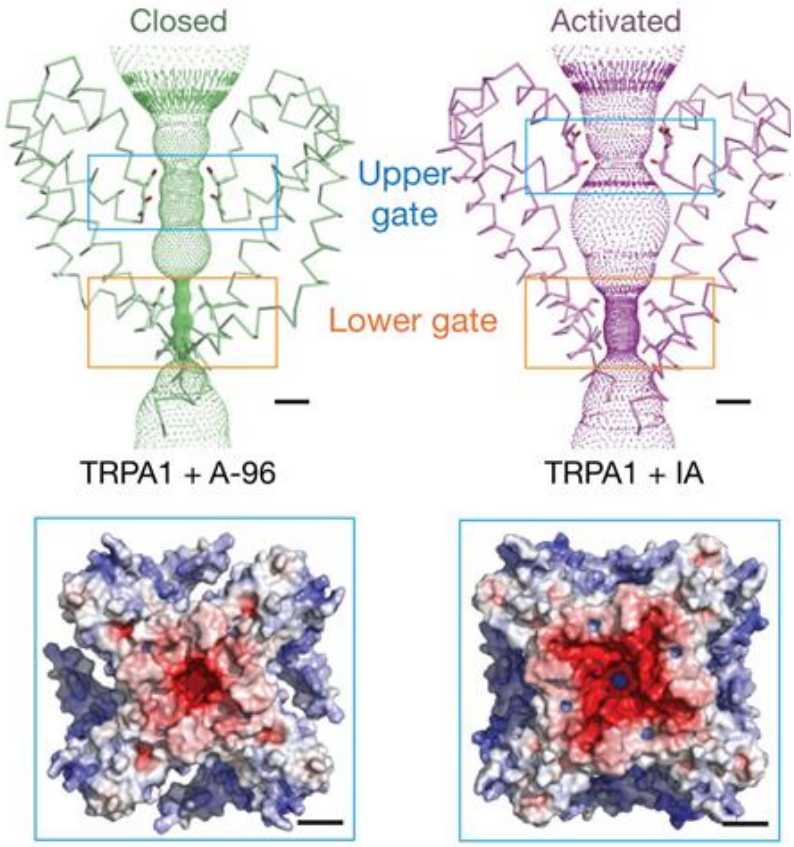


Figure 1-9: Views of the internal and external (charged) surface of TRPA1 pore region in the open (activated) and closed states. (Reprinted with permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature, Irritant-evoked activation and calcium modulation of the TRPA1 receptor, Jianhua Zhao *et al.* (2020)[52])

2.4 Regulation of TRPA1

TRPA1 associates with a number of other proteins, most notably with calmodulin (CaM).[72,73] How CaM binds to TRPA1 is dependent on the Ca^{2+} concentration and the presence of other associated proteins.[73,74] CaM binds to TRPA1 close to its TM domain either on its NTD (residues 581-600) or its CTD (residues 992-1011) in the presence of Ca^{2+}.[73,74] Association of TRPA1 with CaM allows for Ca^{2+} dependent regulation of the channel, potentiating responses at low concentrations and inhibiting

at high concentrations.[69] CaM also facilitates the formation of supramolecular protein complexes involving TRPA1 which are important in signalling cascades. For example, TRPA1 has been found to associate with μ- and δ-opioid receptors, the glutamate N-methyl-d-aspartate (NMDA) and type 1 sigma receptors, with the associations changing upon treatment with a pharmaceutical agent.[73,75] It is unclear what the physiological relevance of these interactions is, however, there appears to be some degree of cross-regulation of these receptors.[73,76]

The TRPA1 gene is co-expressed with that of TRPV1 and so interplay between TRPA1 and TRPV1 seems likely.[59] TRPA1 and TRPV1 co-immunoprecipitate from DRG neurons and human cell expression systems, and the two proteins are shown to interact closely by Förster resonance energy transfer (FRET).[77] Expression of a TRPV1::TRPA1 concatamer forms a functional ion channel that responds to the TRPV1 agonist capsaicin but not TRPA1 agonists.[78] Regulation of TRPA1 by TRPV1 lowers the probability of TRPA1 opening (P_0) at -60 mV, suggesting that each channel cross-desensitises the other.[77]

TRPA1 is regulated by a number of post-translational modifications (PTM). TRPA1 is a substrate for protein kinase A (PKA) with phosphorylation of TRPA1 eliciting a greater response from agonists such as **AITC**.[79–81] PKA is the downstream mediator of bradykinin (a pro-inflammatory nonapeptide) induced hyperalgesia.[81] TRPA1 has been found to associate with PKA anchoring protein (AKAP) which mediates phosphorylation of TRPA1 by PKA.[82] The exact site of modification by PKA has not yet been identified, however S86, S317, S428 and S972 have all been implicated as targets for modification by PKA.[79] TRPA1 is also potentiated by phosphorylation of S449 by cyclin dependent kinase 5 (CDK5).[83] Phosphorylation of S602, however, causes inactivation of TRPA1 but it is unclear which kinase mediates this phosphorylation.[84]

TRPA1 is also modulated by the phosphoinositide PIP_2. PIP_2 is essential for the interaction of the PIP_2 binding site of TRPA1 (residues 1006-1031) with the lipid bilayer and competes with CaM for binding. The PIP_2-dependent regulation of TRPA1 is complicated by differing affinities of PIP_2 for TRPA1 under different Ca^{2+} concentrations.[85] However, an aspect of this interaction at low Ca^{2+} concentration is the prevention of CaM dependent sensitisation of the channel. Upon cleavage of PIP_2 by phospholipase C (PLC), CaM can bind to TRPA1 and potentiate the channel. This may explain why the treatment of rat DRG neurons expressing TRPA1 with PLC inhibitors inhibited sensitisation of TRPA1, while PLC activators sensitised TRPA1.[81]

Another reason for the observed sensitisation of TRPA1 in DRG neurons is that activation of TRPA1 causes an increase in the levels of TRPA1 at the plasma membrane.[86] This appears to be the result of activation of the PKA or PLC pathway as inhibition of these pathways prevents the recruitment of TRPA1 to the membrane.[86] Not only is more TRPA1 present on the plasma membrane after the application of a TRPA1 agonist, but the number of DRG neurons responding to TRPA1 agonists increases in the days after the initial exposure to an agonist. The upregulation of TRPA1 coincides well with the observed hyperalgesia of rats after application of a TRPA1 agonist.[87]

Taken together, the rich interplay of TRPA1 with different kinases produces a variety of channel sensitivities with relevance to nociception under various (pathological) conditions. Additionally, the association of TRPA1 with numerous receptors in the peripheral and central nervous system allows the cross-regulation of pain inducing pathways with TRPA1.

3. Small Molecule Toolkits for TRPA1

3.1 Probing Intracellular Calcium Concentration

Calcium ions are an important secondary messenger within mammalian cells and many signalling pathways are dependent on increases in intracellular Ca^{2+} concentration (hereafter referred to as $[Ca^{2+}]_i$). Microbiologists have therefore sought convenient ways of measuring $[Ca^{2+}]_i$ without disrupting cell viability.

Initial work on measuring Ca^{2+} concentration involved the use of aequorin, a protein isolated from the jellyfish family *Aequorea* found in Friday Harbor in Washington, USA. It was found that aequorin undergoes a light emitting reaction on the addition of Ca^{2+}, with the protein showing strong selectivity for Ca^{2+} and Sr^{2+} and a dissociation constant in the high nanomolar range.[88] Initial work on measuring $[Ca^{2+}]_i$ involved disrupting the cell membrane to allow aequorin to enter the cytoplasm.,[89] however, more recent experiments using aequorin make use of cell lines stably expressing aequorin to avoid disrupting the plasma membrane of the cell.[90]

Despite promising properties, the production of aequorin on a large scale proved difficult. The protein is found in small organs within the jellyfish which can be damaged easily, destroying some of the aequorin in the process, thus requiring them to be caught gently by hand.[91] This coupled with the fact that several thousand jellyfish are required to produce 1 mg of the protein made the case that alternative Ca^{2+} sensing molecules needed to be developed.[91] Modern experiments making use of cell lines stably expressing aequorin still need to be incubated with a membrane permeable derivative of the aequorin co-factor coelenterazine before luminescence can be observed.[90] Thus alternatives to aequorin were sought.

Ca^{2+} is an ion which can be readily chelated by polydentate ligands to give a stable complex. Selectivity for Ca^{2+} over other ions can be achieved by use of derivatives of the ligand 1,2-bis(o-

aminophenoxy)-ethane-*N*,*N*,*N*',*N*'-tetraacetic acid (BAPTA, **7**). Upon binding Ca^{2+}, the UV spectrum of the ligand changes profoundly as a result of loss of delocalisation of the nitrogen lone pair into the π_g* orbital on the phenyl ring. This manifests in a decrease in extinction coefficient at 255 nm with increasing Ca^{2+} concentration.[92] When coupled to a fluorescent moiety as in Fluo-3 and Fluo-4, the fluorescence intensity increases upon Ca^{2+} binding, allowing for the relative change in Ca^{2+} concentration to be measured.[93] Fluo-3 (**8**) and Fluo-4 (**9**) are typically used as acetoxymethyl (AM) esters (shown in red) which increases cell permeability. The AM ester is cleaved by esterases within the cell, allowing Ca^{2+} binding (fig. 1-10).

Although TRPA1 is primarily a non-selective cation permeable ion channel, its activity can be assayed using calcium sensitive dyes. When an agonist is applied to cells expressing functional TRPA1 preloaded with **9**, an increase in fluorescence can be observed due to an increase in $[Ca^{2+}]_i$. This allows for the construction of a dose-response curve from which half-maximal effect concentration (EC_{50}) can be calculated.[49] Cell lines expressing TRPA1 may also be pre-incubated with an antagonist and a fixed concentration of agonist added. This allows for a concentration dependent decrease in response to be observed. An IC_{50} for the antagonist may then be calculated.

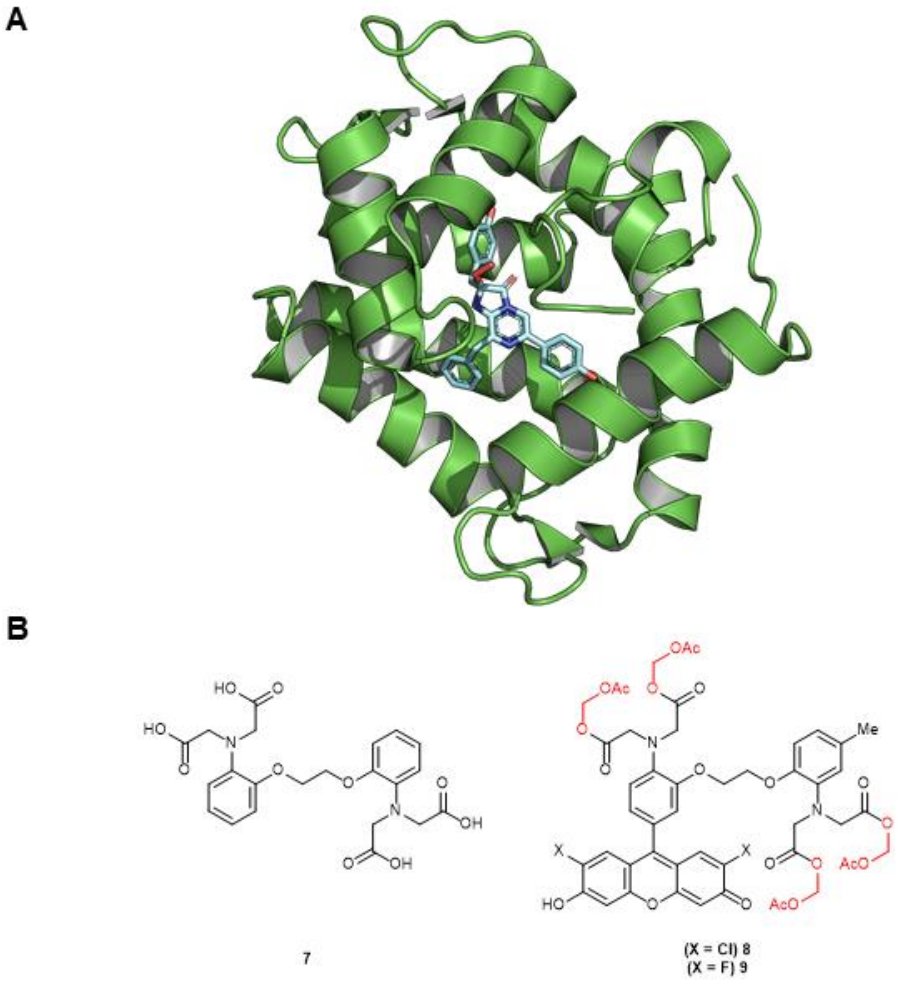


Figure 1-10: Structures of calcium sensitive molecules used to study Ca^{2+} concentration in cells. A) Aequorin (PDB accession code: 1E3J) with the peroxide of coelenterazine bound. Decomposition of this peroxide gives rise to luminescence.[126] **B)** Structures of BAPTA (**7**) and Ca^{2+} sensitive dyes. NB: Acetoxymethyl (AM) ester group shown in red.

3.2 Commonly Used Ligands

Although TRPA1 responds to a wide variety of stimuli, a select few ligands are used to study TRPA1 *in-vivo* or *ex-vivo* neurons. The most selective agonist for TRPA1 is **JT010** (see fig. 1-3) which shows excellent selectivity and potency in inducing activity over other TRP channels with an EC_{50} value of 0.65 nM.[51] Pulldown studies from HEK293 overproducing TRPA1 using biotinylated **JT010** showed a high level of selectivity for TRPA1 once the background protein binding of avidin agarose beads was discounted.[51] Although **JT010** has been used in relatively few studies, it has been used as an agonist to determine the confirmation of the electrophile binding domain of TRPA1[68] and has been

used for intradermal injection in humans to assay pain-relieving compounds acting *via* TRPA1.[94] However, to date, no extensive studies on the interactions of **JT010** with the human proteome have been reported.

The more commonly used TRPA1 agonists are cinnamaldehyde and mustard oil, of which the principal component is **AITC** (see fig. 1-3).[55] **AITC** is less selective, activating TRPV1 at higher concentrations which inhibit TRPA1.[95] The EC_{50} value of **AITC** for mTRPV1 is 3.0 ± 0.6 mM[95] whereas it is 22 ± 3 μM for mTRPA1,[54] thus showing 100-fold selectivity. Off-target effects of **AITC** include the inhibition of cytochrome P450 2B1 with an IC_{50} value of 78 μM[96] and disruption of the mitochondrial membrane potential, damaging mitochondria at 100 μM in cancerous UM-UC-3 cells.[97] It appears that cinnamaldehyde modulates toll-like receptor activation[98] and inhibits cytokine production[99] in macrophages, though it is unclear what relevance this has to neuronal cells where TRPA1 would be expressed.

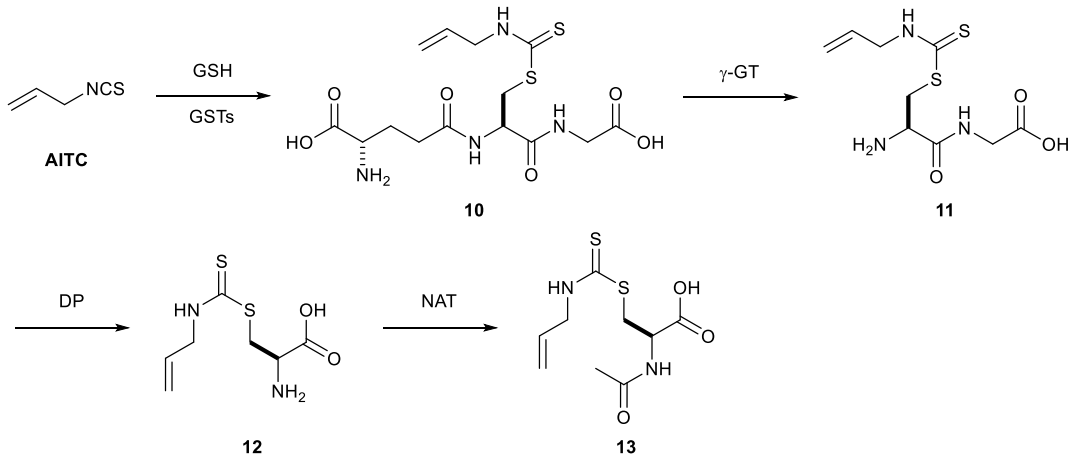


Figure 1-11: Metabolism of allyl isothiocyanate (AITC) *via* the mercapturic pathway. GSH = glutathione, GST = glutathione S-transferase, γ-GT = γ-glutamyltransferase, DP = dipeptidase, NAT = N-acetyltransferase

Both **AITC** and cinnamaldehyde are metabolised via the mercapturic pathway.[100] **AITC** is rapidly and reversibly conjugated to glutathione (GSH) with the reaction being catalysed by glutathione *S*-transferase (GST) M1-1 and GST P1-1.[101,102] The resulting *S*-alkylated GSH derivative **10** undergoes

hydrolysis on the isopeptide bond of glutamate which is catalysed by γ-glutamyltransferase (γ-GT) to give **11**. A dipeptidase (DP) cleaves **11** to give the cysteine derivative **12** which is then acetylated by N-acetyltransferase (NAT) to give **13** (fig. 1-11).[100]

The antagonists **A-967079** and **HC-030031** (see fig. 1-4) both show excellent selectivity for TRPA1 over other receptors.[103,104] Extensive studies on the interaction of these molecules with the rest of the human proteome have not been reported. An approach to study these interactions could involve the generation of a photoaffinity probe or pull-down probe which can capture the proteins it reacts with which can then be assigned by proteomics.

4. Hypoxia

The role of oxygen in eukaryotic life is essential. While the atmospheric concentration of oxygen in the atmosphere is 20.9 %, most tissues in metazoans have a much lower concentration of oxygen, around 2 – 9 %. Below this "physiological normoxia", tissues evoke a hypoxic response.[105] Mechanisms for organisms to adapt to a low oxygen environment have evolved, including the increased uptake of glucose, upregulation of glycolytic enzymes, suppression of oxidative-phosphorylation and metabolic rate, and physiological changes to increase the efficiency of transport of O_2 (*e.g.* erythropoesis and angiogenesis).[106]

In humans and other mammals, the chronic hypoxic (*i.e.* on a timescale of hours) response is substantially mediated by hypoxia-inducible transcription Factor (HIF).[107] HIF is a heterodimer which contains the basic-helix-loop-helix (bHLH) and per-Arnt-Sim (PAS) domains which facilitate its binding to hypoxia response elements (HREs) in DNA which have a consensus sequence of (G/A)CGTG.[108] HREs are present in the promoter regions of specific genes and binding of the HIF heterodimer causes increased levels of transcription upon association of the complex with the transcription coactivator p300.[109] HIF is comprised of an α- and β- subunit. There are three human

HIF-α homologues, HIF-1α which is ubiquitously expressed, and HIF-2α and HIF-3α which show tissue specific expression.

The hydroxylation of human HIF-1α on P402 and P564, and the analogous residues in HIF-2α, is carried out by a set of Fe^{2+} and 2-oxoglutarate (2-OG) dependent enzymes, the HIF prolyl hydroxylases (PHDs). In addition to proline hydroxylation, hydroxylation of HIF-1α at the β-carbon of N803 is carried out by factor inhibiting HIF (FIH).[110] The hydroxylation of proline residues in the oxygen-dependent degradation domain (ODD) of HIF-α facilitates the binding of the HIF-α isoforms to the von Hippel-Lindau factor (pVHL), subsequent poly-ubiquitination and degradation via the proteosome. Hydroxylation of N803 in the C-terminal transactivation domain (CTAD) of HIF-1α prevents the recruitment of p300 (fig. 1-12).[109] In this manner, the HIF-α subunit is considered to be unstable and the HIF-β subunit stable.[111,112]

The role of FIH and PHDs in the cellular response to hypoxia is more complex than just the hydroxylation of the HIF-α subunits. It is reported that the PHDs may also hydroxylate proline residues in other proteins that contain the consensus sequence LXXLAP (where X is any amino acid) which is found in the ODD of HIF-1α,[113] although the strength of association of this consensus with hydroxylation is unclear[114] and recent work casts doubt on whether many reported non-HIF-α substrates are true substrates for PHDs.[115] FIH mediated hydroxylation of asparagines within the ankyrin repeat domain of proteins has been reported by several groups and is likely widespread.[116] An ankyrin repeat is a 30-34 residue structure that consists of a helix-turn-helix motif followed by a flexible loop.[117] There are a variety of proposed consensus sequences for ankyrin repeats which are also substrates of FIH which are linked by a leucine in the -8 position (relative to the hydroxylated asparagine) and E/D-V/I in the -2 and -1 positions (fig. 1-13).[118–120]

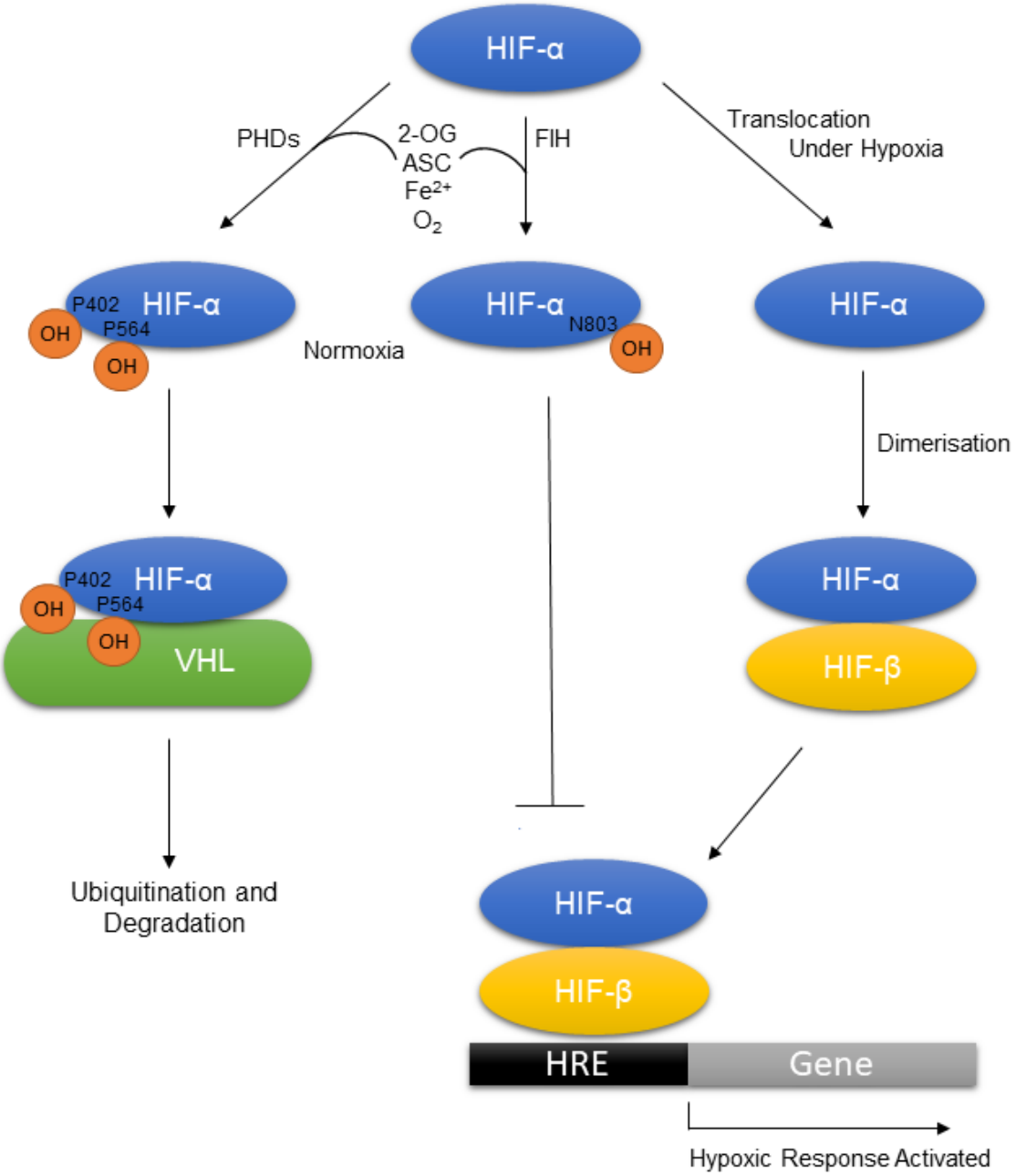


Figure 1-12: Summary of the activation of the hypoxic response *via* HIF. Under normoxia HIF-α is hydroxylated by PHDs or FIH, targeting the protein for degradation or inhibiting dimerisation with HIF-β respectively. Under hypoxia, HIF-α undergoes translocation to the nucleus where it binds to HIF-β, allowing the heterodimer to bind to hypoxic response elements (HRE) in DNA and promote transcription of genes involved in the hypoxic response. 2-OG = 2-oxoglutarate, ASC = L-ascorbic acid. VHL = von Hippel-Lindau factor.

TRPA1 is an oxygen sensing protein, displaying "U"-shaped calcium influx with a minimum around 18% O_2 which is close to the value of atmospheric oxygen concentration. It is proposed that the increased sensitivity to agonists under hyperoxia is the result of oxidation of cysteine residues on the

A

Protein	Sequence
HIF-1α	ESGLPQLTSYD**CEVN**API
P105 (195-215)	LPCLLLLVAAG**ADVN**AQEQ
IKBα	LGIVELLVSLG**ADVN**AQEP
IKBα	PDLVSLLLKCG**ADVN**RVTY
Notch-1	EGMLEDLINSH**ADVN**AVDD
Notch-1	SDAAKRLLEAS**ADAN**IQDN
1CA	LEVVKLLLEAG**ADVN**AQDK
TRPV3	GDIAALLIAAG**ADVN**AHAK
FIH Consensus	------L----$\pi^{D}_{E}\varphi$N----
TRPA1*	HELADYLISVG**ADIN**KIDS

B

Protein	Sequence
HIF-1α (NODD)	KEPDALTL**LAP**-AAGD
HIF-2α (NODD)	EEPEELAQ**LAP**-TPGD
HIF-1α (CODD)	DTDLDLEM**LAP**YIPMD
HIF-2α (CODD)	FSELDLET**LAP**YIPMD
PHD2 Consensus	-----L--**LAP**-----
TRPA1	QQPYGLKNLRP-EFMQ

Figure 1-13: Substrates of the HIF-hydroxylases show conserved residues. A) Selected reported FIH substrates including TRPA1 ankyrin 8 as a putative substrate. π = small uncharged residue, ϕ = small hydrophobic residue. **B)** Selected reported PHD substrates.

N-terminal cytosolic domain to give cysteine sulfinic acids.[121] The changing sensitivity of TRPA1 with decreasing oxygen concentration has been reported to be the result of hydroxylation on P394 by PHD2,[121,122] although this report has not been validated by biochemical studies on PHDs.[115] However,

it is also possible that TRPA1 is a substate for FIH given several of its ankyrin repeats fit the FIH consensus sequence and that membrane proteins such as Notch-1 serve to regulate the level of circulating FIH in the cytoplasm.[123] The potential of FIH mediated hydroxylation of TRPA1 ankyrins was examined in the work of this thesis.

5. Thesis Objectives

The principal aims of the work described in this thesis can be divided into three categories:

- To investigate the reaction between TRPA1 and HCHO
- To investigate the binding site of **HC-030031** and **GRC-17536** in TRPA1
- To investigate the nature of the potential interaction between TRPA1 and FIH

6. Chapter I Bibliography:

(1) Terman, G. W.; Bonica, J. J. Spinal Mechanisms and Their Modulation. In *Bonica's Management of Pain 3rd Edition*; Loeser, J. D., Butler, S. H., Chapman, R. C., Turk, D. C., Eds.; Lippincott Williams & Wilkins, 2001; p 99.

(2) Dubin, A. E.; Patapoutian, A. Nociceptors: The Sensors of the Pain Pathway. *Journal of Clinical Investigation*. American Society for Clinical Investigation November 1, 2010, pp 3760–3772. https://doi.org/10.1172/JCI42843.

(3) Colloca, L.; Ludman, T.; Bouhassira, D.; Baron, R.; Dickenson, A. H.; Yarnitsky, D.; Freeman, R.; Truini, A.; Attal, N.; Finnerup, N. B.; Eccleston, C.; Kalso, E.; Bennett, D. L.; Dworkin, R. H.; Raja, S. N. Neuropathic Pain. *Nature Reviews Disease Primers* **2017**, *3* (1), 1–19. https://doi.org/10.1038/nrdp.2017.2.

(4) Purves, D.; Augustine, G. J.; Fitzpatrick, D.; Katz, L. C.; LaMantia, A.; McNamara, J. O.; Williams, S. M. *Neuroscience*; Sinauer Associates: Sunderland MA, 2001.

(5) Siegel, A. *Essential Neuroscience*; Lippiincott Williams & Wilkins: Baltimore, MD, 2006.

(6) Bean, B. P. The Action Potential in Mammalian Central Neurons. *Nature Reviews Neuroscience*. Nature Publishing Group June 2007, pp 451–465. https://doi.org/10.1038/nrn2148.

(7) Pirahanchi, Y.; Aeddula, N. R. *Physiology, Sodium Potassium Pump (Na+ K+ Pump)*; StatPearls Publishing, 2019.

(8) Blaustein, M. P.; Lederer, W. J. Sodium/Calcium Exchange: Its Physiological Implications. *Physiological Reviews*. American Physiological Society 1999, pp 763–854. https://doi.org/10.1152/physrev.1999.79.3.763.

(9) Carafoli, E. Calcium Pump of the Plasma Membrane. *Physiological Reviews*. 1991, pp 129–153. https://doi.org/10.1152/physrev.1991.71.1.129.

(10) Hodgkin, A. L.; Huxley, A. F. Currents Carried by Sodium and Potassium Ions through the Membrane of the Giant Axon of Loligo. *The Journal of Physiology* **1952**, *116* (4), 449–472. https://doi.org/10.1113/jphysiol.1952.sp004717.

(11) Hodgkin, A. L.; Huxley, A. F. A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve. *The Journal of Physiology* **1952**, *117* (4), 500–544. https://doi.org/10.1113/jphysiol.1952.sp004764.

(12) Hodgkin, A. L. Evidence for Electrical Transmission in Nerve: Part II. *The Journal of Physiology* **1937**, *90* (2), 211–232. https://doi.org/10.1113/jphysiol.1937.sp003508.

(13) Hodgkin, A. L. Evidence for Electrical Transmission in Nerve: Part I. *The Journal of Physiology* **1937**, *90* (2), 183–210. https://doi.org/10.1113/jphysiol.1937.sp003507.

(14) Neishabouri, A.; Faisal, A. A. Saltatory Conduction in Unmyelinated Axons: Clustering of Na+ Channels on Lipid Rafts Enables Micro-Saltatory Conduction in C-Fibers. *Frontiers in Neuroanatomy* **2014**, *8* (OCT), 109. https://doi.org/10.3389/fnana.2014.00109.

(15) Dubin, A. E.; Patapoutian, A. Nociceptors: The Sensors of the Pain Pathway. *Journal of Clinical Investigation*. American Society for Clinical Investigation November 1, 2010, pp 3760–3772. https://doi.org/10.1172/JCI42843.

(16) Julius, D.; Basbaum, A. I. Molecular Mechanisms of Nociception. *Nature*. Nature Publishing Group September 13, 2001, pp 203–210. https://doi.org/10.1038/35093019.

(17) Minke, B. The History of the Drosophila TRP Channel: The Birth of a New Channel Superfamily. *Journal of Neurogenetics*. NIH Public Access December 2010, pp 216–233. https://doi.org/10.3109/01677063.2010.514369.

(18) Cosens, D. J.; Manning, A. Abnormal Electroretinogram from a Drosophila Mutant. *Nature*. Nature Publishing Group 1969, pp 285–287. https://doi.org/10.1038/224285a0.

(19) Montell, C.; Rubin, G. M. Molecular Characterization of the Drosophila Trp Locus: A Putative Integral Membrane Protein Required for Phototransduction. *Neuron* **1989**, *2* (4), 1313–1323. https://doi.org/10.1016/0896-6273(89)90069-X.

(20) Gaudet, R. *Structural Insights into the Function of TRP Channels*; CRC Press/Taylor & Francis, 2007.

(21) Himmel, N. J.; Cox, D. N. Transient Receptor Potential Channels: Current Perspectives on Evolution. *Proceedings of the Royal Society B: Biological Sciences*. Royal Society Publishing August 26, 2020. https://doi.org/10.1098/rspb.2020.1309.

(22) Wang, H.; Cheng, X.; Tian, J.; Xiao, Y.; Tian, T.; Xu, F.; Hong, X.; Zhu, M. X. TRPC Channels: Structure, Function, Regulation and Recent Advances in Small Molecular Probes. *Pharmacology and Therapeutics*. Elsevier Inc. May 1, 2020, p 107497. https://doi.org/10.1016/j.pharmthera.2020.107497.

(23) Jardín, I.; López, J. J.; Diez, R.; Sánchez-Collado, J.; Cantonero, C.; Albarrán, L.; Woodard, G. E.; Redondo, P. C.; Salido, G. M.; Smani, T.; Rosado, J. A. TRPs in Pain Sensation. *Frontiers in Physiology*. Frontiers Media S.A. June 9, 2017, p 392. https://doi.org/10.3389/fphys.2017.00392.

(24) Pérez De Vega, M. J.; Gómez-Monterrey, I.; Ferrer-Montiel, A.; González-Muñiz, R. Transient Receptor Potential Melastatin 8 Channel (TRPM8) Modulation: Cool Entryway for Treating Pain and Cancer. *Journal of Medicinal Chemistry*. American Chemical Society November 23, 2016, pp 10006–10029. https://doi.org/10.1021/acs.jmedchem.6b00305.

(25) Kobayashi, K.; Fukuoka, T.; Obata, K.; Yamanaka, H.; Dai, Y.; Tokunaga, A.; Noguchi, K. Distinct Expression of TRPM8, TRPA1, and TRPV1 MRNAs in Rat Primary Afferent Neurons with Aδ/C-Fibers and Colocalization with Trk Receptors. *Journal of Comparative Neurology* **2005**, *493* (4), 596–606. https://doi.org/10.1002/cne.20794.

(26) Proudfoot, C. J.; Garry, E. M.; Cottrell, D. F.; Rosie, R.; Anderson, H.; Robertson, D. C.; Fleetwood-Walker, S. M.; Mitchell, R. Analgesia Mediated by the TRPM8 Cold Receptor in Chronic Neuropathic Pain. *Current Biology* **2006**, *16* (16), 1591–1605. https://doi.org/10.1016/j.cub.2006.07.061.

(27) Ramachandran, R.; Hyun, E.; Zhao, L.; Lapointe, T. K.; Chapman, K.; Hirot, C. L.; Ghosh, S.; McKemy, D. D.; Vergnolle, N.; Beck, P. L.; Altier, C.; Hollenberg, M. D. TRPM8 Activation Attenuates Inflammatory Responses in Mouse Models of Colitis. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110* (18), 7476–7481. https://doi.org/10.1073/pnas.1217431110.

(28) Parks, D. J.; Parsons, W. H.; Colburn, R. W.; Meegalla, S. K.; Ballentine, S. K.; Illig, C. R.; Qin, N.; Liu, Y.; Hutchinson, T. L.; Lubin, M. lou; Stone, D. J.; Baker, J. F.; Schneider, C. R.; Ma, J.; Damiano, B. P.; Flores, C. M.; Player, M. R. Design and Optimization of Benzimidazole-Containing Transient Receptor Potential Melastatin 8 (TRPM8) Antagonists. *Journal of Medicinal Chemistry* **2011**, *54* (1), 233–247. https://doi.org/10.1021/jm101075v.

(29) Zuo, X.; Ling, J. X.; Xu, G.-Y.; Gu, J. G. Operant Behavioral Responses to Orofacial Cold Stimuli in Rats with Chronic Constrictive Trigeminal Nerve Injury: Effects of Menthol and Capsazepine. *Molecular pain* **2013**, *9* (1), 28. https://doi.org/10.1186/1744-8069-9-28.

(30) Caterina, M. J.; Schumacher, M. A.; Tominaga, M.; Rosen, T. A.; Levine, J. D.; Julius, D. The Capsaicin Receptor: A Heat-Activated Ion Channel in the Pain Pathway. *Nature* **1997**, *389* (6653), 816–824. https://doi.org/10.1038/39807.

(31) Davis, J. B.; Gray, J.; Gunthorpe, M. J.; Hatcher, J. P.; Davey, P. T.; Overend, P.; Harries, M. H.; Latcham, J.; Clapham, C.; Atkinson, K.; Hughes, S. A.; Rance, K.; Grau, E.; Harper, A. J.; Pugh, P. L.; Rogers, D. C.; Bingham, S.; Randall, A.; Sheardown, S. A. Vanilloid Receptor-1 Is Essential for Inflammatory Thermal Hyperalgesia. *Nature* **2000**, *405* (6783), 183–187. https://doi.org/10.1038/35012076.

(32) Tékus, V.; Bölcskei, K.; Kis-Varga, Á.; Dézsi, L.; Szentirmay, É.; Visegrády, A.; Horváth, C.; Szolcsányi, J.; Petho, G. Effect of Transient Receptor Potential Vanilloid 1 (TRPV1) Receptor Antagonist Compounds SB705498, BCTC and AMG9810 in Rat Models of Thermal Hyperalgesia Measured with an Increasing-Temperature Water Bath. *European Journal of Pharmacology* **2010**, *641* (2–3), 135–141. https://doi.org/10.1016/j.ejphar.2010.05.052.

(33) Bautista, D. M.; Jordt, S. E.; Nikai, T.; Tsuruda, P. R.; Read, A. J.; Poblete, J.; Yamoah, E. N.; Basbaum, A. I.; Julius, D. TRPA1 Mediates the Inflammatory Actions of Environmental Irritants and Proalgesic Agents. *Cell* **2006**, *124* (6), 1269–1282. https://doi.org/10.1016/j.cell.2006.02.023.

(34) Vandewauw, I.; Owsianik, G.; Voets, T. Systematic and Quantitative MRNA Expression Analysis of TRP Channel Genes at the Single Trigeminal and Dorsal Root Ganglion Level in Mouse. *BMC Neuroscience* **2013**, *14* (1), 21. https://doi.org/10.1186/1471-2202-14-21.

(35) Taylor-Clark, T. E.; McAlexander, M. A.; Nassenstein, C.; Sheardown, S. A.; Wilson, S.; Thornton, J.; Carr, M. J.; Undem, B. J. Relative Contributions of TRPA1 and TRPV1 Channels in the Activation of Vagal Bronchopulmonary C-Fibres by the Endogenous Autacoid 4-Oxononenal. *The Journal of Physiology* **2008**, *586* (14), 3447–3459. https://doi.org/10.1113/jphysiol.2008.153585.

(36) Burt, A. M. *Textbook of Neuroanatomy*, 1st ed.; W.B. Saunders: Philadelphia, 1993.

(37) Cattaruzza, F.; Spreadbury, I.; Miranda-Morales, M.; Grady, E. F.; Vanner, S.; Bunnett, N. W. Transient Receptor Potential Ankyrin-1 Has a Major Role in Mediating Visceral Pain in Mice. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **2010**, *298* (1), G81–G91. https://doi.org/10.1152/ajpgi.00221.2009.

(38) Nagata, K.; Duggan, A.; Kumar, G.; García-Añoveros, J. Nociceptor and Hair Cell Transducer Properties of TRPA1, a Channel for Pain and Hearing. *Journal of Neuroscience* **2005**, *25* (16), 4052–4061. https://doi.org/10.1523/JNEUROSCI.0013-05.2005.

(39) Sikandar, S.; Dickenson, A. H. Visceral Pain: The Ins and Outs, the Ups and Downs. *Current Opinion in Supportive and Palliative Care*. Europe PMC Funders March 2012, pp 17–26. https://doi.org/10.1097/SPC.0b013e32834f6ec9.

(40) Corey, D. P.; Garcia-Añoveros, J.; Holt, J. R.; Kwan, K. Y.; Lin, S. Y.; Vollrath, M. A.; Amalfitano, A.; Cheung, E. L. M.; Derfler, B. H.; Duggan, A.; Géléoc, G. S. G.; Gray, P. A.; Hoffman, M. P.; Rehm, H. L.; Tamasauskas, D.; Zhang, D. S. TRPA1 Is a Candidate for the Mechanosensitive Transduction Channel of Vertebrate Hair Cells. *Nature* **2004**, *432* (7018), 723–730. https://doi.org/10.1038/nature03066.

(41) Kunert-Keil, C.; Bisping, F.; Krüger, J.; Brinkmeier, H. Tissue-Specific Expression of TRP Channel Genes in the Mouse and Its Variation in Three Different Mouse Strains. *BMC Genomics* **2006**, *7*, 159. https://doi.org/10.1186/1471-2164-7-159.

(42) Edelmayer, R. M.; Le, L. N.; Yan, J.; Wei, X.; Nassini, R.; Materazzi, S.; Preti, D.; Appendino, G.; Geppetti, P.; Dodick, D. W.; Vanderah, T. W.; Porreca, F.; Dussor, G. Activation of TRPA1 on Dural Afferents: A Potential Mechanism of Headache Pain. *Pain* **2012**, *153* (9), 1949–1958. https://doi.org/10.1016/j.pain.2012.06.012.

(43) Mukhopadhyay, I.; Gomes, P.; Aranake, S.; Shetty, M.; Karnik, P.; Damle, M.; Kuruganti, S.; Thorat, S.; Khairatkar-Joshi, N. Expression of Functional TRPA1 Receptor on Human Lung Fibroblast and Epithelial Cells. *Journal of Receptors and Signal Transduction* **2011**, *31* (5), 350–358. https://doi.org/10.3109/10799893.2011.602413.

(44) Atoyan, R.; Shander, D.; Botchkareva, N. v. Non-Neuronal Expression of Transient Receptor Potential Type A1 (TRPA1) in Human Skin. *Journal of Investigative Dermatology*. Elsevier September 1, 2009, pp 2312–2315. https://doi.org/10.1038/jid.2009.58.

(45) Egbuniwe, O.; Grover, S.; Duggal, A. K.; Mavroudis, A.; Yazdi, M.; Renton, T.; di Silvio, L.; Grant, A. D. TRPA1 and TRPV4 Activation in Human Odontoblasts Stimulates ATP Release. *Journal of Dental Research* **2014**, *93* (9), 911–917. https://doi.org/10.1177/0022034514544507.

(46) Macpherson, L. J.; Dubin, A. E.; Evans, M. J.; Marr, F.; Schultz, P. G.; Cravatt, B. F.; Patapoutian, A. Noxious Compounds Activate TRPA1 Ion Channels through Covalent Modification of Cysteines. *Nature* **2007**, *445* (7127), 541–545. https://doi.org/10.1038/nature05544.

(47) Taylor-Clark, T. E.; Undem, B. J.; MacGlashan, D. W.; Ghatta, S.; Carr, M. J.; McAlexander, M. A. Prostaglandin-Induced Activation of Nociceptive Neurons via Direct Interaction with Transient

Receptor Potential A1 (TRPA1). *Molecular Pharmacology* **2007**, *73* (2), 274–281. https://doi.org/10.1124/mol.107.040832.

(48) Benedetti, A.; Comporti, M.; Esterbauer, H. Identification of 4-Hydroxynonenal as a Cytotoxic Product Originating from the Peroxidation of Liver Microsomal Lipids. *Biochimica et Biophysica Acta - Lipids and Lipid Metabolism* **1980**, *620* (2), 281–296. https://doi.org/10.1016/0005-2760(80)90209-X.

(49) Lindsay, C. D.; Green, C.; Bird, M.; Jones, J. T. A.; Riches, J. R.; McKee, K. K.; Sandford, M. S.; Wakefield, D. A.; Timperley, C. M. Potency of Irritation by Benzylidenemalononitriles in Humans Correlates with TRPA1 Ion Channel Activation. *Royal Society Open Science* **2015**, *2* (1), 140160. https://doi.org/10.1098/rsos.140160.

(50) Gijsen, H. J. M.; Berthelot, D.; Zaja, M.; Brône, B.; Geuens, I.; Mercken, M. Analogues of Morphanthridine and the Tear Gas Dibenz[b,f][1,4]Oxazepine (CR) as Extremely Potent Activators of the Human Transient Receptor Potential Ankyrin 1 (TRPA1) Channel. *Journal of Medicinal Chemistry* **2010**, *53* (19), 7011–7020. https://doi.org/10.1021/jm100477n.

(51) Takaya, J.; Mio, K.; Shiraishi, T.; Kurokawa, T.; Otsuka, S.; Mori, Y.; Uesugi, M. A Potent and Site-Selective Agonist of TRPA1. *Journal of the American Chemical Society* **2015**, *137* (50), 15859–15864. https://doi.org/10.1021/jacs.5b10162.

(52) Zhao, J.; Lin King, J. v.; Paulsen, C. E.; Cheng, Y.; Julius, D. Irritant-Evoked Activation and Calcium Modulation of the TRPA1 Receptor. *Nature* **2020**, *585* (7823), 141–145. https://doi.org/10.1038/s41586-020-2480-9.

(53) McNamara, C. R.; Mandel-Brehm, J.; Bautista, D. M.; Siemens, J.; Deranian, K. L.; Zhao, M.; Hayward, N. J.; Chong, J. A.; Julius, D.; Moran, M. M.; Fanger, C. M. TRPA1 Mediates Formalin-Induced Pain. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104* (33), 13525–13530. https://doi.org/10.1073/pnas.0705924104.

(54) Bandell, M.; Story, G. M.; Hwang, S. W.; Viswanath, V.; Eid, S. R.; Petrus, M. J.; Earley, T. J.; Patapoutian, A. Noxious Cold Ion Channel TRPA1 Is Activated by Pungent Compounds and Bradykinin. *Neuron* **2004**, *41* (6), 849–857. https://doi.org/10.1016/S0896-6273(04)00150-3.

(55) Jordt, S. E.; Bautista, D. M.; Chuang, H. H.; McKemy, D. D.; Zygmunt, P. M.; Högestätt, E. D.; Meng, I. D.; Julius, D. Mustard Oils and Cannabinoids Excite Sensory Nerve Fibres through the TRP Channel ANKTM1. *Nature* **2004**, *427* (6971), 260–265. https://doi.org/10.1038/nature02282.

(56) Paulsen, C. E.; Armache, J. P.; Gao, Y.; Cheng, Y.; Julius, D. Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms. *Nature* **2015**, *520* (7548), 511–517. https://doi.org/10.1038/nature14367.

(57) Woll, K. A.; Skinner, K. A.; Gianti, E.; Bhanu, N. v.; Garcia, B. A.; Carnevale, V.; Eckenhoff, R. G.; Gaudet, R. Sites Contributing to TRPA1 Activation by the Anesthetic Propofol Identified by Photoaffinity Labeling. *Biophysical Journal* **2017**. https://doi.org/10.1016/j.bpj.2017.08.040.

(58) Gupta, R.; Saito, S.; Mori, Y.; Itoh, S. G.; Okumura, H.; Tominaga, M. Structural Basis of TRPA1 Inhibition by HC-030031 Utilizing Species-Specific Differences. *Scientific Reports* **2016**, *6* (October), 37460. https://doi.org/10.1038/srep37460.

(59) Jordt, S. E.; Bautista, D. M.; Chuang, H. H.; McKemy, D. D.; Zygmunt, P. M.; Högestätt, E. D.; Meng, I. D.; Julius, D. Mustard Oils and Cannabinoids Excite Sensory Nerve Fibres through the TRP Channel ANKTM1. *Nature* **2004**, *427* (6971), 260–265. https://doi.org/10.1038/nature02282.

(60) de Petrocellis, L.; Vellani, V.; Schiano-Moriello, A.; Marini, P.; Magherini, P. C.; Orlando, P.; di Marzo, V. Plant-Derived Cannabinoids Modulate the Activity of Transient Receptor Potential Channels of Ankyrin Type-1 and Melastatin Type-8. *Journal of Pharmacology and Experimental Therapeutics* **2008**, *325* (3), 1007–1015. https://doi.org/10.1124/jpet.107.134809.

(61) Andersson, D. A.; Gentry, C.; Alenmyr, L.; Killander, D.; Lewis, S. E.; Andersson, A.; Bucher, B.; Galzi, J. L.; Sterner, O.; Bevan, S.; Högest, E. D.; Zygmunt, P. M. TRPA1 Mediates Spinal Antinociception Induced by Acetaminophen and the Cannabinoid Δ9-Tetrahydrocannabiorcol. *Nature Communications* **2011**, *2* (1), 551. https://doi.org/10.1038/ncomms1559.

(62) Ryckmans, T.; Aubdool, A. A.; Bodkin, J. v.; Cox, P.; Brain, S. D.; Dupont, T.; Fairman, E.; Hashizume, Y.; Ishii, N.; Kato, T.; Kitching, L.; Newman, J.; Omoto, K.; Rawson, D.; Strover, J. Design and Pharmacological Evaluation of PF-4840154, a Non-Electrophilic Reference Agonist of the TrpA1 Channel. *Bioorganic & Medicinal Chemistry Letters* **2011**, *21* (16), 4857–4859. https://doi.org/10.1016/J.BMCL.2011.06.035.

(63) Ton, H. T.; Phan, T. X.; Abramyan, A. M.; Shi, L.; Ahern, G. P. Identification of a Putative Binding Site Critical for General Anesthetic Activation of TRPA1. *Proceedings of the National Academy of Sciences* **2017**, *114* (14), 3762–3767. https://doi.org/10.1073/pnas.1618144114.

(64) Zurborg, S.; Yurgionas, B.; Jira, J. A.; Caspani, O.; Heppenstall, P. A. Direct Activation of the Ion Channel TRPA1 by Ca2+. *Nature neuroscience* **2007**, *10* (3), 277–279. https://doi.org/10.1038/nn1843.

(65) Paulsen, C. E.; Armache, J.-P.; Gao, Y.; Cheng, Y.; Julius, D. Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms. *Nature* **2015**, *520* (7548), 511–517. https://doi.org/10.1038/nature14367.

(66) Zurborg, S.; Yurgionas, B.; Jira, J. A.; Caspani, O.; Heppenstall, P. A. Direct Activation of the Ion Channel TRPA1 by Ca2+. *Nature Neuroscience* **2007**, *10* (3), 277–279. https://doi.org/10.1038/nn1843.

(67) Li, J.; Mahajan, A.; Tsai, M. D. Ankyrin Repeat: A Unique Motif Mediating Protein-Protein Interactions. *Biochemistry*. Biochemistry December 26, 2006, pp 15168–15178. https://doi.org/10.1021/bi062188q.

(68) Suo, Y.; Wang, Z.; Zubcevic, L.; Hsu, A. L.; He, Q.; Borgnia, M. J.; Ji, R. R.; Lee, S. Y. Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. *Neuron* **2020**, *105* (5), 882-894.e5. https://doi.org/10.1016/j.neuron.2019.11.023.

(69) Wang, Y. Y.; Chang, R. B.; Waters, H. N.; McKemy, D. D.; Liman, E. R. The Nociceptor Ion Channel TRPA1 Is Potentiated and Inactivated by Permeating Calcium Ions. *Journal of Biological Chemistry* **2008**, *283* (47), 32691–32703. https://doi.org/10.1074/jbc.M803568200.

(70) Aryal, P.; Sansom, M. S. P.; Tucker, S. J. Hydrophobic Gating in Ion Channels. *Journal of Molecular Biology*. Academic Press January 16, 2015, pp 121–130. https://doi.org/10.1016/j.jmb.2014.07.030.

(71) Rao, S.; Klesse, G.; Lynch, C. I.; Tucker, S. J.; Sansom, M. S. P. Molecular Simulations of Hydrophobic Gating of Pentameric Ligand Gated Ion Channels: Insights into Water and Ions. *Journal of Physical Chemistry B*. American Chemical Society February 4, 2021, pp 981–994. https://doi.org/10.1021/acs.jpcb.0c09285.

(72) Hasan, R.; Leeson-Payne, A. T. S.; Jaggar, J. H.; Zhang, X. Calmodulin Is Responsible for Ca2+-Dependent Regulation of TRPA1 Channels. *Scientific Reports* **2017**, *7* (1), 45098. https://doi.org/10.1038/srep45098.

(73) Cortés-Montero, E.; Rodríguez-Muñoz, M.; Ruiz-Cantero, M. D. C.; Cobos, E. J.; Sánchez-Blázquez, P.; Garzón-Niño, J. Calmodulin Supports Trpa1 Channel Association with Opioid Receptors and Glutamate Nmda Receptors in the Nervous Tissue. *International Journal of Molecular Sciences* **2021**, *22* (1), 1–21. https://doi.org/10.3390/ijms22010229.

(74) Hasan, R.; Leeson-Payne, A. T. S.; Jaggar, J. H.; Zhang, X. Calmodulin Is Responsible for Ca2+-Dependent Regulation of TRPA1 Channels. *Scientific Reports* **2017**, *7*. https://doi.org/10.1038/srep45098.

(75) Cortés-Montero, E.; Sánchez-Blázquez, P.; Onetti, Y.; Merlos, M.; Garzón, J. Ligands Exert Biased Activity to Regulate Sigma 1 Receptor Interactions with Cationic TRPA1, TRPV1, and TRPM8 Channels. *Frontiers in Pharmacology* **2019**, *10* (JUN). https://doi.org/10.3389/fphar.2019.00634.

(76) Cortés-Montero, E.; Sánchez-Blázquez, P.; Onetti, Y.; Merlos, M.; Garzón, J. Ligands Exert Biased Activity to Regulate Sigma 1 Receptor Interactions With Cationic TRPA1, TRPV1, and TRPM8 Channels. *Frontiers in Pharmacology* **2019**, *10* (JUN), 634. https://doi.org/10.3389/fphar.2019.00634.

(77) Staruschenko, A.; Jeske, N. A.; Akopian, A. N. Contribution of TRPV1-TRPA1 Interaction to the Single Channel Properties of the TRPA1 Channel. *Journal of Biological Chemistry* **2010**, *285* (20), 15167–15177. https://doi.org/10.1074/jbc.M110.106153.

(78) Fischer, M. J. M.; Balasuriya, D.; Jeggle, P.; Goetze, T. A.; McNaughton, P. A.; Reeh, P. W.; Edwardson, J. M. Direct Evidence for Functional TRPV1/TRPA1 Heteromers. *Pflugers Archiv European Journal of Physiology* **2014**, *466* (12), 2229–2241. https://doi.org/10.1007/s00424-014-1497-z.

(79) Meents, J. E.; Fischer, M. J. M.; McNaughton, P. A. Sensitization of TRPA1 by Protein Kinase A. *PLoS ONE* **2017**, *12* (1). https://doi.org/10.1371/journal.pone.0170097.

(80) Miyano, K.; Shiraishi, S.; Minami, K.; Sudo, Y.; Suzuki, M.; Yokoyama, T.; Terawaki, K.; Nonaka, M.; Murata, H.; Higami, Y.; Uezono, Y. Carboplatin Enhances the Activity of Human Transient Receptor Potential Ankyrin 1 through the Cyclic AMP-Protein Kinase A-A-Kinase Anchoring Protein (AKAP) Pathways. *International Journal of Molecular Sciences* **2019**, *20* (13). https://doi.org/10.3390/ijms20133271.

(81) Wang, S.; Dai, Y.; Fukuoka, T.; Yamanaka, H.; Kobayashi, K.; Obata, K.; Cui, X.; Tominaga, M.; Noguchi, K. Phospholipase C and Protein Kinase A Mediate Bradykinin Sensitization of TRPA1: A Molecular Mechanism of Inflammatory Pain. *Brain* **2008**, *131* (5), 1241–1251. https://doi.org/10.1093/brain/awn060.

(82) Zimova, L.; Barvikova, K.; Macikova, L.; Vyklicka, L.; Sinica, V.; Barvik, I.; Vlachova, V. Proximal C-Terminus Serves as a Signaling Hub for TRPA1 Channel Regulation via Its Interacting Molecules and Supramolecular Complexes. *Frontiers in Physiology* **2020**, *11*, 189. https://doi.org/10.3389/fphys.2020.00189.

(83) Sulak, M. A.; Ghosh, M.; Sinharoy, P.; Andrei, S. R.; Damron, D. S. Modulation of TRPA1 Channel Activity by Cdk5 in Sensory Neurons. *Channels* **2018**, *12* (1), 65–75. https://doi.org/10.1080/19336950.2018.1424282.

(84) Barvikova, K.; Barvik, I.; Sinica, V.; Zimova, L.; Vlachova, V. Phospho-Mimetic Mutation at SER602 Inactivates Human TRPA1 Channel. *International Journal of Molecular Sciences* **2020**, *21* (21), 1–14. https://doi.org/10.3390/ijms21217995.

(85) Macikova, L.; Sinica, V.; Kadkova, A.; Villette, S.; Ciaccafava, A.; Faherty, J.; Lecomte, S.; Alves, I. D.; Vlachova, V. Putative Interaction Site for Membrane Phospholipids Controls Activation of TRPA1 Channel at Physiological Membrane Potentials. *The FEBS Journal* **2019**, *286* (18), 3664–3683. https://doi.org/10.1111/febs.14931.

(86) Schmidt, M.; Dubin, A. E.; Petrus, M. J.; Earley, T. J.; Patapoutian, A. Nociceptive Signals Induce Trafficking of TRPA1 to the Plasma Membrane. *Neuron* **2009**, *64* (4), 498–509. https://doi.org/10.1016/j.neuron.2009.09.030.

(87) Obata, K.; Katsura, H.; Mizushima, T.; Yamanaka, H.; Kobayashi, K.; Dai, Y.; Fukuoka, T.; Tokunaga, A.; Tominaga, M.; Noguchi, K. TRPA1 Induced in Sensory Neurons Contributes to Cold Hyperalgesia after Inflammation and Nerve Injury. *Journal of Clinical Investigation* **2005**, *115* (9), 2393–2401. https://doi.org/10.1172/JCI25437.

(88) Shimomura, O.; Johnson, F. H.; Saiga, Y. Microdetermination of Calcium by Aequorin Luminescence. *Science* **1963**, *140* (3573), 1339–1340.

(89) Snowdowne, K. W.; Borle, A. B. Measurement of Cytosolic Free Calcium in Mammalian Cells with Aequorin. *American Journal of Physiology - Cell Physiology* **1984**, *16* (3). https://doi.org/10.1152/ajpcell.1984.247.5.c396.

(90) Ungrin, M. D.; Singh, L. M. R.; Stocco, R.; Sas, D. E.; Abramovitz, M. An Automated Aequorin Luminescence-Based Functional Calcium Assay for G- Protein-Coupled Receptors. *Analytical Biochemistry* **1999**, *272* (1), 34–42. https://doi.org/10.1006/abio.1999.4145.

(91) Shimomura, O.; Johnson, F. H.; Saiga, Y. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. *Journal of Cellular and Comparative Physiology* **1962**, *59* (3), 223–239. https://doi.org/10.1002/jcp.1030590302.

(92) Tsien, R. Y. New Calcium Indicators and Buffers with High Selectivity against Magnesium and Protons: Design, Synthesis, and Properties of Prototype Structures. *Biochemistry* **1980**, *19* (11), 2396–2404. https://doi.org/10.1021/bi00552a018.

(93) Minta, A.; Kao, J. P.; Tsien, R. Y. Fluorescent Indicators for Cytosolic Calcium Based on Rhodamine and Fluorescein Chromophores. *Journal of Biological Chemistry* **1989**, *264* (14), 8171–8178.

(94) Heber, S.; Gold-Binder, M.; Ciotu, C. I.; Witek, M.; Ninidze, N.; Kress, H. G.; Fischer, M. J. M. A Human Trpa1-Specific Pain Model. *Journal of Neuroscience* **2019**, *39* (20), 3845–3855. https://doi.org/10.1523/JNEUROSCI.3048-18.2019.

(95) Everaerts, W.; Gees, M.; Alpizar, Y. A.; Farre, R.; Leten, C.; Apetrei, A.; Dewachter, I.; van Leuven, F.; Vennekens, R.; de Ridder, D.; Nilius, B.; Voets, T.; Talavera, K. The Capsaicin Receptor TRPV1 Is a Crucial Mediator of the Noxious Effects of Mustard Oil. *Current Biology* **2011**, *21* (4), 316–321. https://doi.org/10.1016/j.cub.2011.01.031.

(96) Conaway, C. C.; Jiao, D.; Chung, F. L. Inhibition of Rat Liver Cytochrome P450 Isozymes by Isothiocyanates and Their Conjugates: A Structure-Activity Relationship Study. *Carcinogenesis* **1996**, *17* (11), 2423–2427. https://doi.org/10.1093/carcin/17.11.2423.

(97) Tang, L.; Zhang, Y. Mitochondria Are the Primary Target in Isothiocyante-Induced Apoptosis in Human Bladder Cancer Cells. *Molecular Cancer Therapeutics* **2005**, *4* (8), 1250–1259. https://doi.org/10.1158/1535-7163.MCT-05-0041.

(98) Youn, H. S.; Lee, J. K.; Choi, Y. J.; Saitoh, S. I.; Miyake, K.; Hwang, D. H.; Lee, J. Y. Cinnamaldehyde Suppresses Toll-like Receptor 4 Activation Mediated through the Inhibition of Receptor Oligomerization. *Biochemical Pharmacology* **2008**, *75* (2), 494–502. https://doi.org/10.1016/j.bcp.2007.08.033.

(99) Chao, L. K.; Hua, K. F.; Hsu, H. Y.; Cheng, S. S.; Lin, I. F.; Chen, C. J.; Chen, S. T.; Chang, S. T. Cinnamaldehyde Inhibits Pro-Inflammatory Cytokines Secretion from Monocytes/Macrophages through Suppression of Intracellular Signaling. *Food and Chemical Toxicology* **2008**, *46* (1), 220–231. https://doi.org/10.1016/j.fct.2007.07.016.

(100) Hanna, P. E.; Anders, M. W. The Mercapturic Acid Pathway. *Critical Reviews in Toxicology*. Taylor and Francis Ltd November 26, 2019, pp 819–929. https://doi.org/10.1080/10408444.2019.1692191.

(101) Kolm, R. H.; Danielson, U. H.; Zhang, Y.; Talalayt, P.; Mannervik, B. *Isothiocyanates as Substrates for Human Glutathione Transferases: Structure-Activity Studies*; 1995; Vol. 311.

(102) Zhang, Y. S.; Kolm, R. H.; Mannervik, B.; Talalay, P. Reversible Conjugation of Isothiocyanates with Glutathione Catalyzed by Human Glutathione Transferases. *Biochemical and Biophysical Research Communications* **1995**, *206* (2), 748–755. https://doi.org/10.1006/bbrc.1995.1106.

(103) Chen, J.; Joshi, S. K.; Didomenico, S.; Perner, R. J.; Mikusa, J. P.; Gauvin, D. M.; Segreti, J. A.; Han, P.; Zhang, X. F.; Niforatos, W.; Bianchi, B. R.; Baker, S. J.; Zhong, C.; Simler, G. H.; McDonald, H. A.; Schmidt, R. G.; McGaraughty, S. P.; Chu, K. L.; Faltynek, C. R.; Kort, M. E.; Reilly, R. M.; Kym, P. R. Selective Blockade of TRPA1 Channel Attenuates Pathological Pain without Altering Noxious Cold Sensation or Body Temperature Regulation. *Pain* **2011**, *152* (5), 1165–1172. https://doi.org/10.1016/j.pain.2011.01.049.

(104) Eid, S. R.; Crown, E. D.; Moore, E. L.; Liang, H. A.; Choong, K. C.; Dima, S.; Henze, D. A.; Kane, S. A.; Urban, M. O. HC-030031, a TRPA1 Selective Antagonist, Attenuates Inflammatory- and Neuropathy-Induced Mechanical Hypersensitivity. *Molecular Pain* **2008**, *4*, 48. https://doi.org/10.1186/1744-8069-4-48.

(105) Simon, M. C.; Keith, B. The Role of Oxygen Availability in Embryonic Development and Stem Cell Function. *Nature Reviews Molecular Cell Biology*. Howard Hughes Medical Institute April 2008, pp 285–296. https://doi.org/10.1038/nrm2354.

(106) Nakayama, K. Cellular Signal Transduction of the Hypoxia Response. *Journal of Biochemistry* **2009**, *146* (6), 757–765. https://doi.org/10.1093/jb/mvp167.

(107) Saxena, K.; Jolly, M. K. Acute vs. Chronic vs. Cyclic Hypoxia: Their Differential Dynamics, Molecular Mechanisms, and Effects on Tumor Progression. *Biomolecules*. MDPI AG August 1, 2019. https://doi.org/10.3390/biom9080339.

(108) Kaelin, W. G.; Ratcliffe, P. J. Oxygen Sensing by Metazoans: The Central Role of the HIF Hydroxylase Pathway. https://doi.org/10.1016/j.molcel.2008.04.009.
(109) Lisy, K.; Peet, D. J. Turn Me on: Regulating HIF Transcriptional Activity. *Cell Death and Differentiation*. Nature Publishing Group January 18, 2008, pp 642–649. https://doi.org/10.1038/sj.cdd.4402315.
(110) Lando, D.; Peet, D. J.; Gorman, J. J.; Whelan, D. A.; Whitelaw, M. L.; Bruick, R. K. FIH-1 Is an Asparaginyl Hydroxylase Enzyme That Regulates the Transcriptional Activity of Hypoxia-Inducible Factor. *Genes and Development* **2002**, *16* (12), 1466–1471. https://doi.org/10.1101/gad.991402.
(111) Schofield, C. J.; Ratcliffe, P. J. Oxygen Sensing by HIF Hydroxylases. *Nature Reviews Molecular Cell Biology* **2004**, *5* (5), 343–354. https://doi.org/10.1038/nrm1366.
(112) Schofield, C. J.; Ratcliffe, P. J. Signalling Hypoxia by HIF Hydroxylases. *Biochemical and Biophysical Research Communications* **2005**, *338* (1), 617–626. https://doi.org/10.1016/j.bbrc.2005.08.111.
(113) Huang, J.; Zhao, Q.; Mooney, S. M.; Lee, F. S. Sequence Determinants in Hypoxia-Inducible Factor-1α for Hydroxylation by the Prolyl Hydroxylases PHD1, PHD2, and PHD3. *Journal of Biological Chemistry* **2002**, *277* (42), 39792–39800. https://doi.org/10.1074/jbc.M206955200.
(114) Zurlo, G.; Guo, J.; Takada, M.; Wei, W.; Zhang, Q. New Insights into Protein Hydroxylation and Its Important Role in Human Diseases. *Biochimica et Biophysica Acta - Reviews on Cancer*. Elsevier B.V. December 1, 2016, pp 208–220. https://doi.org/10.1016/j.bbcan.2016.09.004.
(115) Cockman, M. E.; Lippl, K.; Tian, Y. M.; Pegg, H. B.; Figg, W. D.; Abboud, M. I.; Heilig, R.; Fischer, R.; Myllyharju, J.; Schofield, C. J.; Ratcliffe, P. J. Lack of Activity of Recombinant HIF Prolyl Hydroxylases (PHDs) on Reported Non-HIF Substrates. *eLife* **2019**, *8*. https://doi.org/10.7554/eLife.46490.
(116) Cockman, M. E.; Webb, J. D.; Kramer, H. B.; Kessler, B. M.; Ratcliffe, P. J. Proteomics-Based Identification of Novel Factor Inhibiting Hypoxia-Inducible Factor (FIH) Substrates Indicates Widespread Asparaginyl Hydroxylation of Ankyrin Repeat Domain-Containing Proteins. *Molecular and Cellular Proteomics* **2009**, *8* (3), 535–546. https://doi.org/10.1074/mcp.M800340-MCP200.
(117) Li, J.; Mahajan, A.; Tsai, M. D. Ankyrin Repeat: A Unique Motif Mediating Protein-Protein Interactions. *Biochemistry*. Biochemistry December 26, 2006, pp 15168–15178. https://doi.org/10.1021/bi062188q.
(118) Binz, H. K.; Stumpp, M. T.; Forrer, P.; Amstutz, P.; Plückthun, A. Designing Repeat Proteins: Well-Expressed, Soluble and Stable Proteins from Combinatorial Libraries of Consensus Ankyrin Repeat Proteins. *Journal of Molecular Biology* **2003**, *332* (2), 489–503. https://doi.org/10.1016/S0022-2836(03)00896-9.
(119) Kelly, L.; McDonough, M. a; Coleman, M. L.; Ratcliffe, P. J.; Schofield, C. J. Asparagine Beta-Hydroxylation Stabilizes the Ankyrin Repeat Domain Fold. *Molecular bioSystems* **2009**, *5* (1), 52–58. https://doi.org/10.1039/b815271c.
(120) Mosavi, L. K.; Minor, D. L.; Peng, Z. yu. Consensus-Derived Structural Determinants of the Ankyrin Repeat Motif. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99* (25), 16029–16034. https://doi.org/10.1073/pnas.252537899.
(121) Takahashi, N.; Kuwaki, T.; Kiyonaka, S.; Numata, T.; Kozai, D.; Mizuno, Y.; Yamamoto, S.; Naito, S.; Knevels, E.; Carmeliet, P.; Oga, T.; Kaneko, S.; Suga, S.; Nokami, T.; Yoshida, J.; Mori, Y. TRPA1 Underlies a Sensing Mechanism for O2. *Nature Chemical Biology* **2011**, *7* (10), 701–711. https://doi.org/10.1038/nchembio.640.
(122) Miyake, T.; Nakamura, S.; Zhao, M.; So, K.; Inoue, K.; Numata, T.; Takahashi, N.; Shirakawa, H.; Mori, Y.; Nakagawa, T.; Kaneko, S. Cold Sensitivity of TRPA1 Is Unveiled by the Prolyl Hydroxylation Blockade-Induced Sensitization to ROS. *Nature Communications* **2016**, *7* (1), 12840. https://doi.org/10.1038/ncomms12840.
(123) Coleman, M. L.; McDonough, M. A.; Hewitson, K. S.; Coles, C.; Mecinović, J.; Edelmann, M.; Cook, K. M.; Cockman, M. E.; Lancaster, D. E.; Kessler, B. M.; Oldham, N. J.; Ratcliffe, P. J.; Schofield, C.

J. Asparaginyl Hydroxylation of the Notch Ankyrin Repeat Domain by Factor Inhibiting Hypoxia-Inducible Factor. *Journal of Biological Chemistry* **2007**, *282* (33), 24027–24038. https://doi.org/10.1074/jbc.M704102200.

(124) Bean, B. P. The Action Potential in Mammalian Central Neurons. *Nature Reviews Neuroscience*. Nature Publishing Group June 2007, pp 451–465. https://doi.org/10.1038/nrn2148.

(125) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Molecular Systems Biology* **2011**, *7* (1), 539. https://doi.org/10.1038/msb.2011.75.

(126) Shimomura, O.; Johnson, F. H. Peroxidized Coelenterazine, the Active Group in the Photoprotein Aequorin. *Proceedings of the National Academy of Sciences of the United States of America,* **1978**, 75 (6), 2611-2615. https://doi.org/10.1073/pnas.75.6.2611.

Chapter II:

Studies on the Electrophilic Activation of TRPA1

1. Previous Work

1.1 Reaction of Formaldehyde with Amino Acids

Previous work by Kamps *et al.* has shown that a large molar excess of formaldehyde (HCHO) will react to form stable covalent adducts with most amino acids containing a nucleophilic sidechain in non-buffered D_2O.[1] The extent of reaction is condition dependant, including the pH of the reaction mixture, with reactions occurring in an apparently reversible or irreversible manner, depending on the residue involved. Of the work described in Kamps *et al.*, the reactions of cysteine and lysine, the residues which allow TRPA1 to sense electrophiles, are of particular relevance to the work described in this thesis.[1] Threonine and histidine, which contain nucleophilic sidechains, are among the neighbouring residues of C621, C641 and C665 in the electrophile binding domain of TRPA1, and thus may contribute to the high reactivity of these sites.

Cysteine reacts initially with HCHO to give a hemi-thioacetal adduct **15** (an *S*-hydroxymethyl adduct). When cysteine is the N-terminal residue of a protein or peptide, the hemi-thioacetal can undergo intramolecular cyclisation to give a thiazolidine **16**, also known as a thioproline residue (fig. 2-1). Thioproline seems relatively stable in the presence of other electrophiles, aldehyde scavengers such as 1,3-diketocyclohexane, and to lyophilisation.[1] However, it is not clear how stable the cysteine hemi-thioacetal is in the presence of a ubiquitous biological nucleophile such as glutathione (GSH), something the work described in this chapter aimed to explore.

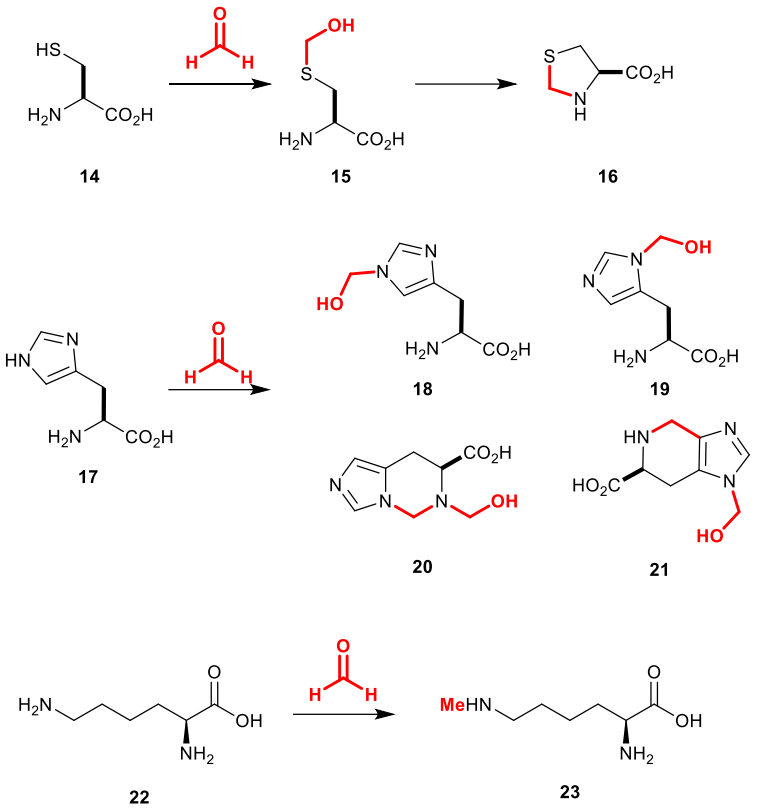


Figure 2-1: Reaction of cysteine, histidine and lysine with HCHO occurs to give diverse range of products as reported by Kamps *et al.*[1]

1.2 Reaction of Formaldehyde with Peptides

The diversity in reaction of peptides with HCHO is exemplified by the reaction of GSH with HCHO. GSH reacts rapidly with HCHO to give *S*-(hydroxymethyl)glutathione (HMG) *via* a non-enzymatic process.[2] In protein crystallographic work, it was shown that HMG can react with an additional HCHO molecule to give a 7,7-membered bicyclic product ($BiGF_2$).[3] An additional minor product, distinct from HMG and $BiGF_2$, was identified under mildly acidic conditions, however the authors could not assign a structure to this compound. The compound was later identified by NMR studies as PGF (fig. 2-2).[4] HMG appears to rearrange to give a 10-membered cyclic product (MGF) where the free amino group of glutamate is linked *via* a formaldehyde derived methylene group to the cysteine thiol group (fig. 2-2).

MGF is reminiscent of $BiGF_2$ but lacks a methylene bridge between the cysteine amide bond and the glutamate amine, suggesting $BiGF_2$ could form from the reaction of MGF and HCHO. PGF could result from the rearrangement of MGF through the acid catalysed transamination of the cysteine amide bond. The 10-membered ring of MGF likely adopts a boat-chair-boat conformation, where the amide carbonyl and amine lie in close proximity and in a favourable geometry for 5-membered ring formation, which is more thermodynamically stable (fig. 2-2). The pH dependency of the rearrangement of MGF to PGF is presumably due to balancing the requirement for a nucleophilic amine (favoured at high pH) and an activated (protonated) carbonyl group (favoured at low pH).

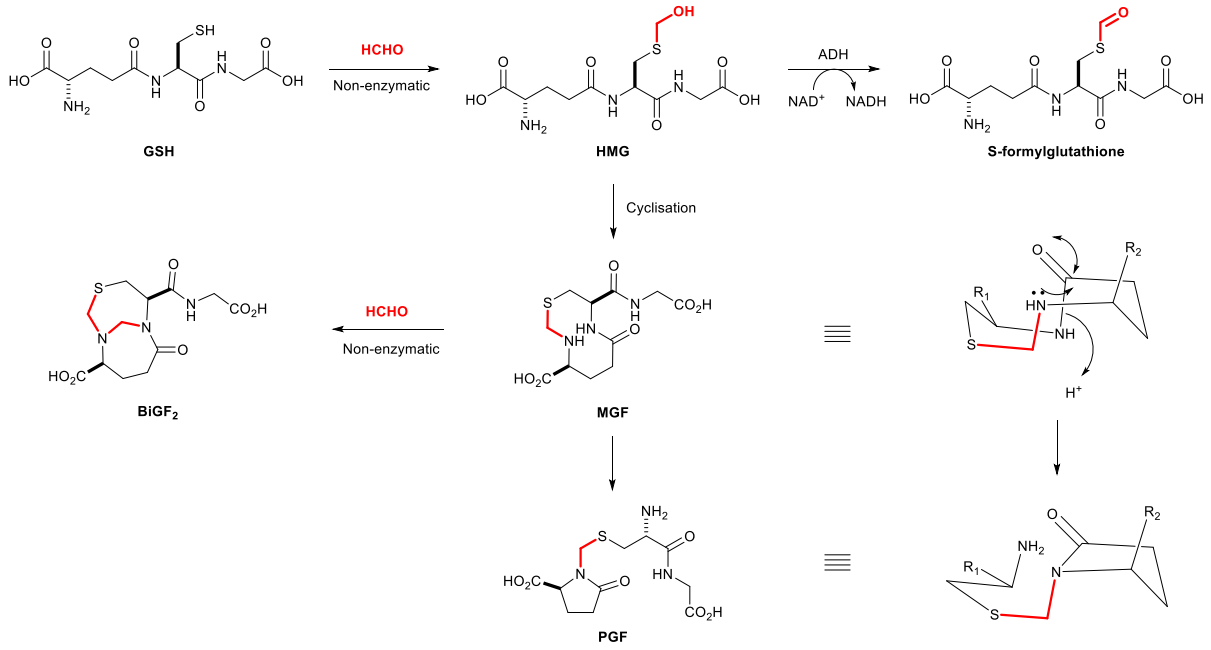


Figure 2-2: Reaction of GSH with HCHO gives cyclic products. HCHO derived groups are shown in red. ADH = type III alcohol dehydrogenase. NB: conversion of GSH to HMG is catalysed by glutathione-dependent formaldehyde activating enzyme (GFA) *in vivo*.[34]

It is unclear what the biological relevance of the cyclic GSH derivatives is. The accumulation of MGF and its derivatives will likely be prevented by the rapid conversion of HMG to *S*-formylglutathione in cells as catalysed by alcohol dehydrogenase enzymes and the fact that these intermediates form relatively slowly at non-physiological pD/pH.[5] Therefore, in the work

described in this thesis, particular attention was drawn to relatively rapid reactions between peptides and HCHO under physiologically relevant conditions.

2. Reaction of Peptides with HCHO

A series of short cysteine containing peptides was designed to investigate the reaction of cysteine with HCHO as a model for the reaction of TRPA1 with HCHO. It was hypothesised that protection of the peptide termini would more closely mimic the reaction of an internal cysteine residue in a protein such as C621 in TRPA1. In the first instance, peptides containing cysteine with no other obviously nucleophilic groups were evaluated for their ability to form an *S*-hydroxymethyl group in the presence of HCHO. Peptides containing cysteine and other nucleophilic residues which are present in the TRPA1 electrophile binding domain (EBD) or in other HCHO sensing proteins in close proximity to a reactive cysteine were then evaluated. Peptides were synthesised by solid phase peptide synthesis (SPPS) using rink amide MHBA resin (see Chapter 4 §2 for an overview of SPPS). N-terminal acetylation was realised using acetic anhydride. Global deprotection and cleavage of the peptide from the resin was carried out using CF_3CO_2H, giving rise to N-terminal acetylated and C-terminal amidated peptides.

2.1 Acetyl-Alanine-Cysteine-Alanine (Ac-ACA-NH_2)

Acetyl-alanine-cysteine-alanine (Ac-ACA-NH_2) was chosen as a convenient starting point to examine the reaction between a typical cysteine containing peptide and HCHO. Previous work had explored the reaction of GSH with HCHO, which is atypical because the glutamate residue is linked *via* its sidechain.[4] A solution of HCHO was prepared by boiling paraformaldehyde in D_2O in a sealed vial to give a stock solution which was further diluted with D_2O and stored in sealed vials at rt. The general procedure for reacting a peptide with HCHO involved the

following steps. The peptide was first dissolved in D_2O and the concentration of the peptide determined by calibration with 1 mg/mL sodium 3-(trimethylsilyl)-2,2,3,3-tetradeuteropropanoate (TSP) by ^{1}H-NMR. This was carried out by mixing 16 µL peptide, 8 µL TSP and 136 µL of deuterated Tris buffer (Tris-d_{11}) in D_2O in a 3 mm Bruker Match NMR tube. This spectrum also served as a reference for unreacted starting material. Once the concentration of the peptide was known, a portion of the stock solution of the peptide was added to Tris-d_{11} in D_2O in an Eppendorf vial to give a final concentration of 1 mM peptide after the addition of HCHO. HCHO was added as a solution in D_2O and a reaction timer was started. The mixture was briefly vortexed, pipetted into a 3 mm Bruker Match NMR tube and any air bubbles in the tube were displaced by brief centrifugation with a hand centrifuge. The tube was then loaded into the spectrometer, the spectrometer was locked to the deuterium resonance of D_2O and shimmed before the first spectrum was acquired. The process typically took between 5 and 7 minutes from adding HCHO to acquiring the first spectrum.

It was found that 1 mM Ac-ACA-NH_2 rapidly reacted with 8 mM HCHO in buffered D_2O to give only the *S*-hydroxymethyl species (fig. 2-3). Complete consumption of the starting material to give a single identifiable product occurred within the time taken to acquire the first spectrum of the reaction mixture. The presence of an *S*-hydroxymethyl species is evidenced by the presence of a new resonance at δ_H 4.71 / δ_C 65.4 ppm in the HSQC spectrum of the reaction mixture. The phase of the peak suggests the new resonance originates from a CH_2 group. The HMBC spectrum shows a correlation between this new resonance and the cysteine CH_2 group. There is good agreement when comparing the chemical shift of the new CH_2 group in Ac-ACA-NH_2 (δ_C 65.4 ppm) to *S*-hydoxymethyl cysteine (δ_C 65.1 ppm).[1] The *S*-hydroxymethyl moiety of GSH derived HMG has a slightly higher chemical shift (δ_C 66.6 ppm) compared to that observed in Ac-ACA-NH_2 (fig. 2-3).[6]

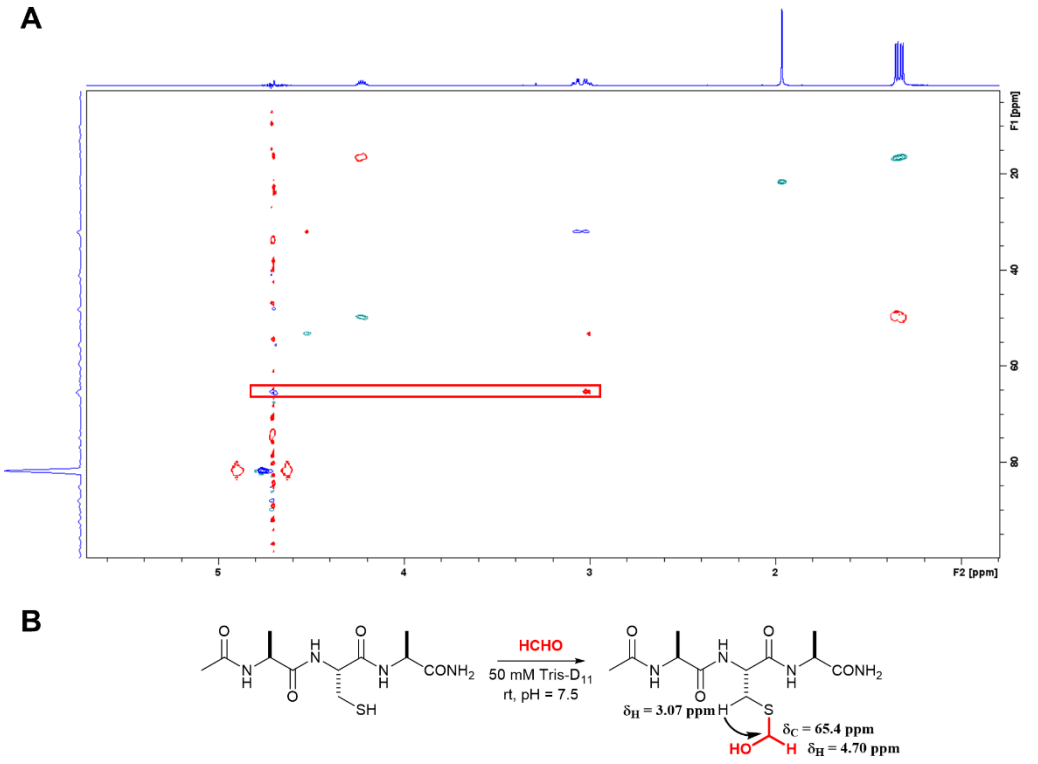


Figure 2-3: NMR spectra of the reaction between Ac-ACA-NH$_2$ and HCHO. A) HSQC (green/blue) and HMBC (red), show the presence of a new CH_2 peak normally hidden under the HOD peak in ^{1}H NMR. The red box shows the new peak correlates to the cysteine CH_2 group. NB. the peak at 4.76/81.8 ppm is unreacted HCHO. **B)** reaction scheme showing the formation of the *S*-hydroxymethyl adduct of Ac-ACA-NH$_2$.

2.2 Acetyl-Alanine-Cysteine-Alanine-Cysteine-Alanine (Ac-ACACA-NH$_2$)

Similarly to the results with Ac-ACA-NH$_2$, it was found that under the same conditions Ac-ACACA-NH$_2$ reacts with HCHO to give rise to a single identifiable species. A peak, presumably derived from HCHO, was observed in the HSQC of the product with a phase corresponding to a CH_2 group at δ_H 4.72 / δ_C 65.4. Both cysteine C(β)H$_2$ resonances overlapped in both the ^{1}H and ^{13}C dimension of the HSQC spectrum (fig. 2-4 a, b), meaning that it was not possible to distinguish between the *S*-hydroxymethyl adduct and a cyclic thioacetal by HMBC. Given the similarity in chemical shift to the *S*-hydroxymethyl adduct of Ac-ACA-NH$_2$ and the

fact that a thioacetal would have a much lower chemical shift than the *S*-hydroxymethyl adduct

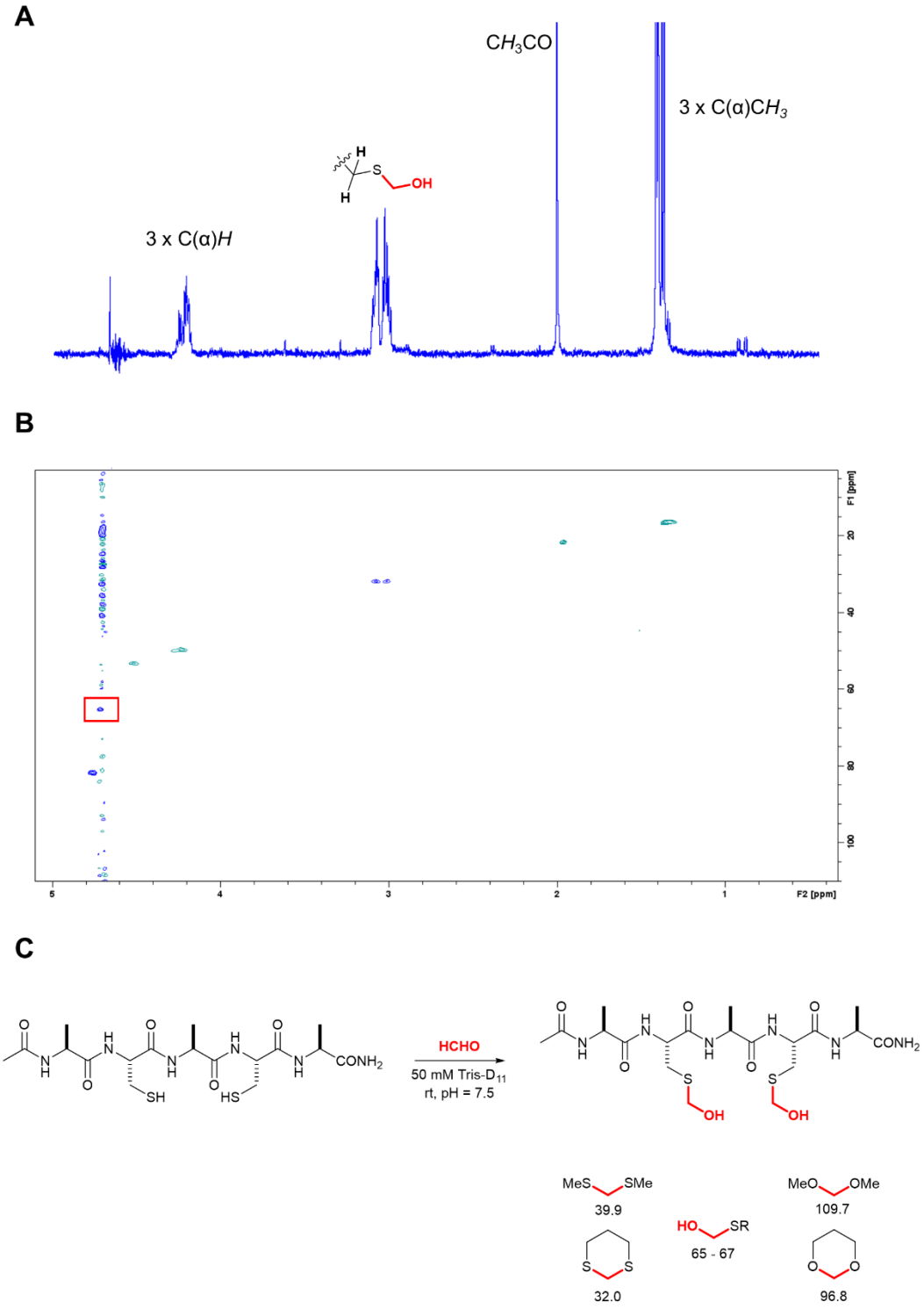


Figure 2-4: Examining the reaction of Ac-ACACA-NH_2 with HCHO by NMR: A) ^{1}H-ESPE spectrum of the reaction mixture; **B)** ^{1}H-^{13}C HSQC spectrum with new CH_2 resonance highlighted (red box); **C)** Reaction scheme including literature ^{13}C chemical shift values for thioacetals and acetals derived from HCHO.

on account of the increased electron density (and hence shielding of the nucleus) on the

bridging carbon, the product was provisionally assigned as the *S*-hydroxymethyl adduct in which both cysteine residues have reacted with HCHO (fig. 2-4 c).

2.3 Acetyl-Alanine-Cysteine-Histidine-Aspartate-Alanine (Ac-ACHDA-NH_2)

Ac-ACHDA-NH_2 was chosen as the next a substrate to be studied for two reasons: firstly cysteine, histidine and aspartic acid residues form the catalytic triad of most cysteine proteases, whereby the histidine and aspartate residues facilitate proton transfers between the nucleophilic cysteine and the substrate.[7] Secondly, Ac-ACHDA-NH_2 contains a histidine in close proximity to cysteine which mimics the electrophile binding domain of TRPA1 which contains C621 and H614.

Addition of 25 mM HCHO to 5 mM Ac-ACHDA-NH_2 in 50 mM Tris-d_{11} gave the *S*-hydroxymethyl species of Ac-ACHDA-NH_2. A new peak corresponding to a methylene group appeared in the HSQC spectrum at δ_H 4.65 / δ_C 65.4 which correlates with the cysteine CH_2 group in the HMBC spectrum (fig. 2-5). Similar to the reaction between Ac-ACA-NH_2 and HCHO, the reaction was complete within the time taken to acquire the first proton spectrum of the reaction mixture. The composition of the reaction mixture did not change within 90 minutes of the first spectrum being acquired. The NMR analyses indicate selective reaction of HCHO with the cysteine residue of Ac-ACHDA-NH_2, at least to give a stable adduct.

Interestingly, HCHO was not observed to react with the imidazole ring of histidine of Ac-ACHDA-NH_2, at least as a major product (*vide infra*) under the tested conditions. The reaction

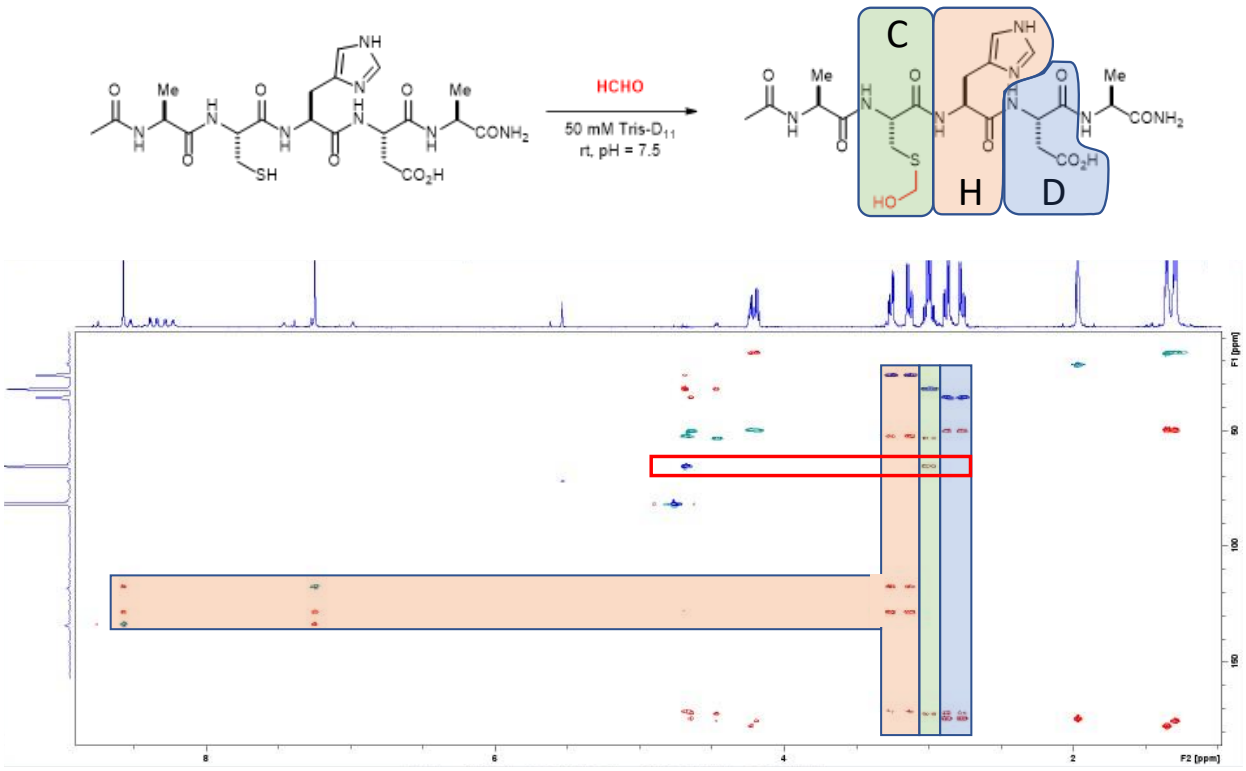


Figure 2-5: NMR spectra of the reaction mixture of Ac-ACHDA-NH_2 and HCHO. HSQC (blue/green) shows a new CH_2 peak which shows a HMBC (red) correlation (red box) to the cysteine CH_2 group. Signals arising from the CH_2 group are highlighted for aspartate (blue) and cysteine (green). For histidine (orange), signals arising from the imidazole ring and CH_2 group are highlighted.

of histidine with HCHO was studied by Kamps *et al.*[1] In this work, histidine was reacted in unbuffered D_2O with a large molar excess of HCHO over a time course of days to give multiple products which contain an *N*-hydroxymethyl-imidazole moiety, including **18** and **19** (see fig. 2-1).[1] These conditions contrast with the conditions used to study the reaction of Ac-ACHDA-NH_2 with HCHO, which makes use of a shorter reaction time and smaller molar excess of HCHO. Additionally, the reaction is buffered by Tris-d_{11}, which is able to react with HCHO to give *N*-hydroxymethyl-Tris (**24**), a reaction which may limit the concentration of available HCHO.[8] It is unclear whether a Tris-HCHO adduct is more thermodynamically stable than a histidine-HCHO adduct and what the kinetics of converting between the two adducts are (fig. 2-6).

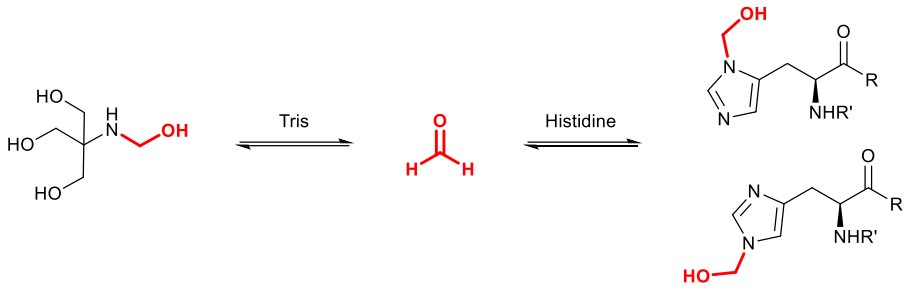
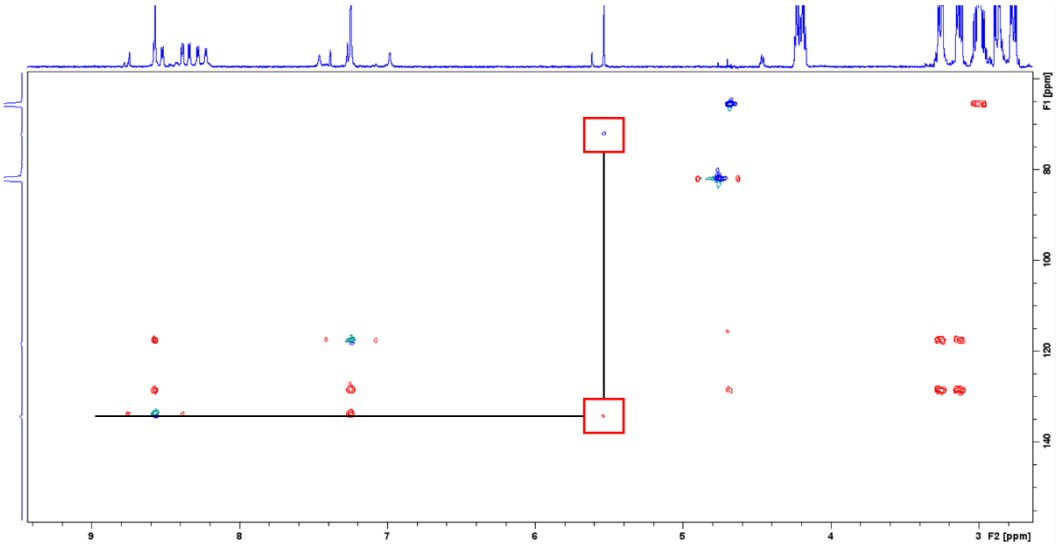


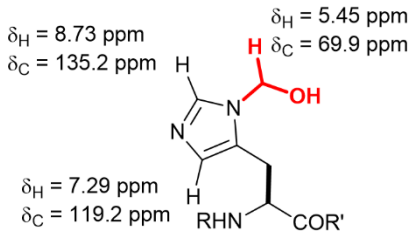
Figure 2-6: Equilibrium formed in mixtures containing Tris buffer, histidine and HCHO.

A minor product with a CH_2 resonance at δ_H 5.54 / δ_C 71.9, a frequency similar to that observed previously for an *N*-hydroxymethyl group attached to the imidazole ring of histidine,[1] was observed in the reaction mixture (fig. 2-7 a, b, c). Two resonances were observed for this species (fig. 2-7 a) though it is not clear whether these are the result of regioisomers in the reaction mixture as Kamps *et al.* suggest or whether the second resonance arises from a separate species.[1] The minor species comprised less than 10% of the reaction mixture and no other resonances were detected in the HMBC or HSQC that could have arisen from the minor species making it impossible to distinguish it from other impurities/artefacts. However, based on the similarity in reaction substrate and δ_C resonances it is likely that this minor component of the reaction mixture contains an *N*-hydroxymethyl group (fig. 2-7 c). It is unclear whether this minor component of the reaction mixture also contains an *S*-hydroxymethyl fragment on account of there being no additional peak in the region highlighted (fig. 2-7 a). However, this does not exclude the possibility that both the major and minor products have similar resonances for this group and thus the minor product resonance is hidden by the more intense signal from the major product.

A

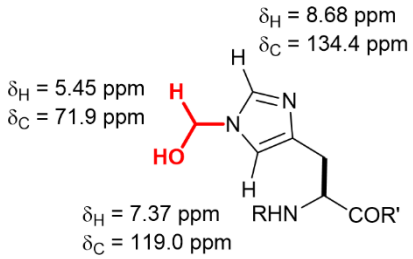


B



	Kamps *et al.*	Observed
N-CH_2-OH	5.45, 69.9	5.54, 71.9
H-C(2)	8.73, 135.2	~8.57, 134.0

C



	Kamps *et al.*	Observed
N-CH_2-OH	5.45, 71.9	5.54, 71.9
H-C(2)	8.73, 134.4	~8.57, 134.0

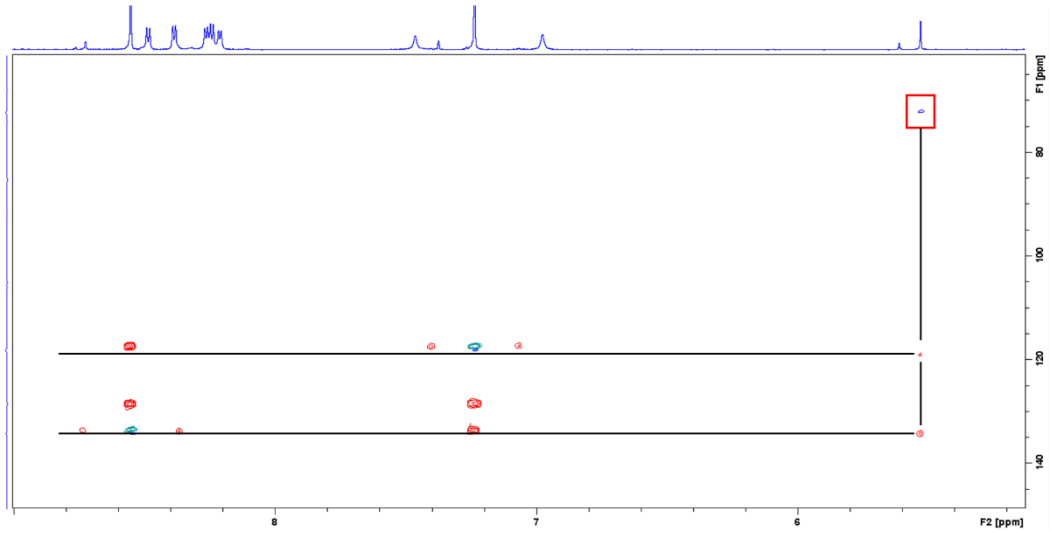
Figure 2-7: Exploring the existence of a minor product in the reaction of Ac-ACHDA-NH_2 with HCHO *via* NMR. A) Overlaid HSQC (blue/green) and HMBC (red). A HCHO derived resonance and its correlation shown in red boxes. **B)** Selected NMR data for *N*π-hydroxymethyl-histidine and comparison with observed values. **C)** Selected NMR data for *N*τ-hydroxymethyl-histidine and comparison with observed values

2.4 Acetyl-Alanine-Serine-Histidine-Aspartate-Alanine (Ac-ASHDA-NH_2)

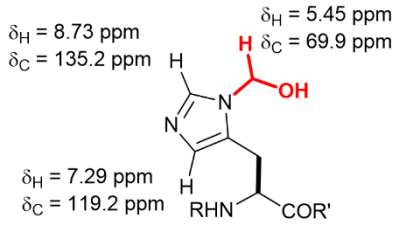
To further explore the minor product that is formed between the HCHO and Ac-ACHDA-NH_2, Ac-ASHDA-NH_2 was chosen on account of the reduced nucleophilicity and structural similarity of serine to cysteine. It was hoped that these two features would allow for the formation of more of the analogous product from the Ac-ACHDA-NH_2 reaction and thus the minor product from that reaction could be assigned by comparison with the product from this reaction. However, the majority of the reaction mixture upon addition of 25 mM HCHO to Ac-ASHDA-NH_2 (5 mM) in 50 mM Tris-d_{11} was comprised of unreacted Ac-ASHDA-NH_2.

Similar to the reaction to give the minor product observed with Ac-ACHDA-NH_2 (see §2.3), a new resonance corresponding to a methylene group was observed in the HSQC spectrum of the reaction mixture at δ_H 5.54 / δ_C 71.9 (fig. 2-8 a) which did not correlate with any resonances originating in the starting material by HMBC. However, two correlations to carbon atoms at δ_C 119.0 and 134.2 were observed from the protons at δ_H 5.54 by HMBC (fig 2-8 a). These carbon resonances are similar to those originating from the imidazole ring carbon atoms in the starting material (δ_C 117.3 and 133.4) and are extremely similar to the previously reported chemical shifts for *N*τ-hydroxymethyl-histidine (δ_C 119.0 and 134.4). Taking these observations together, the product, albeit formed at low levels, of the reaction of Ac-ASHDA-NH_2 and Ac-ACHDA-NH_2 (minor product) with HCHO was assigned as containing a *N*τ-hydroxymethyl-histidine fragment (fig. 2-8 c). The results with Ac-ASHDA-NH_2 and Ac-ACHDA-NH_2 also highlight the preferred reaction of HCHO with cysteine over other residues.

A

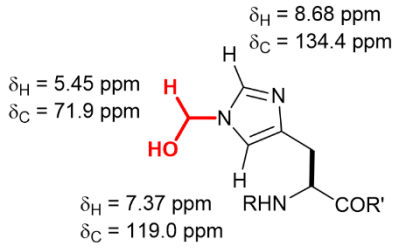


B



	Kamps *et al.*	Observed
N-CH_2-OH	5.45, 69.9	5.53, 72.0
H-C(2)	8.73, 135.2	?, 134.2
H-C(4)	7.29, 119.2	?, 119.0

C



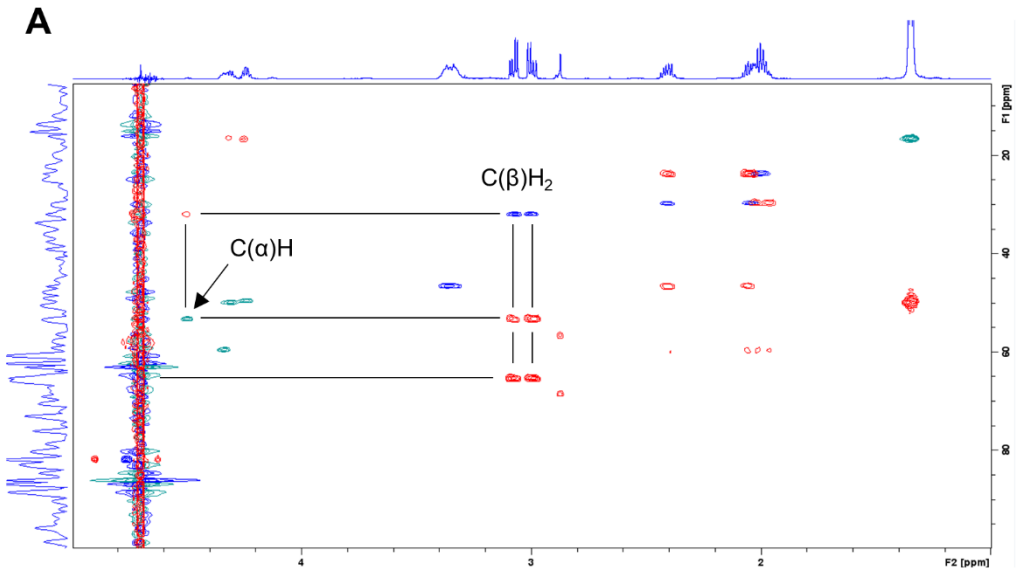
	Kamps *et al.*	Observed
N-CH_2-OH	5.45, 71.9	5.53, 72.0
H-C(2)	8.73, 134.4	?, 134.2
H-C(4)	7.37, 119.0	?, 119.0

Figure 2-8: Potential assignment of product of reaction between Ac-ASHDA-NH_2 with HCHO. A) Overlaid HSQC (blue/green) and HMBC (red). A HCHO derived resonance and its correlation shown in red boxes. **B)** Selected NMR data for *N*π-hydroxymethyl-histidine and comparison with observed values. **C)** Selected NMR data for *N*τ-hydroxymethyl-histidine and comparison with previously observed values.

2.5 Proline-Alanine-Cysteine-Alanine (PACA-NH_2)

To link the work on peptides with internal cysteines to a larger variety of HCHO sensing proteins such as the bacterial protein formaldehyde-induced regulator (FrmR) which contains a nucleophilic cysteine in close proximity to the N-terminal proline, the reaction of HCHO with further peptides, including PACA-NH_2, was explored.[9] While not directly relevant to the work

on TRPA1, this topic was explored to gain potential insight into the reaction mechanism of other HCHO sensing proteins. Upon mixing 5 mM PACA-NH_2 with 15 mM HCHO, a new correlation between the cysteine CH_2 group and a ^{13}C resonance at 65.4 ppm was observed in the HMBC spectrum of the reaction mixture (fig. 2-9 a). The peak at 65.4 ppm in the HSQC spectrum was obscured by the residual HOD signal, however, it was assigned by comparison to previously observed resonances as originating from a product containing *S*-hydroxymethyl-cysteine fragment (fig 2-9 b).



B

	Ac-ACA-NH_2	Ac-ACACA-NH_2	Ac-ACHDA-NH_2	PACA-NH_2	Free Cysteine
S-CH_2-OH	65.4	65.4	65.4	65.4	65.1

Figure 2-9: Assigning *S*-hydroxymethyl-PACA-NH_2 by similarity to other peptides described in this thesis. A) Overlaid HSQC (blue/green) and HMBC (red) spectra of the reaction mixture between PACA-NH_2 and HCHO. B) Comparing previously observed chemical shifts for the *S*-hydroxymethyl carbon with that observed in PACA-NH_2. The reaction of cysteine with HCHO is described in Kamps *et al.*[1]

2.6 Proline-Cysteine-Alanine-Alanine (PCAA-NH_2)

Addition of 15 mM HCHO to the peptide PCAA-NH_2 (5 mM) in 50 mM Tris-d_{11} also gave the anticipated *S*-hydroxymethyl species. The evidence for this was a clear correlation in the HMBC between the cysteine C(β)H_2 protons and a carbon atom at 65.4 ppm (fig. 2-10). This chemical shift is analogous to previously observed values for the *S*-hydroxymethyl species (fig. 2-9 b).

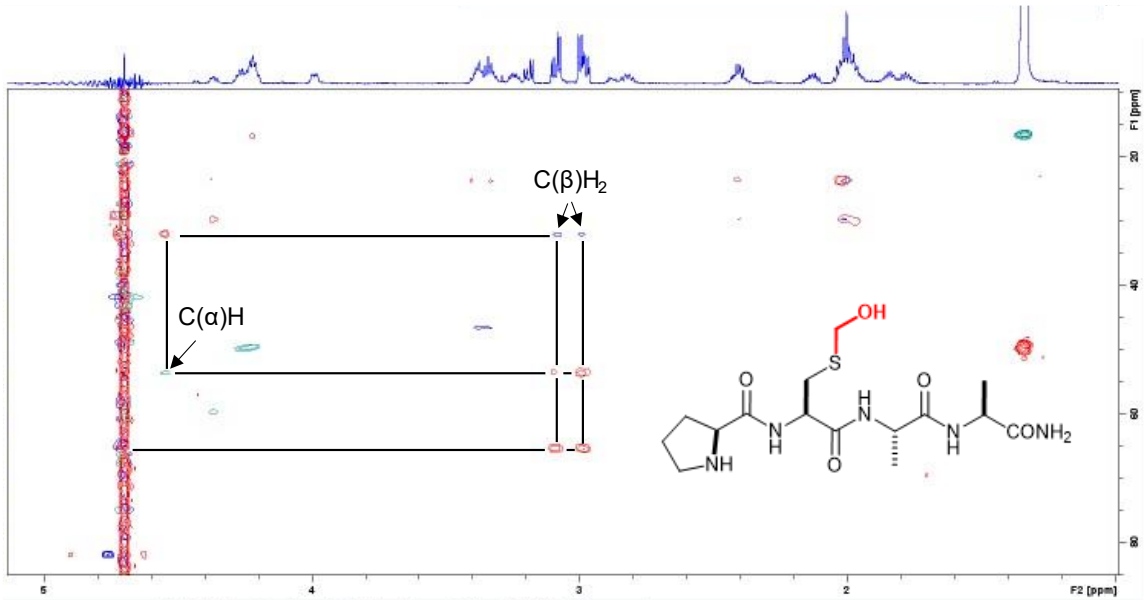


Figure 2-10: Overlaid HSQC (blue/green) and HMBC (red) spectra of the reaction mixture between PCAA-NH_2 and HCHO. Spectra show the initial composition of the reaction mixture over the first two hours of reaction. The data is annotated with the correlations for the cysteine C(β)$_2$ protons.

After acquiring HSQC and HMBC spectra of the initial reaction mixture, it was found that the composition of the reaction mixture of PCAA-NH_2 and HCHO was slowly changing over time. The *S*-hydroxymethyl species underwent a further slow reaction over the period of 7 days to give a 5,5-bicyclic structure with a methylene bridge between the free N-terminus and the nitrogen atom of the amide bond in addition to retaining the *S*-hydroxymethyl moiety (fig. 2-11). This structure was assigned due to three key correlations in the HMBC spectrum between: 1) the new CH_2 resonance (methylene bridge) derived from HCHO and the proline carbonyl group, 2) the cysteine C(α)H and the methylene bridge, and finally 3) the methylene bridge and the C(δ)H_2 group of the piperazine ring of proline. An analogous structure was observed in the

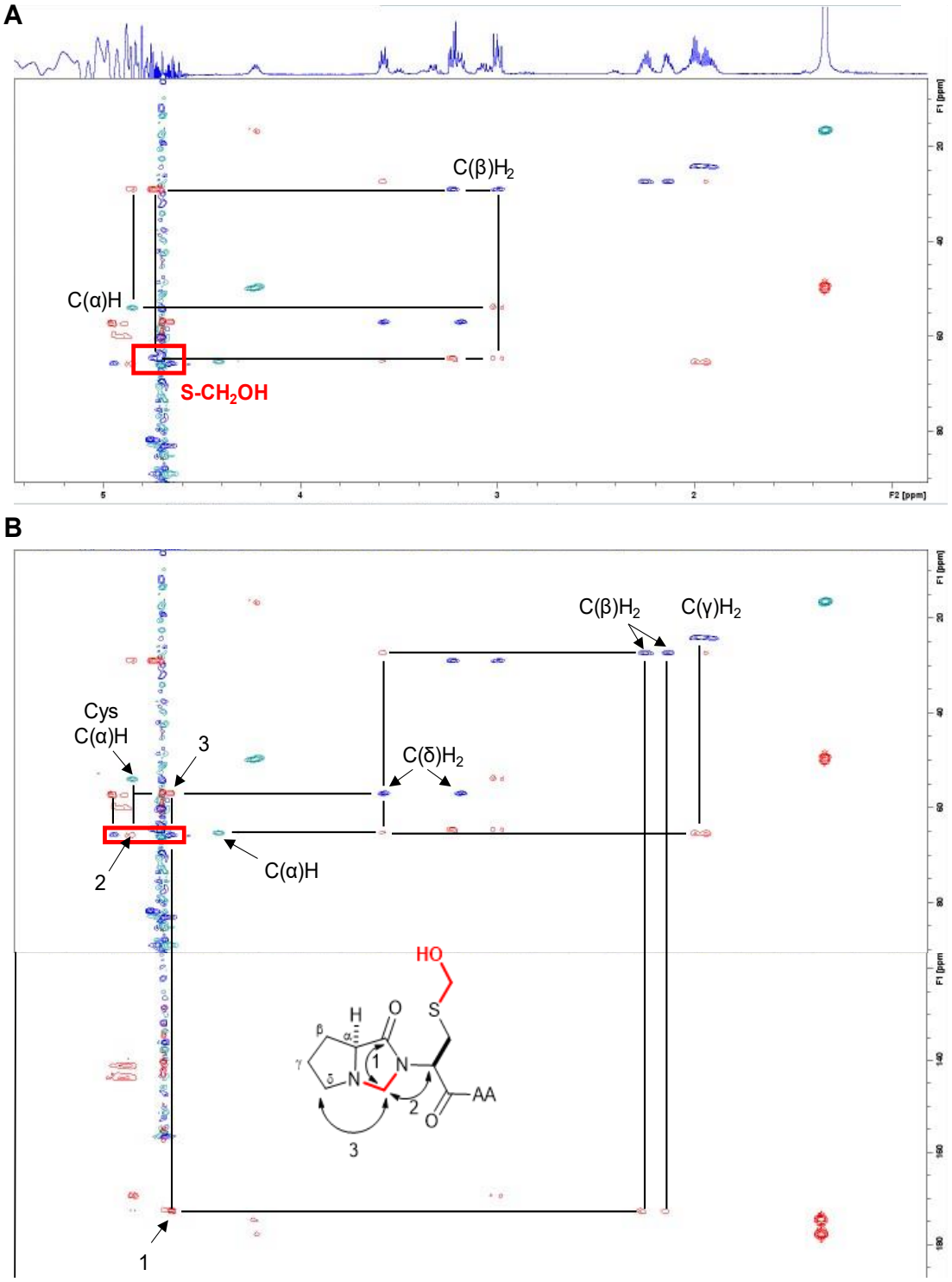


Figure 2-11: Overlaid HSQC (blue/green) and HMBC (red) of the final product of reaction of PCAA-NH_2 and HCHO. A) Correlations from cysteine demonstrating the *S*-hydroxymethyl fragment. **B)** Correlations from proline allowing the assignment of a methylene bridge.

reaction of the N-terminal residues of histone H2B by Mr Paulo Spingardi.[10] Interestingly, the chemical shift of the *S*-hydroxymethyl moiety in this structure is slightly lower (64.6 ppm) than that observed for other peptides in this thesis (65.4 ppm). The methylene bridge has a chemical

shift of 65.8 ppm in PCAA-NH_2 which is slightly lower than that observed in the reaction of a peptide derived from the N-terminal tail of histone H2B (PEPAKSAPAPKKGKK, 66.4 ppm) and of the dipeptide PE (66.6 ppm) with HCHO.[10] In the reaction of a dipeptide containing pipecolic acid followed by glutamic acid (Pip-Glu) with HCHO the observed chemical shift for the methylene bridge is 65.4 ppm (fig. 2-12).

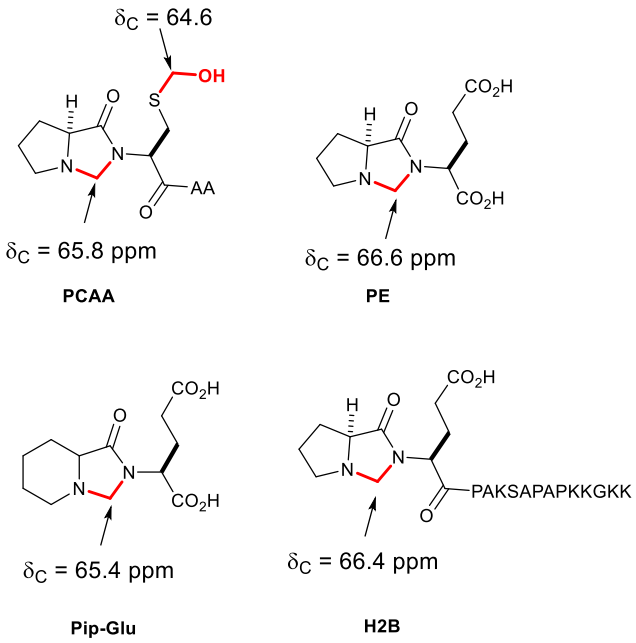


Fig 2-12: Comparison of the chemical shift of the methylene bridge carbon atom in different peptides. The reactions of PE, Pip-Glu and a histone H2B fragment with HCHO are described elsewhere.[10]

Taking these results together, the chemical shift of the methylene bridge between the N-terminal amino group and neighbouring amide bond appears to be much more variable than the chemical shift of the *S*-hydroxymethyl moieties observed in this work. The difference in observed chemical shift of the methylene bridge may be a result of varying electron density on the methylene bridge carbon as a result of minorly varying inductive withdrawal effects of the sidechain of the second amino acid in the sequence. The difference could also be explained by differences in secondary structure or other "through-space effects" affecting the local magnetic field of the methylene bridge carbon. Further investigation is needed to determine whether the

S-hydroxymethyl group carbon moving to a lower chemical shift upon the formation of a methylene bridge on the neighbouring amino acid is diagnostic or not. However, this was beyond the scope of this thesis and was not explored further. It is not immediately obvious why PCAA-NH_2 forms a bicyclic structure on reaction with HCHO while PACA-NH_2 does not. An analogous species may exist for PACA-NH_2, although the rate of reaction may be slower, meaning it was not detected in the short reaction time employed.

3. Reaction of Peptides and HCHO with Glutathione

3.1 Evaluating the Stability of Peptide-HCHO Adducts with Glutathione

As judged by Ca^{2+} influx and patch-clamp assays, TRPA1 is rapidly activated by electrophiles. However, it recovers quickly after activation and can be activated again in a short period of time.[11] This suggests the covalent bond formed between the cysteine(s) of the electrophile sensing domain and an agonist is highly reversible. In addition to this, cells treated with electrophiles appear to be depleted of glutathione (GSH) which acts as a ubiquitous nucleophile and could well comprise a cellular "sink" for HCHO and other electrophiles after their reaction with TRPA1 (and other electrophile sensing proteins).[12] To this end, GSH was used to investigate the stability of the peptide-HCHO adducts.

3.2 Effect of GSH Concentration on Ac-ACA-NH_2-HCHO Adduct

Given Ac-ACA-NH_2 forms the *S*-hydroxymethyl species rapidly when reacted with HCHO and does not undergo further reaction under the tested conditions, Ac-ACA-NH_2 was chosen as a biologically relevant substrate to explore the effects of GSH concentration on the rate of loss of HCHO from the peptide. 2 mM Ac-ACA-NH_2 was pre-incubated with 4 mM HCHO for 5 minutes before the addition of differing concentrations of GSH. In the case of addition of

an equimolar amount of GSH and Ac-ACA-NH_2, an equilibrium was rapidly reached within the time taken to acquire the first proton NMR spectrum of the reaction mixture. The *S*-hydroxymethyl-glutathione (HMG) concentration (~ 2 mM) was higher than that of the Ac-ACA-NH_2-HCHO concentration (~ 1.55 mM) suggesting that HMG is the thermodynamically favoured product and thus can act as a "sink" for electrophiles (fig. 2-13 a). Unfortunately, it was not possible to determine the concentration of free GSH or Ac-ACA-NH_2 due to the resonances for the cysteine CH_2 group in each species overlapping with one another and no other resonance arising from either species changing upon formation of the HCHO adduct (fig. 2-14).

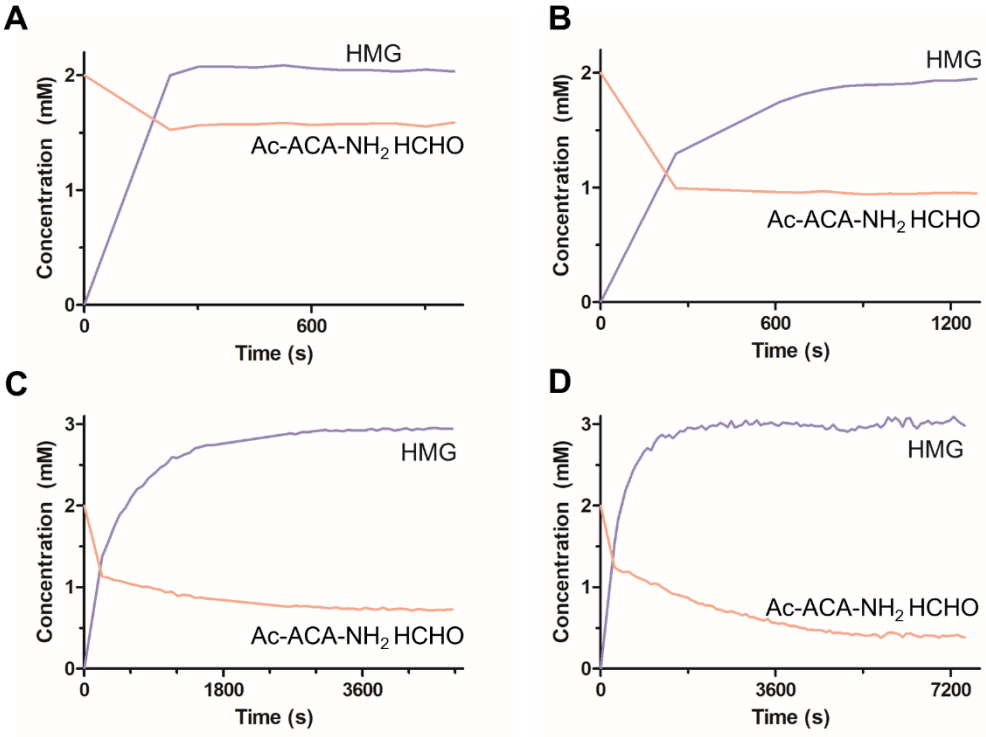


Figure 2-13: Effect of GSH on concentration of Ac-ACA-NH_2 HCHO adduct (red) to form HMG (blue) as assayed by ^{1}H-NMR. Conditions: 4 mM HCHO, 2 mM Ac-ACA-NH_2, 50 mM Tris-d_{11} pH 7.5, pre-incubated for 5 minutes before addition of: **A)** 2 mM GSH, **B)** 4 mM GSH, **C)** 10 mM GSH, **D)** 20 mM GSH.

Upon increasing the GSH concentration, it became apparent that the rate of sequestration of HCHO by GSH was much faster than the loss of HCHO from the Ac-ACA-NH_2 HCHO adduct (fig. 2-13). This is most clearly shown upon addition of a 10 molar excess of GSH compared

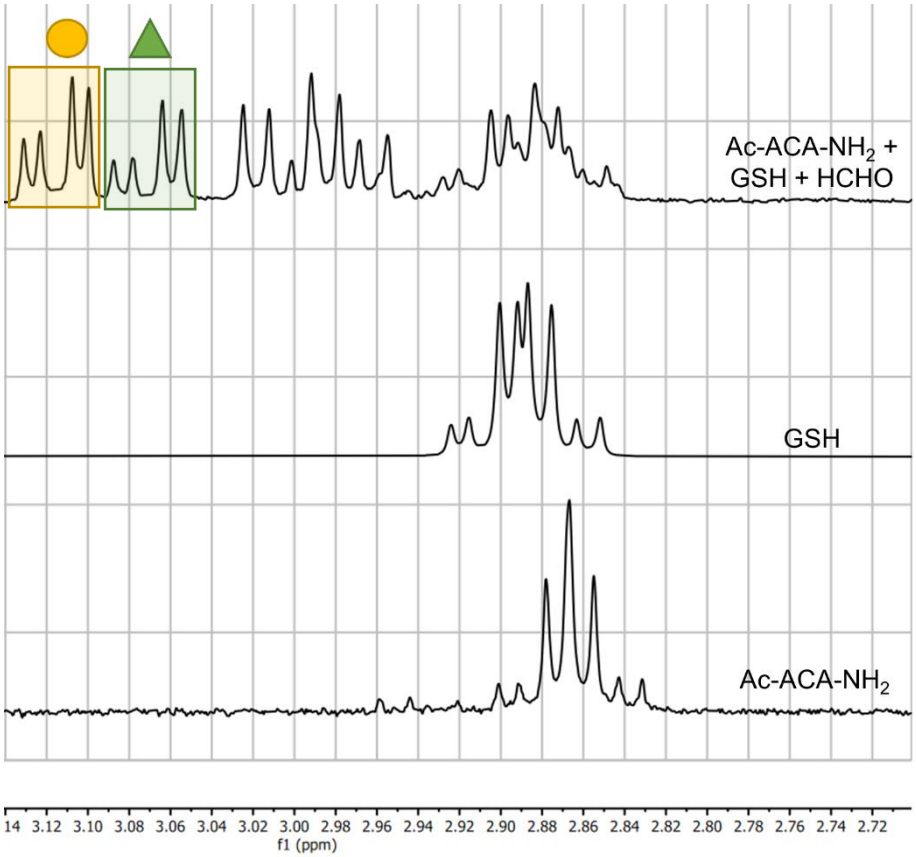


Figure 2-14: Section of ^{1}H-ESPE NMR spectra of Ac-ACA-NH_2 and GSH in the absence of HCHO, and a mixture of Ac-ACA, HCHO and GSH. The cysteine C(β)H_2 protons of free Ac-ACA-NH_2 and GSH overlap complicating analysis. However, the shift of one of the diastereotopic C(β)H_2 protons of the Ac-ACA-NH_2-HCHO adduct (yellow circle) has a chemical shift distinct from that of HMG (green triangle), allowing for their concentration to be calculated.

to Ac-ACA-NH_2 (fig. 2-13 d), wherein the concentration of the GSH-HCHO adduct levels out at about 3 mM at approximately 1800 seconds (30 minutes) after mixing, while the concentration of the Ac-ACA-NH_2-HCHO adduct continues to fall until approximately 5400 second (90 minutes), where it levels out at 0.4 mM. This suggests that: 1) loss of HCHO from a peptide-HCHO adduct is a rate limiting step, 2) there are multiple species in equilibria, 3) the reaction mechanism is not simply GSH reacting with the Ac-ACA-NH_2-HCHO adduct to give HMG and Ac-ACA-NH_2. This suggests that HCHO may be lost in a unimolecular process, followed by sequestration of HCHO with GSH to give HMG.

To investigate whether a HCHO adduct of Tris buffer was involved in the equilibrium, a solution of 2 mM Ac-ACA-NH_2 was mixed with 4 mM HCHO in 50 mM phosphate buffered D_2O before the addition of 20 mM GSH. Unfortunately, analysis of the reaction mixture was complicated by overlapping signals and signal drift of the cysteine CH_2 of the HCHO adduct of Ac-ACA-NH_2 and HMG. Analysis was attempted using linewidth narrowing using a Gaussian window function (gb = 0.3), line broadening set to -2 Hz (lb = -2), zero-filling (si = 128000) before applying the Fourier transform to the FID. Even with this line narrowing, kinetic analysis was not possible (fig. 2-15).

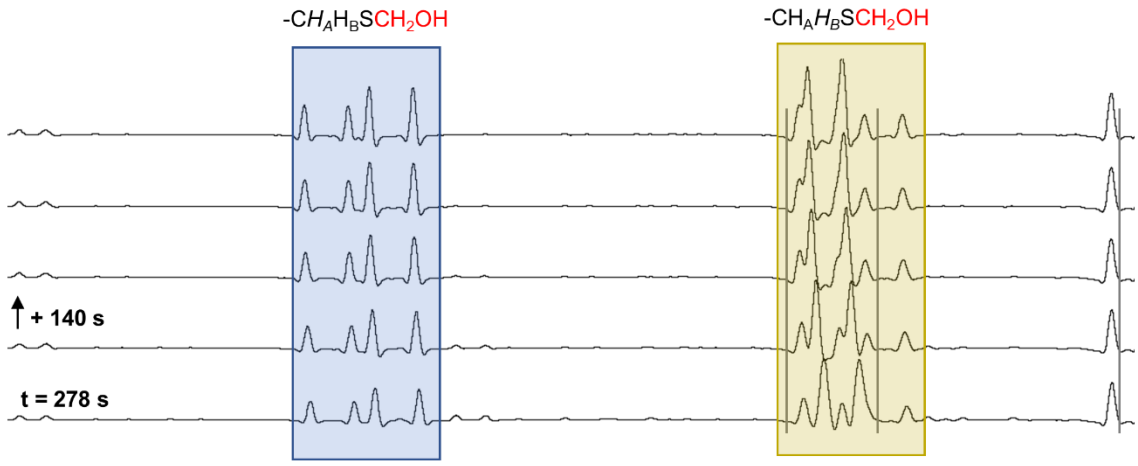


Figure 2-15: Sections of zoomed and vertically stacked ^{1}H-NMR spectra of 2 mM Ac-ACA-NH_2, 4 mM HCHO, 20 mM GSH in 50 mM phosphate buffer (pH 7.5). The resonance of one cysteine C(β)H_2 proton has the same chemical shift for both HMG and the Ac-ACA-NH_2 HCHO adduct (blue box) and the other diastereotopic proton (yellow box) overlaps and eventually coalesces. **NB**: vertical lines added for clarity of signal drift of cysteine CH_2 resonance compared to non-drifting resonances.

3.3 Effect of 20 mM GSH on Peptide-HCHO Adducts

Similar to the results with the Ac-ACA-NH_2 peptide, GSH facilitated the slow and incomplete elimination of HCHO from several other peptide-HCHO adducts (fig. 2-16). Interestingly, the cysteine rich Ac-ACACA-NH_2 seemed to form the most stable adduct (fig. 2-16 a), as judged by a higher concentration of the adduct in the reaction mixture after 3600 s (1 h) compared to the other peptides. The reason for this increased stability compared to other peptides is not

immediately apparent, although the increased hydrophobicity of the peptide may cause it to aggregate in solution, and limit the mixing and formation of an equilibrium between the different species in the reaction mixture. Alternatively, the presence of a second cysteine, which can also form an *S*-hydroxymethyl group, may mean that an extra equivalent of HCHO must be lost from the peptide, thus slowing the observed rate of reaction.

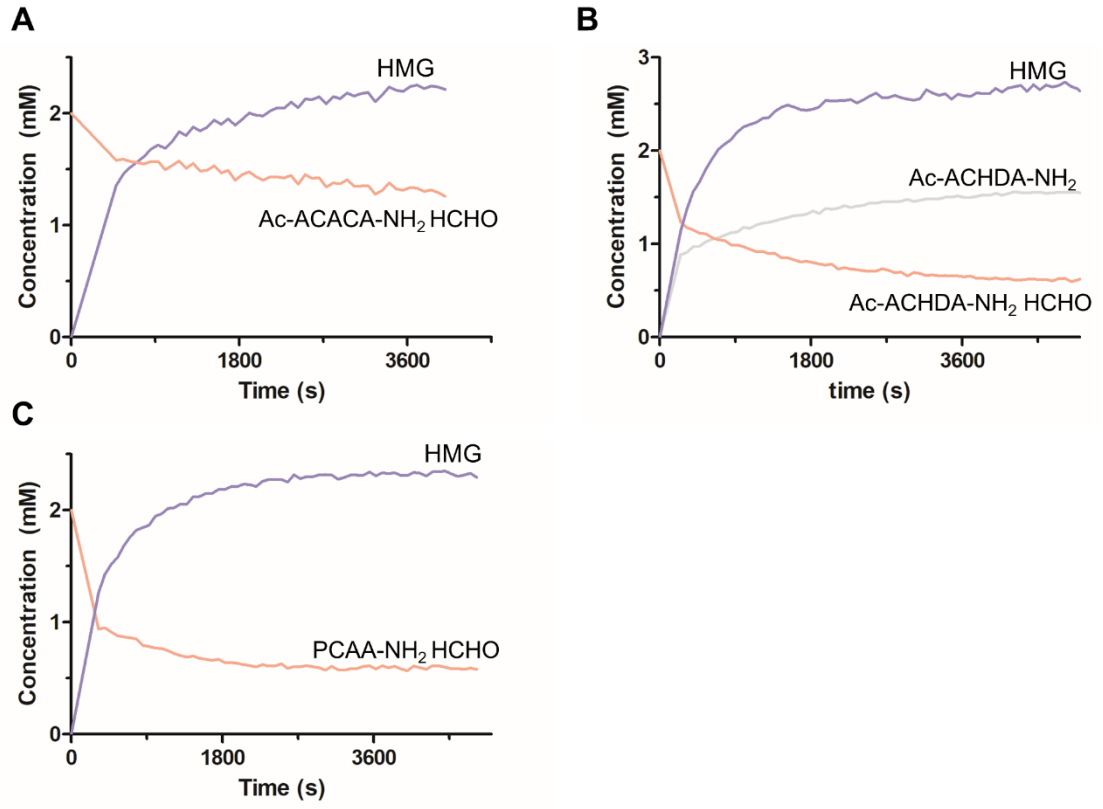


Figure 2-16: Reaction of peptide-HCHO adducts (2 mM peptide, 4 mM HCHO) with 20 mM GSH in 50 mM Tris-d_{11} in D_2O (pH 7.5) as judged by ^{1}H-NMR. Graphs show concentration in mM of HMG (blue), peptide-HCHO adduct (red), and of the peptide (grey). **A)** Ac-ACACA-NH_2, **B)** Ac-ACHDA-NH_2, **C)** PCAA-NH_2.

The results of these experiments do not fully correlate with the observation that TRPA1 can be re-activated within minutes after exposure to an electrophile given the slow and incomplete loss of HCHO over the timescale of hours. Assuming that the cysteines of TRPA1 react with HCHO in an analogous manner as observed with peptides, *i.e.* by reversibly forming *S*-hydroxymethyl moieties, there are possible explanations for this observation: 1) the local concentration of GSH (1-2 mM in most cells)[13,14] is in a sufficiently large excess compared to

that of the TRPA1-electrophile adduct. It is likely this molar excess is greater than 1000x and thus this would not have been feasible to study by ^{1}H-NMR, or 2) proteins (or other molecules) associate with TRPA1 and catalyse the elimination of HCHO, or other electrophiles, from the cysteine thiol group. However, the diversity in electrophilic moieties that TRPA1 can bind, and hence elimination mechanism, is so great that of these possibilities the first explanation is perhaps more likely.

4. Modelling TRPA1 Reactivity

4.1 Engineering a Soluble Electrophile Binding Domain

In order to further explore the interaction between TRPA1 and electrophiles, I sought to create a highly soluble, easy to express and high yielding protein construct that contains the TRPA1 electrophile binding domain (EBD). Previous work in the group by Dr. R. Hopkinson in collaboration with the Oxford Protein Production Facility (OPPF) had failed to identify any constructs from the TRPA1 gene containing the EBD which showed good levels of expression or solubility from *Escherichia coli*. Strategies to get around these issues are to switch to a eukaryotic expression system (*e.g.* yeast, human cell lines (HEK293) or insect cells (Sf9)) to ensure the protein is properly folded, increasing its solubility. Additionally, it was envisaged that a fusion protein (*e.g.* thioredoxin, maltose-binding protein) which is easily expressed and has high solubility may be added to the N- or C-terminus of the target protein, thus counteracting its low expression/solubility. The fusion protein approach was not explored due to concerns that the fusion protein would interfere with electrophile binding, either from cysteines/lysines on the fusion protein binding electrophiles, or the fusion protein increasing the steric bulk around the reactive cysteines in the EBD. Furthermore, target proteins with low solubility may aggregate and precipitate after cleavage of the fusion protein.

An alternative strategy using *E. coli* was used for several reasons: 1) previous experience in the group in working with *E. coli*, 2) expression from human cell lines are typically low yielding and expensive to maintain, 3) expression from yeast/insect cells was not available "in-house" for the group. Previous studies in the group have shown that synthetic consensus ankyrin repeats[15] can manifest a high level of expression, solubility and are correctly folded when expressed in *E. coli*. Given that the EBD of TRPA1 sits adjacent to the ankyrin repeat domain (ARD), it was envisaged that capping the EBD construct with synthetic ankyrins would improve solubility, aid folding and mimic the protein environment surrounding the EBD. To this end, two different constructs were designed based on the 3CA and 5CA proteins,[16] substituting the central ankyrin repeat with the TRPA1 electrophile binding domain, named Benkyrin 3 and 5 respectively (BK3/5).

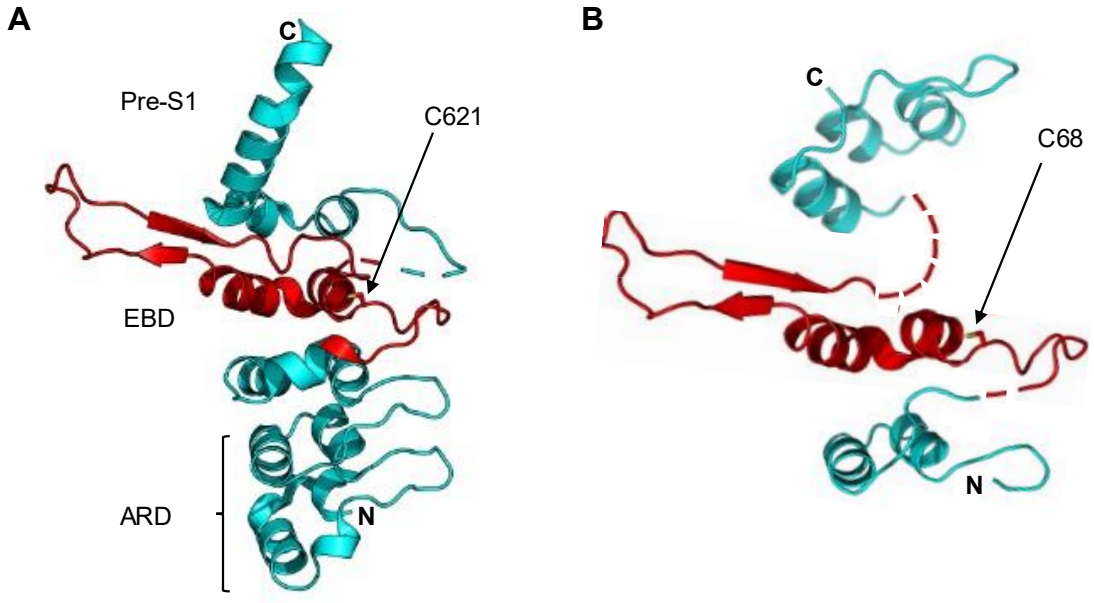


Figure 2-7: Views comparing the electrophile binding domain (EBD, red) of TRPA1 with that of BK3. A) Views from a cryo-EM structure of TRPA1 (PDB accession code 6PQQ). The ankyrin repeat domain (ARD) and pre-S1 helix are shown in blue. **B)** Model of BK3 built by modelling the ankyrin end caps (blue) with Phyre2.[35]

4.2 Cloning and Protein Expression Trials

The gene encoding for BK3 was prepared by Gene Art (Invitrogen) and amplified using polymerase chain-reaction (PCR) with forward and reverse primers that extended the target gene to include the ligation-independent cloning (LIC) site at both the 5′- and 3′-end of the gene.[17] An intense band at approximately 500 base pairs (b.p.) was observed in all of the reaction mixtures which correlates to the BK3 gene with LIC ends (fig. 2-18). This band was cut from the agarose gel and the DNA extracted using a GeneJet gel extraction kit (Thermo Scientific). Further treatment of the isolated DNA with T4 polymerase, utilising its 3′-endonuclease activity to create overhangs, generated an insert suitable for a LIC reaction with the plasmid pNIC28-Bsa4. The pNIC28-Bsa4 plasmid was cut with BsaI and treated with T4 polymerase (kindly performed by Dr. G. Langley) to generate complimentary LIC overhangs to the BK3 gene. The plasmid and gene were mixed and incubated at rt before transformation into β-10 *E. coli* (NEB) and plating onto kanamycin-sucrose plates. Kanamycin-sucrose agar plates were used as a negative selection for the untreated pNIC28-Bsa4 plasmid. pNIC28-Bsa4 contains the SacB gene which, when transformed into *E. coli*, gives rise to a protein that converts sucrose to levan, a polysaccharide which is a toxic to *E. coli* when it accumulates in the periplasm of the cell.[18]

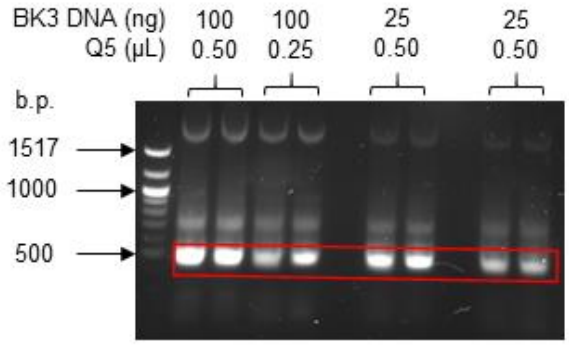


Figure 2-18: Agarose gel analysis of PCR mixture amplifying BK3 DNA insert with LIC sites at 5′- and 3′- terminus. Reactions carried out with differing amounts of template DNA and Q5 polymerase. **Reaction conditions**: 0.5 mM primers, 1 mM dNTP mix, BK3 template DNA, Q5 polymerase, Q5 buffer. Bands at 500 bp excised (red square).

Colonies were grown in 2-YT media supplemented with kanamycin at 37 °C overnight before the cells were harvested and the BK3-pNIC28-Bsa4 plasmid was extracted using a GeneJet plasmid miniprep kit (Thermo Scientific). Sequencing of several colonies using the T7 promoter (T7F) and terminator (T7R) on the plasmid backbone for the forward and reverse primers respectively confirmed the presence of the intact BK3 gene in the correct orientation for protein production.

In order to optimise protein production conditions for BK3, the BK3-pNIC28-Bsa4 plasmid was transformed into BL21(DE3) *E. coli* which produce the T7 RNA polymerase under control of the lacUV5 promoter.[19] Under normal fermentation conditions, the Lac-repressor binds to the lacUV5 promoter, preventing transcription of the downstream gene.[19] If isopropyl β-d-1-thiogalactopyranoside (IPTG) is added to the culture media, it binds to the Lac-repressor and releases it from the lacUV5 promoter, allowing expression of the T7 polymerase and subsequent transcription of the target gene under control of the T7 promoter.[20] In this way, expression of the target gene can be finely controlled by the addition of IPTG to the culture media when using BL21(DE3).

The first two variables optimised for in protein production trials were the fermentation temperature and IPTG concentration. A higher temperature contributes to a greater metabolic rate of *E. coli* but could negatively affect protein structure. Transformed BL21(DE3) were grown at 37 °C overnight in 2-YT media supplemented with kanamycin before being inoculated into fresh 2-YT media, grown at 37 °C to an OD_{600} of 0.6, cooled to 18 or 30 °C and induced with 0, 50 or 500 µM IPTG. These cultures were allowed to grow overnight before being harvested and lysed by sonication in binding buffer (20 mM Na_2HPO_4, 500 mM NaCl,

5mM imidazole, pH 7.4). Analysis by SDS-PAGE showed good levels of production of a protein at ~20 kDa, which is in close agreement with the expected mass of 20.4 kDa for BK3, after induction by 50 and 500 µM IPTG at 18 °C. However, most BK3 was confined to the insoluble fraction after clearing the lysate by centrifugation. The same results were seen for induction at 30 °C, however, there appeared to be a significant level of BK3 synthesis without induction by IPTG suggesting poor control over protein production levels (fig. 2-19). Therefore, induction at 18 °C with 500 µM IPTG was taken to be the optimised condition.

With the aim of optimising protein solubility, cell lysis was carried out by sonication in different buffers. Despite initial discouraging results showing that BK3 was insoluble in binding buffer, BK3 appeared soluble in a range of buffers over the pH range 7.5 to 8.5 and in the presence of chaotropic agents or not (fig. 2-20).[21] Tris pH 7.5 was chosen as an appropriate buffer for BK3 on account of solubilising BK3 and being compatible with a wide range of other proteins.

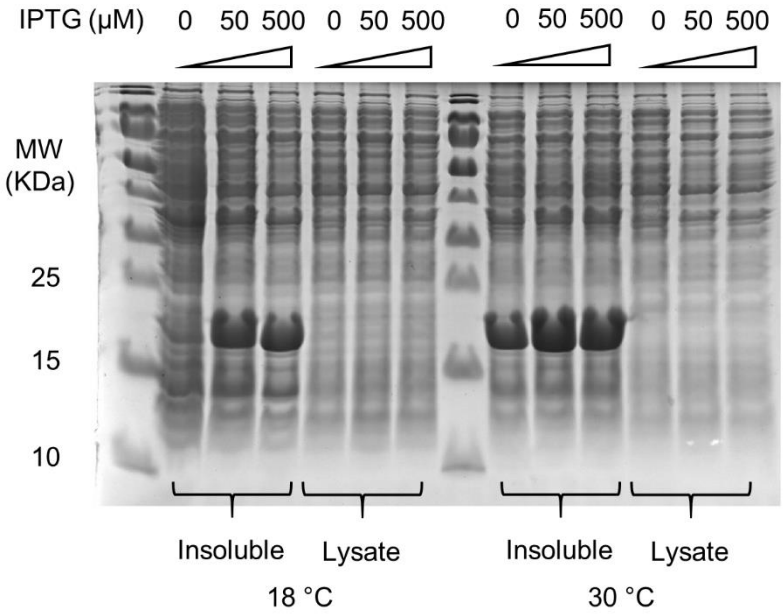


Figure 2-19: SDS-PAGE analysis of insoluble and soluble fractions of BK3 expressing BL21(DE3) lysate. BK3 production was induced at different temperatures and with different concentrations of IPTG. Approximate molecular weight (MW) is shown using PageRuler pre-stained protein ladder (Thermo Scientific)

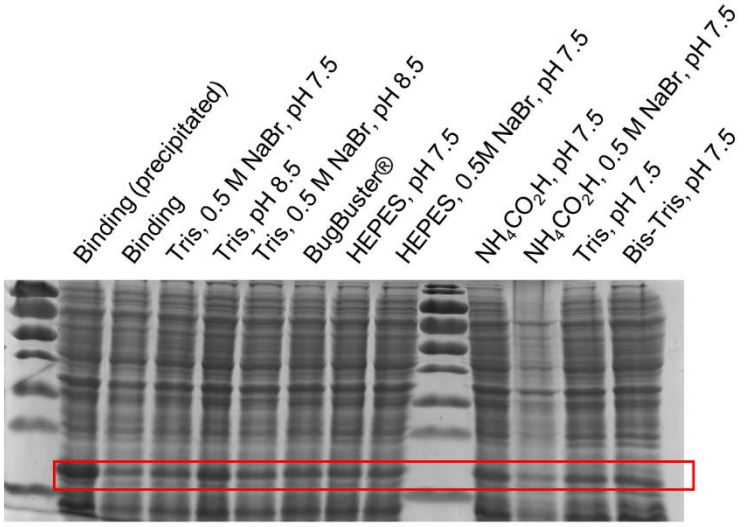


Figure 2-20: SDS-PAGE analysis of BK3 expressing BL21(DE3) lysed by sonication in different buffers. Concentration of buffer component is 50 mM unless stated otherwise. Approximate molecular weight (MW) is shown using PageRuler pre-stained protein ladder (Thermo Scientific)

4.3 Purification of BK3

Large scale purification of BK3 from 8 L of 2-YT media induced with 0.5 mM IPTG at 18 °C overnight was attempted following a two-step chromatographic procedure. First, cells were harvested by centrifugation and lysed by sonication in 50 mM Tris pH 7.5 supplemented with the protease inhibitor phenylmethylsulfonyl fluoride and DNase on ice. The resulting lysate was cleared by centrifugation and syringe filtration before being loaded onto a nickel affinity column (HisTrap, GE). The HisTrap was washed with 30 column volumes (CV) of 25 mM imidazole in 50 mM Tris pH 7.5. Proteins were eluted using a gradient from 25 to 250 mM imidazole in 50 mM Tris pH 7.5 over 20 CV. Fractions were analysed by SDS-PAGE, and those containing BK3 were pooled and concentrated. The pooled fractions were loaded onto a gel filtration column (S75 300 mL) and eluted with 50 mM Tris pH 7.5. The protein mixture eluted as a single peak on the UV trace of the FPLC but fortunately SDS-PAGE analysis revealed that BK3 separated from a heavier protein and could be pooled and concentrated (fig. 2-21).

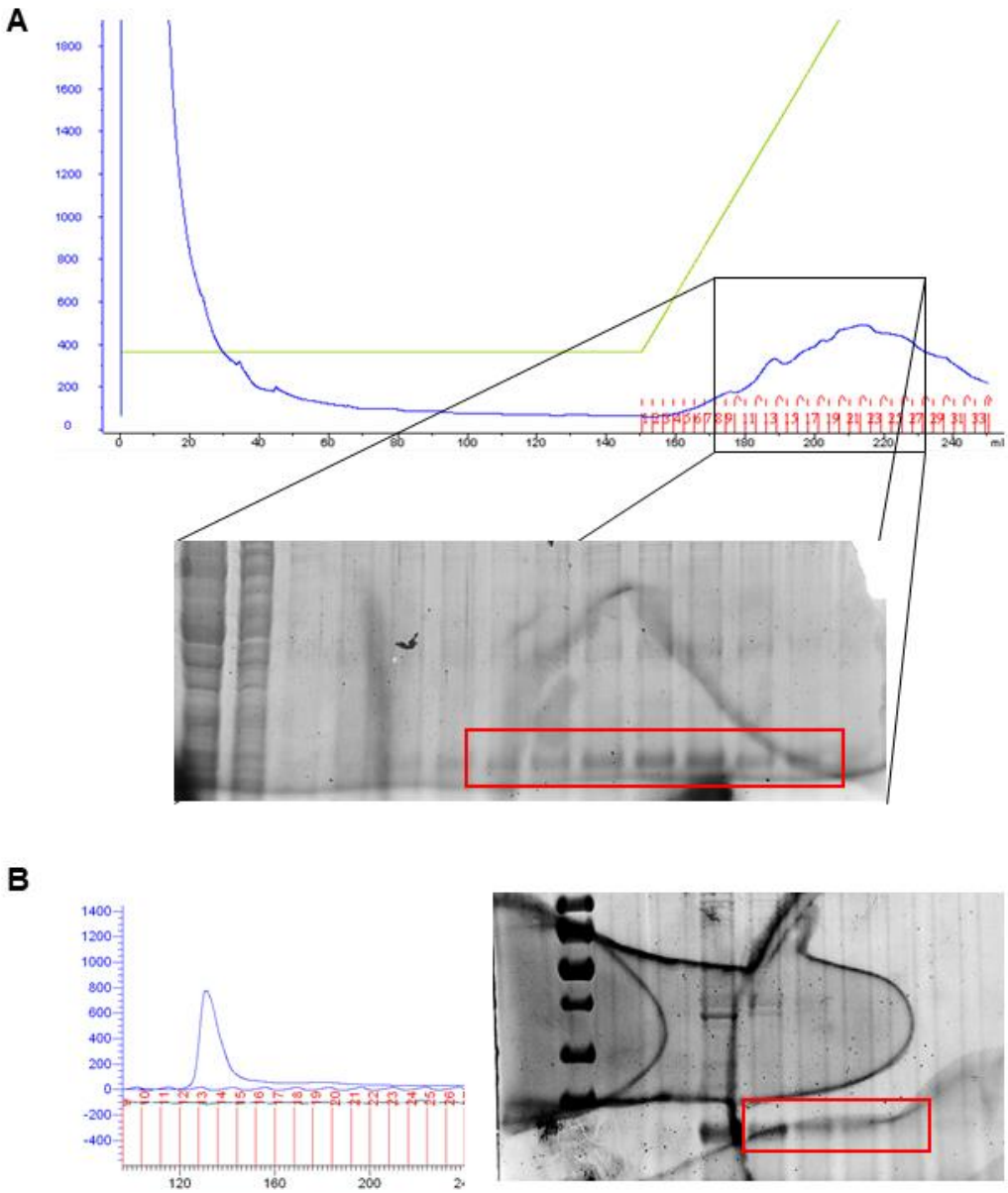


Figure 2-21: FPLC traces and SDS-PAGE analysis of BK3 purification. Fractions in red box were pooled and concentrated. **A)** HisTrap (nickel affinity chromatography) **B)** S75 Superdex 300 mL (gel filtration chromatography). Approximate molecular weight (MW) is shown using PageRuler pre-stained protein ladder (Thermo Scientific)

4.4 Mass-Photometry of BK3 (performed by Ms. Anna Olerinyova)

Mass photometry is a new technique that has been developed by the Kukura group (University of Oxford) that uses light scattering to measure the mass of biological polymers in solution.[22] Light scattering is proportional to the refractive index and volume of the protein. Thus, given the specific volume of amino acids scales with molecular weight and the refractive index of

proteins varies little, it can be assumed that scattering is proportional to the mass of the protein.[22] Mass photometry can be performed under non-denaturing conditions, allowing the molecular weight of protein complexes to be determined and thus their stoichiometry.[23]

BK3 was subjected to mass photometry analysis (fig. 2-25) to investigate its stoichiometry in solution. Given TRPA1 is a tetramer in its native state (see Chapter I §2.3), a protein mimicking it should ideally be a tetramer in solution. The mass photometry spectrum implies that the majority of the protein in solution was a dimer due to the intense peak at 40 kDa. The main peak obscures an additional peak on its shoulder which is difficult to resolve due to the low resolution of mass photometry but its presence suggests the formation of a trimer. In addition, peaks corresponding to a hexamer (120 kDa) and dodecamer (240 kDa) were also observed suggesting that BK3 is polydisperse in solution.

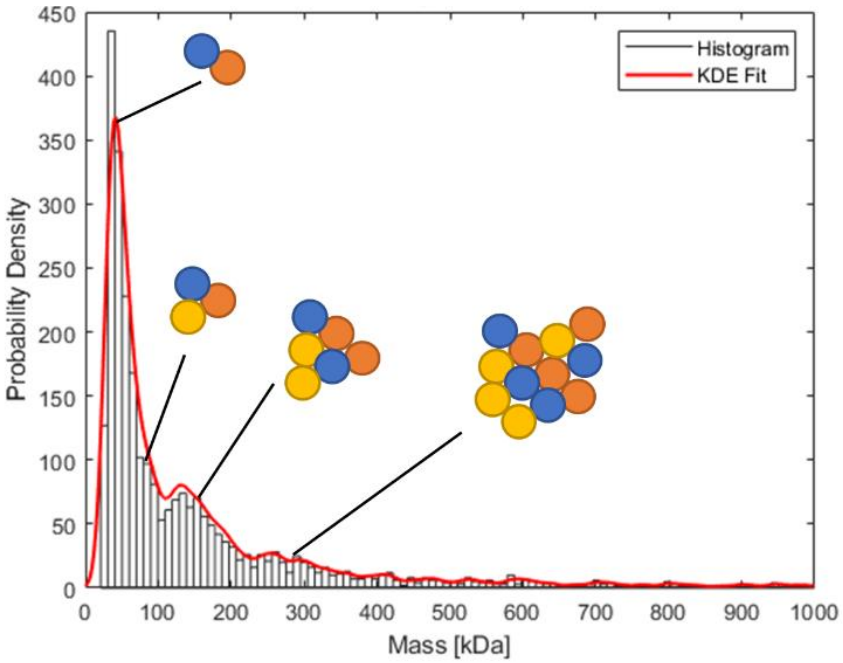


Figure 2-22: Mass photometry spectrum of BK3 in 50 mM Tris pH 7.5. Oligomerisation state of BK3 is represented by spheres. KDE = kernel density estimation. (Experiment kindly performed and data analysed by Ms. Olerinyova)

4.5 Reaction of BK3 with Electrophiles

To investigate how BK3 (and, potentially, by implication TRPA1) reacts with electrophiles, BK3 was mixed with electrophiles and subjected to denaturing mass spectrometry (MS). Denaturing MS disrupts the structure of proteins through the use of acidic buffers which contain organic solvent to aid ionisation of the protein, allowing for a mass spectrum to be acquired.[24] The mass spectrum of BK3 without any electrophiles added shows that the protein carries no post-translational modifications and has an observed mass (20377 Da) in agreement with its predicted mass (20379 Da). When mixed with iodoacetamide (**IAA**) and the riot control agent **CS**[25] at rt, BK3 reacted to give a products apparently arising from the single and double addition of an electrophile. This mirrors the reaction of C621 and C665 with electrophiles within the electrophile binding domain of TRPA1, resulting in TRPA1 activation (see Chapter 1 §2.3). BK3 underwent several reactions with 3-bromo-1,1,1-trifluoroacetone (**BTFA**), apparently with all six cysteines in the protein (fig. 2-22).

Attempts to observe the reversible binding of electrophiles such as HCHO or **CR** by denaturing MS were unsuccessful and only unreacted BK3 was observed. This is possibly due to the ligand being displaced as BK3 undergoes denaturing when introduced to the mass spectrometer. Thus, the interaction of BK3 with reversibly binding electrophiles was examined by NMR. Another benefit of examining the interaction by a non-denaturing, in-solution method such as NMR is that the effect of protein structure on ligand binding can be investigated. TRPA1 agonists are generally lipophilic and it was envisaged that solutions of the agonists in organic solvents (*eg.* DMSO) could affect the structure of BK3 and disrupt binding of the agonists.

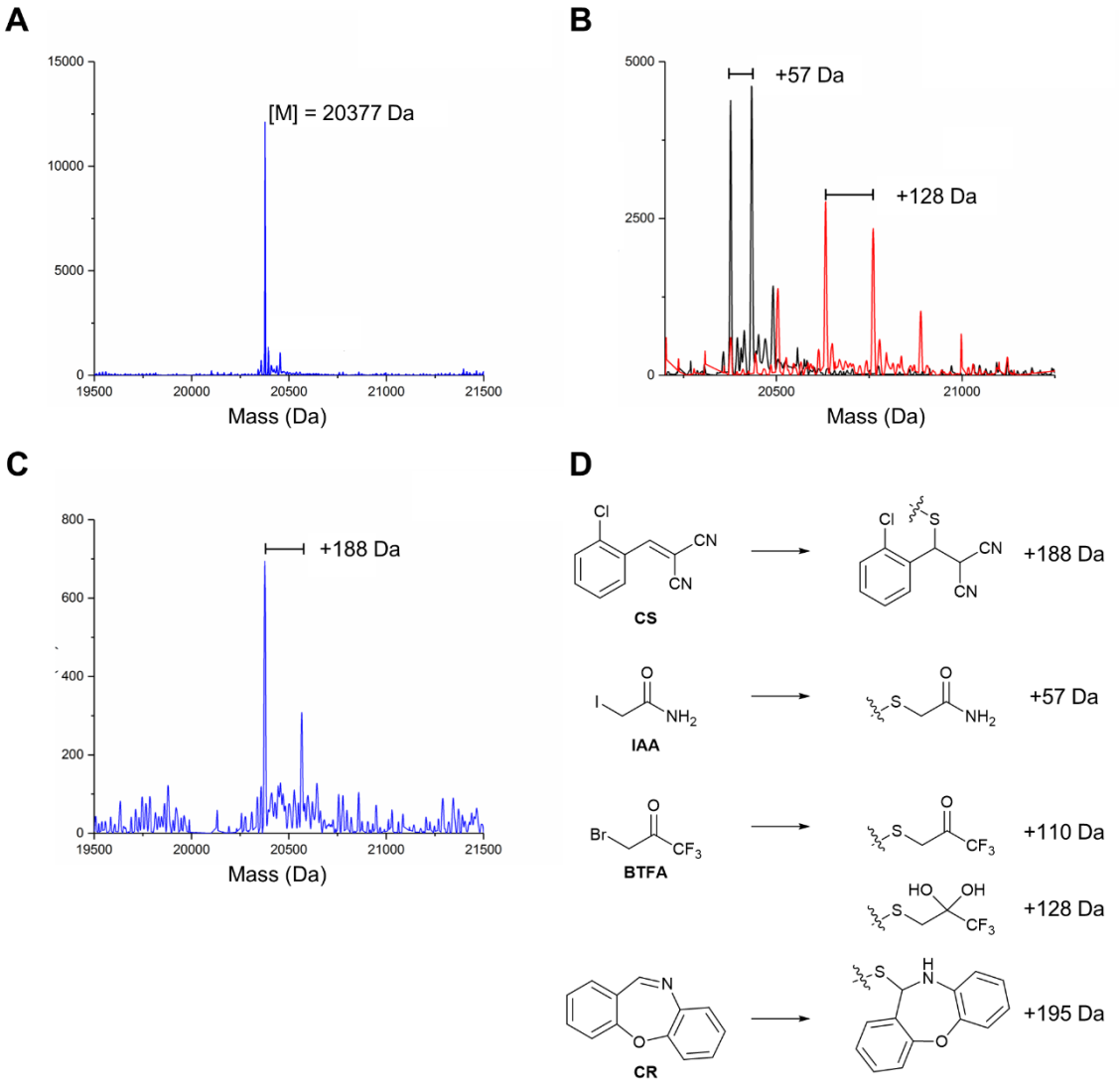


Figure 2-23: Reactions of BK3 with electrophiles monitored by denaturing MS. A) BK3 **B)** BK3 + IAA (black), BK3 + BTFA (red). **C)** BK3 + CS tear gas. **D)** reaction schemes for cysteine reacting with electrophiles and predicted mass shift. **Reaction conditions:** 16 µM BK3, 0-0.5 mM IAA or CS, 0-1 mM BTFA, 50 mM sodium phosphate buffer (pH 7.5).

To investigate the effect that different concentrations of DMSO have on BK3, the binding of **CR** to BK3 was examined using the Carr-Purcell-Meiboom-Gill method with refocusing pulses to suppress J modulation (CPMG-PROJECT).[26] The CPMG pulse sequence can be used to supress broadened signals arising from protein-bound ligands, this is a result of the bound and free ligands having different spin-spin (T2) relaxation rates. This means that the integral of ligand resonances in a spectrum acquired with the CPMG-PROJECT experiment correspond

to the concentration of the free ligand in solution.[27,28] As a result, the dissociation constant (k_D) can be calculated for a given reaction.

Upon mixing 500 µM **CR** with BK3 in varying concentrations of DMSO or with BK3 that was denatured by heating to 95 °C for 5 minutes, binding of CR was observed. The data were fit using a non-linear method previously described in the literature.[29] It was found that the binding curves for the heat denatured BK3 and 5% DMSO (v/v) were similar and showed much less tighter binding than BK3 in 1% DMSO (v/v) (fig. 2-23). The k_D for CR gas in 1% DMSO is 3.49 ± 1.20 µM which is significantly different from that observed for 5% DMSO (16.0 ± 5.83 µM) and the heat denatured BK3 (11.6 ± 5.66 µM). This result indicates that BK3 denatures when exposed to high concentrations of DMSO, potentially limiting the ability to study highly lipophilic TRPA1 agonists due to needing a relatively high concentration (0.05 - 0.5 mM) of a ligand to observe an intense signal and generate a binding curve.

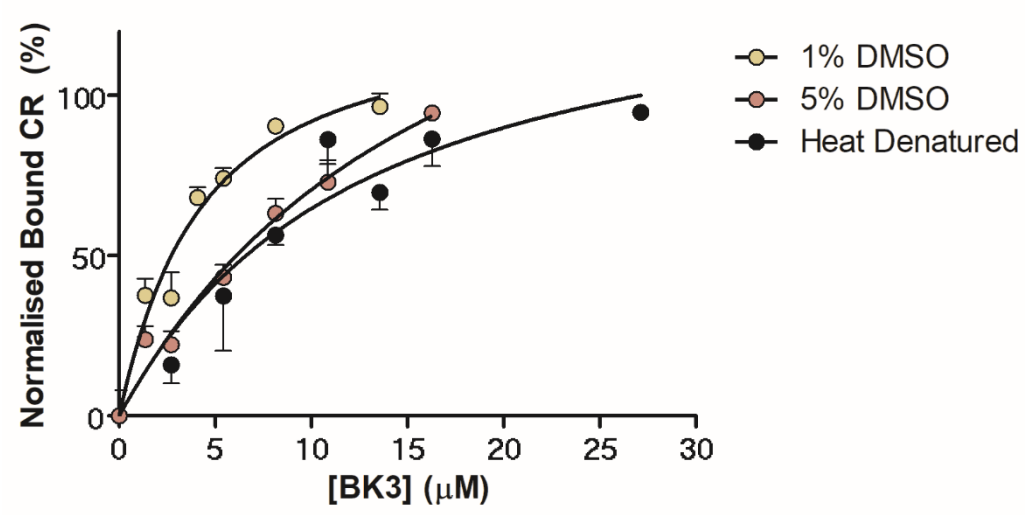


Figure 2-24: CPMG-PROJECT generated binding curve of CR to BK3. Reaction conditions: 0.5 mM CR tear gas, DMSO (1% or 5% v/v), BK3, 50 mM Na_2HPO_4 pH 7.5.

4.5 Tryptic Digests of BK3

To further characterise the reaction of BK3 with electrophiles, tryptic digests were conducted on reacted BK3. BK3 was first denatured with 6M urea then reduced using DTT. The reaction mixture was then split in two with one half being treated with IAA and the other not alkylated. The mixture was then diluted with 50 mM Tris pH 8 to bring the urea concentration below 1M. Trypsin/Lys-C mix was added and the mixture was incubated overnight at 37 °C.[30] The digest was quenched by the addition of CF_3CO_2H and peptides were extracted from the digestion using C-18 ZipTip purification. Peptides were eluted with CHCA matrix directly onto the MALDI target. Analysis of the resulting peptide mass fingerprint (PMF) showed 4 out of 5 cysteines in BK3 were covered by the digestion, with each cysteine being labelled by IAA when alkylated under reducing conditions (fig. 2-24). The reactive cysteine in the EBD of TRPA1 is C621 which corresponds to C68 in BK3. MS/MS was attempted on the peptide fragment containing C68, however, poor coverage of the peptide was observed. In the interests of time, increasing coverage using LC-MS-MS was not pursued due to instrument down-time compounded by the COVID-19 pandemic.

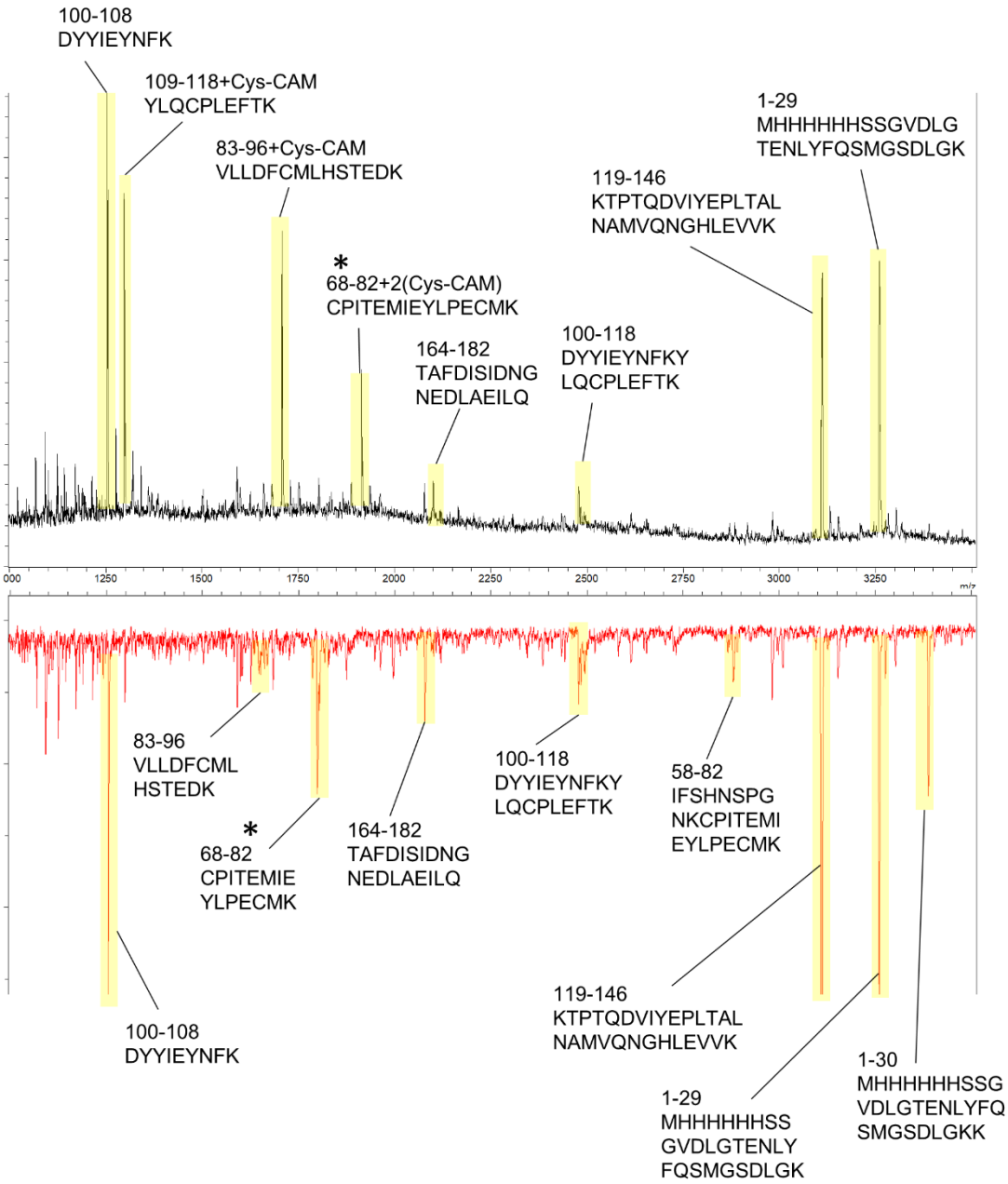


Figure 2-25: MALDI-MS of BK3 treated with IAA (black) or not alkylated (red) digested with trypsin/Lys-C mixture. Mass and residue numbers are given for peptides arising in the spectrum. The presumed reactive site of BK3 (C68) is marked with a star. Cys-CAM represents cysteine modified with IAA to give the SCH_2CONH_2 group.

5. Conclusions and Further Work

The work described in this chapter focused on investigating the nature of the interaction between cysteine containing peptides and HCHO. Cysteine residues undergo rapid reaction with HCHO to give the *S*-hydroxymethyl species. It was found the formation of a resonance at

65.4 ppm in the ^{13}C spectrum of a peptide HCHO mixture is diagnostic of a *S*-hydroxymethyl group (see fig. 2-9). Work on proline containing peptides showed similar reactions (see fig. 2-11) to that observed previously in the group, further expanding our knowledge of how HCHO reacts with N-terminal proline containing biomolecules.

Evidence was found suggesting that GSH is able to act as a "sink" for HCHO and that HCHO adducts in peptides can be removed through an equilibrium GSH to give HMG and the free peptides (see §3.2 and §3.3). Further work would comprise studying this interaction at lower concentrations of the peptide and GSH which are more comparable to that found in cells. Additionally, the reaction of peptides with a greater range of electrophiles that are relevant to TRPA1 such as **HNE**, cinnamaldehyde and **CR** could be investigated. From the work in this chapter, it is possible to hypothesise that the rapidly reversible covalent modification of TRPA1 is due to GSH but experiments at lower concentrations would have to be performed.

Beyond peptides, a recombinant model of the TRPA1 EBD (BK3) was designed and recombinantly produced (see §4.2 and §4.3). It was found that BK3 underwent reaction with several different electrophiles and the reaction could be interrogated by NMR in the case of a reversible reaction, such as with **CR** (see fig. 2-24), and MS in the case of an irreversible reaction (see fig. 2-23). Work on tryptic digests of BK3 has shown the reactive site could be identified by peptide mass fingerprinting (see fig. 2-25) combined with MS/MS of the relevant peptide although this was not achieved in the work contained in this thesis due to the COVID-19 pandemic. Further work would comprise the identification of the site of reaction of a range of electrophiles.

Crystallography on BK3 was attempted but unsuccessful, therefore it is hard to know what comparisons can be drawn between TRPA1 and BK3. BK3 may serve as a model protein for testing the reactivity of compounds with cysteine which opens up new avenues in the discovery of selective cysteine modifying probes, for example. Structural work on BK3 to investigate why C621 of TRPA1 binds electrophiles selectively and how this leads to channel opening was halted on account of the publication of a series of high resolution cryo-EM structures of TRPA1 in complex with various ligands.[31,32] The authors of these structures were able to propose a mechanism by which modification of C621 might lead to channel opening (see Chapter I §2.3). These structures revealed that the reactivity of C621 may be the result of electrostatic interactions between C621, its neighbouring helix and the π-system of F612 (see Chapter I §2.3). These interactions and mechanisms are not immediately obvious when looking at the structure of TRPA1 by Julius *et al*.[33] and are unlikely to have been revealed without studying the TRPA1 EBD in its native environment. However, to further verify BK3 as a TRPA1 mimic, substitution of residues surrounding C68 (C621 in TRPA1) and disruption of key interactions could be undertaken, but due to time limitations this was not explored.

6. Chapter II Bibliography

(1) Kamps, J. J. A. G.; Hopkinson, R. J.; Schofield, C. J.; Claridge, T. D. W. How Formaldehyde Reacts with Amino Acids. *Communications Chemistry* **2019**, *2* (1). https://doi.org/10.1038/s42004-019-0224-2.

(2) Mason, R. P.; Sanders, J. K. M.; Crawford, A.; Hunter, B. K. Formaldehyde Metabolism by Escherichia Coli. Detection by in Vivo ^{13}C NMR Spectroscopy of S-(Hydroxymethyl)Glutathione as a Transient Intracellular Intermediate. *Biochemistry* **1986**, *25* (16), 4504–4507. https://doi.org/10.1021/bi00364a008.

(3) Bateman, R.; Rauh, D.; Shokat, K. M. Glutathione Traps Formaldehyde by Formation of a Bicyclo[4.4.1]Undecane Adduct. *Organic and Biomolecular Chemistry* **2007**, *5* (20), 3363–3367. https://doi.org/10.1039/b707602a.

(4) Hopkinson, R. J.; Barlow, P. S.; Schofield, C. J.; Claridge, T. D. W. Studies on the Reaction of Glutathione and Formaldehyde Using NMR. *Organic and Biomolecular Chemistry* **2010**, *8* (21), 4915–4920. https://doi.org/10.1039/c0ob00208a.

(5) Holmquist, B.; Vallee, B. L. Human Liver Class III Alcohol and Glutathione Dependent Formaldehyde Dehydrogenase Are the Same Enzyme. *Biochemical and Biophysical Research Communications* **1991**, *178* (3), 1371–1377. https://doi.org/10.1016/0006-291X(91)91045-E.

(6) Naylor, S.; Mason, R. P.; Sanders, J. K.; Williams, D. H.; Moneti, G. Formaldehyde Adducts of Glutathione. Structure Elucidation by Two-Dimensional n.m.r. Spectroscopy and Fast-Atom-Bombardment Tandem Mass Spectrometry. *The Biochemical journal* **1988**, *249* (2), 573–579. https://doi.org/10.1042/bj2490573.

(7) Buller, A. R.; Townsend, C. A. Intrinsic Evolutionary Constraints on Protease Structure, Enzyme Acylation, and the Identity of the Catalytic Triad. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110* (8), E653. https://doi.org/10.1073/pnas.1221050110.

(8) Henry, L.; Leung, I. K. H.; Claridge, T. D. W.; Schofield, C. J. γ-Butyrobetaine Hydroxylase Catalyses a Stevens Type Rearrangement. *Bioorganic and Medicinal Chemistry Letters* **2012**, *22* (15), 4975–4978. https://doi.org/10.1016/j.bmcl.2012.06.024.

(9) Denby, K. J.; Iwig, J.; Bisson, C.; Westwood, J.; Rolfe, M. D.; Sedelnikova, S. E.; Higgins, K.; Maroney, M. J.; Baker, P. J.; Chivers, P. T.; Green, J. The Mechanism of a Formaldehyde-Sensing Transcriptional Regulator. *Scientific Reports* **2016**, *6* (1), 1–15. https://doi.org/10.1038/srep38879.

(10) Spingardi, P. Studies on the Reaction of Formaldehyde with Peptides and Related Biomolecules, Part II Thesis, Department of Chemistry, University of Oxford, 2016.

(11) Hinman, A.; Chuang, H.-H.; Bautista, D. M.; Julius, D. TRP Channel Activation by Reversible Covalent Modification. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 19564–19568. https://doi.org/10.1073/pnas.0609598103.

(12) Boyland, E.; Chasseaud, A. L. F. Enzyme-Catalysed Conjugations of Glutathione with Unsaturated Compounds. *The Biochemical journal* **1967**, 104 (1), 95-102. https://doi.org/10.1042/bj1040095.

(13) Meister, A. Glutathione Metabolism and Its Selective Modification. *Journal of Biological Chemistry* **1988**, 263 (33), 17205-17208. https://doi.org/10.1016/S0021-9258(19)77815-6.

(14) Forman, H. J.; Zhang, H.; Rinna, A. Glutathione: Overview of Its Protective Roles, Measurement, and Biosynthesis. *Molecular Aspects of Medicine*. NIH Public Access February 2009, pp 1–12. https://doi.org/10.1016/j.mam.2008.08.006.

(15) Kelly, L.; McDonough, M. a; Coleman, M. L.; Ratcliffe, P. J.; Schofield, C. J. Asparagine Beta-Hydroxylation Stabilizes the Ankyrin Repeat Domain Fold. *Molecular bioSystems* **2009**, *5* (1), 52–58. https://doi.org/10.1039/b815271c.

(16) Hardy, A. P.; Prokes, I.; Kelly, L.; Campbell, I. D.; Schofield, C. J. Asparaginyl β-Hydroxylation of Proteins Containing Ankyrin Repeat Domains Influences Their Stability and Function. *Journal of Molecular Biology* **2009**, *392* (4), 994–1006. https://doi.org/10.1016/j.jmb.2009.07.070.

(17) Aslanidis, C.; de Jong, P. J. Ligation-Independent Cloning of PCR Products (LIC-PCR). *Nucleic Acids Research* **1990**, *18* (20), 6069–6074. https://doi.org/10.1093/nar/18.20.6069.

(18) Li, X. T.; Thomason, L. C.; Sawitzke, J. A.; Costantino, N.; Court, D. L. Positive and Negative Selection Using the TetA-SacB Cassette: Recombineering and P1 Transduction in Escherichia Coli. *Nucleic Acids Research* **2013**, *41* (22), e204. https://doi.org/10.1093/nar/gkt1075.

(19) Studier, F. W.; Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *Journal of Molecular Biology* **1986**, *189* (1), 113–130. https://doi.org/10.1016/0022-2836(86)90385-2.

(20) Mierendorf, R. C.; Morris, B. B.; Hammer, B.; Novy, R. E. Expression and Purification of Recombinant Proteins Using the PET System. In *Molecular Diagnosis of Infectious Diseases*; Humana Press, 2003; Vol. 13, pp 257–292. https://doi.org/10.1385/0-89603-485-2:257.

(21) Pinder, J. C.; Smith, K. S.; Pekrun, A.; Gratzer, W. B. Preparation and Properties of Human Red-Cell Ankyrin. *Biochemical Journal* **1989**, *264* (2), 423–428. https://doi.org/10.1042/bj2640423.

(22) Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E. G.; Collier, M. P.; Chandler, S. A.; Tkachenko, O.; Allen, J.; Crispin, M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Eichmann, C.; Selenko, P.; Frey, L.; Riek, R.; Galpin, M. R.; Struwe, W. B.; Benesch, J. L. P.; Kukura, P. Quantitative Mass Imaging of Single Biological Macromolecules. *Science* **2018**, *360* (6387), 423–427. https://doi.org/10.1126/science.aar5839.

(23) Olerinyova, A.; Sonn-Segev, A.; Gault, J.; Eichmann, C.; Schimpf, J.; Kopf, A. H.; Rudden, L. S. P.; Ashkinadze, D.; Bomba, R.; Frey, L.; Greenwald, J.; Degiacomi, M. T.; Steinhilper, R.; Killian, J. A.; Friedrich, T.; Riek, R.; Struwe, W. B.; Kukura, P. Mass Photometry of Membrane Proteins. *Chem* **2021**, *7* (1), 224–236. https://doi.org/10.1016/j.chempr.2020.11.011.

(24) Kafader, J. O.; Melani, R. D.; Schachner, L. F.; Ives, A. N.; Patrie, S. M.; Kelleher, N. L.; Compton, P. D. Native vs Denatured: An in Depth Investigation of Charge State and Isotope Distributions. *Journal of the American Society for Mass Spectrometry* **2020**, *31* (3), 574–581. https://doi.org/10.1021/jasms.9b00040.

(25) Lindsay, C. D.; Green, C.; Bird, M.; Jones, J. T. A.; Riches, J. R.; McKee, K. K.; Sandford, M. S.; Wakefield, D. A.; Timperley, C. M. Potency of Irritation by Benzylidenemalononitriles in

Humans Correlates with TRPA1 Ion Channel Activation. *Royal Society Open Science* **2015**, *2* (1), 140160. https://doi.org/10.1098/rsos.140160.

(26) Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. Spin Echo NMR Spectra without J Modulation. *Chemical Communications* **2012**, *48* (6), 811–813. https://doi.org/10.1039/c1cc16699a.

(27) Tang, H.; Wang, Y.; Nicholson, J. K.; Lindon, J. C. Use of Relaxation-Edited One-Dimensional and Two Dimensional Nuclear Magnetic Resonance Spectroscopy to Improve Detection of Small Metabolites in Blood Plasma. *Analytical Biochemistry* **2004**, *325* (2), 260–272. https://doi.org/10.1016/j.ab.2003.10.033.

(28) Verdier, L.; Gharbi-Benarous, J.; Bertho, G.; Evrard-Todeschi, N.; Mauvais, P.; Girault, J. P. Dissociation-Equilibrium Constant and Bound Conformation for Weak Antibiotic Binding Interaction with Different Bacterial Ribosomes. *Journal of the Chemical Society. Perkin Transactions 2* **2000**, No. 12, 2363–2371. https://doi.org/10.1039/b007666j.

(29) Quevedo, C. E.; Cruz-Migoni, A.; Bery, N.; Miller, A.; Tanaka, T.; Petch, D.; Bataille, C. J. R.; Lee, L. Y. W.; Fallon, P. S.; Tulmin, H.; Ehebauer, M. T.; Fernandez-Fuentes, N.; Russell, A. J.; Carr, S. B.; Phillips, S. E. V.; Rabbitts, T. H. Small Molecule Inhibitors of RAS-Effector Protein Interactions Derived Using an Intracellular Antibody Fragment. *Nature Communications* **2018**, *9* (1), 1–12. https://doi.org/10.1038/s41467-018-05707-2.

(30) Saveliev, S.; Bratz, M.; Zubarev, R.; Szapacs, M.; Budamgunta, H.; Urh, M. Trypsin/Lys-C Protease Mix for Enhanced Protein Mass Spectrometry Analysis. *Nature Methods* **2013**, *10* (11), i–ii. https://doi.org/10.1038/nmeth.f.371.

(31) Suo, Y.; Wang, Z.; Zubcevic, L.; Hsu, A. L.; He, Q.; Borgnia, M. J.; Ji, R. R.; Lee, S. Y. Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. *Neuron* **2020**, *105* (5), 882-894.e5. https://doi.org/10.1016/j.neuron.2019.11.023.

(32) Zhao, J.; Lin King, J. v.; Paulsen, C. E.; Cheng, Y.; Julius, D. Irritant-Evoked Activation and Calcium Modulation of the TRPA1 Receptor. *Nature* **2020**, *585* (7823), 141–145. https://doi.org/10.1038/s41586-020-2480-9.

(33) Paulsen, C. E.; Armache, J. P.; Gao, Y.; Cheng, Y.; Julius, D. Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms. *Nature* **2015**, *520* (7548), 511–517. https://doi.org/10.1038/nature14367.

(34) Goenrich, M.; Bartoschek, S.; Hagemeier, C. H.; Griesinger, C.; Vorholt, J. A. A Glutathione-Dependent Formaldehyde-Activating Enzyme (Gfa) from Paracoccus Denitrificans Detected and Purified via Two-Dimensional Proton Exchange NMR Spectroscopy. *Journal of Biological Chemistry* **2002**, *277* (5), 3069–3072. https://doi.org/10.1074/jbc.C100579200.

(35) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nature Protocols* **2015**, *10* (6), 845–858. https://doi.org/10.1038/nprot.2015.053.

Chapter III:
Towards a Small Molecule Toolkit for TRPA1

1. Previous Work:

The identification of the binding site of a ligand in a protein can be investigated by using a modified version of the ligand which will form a covalent link with residues in the vicinity of the binding site. The protein can then be digested using a protease and the resulting peptides examined for evidence of covalent modification.[1–4] Forming a covalent bond between a ligand and protein can be achieved in numerous ways. The natural nucleophilicity of cysteine residues can be harnessed to form a covalent bond using chemically labile groups. **JT-010**, for example, makes use of a labile chloroacetyl group to bind to TRPA1.[5]

The use of labile electrophilic groups, such as the chloroacetyl group, which are susceptible to nucleophilic attack, introduces a bias towards binding amino acid residues which contain a nucleophilic sidechain such as cysteine, lysine, or serine. Residues that do not have a reactive side chain will not readily react with an electrophilic probe such as **JT-010**. The binding sites for modulators of ion channels are typically within the transmembrane (TM) region, especially fenestrations on the surface of the TM region or lipid binding pockets.[6,7] These pockets are typically made up of residues with aromatic and hydrophobic side chains. In this case, a covalently reacting probe needs to insert into C-H or X-H (where X is N or O) bonds, or react with the π-system of aromatic side chains. C-H/X-H insertion is typically limited to free radicals and carbene/nitrenes.[1] Electrophiles are able to react *via* electrophilic substitution on aromatic rings whilst carbenes and nitrenes react with aromatic rings *via* addition reactions. Carbenes are highly reactive and indiscriminately react with nearby residues and water, and the same is true of nitrenes albeit with slightly lower reactivity. As a result, it is important to design a variant of a ligand which contains a functional group which, upon irradiation with light, reacts to give a carbene or nitrene, allowing the reactive species to be confined to a binding pocket with less exposure to solvent.[1]

Theophylline-based ligands were chosen to design a small molecule toolkit for TRPA1 as **GRC-17536** (a theophylline-like ligand) shows high affinity for TRPA1 and underwent phase IIa clinical trials for painful diabetic neuropathy.[8] Extensive structure-activity relationships (SAR) for the theophylline-based TRPA1 antagonists have been published, including structures similar to **HC-030031** and **GRC-17536**,[9] potentially allowing the rapid and effective design of a small molecule toolkit for working with TRPA1.[9] The binding site of these compounds has not yet been fully characterised. A previous study[10] attempting to determine the binding site of **HC-030031** (a theophylline-based ligand) proposes that the amide group between the lipophilic (4-iso-propylphenyl moiety) and hydrophilic (theophyllinyl moiety) regions of **HC-030031** is of critical importance, facilitating a hydrogen bond between N855 and the carbonyl group (fig. 3-1). The study also suggests that the hydrophilic region is solvent exposed, with the lipophilic half of the molecule occupying a hydrophobic pocket. However, this explanation seems overly simplistic and does not take into account the published SAR of theophylline-based ligands. The work described in this chapter concerned studies aimed at elucidating this binding site and identifying key amino acid residues involved in interacting with the ligand.

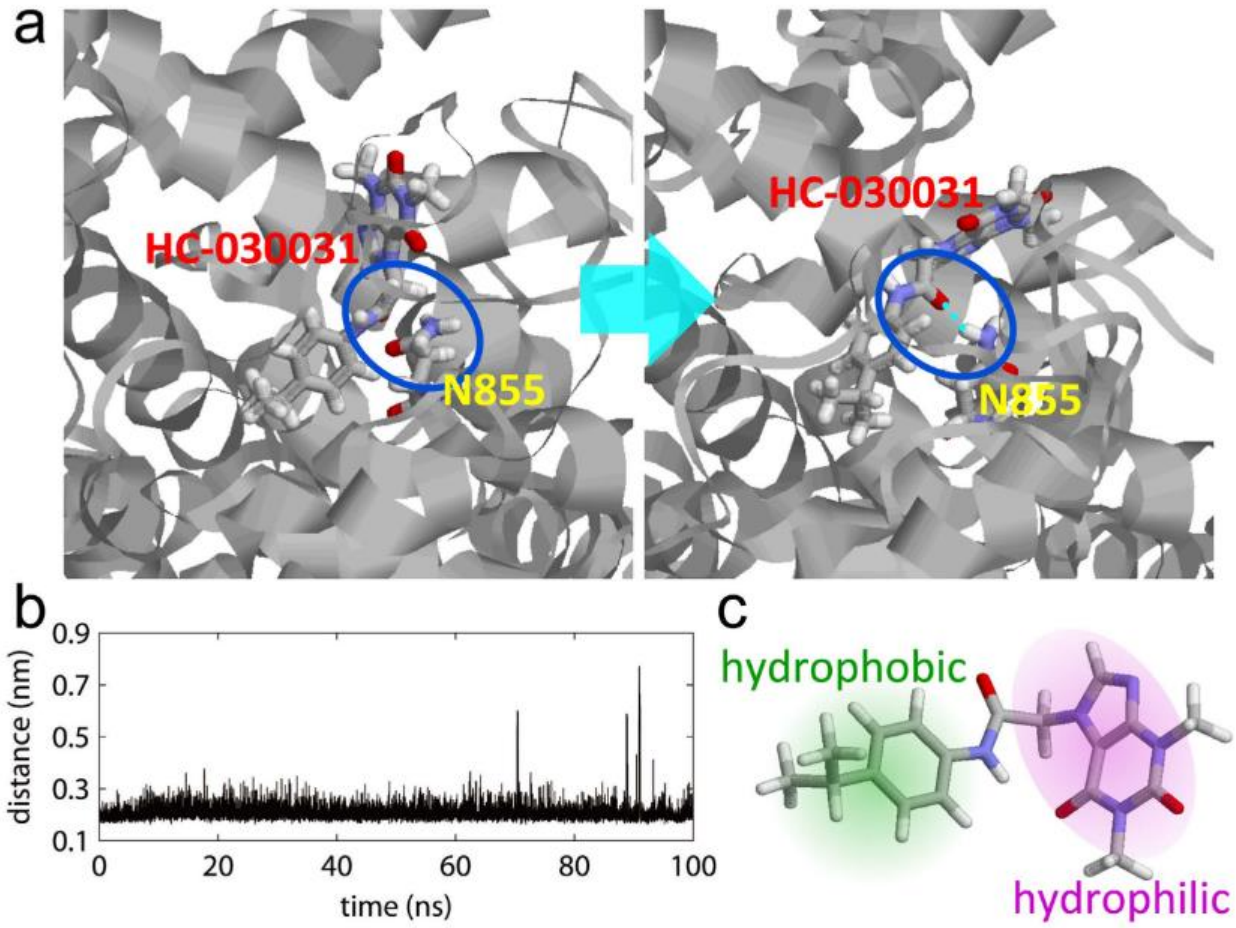


Figure 3-1: Proposed binding site of HC-030031 within TRPA1. (Figure reproduced from Gupta *et al.*[10] under the CC BY 4.0 License): A) Molecular dynamics simulations show the formation of a hydrogen bonding interaction between **HC-030031** and N855. **B)** the distance between O atom of **HC-030031** and the Hδ atom of N855 in the MD simulation remains short, showing the formation of a stable hydrogen bond. **C)** highlighted regions of **HC-030031** important for binding.

The development of small molecules that could capture native TRPA1 while leaving the channel unmodified would be of utility in purifying the protein for further analysis by proteomics to study post-translational modifications[11] or identifying protein-protein interactions made by TRPA1 through pull-down assays. Such probes could also be used to discover "off-target" proteins which bind **GRC-17536**. Fluorescent probes could be useful to study the localisation of TRPA1 within cells or as a stain to determine whether TRPA1 is expressed within cells. In addition to this, the development of a small molecule probe, such as an alkyne tagged probe, which could be modified *via* click chemistry with

commercially available azides to give a bespoke probe suited to a specific application would be of great use to the wider field.[3,4]

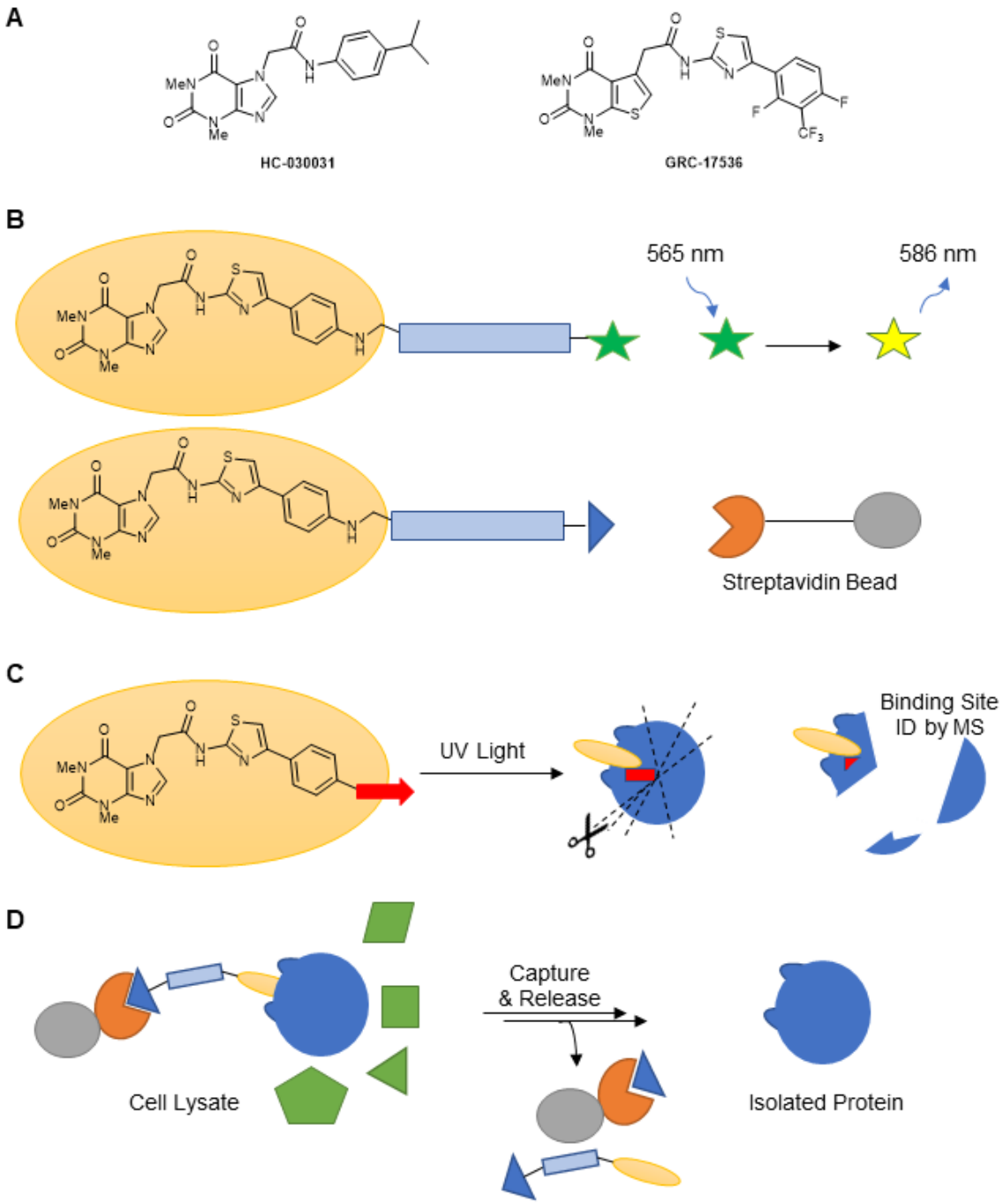


Figure 3-2: Probe design and potential downstream applications. A) Theophylline-based antagonists as scaffolds for the design of probes. **B)** Monofunctional probes incorporating a linker (blue box) between the theophylline warhead and a tag (green star/blue triangle). The tag is able fluoresce or to bind a modified bead (*e.g.* streptavidin beads) **C)** Incorporation of a photolabile moiety (red arrow) allows for the probe to react covalently with a protein of interest under UV light, allowing the binding site to be determined by mass spectrometry. **D)** Proteins interacting with the warhead can be captured from cell lysate, allowing the identification of "off-targets". Protein complexes may also be captured, allowing for the identification of protein-protein interactions.

2. Synthesis of Theophylline-based Ligands

2.1 Tagged Probes

Initial work focused on the construction of an aryl halide intermediate that could undergo a Buchwald reaction with an amine tagged with a florescent moiety or biotin (**29**) to conveniently give an array of functional probes via parallel synthesis (fig. 3-3).

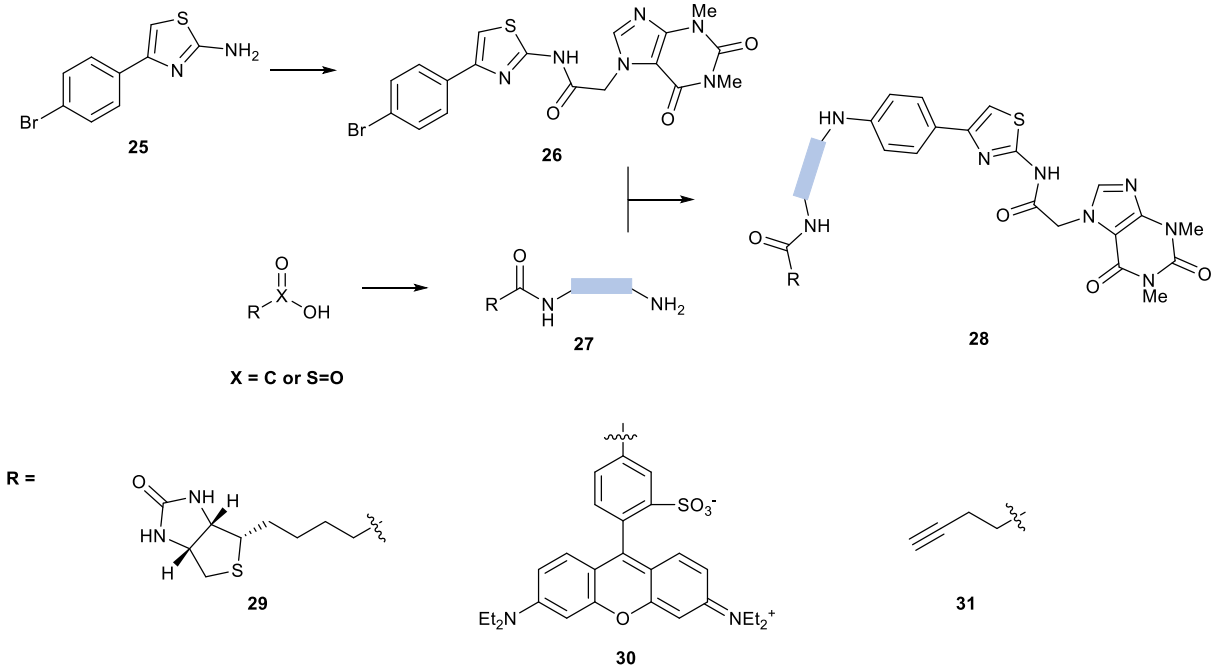


Figure 3-3: Parallel synthesis towards theophylline based functional probes employing the Buchwald reaction as a key step. A variety of tags can be incorporated, for example, biotin (**29**), sulforhodamine B (**30**) or an alkyne (**31**) which can be further derivatised for bespoke applications. The linker (blue rectangle) length and physiochemical properties can easily be modified by using different diamines.

Aryl bromide cores **26** and **32** were constructed from commercially available **25** by amide coupling with theophyline-7-acetyl chloride and 3-methyoxypropionoic acid respectively. *N*-Boc protected **34** was synthesised by the reaction of thiourea with a limiting amount of Boc anhydride, ensuring the addition of only one Boc group to thiourea. This was followed by Hantzsch thiazole condensation with 2,4′-dibromoacetophenone to give **35** (fig. 3-4). **32** and **35** were synthesised to investigate the effects of the theophylline moiety on the coupling reaction as it was envisaged that the electron-rich, heteroaromatic ring may coordinate the Pd^0 catalyst, adversely affecting the efficiency of the reaction.

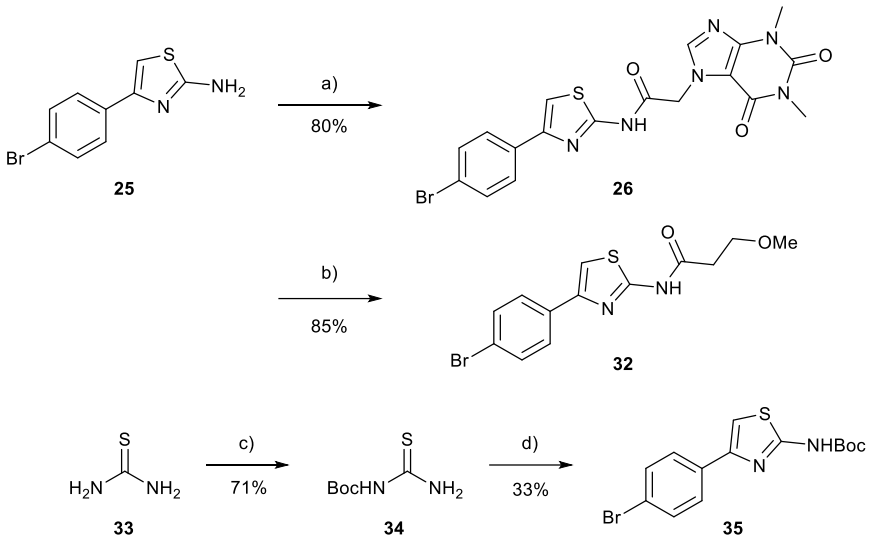


Figure 3-4: Synthesis of aryl bromide cores. Reagents and conditions: a) Theophyline-7-acetyl chloride, pyridine, DMF, rt, 16 h; **b)** 3-methyoxypropionoic acid, HATU, DIPEA, DMF, 100 °C, 96 h; **c)** Boc_2O, NaH, 0 °C to rt, 1 h; **d)** 2,4′-dibromoacetophenone, acetone, reflux, 16 h.

There are few examples of Buchwald-Hartwig reactions of this nature reported in the literature and those that are reported have extremely low yields[12]. For example, as shown in fig. 3-5, **37** is reported to undergo a Buchwald reaction with *N*-methylpiperazine (**38**) and morpholine (**39**) in the presence of the $Pd_2(dba)_3$ / BINAP catalyst system to give **40** and **41** in 5% and 2% yield respectively. Therefore, limited screening of Buchwald reaction and Ulman-type[13] conditions was undertaken to produce the desired products in better yield than those reported previously (fig. 3-6).

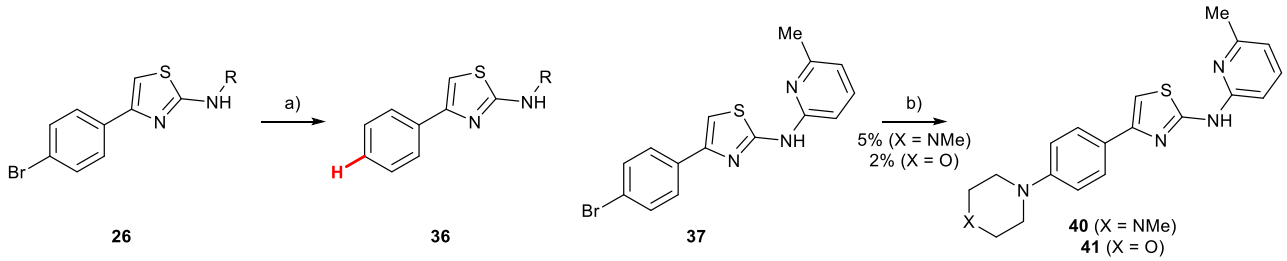
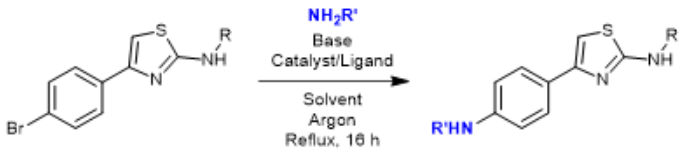


Figure 3-5: Comparison of observed Buchwald reaction outcome with that of a literature outcome. Reagents and conditions: a) KOtBu, $^nBuNH_2$, $Pd_2(dba)_3$, (±)-BINAP, DMF, reflux, 16 h; **b) 38**/**39**, NaOtBu, $Pd_2(dba)_3$, (-)-BINAP, DMF, 140 °C, 1 h, microwave heating;

In some cases, most notably when the reaction was carried out in DMF, the starting material was observed to rapidly undergo dehalogenation to give **36** as a product (fig. 3-5 & fig. 3-6). This reaction confirmed the oxidative insertion of Pd into the C-Br bond and therefore the generation of the Pd^0



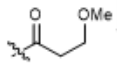
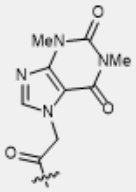
R	Solvent	Base	R'	Catalyst/Ligand	Outcome
	DMSO	K_2CO_3	Me	CuI / Proline.TFA	0% conversion [a]
	THF	KO^{t}Bu	Me	$Pd_2(dba)_3$ / BINAP	No discernible product [a]
H	PhMe	KO^{t}Bu	$(CH_2)_6NH_2$	$Pd_2(dba)_3$ / BINAP	0% conversion [c]
	THF	KO^{t}Bu	Me	$Pd_2(dba)_3$ / BINAP	No discernible product [a]
	DMSO	K2CO3	Me	CuI / Proline.TFA	0% conversion [a]
Boc	Dioxane	Cs_2CO_3	Bu	$Pd_2(dba)_3$ / BINAP	0% conversion [b]
	THF	KO^{t}Bu	Me	$Pd_2(dba)_3$ / BINAP	0% conversion [a]
	THF	KO^{t}Bu	Morpholine	$Pd_2(dba)_3$ / BINAP	0% conversion [a]
	^{t}BuOH	KO^{t}Bu	Me	$Pd_2(dba)_3$ / BINAP	0% conversion [a]
	DMF	KO^{t}Bu	Me	$Pd_2(dba)_3$ / Xphos	Dehalogenated SM [a]
	DMF	KO^{t}Bu	Bu	$Pd_2(dba)_3$ / Xphos	Dehalogenated SM [a,b]
	DMF	KO^{t}Bu	Bu	$Pd(OAc)_2$ / JohnPhos	Dehalogenated SM [a,b]
	Dioxane	Cs_2CO_3	Bu	$Pd(OAc)_2$ / BINAP	0% conversion [c]
	PhMe	KO^{t}Bu	Bu	$Pd_2(dba)_3$/ JohnPhos	0% conversion [a]
	^{t}BuOH	Cs_2CO_3	Bu	BrettPhos Pd G3	0% conversion [c]
	DMA	Cs_2CO_3	Bu	BrettPhos Pd G3	0% conversion [b]
	Dioxane	KO^{t}Bu	Bu	tBu-BrettPhos Pd G3	0% conversion [c]

Figure 3-6: Screening conditions for the Buchwald reaction. Outcome of reaction was judged by [a] proton NMR, [b] LCMS or [c] TLC of the crude reaction mixture. No discernible product is given when the crude NMR was too complex to assign or the starting material (SM) was absent from the acquired spectrum but no peaks corresponding to product could be found.

catalyst in DMF. In other cases, no reaction was observed, including dehalogenation of the starting material, thus the presence of the active catalyst could not be confirmed. With the aim of combatting the dehalogenation side reaction and ensure the presence of the active Pd^0 catalyst, BrettPhos Pd G3

precatalysts were employed. These compounds are palladacycles in complex with a ligand (BrettPhos/^{t}Bu-BrettPhos) which shows high activity in Buchwald-Hartwig reactions.[14] The palladacycle undergoes reductive elimination in the presence of base to give Pd^{0} in complex with the ligand. Unfortunately, this approach also did not give the desired products, thus this synthetic route was abandoned in favour of a stepwise approach.

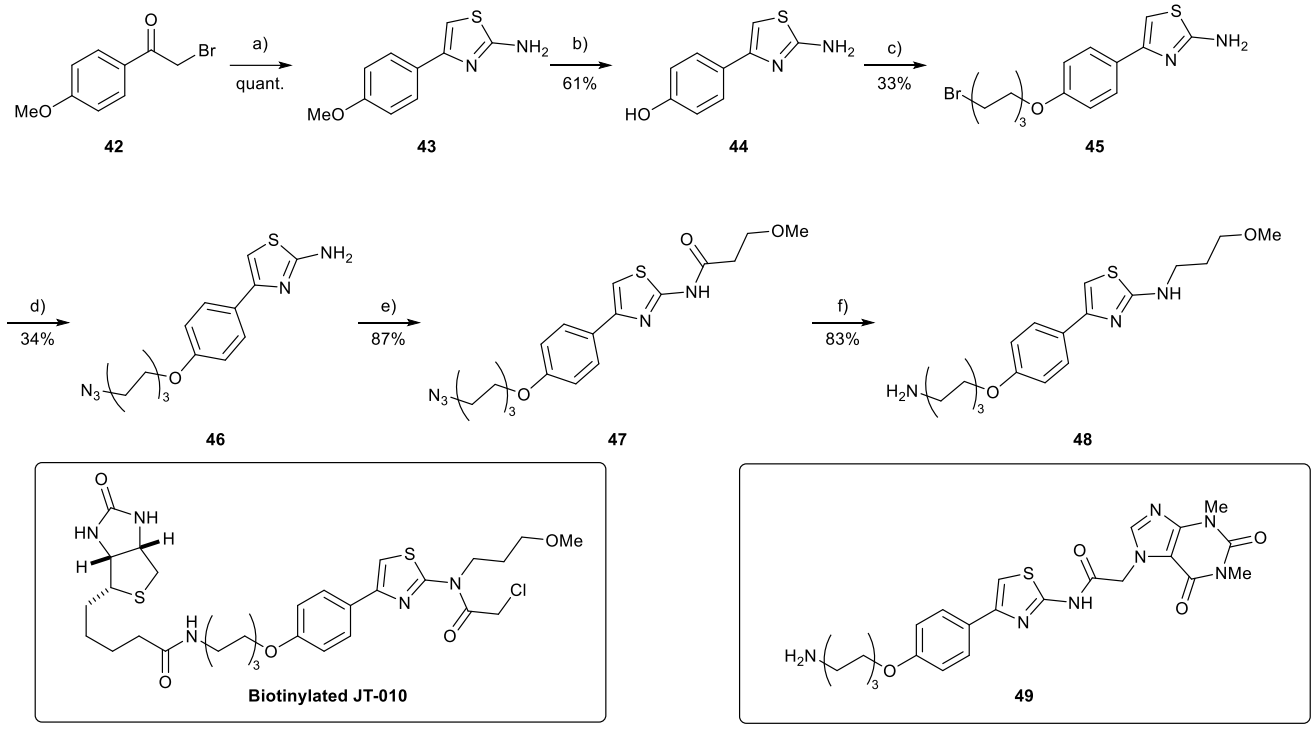


Figure 3-7: Synthesis of intermediates towards biotinylated JT-010. Reagents and conditions: a) thiourea, EtOH, reflux, 2 h; **b)** BBr_3, CH_2Cl_2, 0 °C to rt, 78 h; **c)** 1,6-dibromohexane, K_2CO_3, acetone, reflux, 20 h; **d)** NaN_3, DMSO, rt, 12 h; **e)** 3-methyloxypropionic acid, DIPEA, HCTU, DMF, rt, 14 h; **f)** $LiAlH_4$, THF, 0 °C.

The stepwise route to biotinylated **JT-010**[5] describes the synthesis of **48** which bears a primary amine which could be further derivatised with biotin or fluorescein. Replacing the *N*-aliphatic group on the 2-aminothiazole moiety with theophylline would give **49**, an intermediate suitable for the production of an array of monofunctional non-covalent probes for TRPA1 (fig. 3-7). Condensation of **42** with thiourea gave thiazole **43**. **43** was demethylated using BBr_3 and the resultant phenol **44** was alkylated using 1,6-dibromohexane to give **45**. Azidonation of alkyl bromide **46**, subsequent amide coupling to give **47** and reduction with lithium aluminium hydride gave **48** (fig. 3-7).

Mirroring the synthesis described above, thiazole **43** was formed from the Hantzsch condensation of **42** with thiourea. Subsequent demethylation using BBr_3 gave phenol **44**. Alkylation of **44** with 1,6-dibromohexane gave **45** in excellent regioselectivity, as confirmed by HMBC analysis, presumably as a result of the phenol moiety having a lower pKa than the aminothiazole moiety and thus increased nucleophilicity.[15] At this point, the synthesis of target **53** diverged from the synthesis of the literature compound **49** described above. Gabriel amine synthesis was carried out by *N*-alkylation of potassium phthalamide with **45** to give **50** and subsequent hydrazinolysis gave the primary amine **51** in good yield. Selective Boc protection of the primary alkylamine moiety over the aminothiazole moiety gave **52** and amide coupling with theophylline-7-acetic acid gave **53** (fig. 3-8).

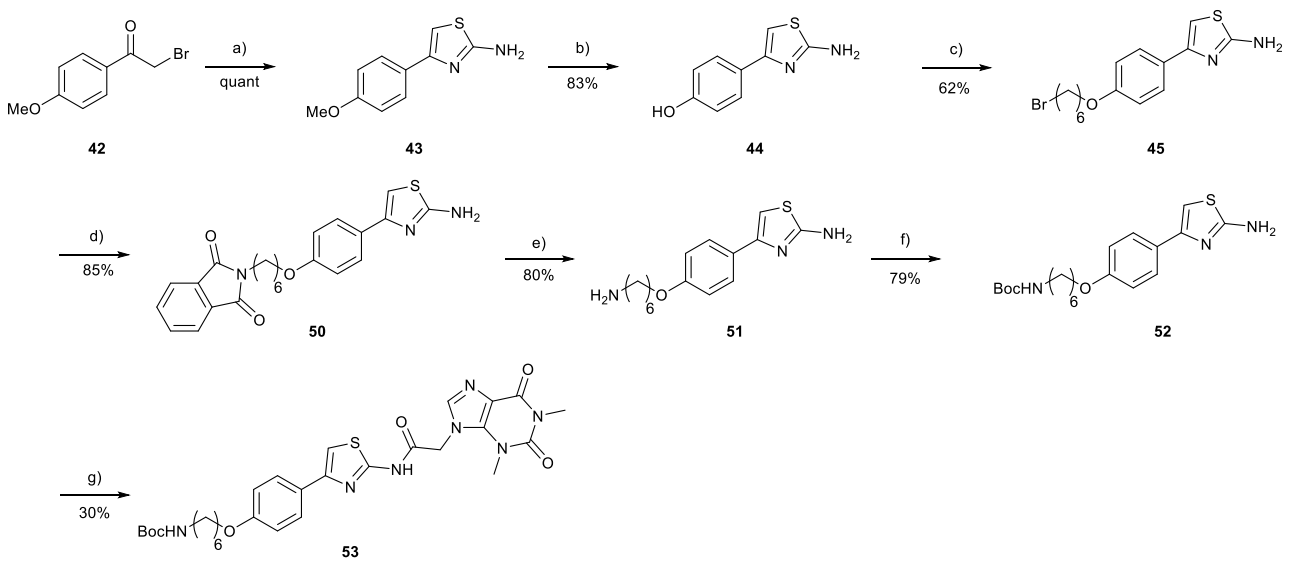


Figure 3-8: Synthesis of key intermediate 53. Reagents and conditions: a) thiourea, EtOH, reflux, 2 h; **b)** BBr_3, CH_2Cl_2, 0 °C to rt, 78 h; **c)** 1,6-dibromohexane, K_2CO_3, acetone, reflux, 20 h; **d)** potassium phthalimide, DMF, 70 °C, 8 h; **e)** hydrazine monohydrate, CH_2Cl_2, MeOH, rt, 16 h; **f)** Boc_2O, MeOH, rt, 1.5 h; **g)** Theophylline-7-acetyl chloride, pyridine, DMF, rt, 16 h.

Sulforhodamine B acid chloride (**54**) was prepared by the reaction of **30** with thionyl chloride. Subsequent condensation of crude acid chloride **54** with methyl 4-aminobutyrate gave **55** in low yield on account of having to chromatographically separate the stereoisomers of **55**. Hydrolysis of **55** gave carboxylic acid **56** in quantitative yield. Deprotection of **52** using CF_3CO_2H in CH_2Cl_2 and subsequent

amide coupling of the crude mixture containing **49** with carboxylic acids derived from sulforhodamine B (**56**) or biotin (**29**) gave **BGSP-001** and **BGSP-002**, respectively (fig. 3-9). Alkylation of phenol **44** with allyl methanesulfonate (**58**) gave **59** and subsequent amide coupling with theopyhline-7-acetic acid gave **BGSP-003** in low yield (fig. 3-10).

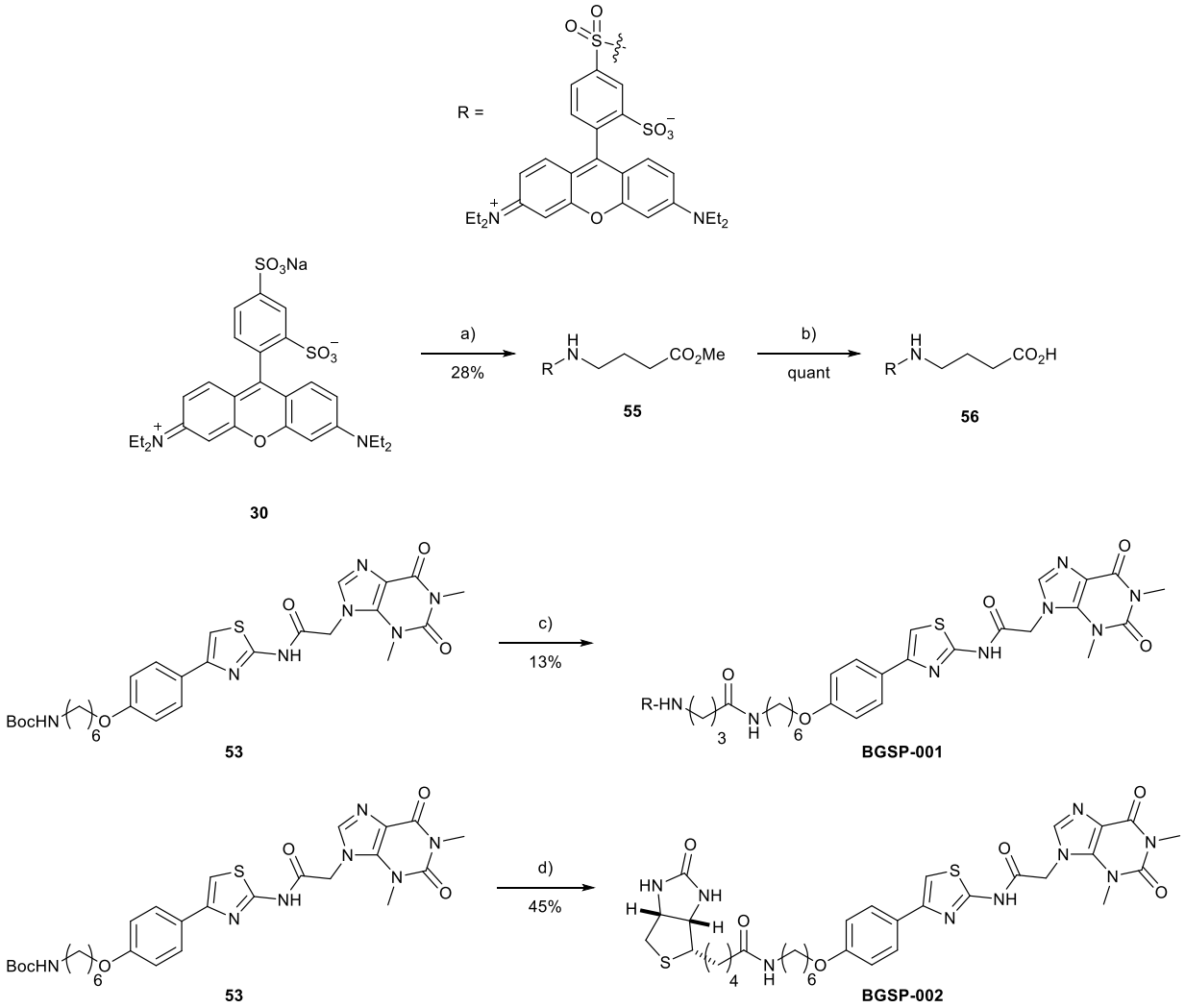


Figure 3-9: Synthesis of BGSP-001 and BGSP-002. Reagents and conditions: a) methyl 4-aminobutyrate hydrochloride, DMAP, DIPEA, 0 °C to rt, 16 h; **b)** NaOH, THF/MeOH, rt, 16 h; **c)** TFA, $CHCl_3$, rt, 1 h then **56**, Oxyma-pure, dicyclohexylcarbodiimide, DMF, rt, 16 h; **d)** TFA, $CHCl_3$, rt, 1 h then biotin-N-hydroxysuccinimide ester, DMF, Et_3N, rt, 16 h.

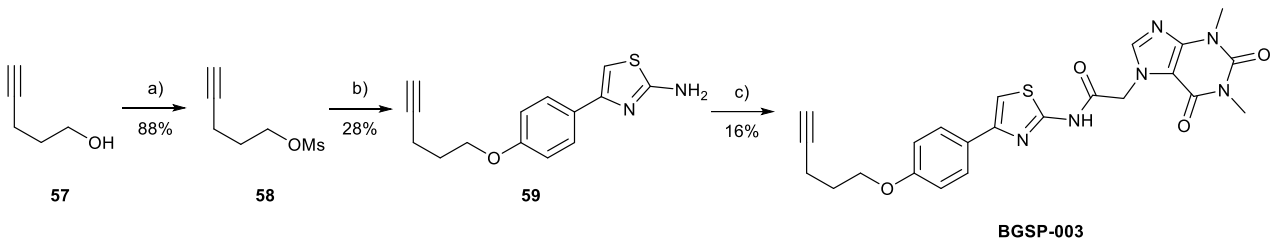


Figure 3-10: Synthesis of BGSP-003. Reagents and conditions: a) MsCl, Et_3N, CH_2Cl_2, 0 °C to rt, 3 h; **b) 44**, K_2CO_3, DMF, 40 °C, 88 h; **c)** theophylline-7-acetyl chloride, pyridine, DMF, rt, 16 h.

2.2 Photoactivatable Ligands

Aryl azides were chosen in part as a synthetic target as it was envisaged that they could be produced *via* the Miyaura borylation of an aryl halide with subsequent Chan-Lam azidonation.[16] The azide group is able to decompose in the presence of ultraviolet (UV) light to give a nitrene species which can react covalently with a variety of amino acid residues within a protein.[1] These proposed photoactivatable aryl azides could be used to investigate the binding site of **HC-030031** and **GRC-17536** after irradiating the TRPA1-ligand complex with UV light, enzymatic digestion and identification of TRPA1 derived peptides that underwent reaction with the nitrene species. To investigate the effects of introducing an azide to the parent structure on the binding affinity of the ligand, the synthesis of a series of aryl azides was attempted (fig. 3-11).

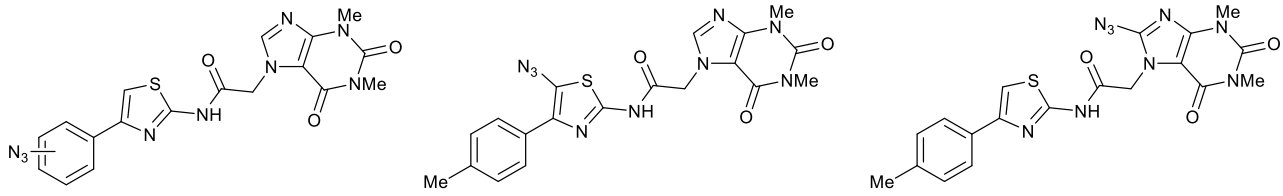


Figure 3-11: Aryl azide targets for photoaffinity labelling of TRPA1. Targets include (left to right): (*ortho*-, *meta*- and *para*-) mono-substituted phenyl azides, an azido thiazole and a theophyllinyl azide.

In the first instance, the synthesis of phenyl azides was attempted. It was found that Miyaura borylation followed by Chan-Lam coupling conveniently gave the *meta*- and *para*- substituted phenyl-azides **61**. Unfortunately, *ortho*-substituted **25** did not undergo Miyuara borylation and instead underwent dehalogenation. In the case of **63**, direct azidonation was carried on bromoaniline (**62**) to

give azidoaniline (**63**) *via* an Ulman reaction. Amide coupling of the resulting azide containing amines and theophyline-7-acetyl chloride followed by HPLC purification gave **3-60**, **3-61** and **4-33** in low yield (fig. 3-12).

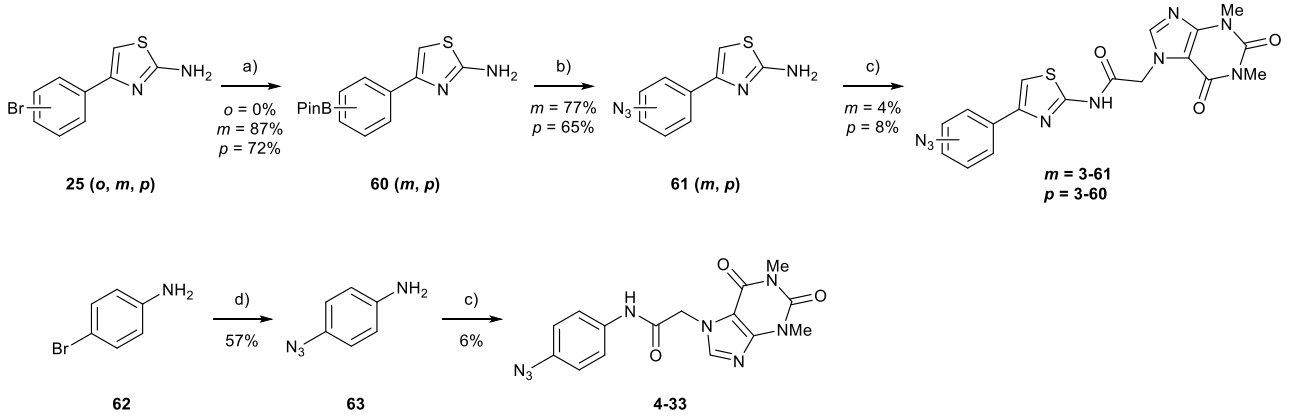


Figure 3-12: Synthesis of substituted phenyl-azides. Reagents and conditions: a) B_2Pin_2, $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$, KOAc, 1,4-dioxane, 100 °C, 16 h; **b)** NaN_3, $Cu(OAc)_2$, MeOH, 50 °C, 5 – 58 h; **c)** theophyline-7-acetic acid, $SOCl_2$, DMF, rt, 1 h then **61/63**, pyridine, DMF, rt, 16 h; **d)** NaN_3, sodium ascorbate, DIPEA, CuI, $EtOH/H_2O$, rt, 15 min then 90 °C, 40 min.

In order to introduce a photolabile azide group on the theophylline moiety, theophylline (**64**) was brominated under acidic conditions then alkylated to give ester **66**. Azidonation *via* an S_NAr reaction proceeded efficiently to give **67**, which was the hydrolysed to give acid **68**, which could be used for amide couplings. Coupling of **68** with **70** or isopropylaniline (**71**) gave rise to **3-64** and **4-35** respectively. Coupling of **70** and **71** with theophylline-7-acetic acid gave **3-55** and **HC-030031** for use as non-photolabile controls (fig. 3-13).

To investigate the hypothesised interaction between N855 and theophylline-based antagonists, the syntheses of dialkyl-diazirine **73** and 5-azidothiazole **76** were attempted. Retrosynthesis of both products identified halogenated thiazoles (**72** and **74**) in the 2- and 5- positions as key intermediates (fig. 3-14). It was found that the regioselectivity of halogenation could be controlled using copper halide salts in different oxidation states as reported (fig. 3-14).[17] A Sandmeyer reaction with CuCl gave the 2-chloro species (**72**) and $CuBr_2$ gave the 5-bromo species (**74**) selectively. Attempts to use Chan-Lam mediated azidonation were unsuccessful on account of **74** rapidly undergoing

dehalogenation when Miyaura borylation was attempted. Interestingly, no 5-azidothiazole compounds are reported in the literature, suggesting this class of molecule is difficult to synthesise, thus this ring system was abandoned as a target.

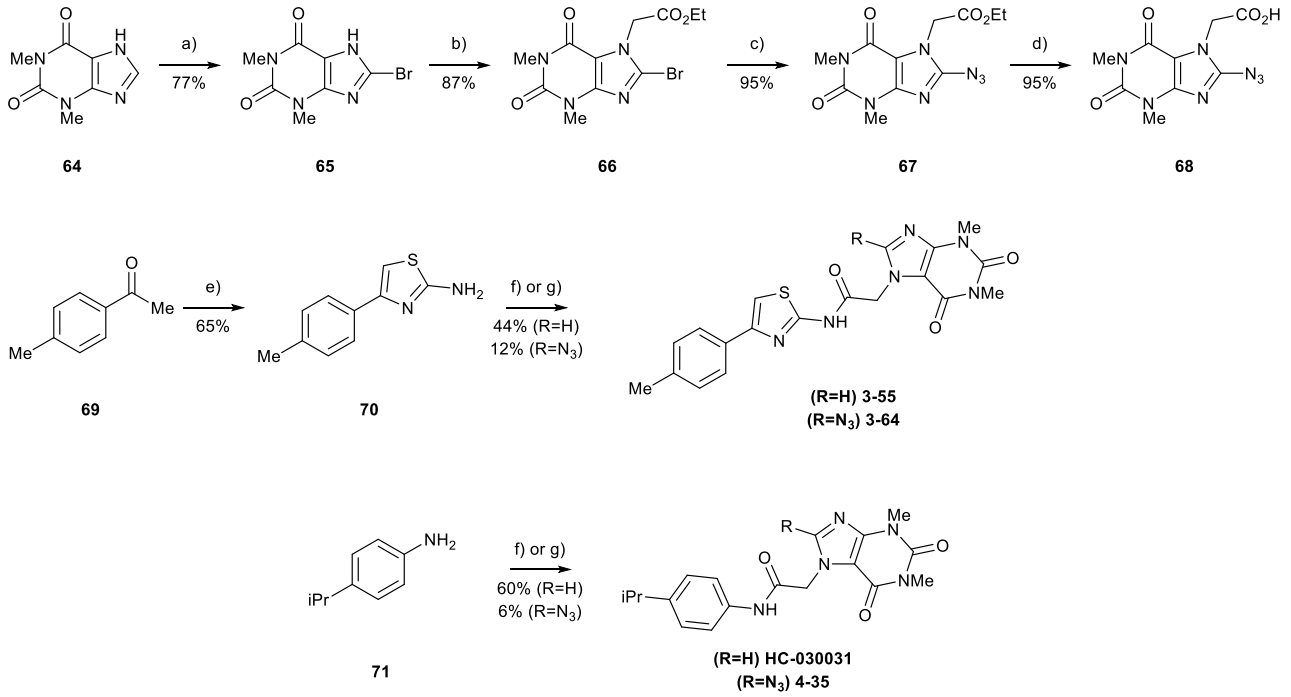


Figure 3-13: Synthesis of substituted theophyllinyl-azides. Reagents and conditions: a) Br_2, $AcOH/H_2O$, 50 °C, 4 h; **b)** K_2CO_3, ethyl bromoacetate, DMF, 70 °C, 4 h; **c)** NaN_3, DMF/H_2O, 50 °C, 2 h; **d)** NaOH, THF/MeOH, rt, 30 min; **e)** thiourea, I_2, EtOH, reflux, 24 h; **f)** theophyline-7-acetic acid, $SOCl_2$, DMF, rt, 1 h then **70**/**71**, pyridine, DMF, rt, 16 h; **g) 68**, $SOCl_2$, DMF, rt, 1 h then **70**/**71**, pyridine, DMF, rt, 16 h;

Synthesis of the corresponding diazirine at the 5-position was attempted by preparation of the 1,3-diketone **77** followed by Hantzsch thiazole synthesis to give **78**. However, the yield of **78** was very low and a second thiazole (**70**) was detected in the reaction mixture by LCMS. The reason for this was not immediately apparent and in the interest of time this reaction was abandoned. Re-examining this reaction, it is possible that **70** is the result of the **77** undergoing a haloform like reaction due to the presence of water and decarboxylation to give **69** which could undergo cyclisation with thiourea to give thiazole **70**. The trifluoromethyl anion acting as a leaving group has been observed previously in the functional group interconversion of an aryl-imine bearing the CF_3 group to a nitrile.[18] Further

evidence for this reaction mechanism is the presence of ketone **69** in the reaction mixture despite the diketone **77** starting material being analytically pure (fig. 3-14).

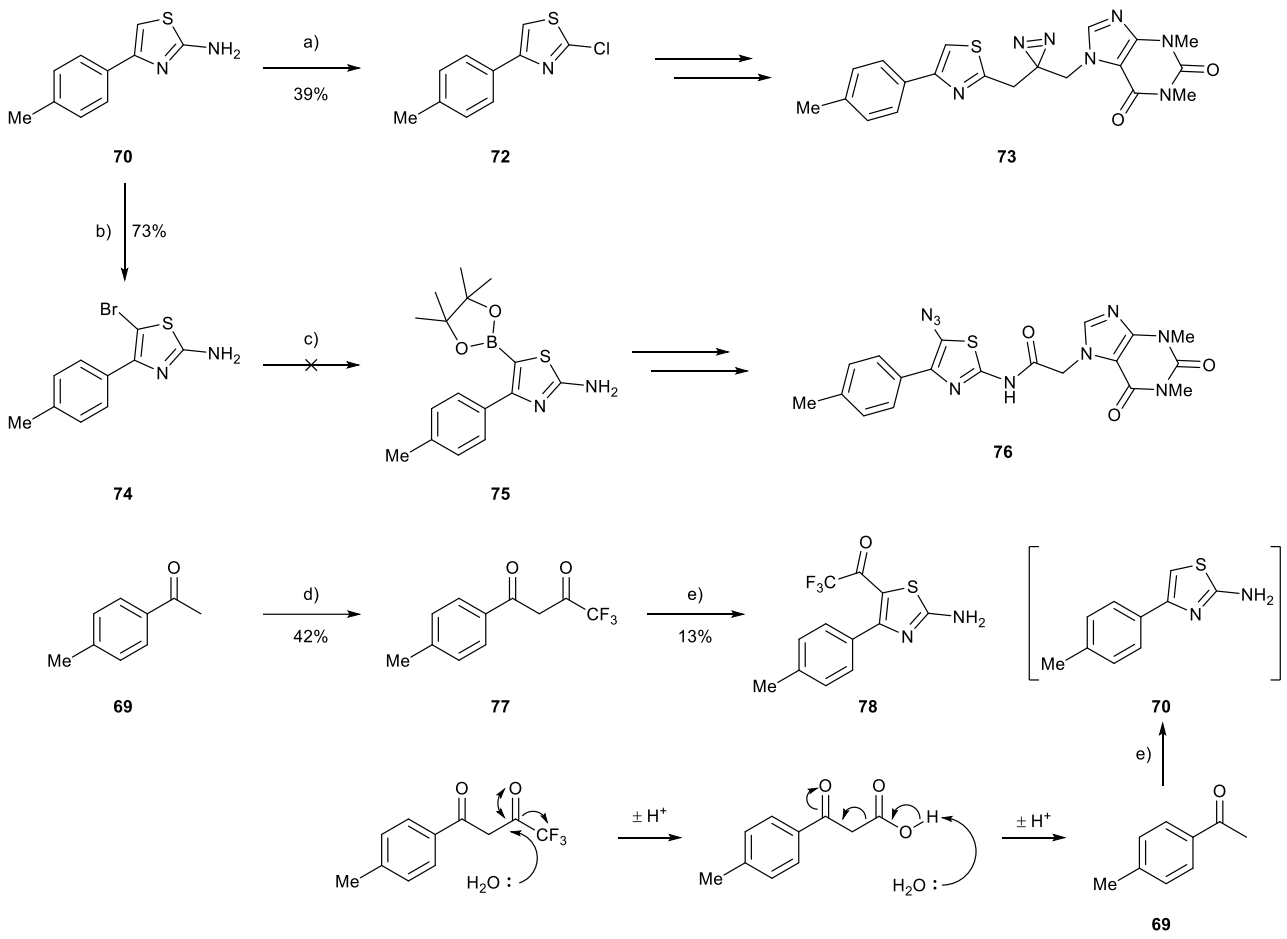


Figure 3-14: Synthesis of intermediates towards substituted thiazole targeting N855. Reagents and conditions: a) ^{t}BuONO, CuCl, MeCN, rt, 4 h; **b)** $CuBr_2$, MeCN, 0 °C, 48h; **c)** B_2Pin_2, KOAc, $Pd(dppf)Cl_2.CH_2Cl_2$, dioxane, 100 °C, 1 h, microwave; **d)** NaH, CF_3CO_2Me, THF, 0 °C, 10 min then **69**, 0 °C to reflux, 2 h; **e)** thiourea, I_2, EtOH, reflux, 16 h; Square brackets show a compound which was not isolated but the presence of which was inferred by LCMS and ^{1}H-NMR analysis.

Synthesis of the dialkyl-diazirine **73** was attempted via ketone **79**. It was proposed that **72** would undergo a facile S_NAr reaction with an enolate derived from **80** to give ketone **79** which could then be converted to the target diazirine **73**. Unfortunately, it was found that **80** instead underwent self-condensation to give **83**, a product derived from an aldol reaction and not giving rise to the desired product. An alternate route to ketone **79** was attempted making use of a Heck reaction with **81**, an alkene derived from the reaction of allyl bromide and theophylline. It was proposed that the resultant alkene **82** could be converted to ketone **79** *via* hydroboration and subsequent oxidation. This route,

however, was also unsuccessful on account of the thiazole undergoing dehalogenation instead of Heck coupling (fig. 3-15).

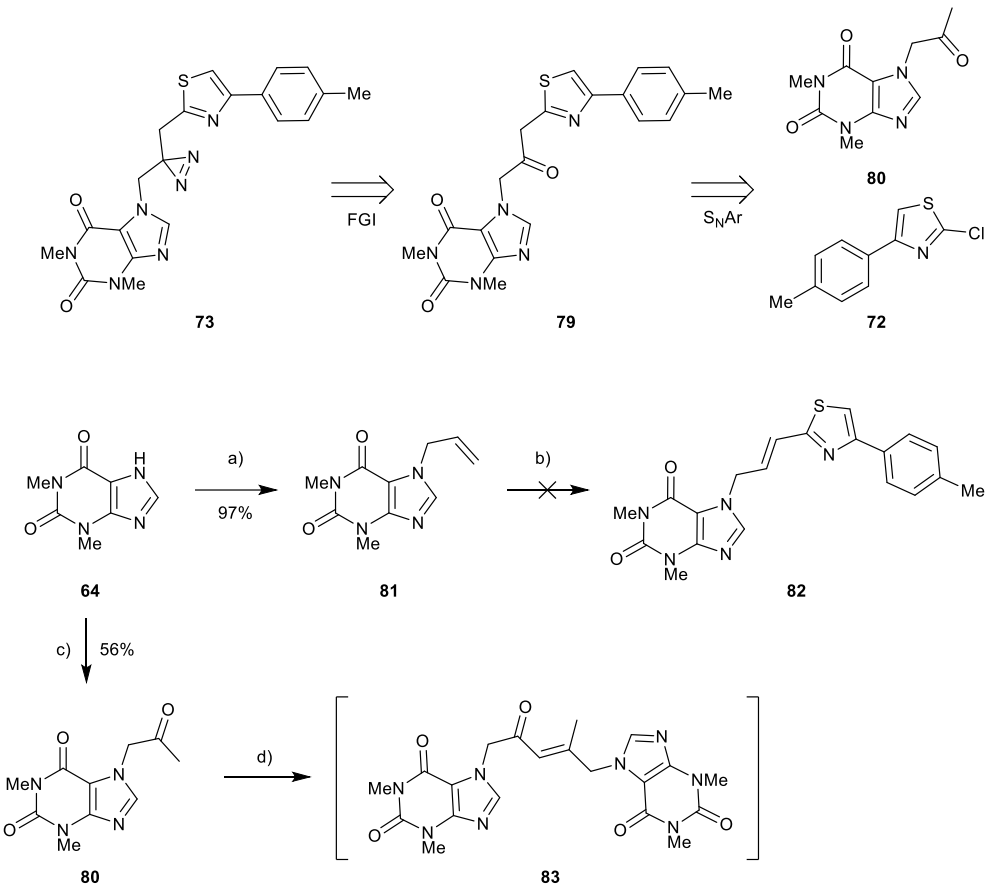


Figure 3-15: Synthesis of intermediates towards diazirine targeting N855. Reagents and conditions: a) allyl bromide, K_2CO_3, DMF, 35 °C, 4 h; **b) 72**, Et_3N, $Pd(OAc)_2$, DMF, 130 °C, 19 h, microwave; **c)** K_2CO_3, chloroacetone, DMF, 120 °C, 1 h; **d)** LiHMDS, **72**, PhMe, 0 °C, 2 h then rt, 16 h; Square brackets show a compound which was not isolated but the presence of which was inferred by LCMS or ^{1}H-NMR analysis.

3. Medium-throughput LCMS Kinetic Solubility Assay

The solubility of the ligands synthesised thus far was evaluated to ensure compatibility with the constraints placed on DMSO concentration for a cell-based assay. Upon mixing 50 mM stocks of the ligands in DMSO with 20 mM HEPES buffer pH 7.5, instant precipitation of the ligands was observed. The solubility of theophylline-based ligands like **GRC-17536** are known to be poor, hence they require formulation as a salt or as a prodrug to improve their solubility.[8] Thus, a method to screen formulation conditions and quantify the concentration of a ligand was required to ensure compatibility of the ligands with the assay.

The Beer-Lambert law, $A = \varepsilon Cl$ (where: A is absorbance, ε is molar attenuation coefficient, C is the concentration and l is the pathlength), shows that the absorbance of light at a particular wavelength is proportional to the concentration of a ligand in solution for a given compound (constant ε) and instrumentation (constant l). Thus concentration-UV-response relationships can be generated and the limiting concentration of a ligand in aqueous solution can be determined.[19] This approach would allow for the estimation of the solubility given that there may be a slight deviation from linearity at high ligand concentration due to changes in refractive index and increasing electrostatic interaction between ligand and solvent molecules.[19] Additionally, it was envisaged that different solvents or additives used in the solubility study may absorb at the same wavelength as the compound of interest, complicating the analysis. In order to counteract this, LCMS was used on account of a ligand eluting from the C18 column in a fixed solvent mixture of H_2O and MeCN. Both of MeCN and H_2O have poor absorptivity at 220-400 nm, the wavelengths typically used to detect organic molecules.[20] Given LCMS/HPLC systems are typically equipped with robotic sample handling, variable wavelength UV detectors (VWD) and automated data analysis, they are an attractive option for determining compound solubility.[21,22]

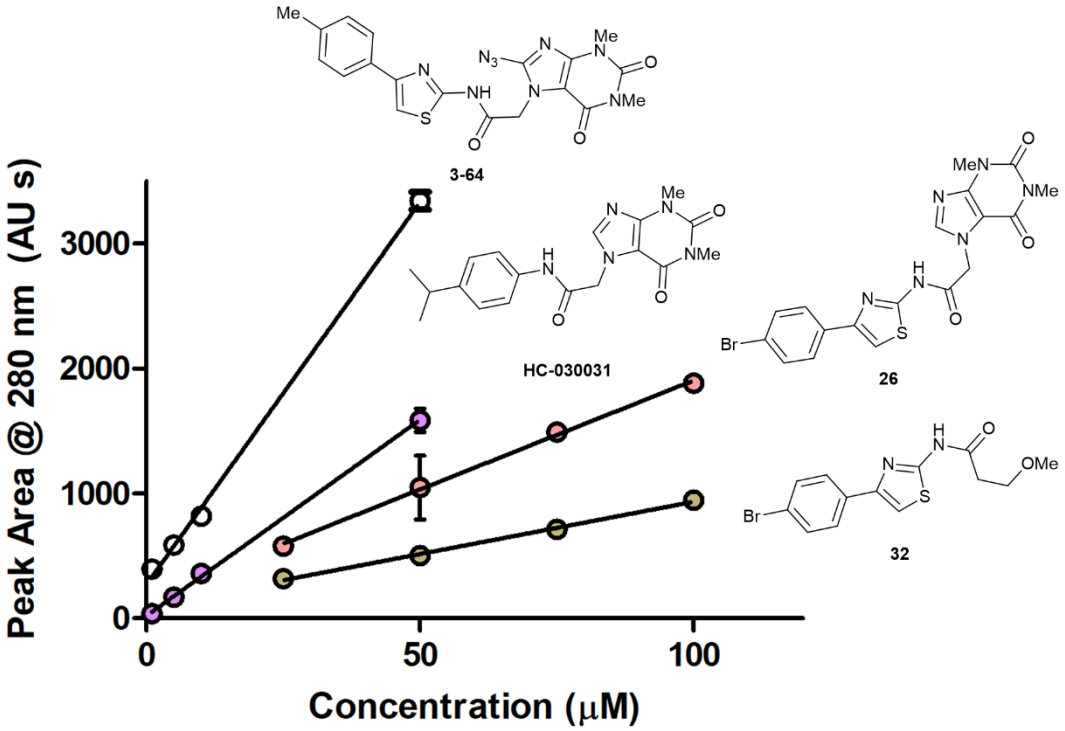
The kinetic solubility of compounds (*i.e.* the maximum concentration of the fastest precipitating form of a compound)[23] was determined instead of the thermodynamic solubility (*i.e.* the limiting concentration of a ligand in solution) given the ligand binding assay can be carried out rapidly, allowing a super-saturated solution of a compound to be permissible in the assay. The method used to determine the kinetic solubility of the theophylline-based ligands was based on previously published work[24] with modifications on account of the different equipment used. The principal difference in the method used was the inclusion of a calibration curve for each compound in methanol due to potential differences in absorptivity for compounds on a VWD compared to the evaporative light scattering detector (ELSD) used in the study. ELSDs produce acceptable measurements of the

absolute amount of analyte present with a generalised calibration curve produced from a series of diverse compounds, this is not the case for a VWD.[22]

A short 3.5 min eluent gradient was established on the LCMS to ensure any compounds used to dissolve the ligands that absorb at 280 nm were separated from the ligand of interest. After integration of the LCMS trace, it was indeed found that there is a large variation in absorptivity for the various compounds that were tested (fig. 3-16 a). Using the calibration curves, an equation describing the relationship between the area of the peak at 280 nm and the concentration of each ligand was determined.

In order to investigate the kinetic solubility of each ligand 2 µL of a 50 mM solution of a ligand in DMSO was added to 198 µL Hank's Balanced Salt Solution supplemented with 20 mM HEPES pH 7.4 to give a final concentration of 1% DMSO (v/v) and a maximum theoretical concentration of the ligand of 500 µM. The mixture was vortexed to ensure proper mixing and then centrifuged to remove any particulate matter. It was found that **HC-030031** was by the most soluble compound of those tested in the buffer. This contrasts with thiazole containing **26** and **3-64** which were not detected in the assay buffer (fig. 3-16 b).

A



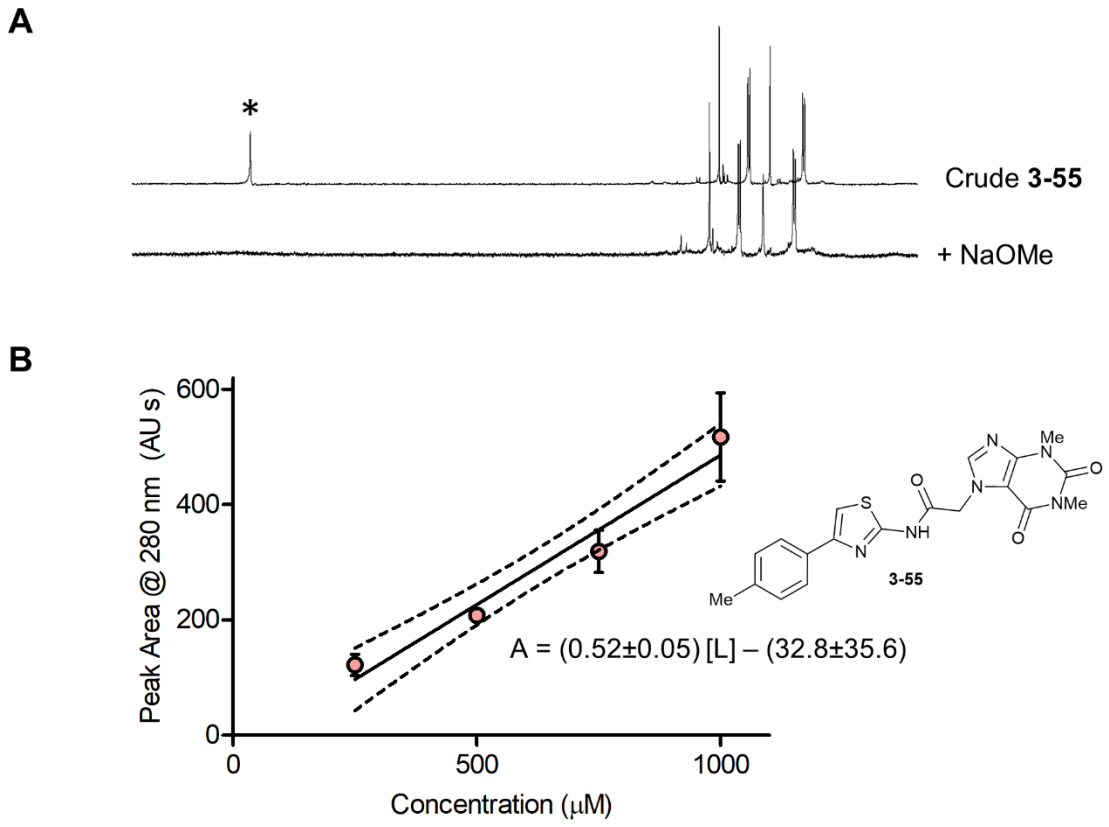
B

Compound	Equation	Peak Area*	Solubility (μM)*
32	A = 8.37[L] + 97.3	107.1	1.16
26	A = 17.4[L] + 161	N.D	N/A
HC-030031	A = 31.4[L] + 20.9	2420	76.5
3-64	A = 31.2[L] + 271	N.D	N/A

Figure 3-16: Determining the kinetic solubility of TRPA1 antagonists in assay buffer. A) Calibration plots in the linear region of the Beer-Lambert law for a range of theophylline-based ligands showing large variability in absorbivity at 280 nm. Error bars represent SD (n=3). **B)** determining the kinetic solubility based on the linear fit of the calibration plot where A is the area under the peak and [L] is the concentration of the ligand. * = assumption that linear region of Beer-Lambert law extends to the calculated value for solubility. N.D = Not Detected. N/A = Not Applicable.

Problems pertaining to the solubility of theophylline-based ligands are well known.[8] A typical approach to improving the solubility of the ligand is to form a salt by reacting the ligand with an acid or base. The pKa of the amide bond in the thiazole containing ligands is predicted to be low, approximately 7.7 in the case of **3-55**,[25] presumably due to the stabilisation of the negative charge on the amide nitrogen by π-conjugation. Therefore, it was thought that the solubility of the ligands could

be improved by formation of a sodium salt of the ligands by reaction of the ligands with NaOMe in MeOH.[21] The deprotonation of **3-55** proceeded smoothly, giving rise to the sodium amide species as judged by the loss of the amide proton in the proton NMR of the reaction mixture. In an analogous fashion to the solubility assay described above, it was found that the solubility of **3-55** was improved and exceeded 500 µM (fig. 3-17).



Peak Area	Solubility (µM)	Limits (µM)*
295.5	631	774 - 513

Figure 3-17: Determining the solubility of the sodium salt of 3-55. A) ^{1}H-NMR of **3-55** with NaOMe. * = amide NH proton. **B)** calibration graph for **3-55** with 95% confidence interval shown with black dashed line. * = upper and lower limit of solubility calculated using uncertainty in gradient and y-intercept. Error bars represent SD (n=3).

It was subsequently found that the sodium salt of **3-55** is highly hygroscopic, giving a solid that has poor solubility after extended storage at -20 °C in a vial filled with air (*i.e.* not flushed with an inert, dry gas). This is likely the result of the salt re-protonating in the presence of atmospheric moisture. In order to get around this issue, stocks of **3-55** were kept in DMSO and added to two equivalents of

aqueous NaOH (*i.e.* to give 500 µM ligand in 1 mM NaOH with 1% DMSO v/v) before use in assays. This was deemed acceptable for use in the cell-based assay due to the small volume this mixture being added which would not affect the pH of the assay due to the assay buffer containing a 50:50 mixture of Eagle's minimal essential medium (EMEM) and Hank's balanced salts solution supplemented with 20 mM HEPES. This meant that the small amount of NaOH added would be buffered out by a mixture of the HEPES, phosphate and bicarbonate equilibria, leading to a minute change in pH.

4. Ligand Binding to Human TRPA1

4.1 Verification of Human TRPA1 Expression in Commercially Available Cell line

Human TRPA1 (hTRPA1) is activated by a variety of ligands with different potencies (see Chapter I §2.2).[26,27] Some of the most potent covalently binding agonists of hTRPA1 are the tear gasses **CR** and **CS**.[28] Other commonly used agonists are cinnamaldehyde and allyl isothiocyanate (**AITC**).[29] Binding efficacies can be determined using a Ca^{2+} fluorescence assay. This assay uses a calcium binding dye (*e.g.* Fluo-4) which undergoes a large increase in fluorescence intensity (FI) upon binding Ca^{2+} (see Chapter I §3.1).[30] Cell lines expressing TRPA1 can be incubated with the acetoxymethyl ester of the dye, allowing the dye to cross the cell membrane (which the unmasked, charged dye could not do), which is then cleaved by esterases to release the dye in the cytosol.[30] This allows for the relative concentration of Ca^{2+} in the cytosol to be measured and thus the activity of TRPA1 can be assayed due to TRPA1 being a Ca^{2+} permeable ion channel.

A typical experiment for collecting a dose-response curve of a certain compound interacting with a TRPA1 involves seeding cells stably expressing TRPA1 at a fixed concentration in a microplate and allowing the cells to grow to a desired count. The medium is aspirated and the cells are incubated with a Ca^{2+} sensitive dye of choice (*e.g.* Fluo-4 AM), allowing for the uptake of the dye and release of the cleaved dye into the cytosol[28]. The microplate is inserted into a plate reader and a background measurement of the FI is taken. The compound of interest is then added and the change in FI over

time is measured. A value for the cellular response can be calculated by dividing the area under the curve of the second portion of the FI measurement by the background FI measurement. Dose-response curves can be generated for particular agonists, allowing for the calculation of an EC_{50} value (the concentration at which 50% of the maximum effect is observed).[28] Using this approach, positive and negative controls are not needed as it is assumed the maximum and minimum response correspond to 100 and 0% of the cellular response.[28] The experiments were carried out using a fluorometric imaging plate reader (FLIPR, Molecular Devices) which transfers a set amount of agonist from a microwell plate to the microplate containing cells and the dye in the dark, preventing any photobleaching or excitation of the dye by light.

A dose-response curve for cinnamaldehyde was constructed to verify the activity of the hTRPA1 PhotoScreen® cell line (Perkin Elmer), a commercially available human embryonic kidney (HEK293) cell line stably expressing hTRPA1. A sigmoidal dose-response curve (fig. 3-18) was observed with a calculated EC_{50} value for cinnamaldehyde of 61.8 – 86.5 μM (95% confidence interval, n = 4). The previously reported EC_{50} value for cinnamaldehyde with hTRPA1 is 61 ± 9 μM,[27] hence the data obtained are in good agreement with the literature.

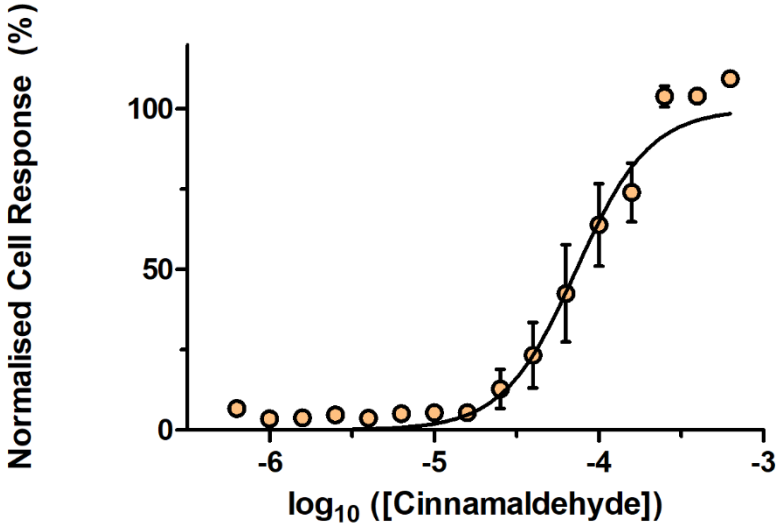


Figure 3-18: Dose-response curve for the hTRPA1 PhotoScreen® cell line in the presence of differing concentrations of cinnamaldehyde. Error bars represent SEM of 4 independent experiments with each value measured in quadruplicate per experiment.

4.2 Determination of Efficacy of the Theophylline Based TRPA1 Antagonists

In order to investigate the half-maximal inhibitory concentration (IC_{50}) values of the theophylline-based ligands, hTRPA1 PhotoScreen® cells were incubated with differing concentrations of the antagonist before cinnamaldehyde (final concentration 250 µM) was added and the change in FI was recorded. The maximum change in FI was normalised to 0% inhibition, while 100% inhibition was taken to be no change from the baseline FI after addition of cinnamaldehyde.

4.2.1 Monofunctional Probes: BGSP- Series

Disappointingly, the **BGSP-** series of monofunctional probes showed evidence of poor binding to hTRPA1. The IC_{50} values of **BGSP-001** and **BGSP-002** could not be determined accurately because they show limited inhibition at 0.25 mM and having a limiting solubility of approximately 0.50 mM. Interestingly, **BGSP-003** exhibited only partial inhibition of cinnamaldehyde evoked Ca^{2+} influx, inhibiting 40% of the baseline response with an IC_{50} value of 16.4 – 27.8 µM (95% confidence interval, n = 3) suggesting the probe does bind to hTRPA1 (fig. 3-19).

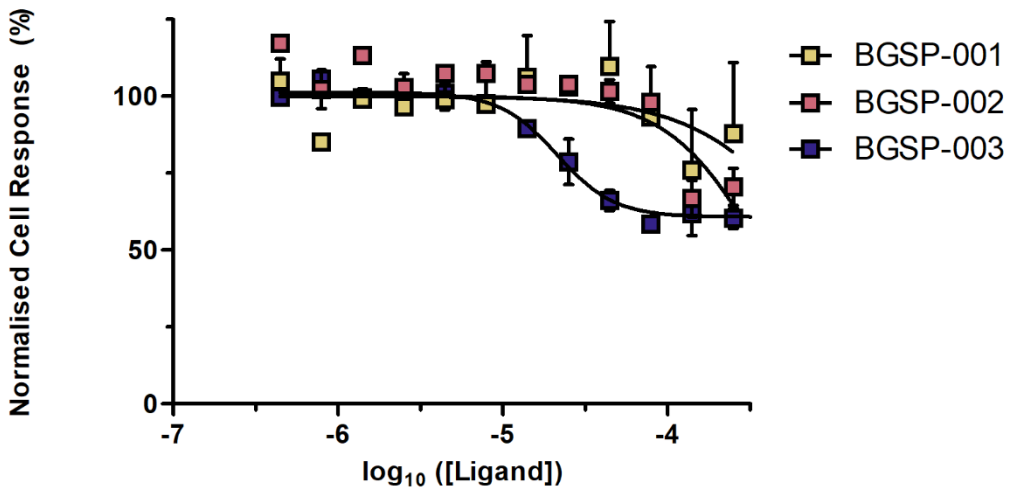


Figure 3-19: Dose-response curves for the hTRPA1 PhotoScreen® cell line in the presence of the monofunctional BGSP- series of compounds on Ca^{2+} influx evoked by 250 µM cinnamaldehyde. Error bars represent SEM of 3 independent experiments. **NB.** Lower portion of error bars for **BGSP-001** and **BGSP-002** are omitted for clarity.

The reason for the partial inhibition of hTRPA1 by **BGSP-003** is not immediately obvious. Assuming that **BGSP-003** binds in a pocket near N855 (as previously suggested for **HC-030031**),[10] the highly hydrophobic nature of pentynyl moiety could mimic the fatty acid chain of the phospholipid 1-

palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (**POPC**). **POPC** occupies the binding site surrounding N855 in the recently published[31] cryo-electron microscopy (cryo-EM) structures of hTRPA1 with **JT-010** bound and the hTRPA1 C621S mutant which is unable to bind electrophiles. A structure with benzyl isothiocyanate (**BITC**) in which **POPC** has been displaced from the binding site surrounding N855 was also reported in the same study. **BITC** is a reversible covalent agonist, unlike **JT-010**, so it is claimed the **BITC**-hTRPA1 structure more closely mimics the open, or a pre-open,[31] state of hTRPA1 while the **JT-010**-hTRPA1 structure is of desensitised hTRPA1. Another recently published structure[32] of hTRPA1 that had reacted with iodoacetamide (**IAA**-hTRPA1), with hTRPA1 presumably in the open state, did not resolve any lipids bound to hTRPA1.

The **POPC** binding site is much more restricted in the **IAA-hTRPA1** structure, with the S5 and pre-S1 helices moving towards one another (fig. 3-20 a). The S4 helix also moves upwards, further compressing the **POPC** binding site (fig. 3-20 b). N855 forms a key interaction with **POPC** in the **JT-010**-hTRPA1 and C621S-TRPA1 structures. However, in the **IAA**-hTRPA1 structure, N855 rotates away from the binding cavity, further ablating **POPC** binding. The **BITC**-hTRPA1 structure has a similarly compressed binding site in which the interfacial helix (IFH) moves downward into the pocket, preventing **POPC** from binding. The highly compressed binding site suggests **POPC** would be displaced from **IAA**-hTRPA1 structure. Given that **POPC** is readily displaced from hTRPA1 upon activation, and that **BGSP-003** could behave in a similar manner to **POPC** given its lipophilic tail, it is possible that activation of hTRPA1 displaces **BGSP-003**. This would allow the channel to transition from a closed to open state, even at high concentrations of **BGSP-003**, meaning **BGSP-003** does not fully inhibit the cinnamaldehyde evoked response.

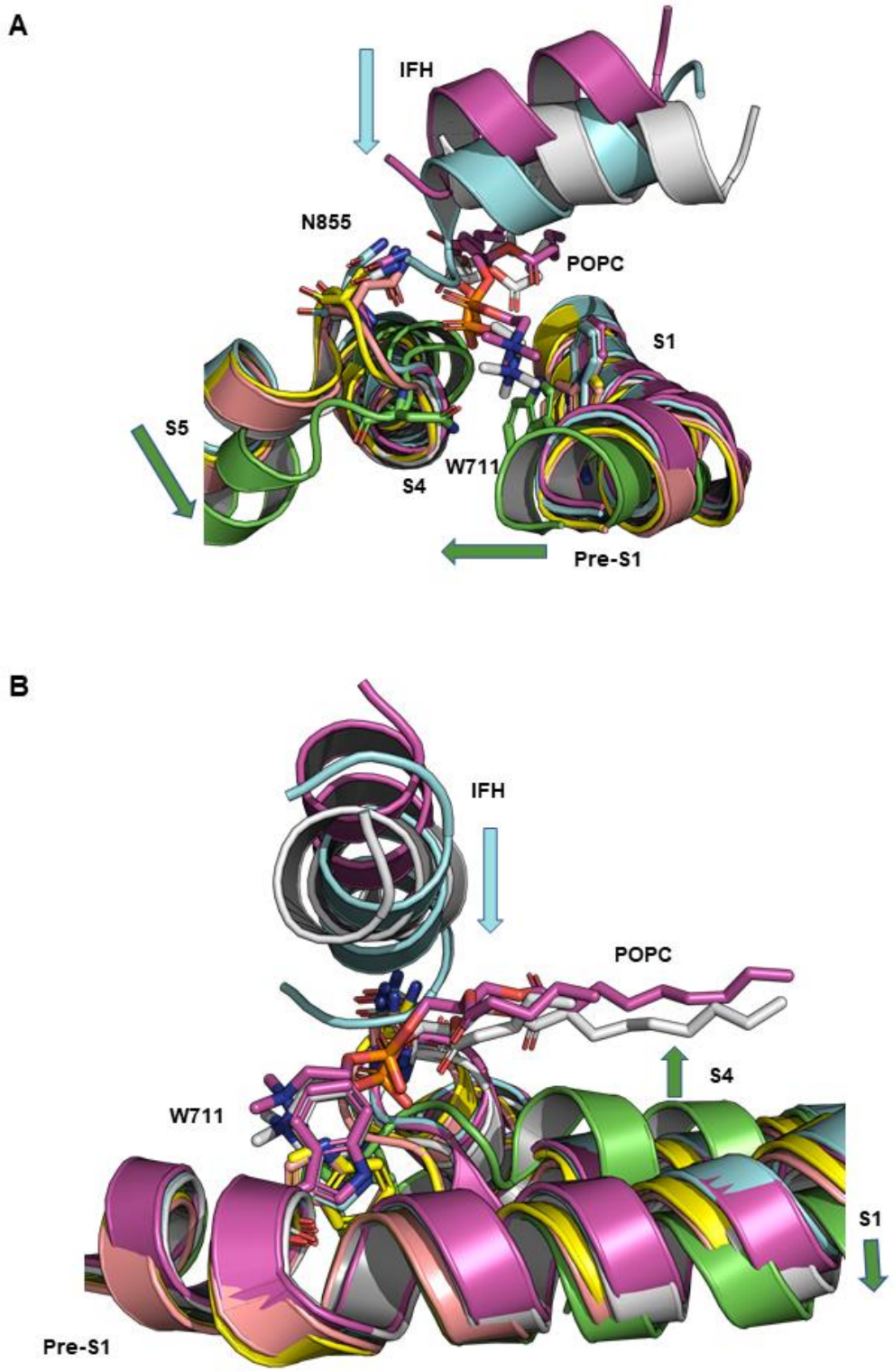


Figure 3-20: Views from the cryo-EM structure of hTRPA1 showing the POPC binding site in various states. PDB accession codes are given in brackets. **Deactivated state: JT-010** (purple, 6PQO), **BIA** (yellow, 6V9V). **Closed state:** C621S mutant (grey, 6PQQ), no agonist (salmon, 6V9W). **Open state: IAA** (green, 6V9X). **Open/pre-open state: BITC** (blue, 6PQP).

The functional use of **BGSP-003** to create bespoke probes with commercially available tagged azides *via* a click reaction will likely be limited by the fact that further derivatisation of the alkyl chain, like in **BGSP-001** and **BGSP-002**, ablates hTRPA1 binding. Thus, further characterisation of the binding site of the theophylline-based probes was sought using the photoactivatable ligands. In this manner, it would be possible to determine what modifications could be made to improve the binding of the **BGSP-** series of compounds.

4.2.2 Photoactivatable Ligands

The measurement of the inhibition of cinnamaldehyde evoked TRPA1 Ca^{2+} influx via FI was carried out as described in the previous section. The assay was verified by the use of a known inhibitor, **HC-030031**, which was found to have an IC_{50} value for hTRPA1 of 11.3 – 13.2 µM (95% confidence interval, n=4, fig. 3-21). This value is slightly higher than the reported IC_{50} value of 4.9 ± 0.1 µM.[33] **HC-030031** also has an IC_{50} of 6.2 ± 0.2 µM, 5.3 ± 0.2 µM and 5.6 µM for **AITC**,[34] formaldehyde,[34] and **CS** tear gas[28] evoked hTRPA1 responses respectively. The reason for the discrepancy is not immediately clear and the IC_{50} values given below are unadjusted, not taking into account the differing IC_{50} values for **HC-030031** reported in this thesis and the literature. All IC_{50} values given below are the 95% confidence interval of 4 independent experiments, with inhibition data at a given concentration collected in quadruplicate per experiment (fig. 3-21).

Of the compounds evaluated (fig. 3-21 and fig. 3-22), **3-60** had the lowest IC_{50} value of 2.17 – 3.09 µM, showing significantly improved binding compared to **HC-030031**. This was closely followed by **3-64** which had an IC_{50} value of 3.61 – 4.59 µM. The only other compound to show improved binding compared to **HC-030031** was **4-35** (**HC-030031** with an 8-azido group) with an IC_{50} value of 8.76 – 11.0 µM. All other compounds showed worse inhibition compared to **HC-030031**. Interestingly, **3-55** has a similar IC_{50} value to **BGSP-003** (4-methyl vs 4-oxypentynyl side chains) of 19.6 – 32.2 µM and 16.4 – 27.8 µM respectively.

The IC_{50} data show that substitution of the theophylline ring at the 8- position with an azide was well tolerated and increased binding affinity when comparing **HC-030031** and **4-33**, and **3-55** and **3-64**. When comparing **3-55**, **4-35** and **HC-030031**, it appears that there is a preference for substrates that have two conjoined hydrophobic moieties (*i.e.* isopropyl group on phenyl ring or *p*-tolyl group on a thiazole ring) over a polar group attached to a hydrophobic moiety (*i.e.* azide group attached to phenyl ring). Comparing **3-60**, **3-61** and **3-55**, it appears that substitution at the 4- position of the phenyl ring is preferred to that at the 3- position and that there is a preference for the polar azide in the 4- position over the more hydrophobic methyl group.

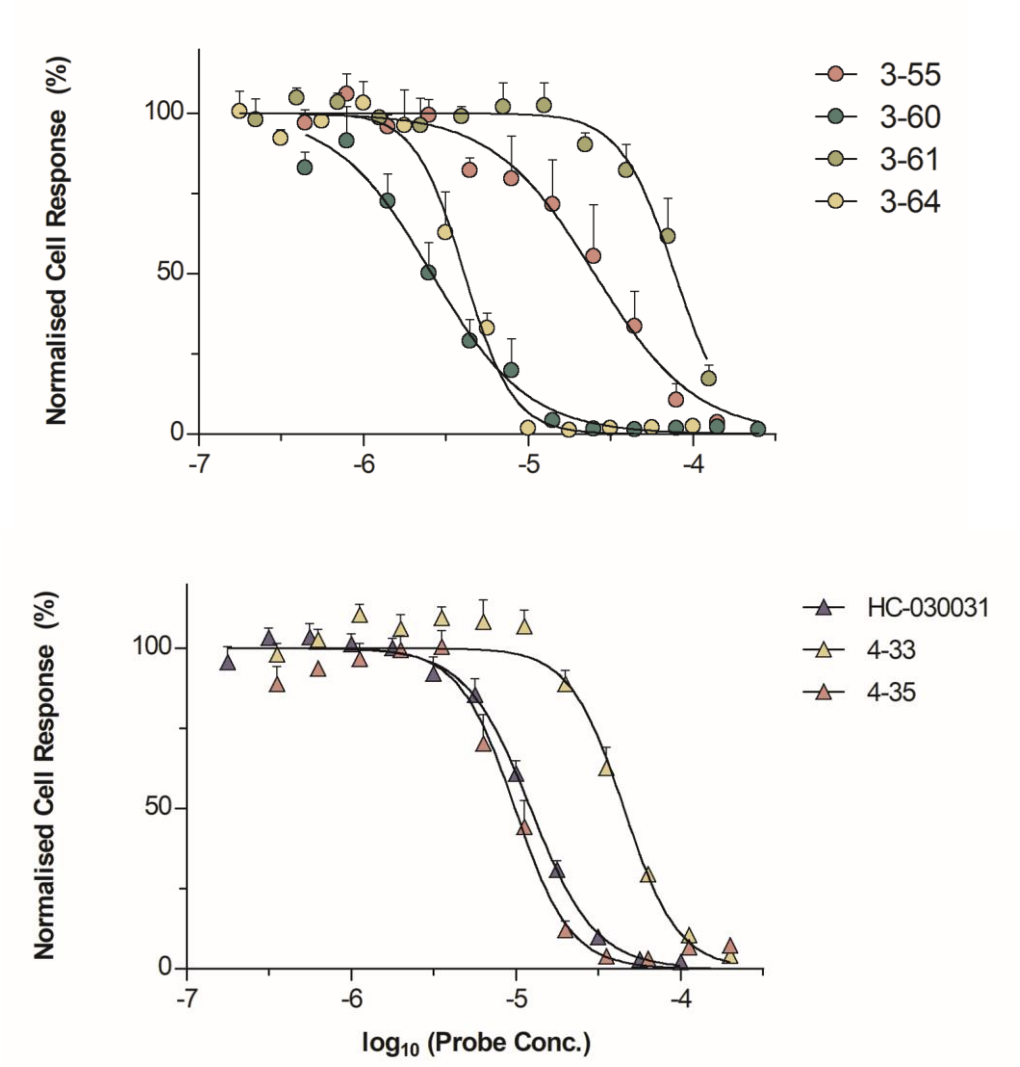
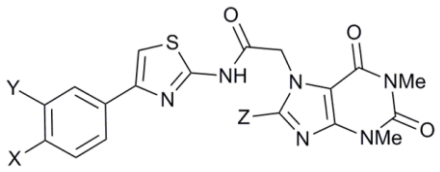
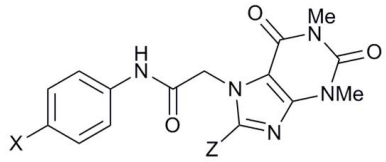


Figure 3-21: Effect of increasing concentration of photolabile probes on Ca^{2+} influx as measured by the change in fluorescence intensity on addition of cinnamaldehyde. Error bars represent SEM of 4 separate experiments.



Compound	X	Y	Z	pIC_{50}	IC_{50} (μM)†
3-55	Me	H	H	4.60± 0.05	19.6 – 32.2
3-60	N_3	H	H	5.59 ± 0.04	2.17 – 3.09
3-61	H	N_3	H	4.11 ± 0.03	67.1 – 88.5
3-64	Me	H	N_3	5.39 ± 0.03	3.61 – 4.59



Compound	X	Z	pIC_{50}	IC_{50} (μM)†
HC-030031	^{i}Pr	H	4.91 ± 0.02	11.3 – 13.2
4-33	N_3	H	4.35 ± 0.02	40.4 – 50.0
4-35	H	N_3	5.01 ± 0.03	8.76 – 11.0

Fig 3-22: Binding efficiencies of a series of theophylline derived probes. † IC_{50} given as a 95% confidence interval.

5. Cryo-Electron Microscopy (by Dr. Mariana Grieben and Dr. Ashley Pike, SGC)

In order to study the binding of **3-60** to hTRPA1 structurally, cryo-electron microscopy (cryo-EM) studies were undertaken by Dr. Mariana Grieben (MG) and Dr. Ashley Pike (AP) from the Prof. Carpenter group at the Structural Genomics Consortium (SGC), University of Oxford. The work in this section was carried out by MG unless otherwise mentioned. Both MG and AP contributed to the refining and model fitting of the Cryo-EM structure of hTRPA1 in complex with **3-60**. The figures, commentary and analysis in this section are my own.

DNA encoding hTRPA1 4-1119 R58T with was cloned into the baculovirus transfer vector pHTBV1.1-LIC to give the transfer vector TRPA1-pHTBV1.1-LIC. Transformation of DH10Bac *E. coli* with TRPA1-pHTBV1.1-LIC gave rise to baculoviral DNA which could be used to infect Sf9

insect cells. DH10Bac *E. coli* contain a host bacmid which undergoes recombination with the transfer vector and helper plasmid which encodes for proteins that carry out the DNA transposition. Baculoviral DNA is generated the transposition of DNA from the TRPA1- pHTBV1.1-LIC donor plasmid between the two Tn7 transposons into the bacmid containing the Tn7 insertion site.[35] Virus amplification was carried out by transfection of Sf9 insect cells with the bacmid using Insect GeneJuice transfection reagent (Merck) followed by incubation for 65 h at 27 °C, removing cells by centrifugation and using the virus containing supernatant to repeat the procedure three times, omitting transfection reagents, to yield a high titre of baculovirus.

Initial work on producing protein from Sf9 cells gave TRPA1 with mixed glycosylation states, as indicated by SDS-PAGE analysis, which was unsuitable for use in cryo-EM, but could be used to study ligand binding by mass spectrometry or other techniques. Using human Expi293F cells it was found that TRPA1 could be produced in an apparently uniform glycosylation state, as judged by SDS-PAGE analysis. Detergent exchange to remove the detergent lauryl maltose neopentyl glycol (LMNG) and introduce the amphipol PMAL-C8 was carried out to aid the acquisition of a cryo-electron micrograph. Free detergents such as LMNG lower the contrast of the electron micrograph and also affect the formation of the thin layer of ice which is required for grid preparation, both of which have a negative outcome on the quality of the resulting micrograph and its analysis.[36] After size exclusion chromatography, TRPA1 was incubated with 100 μM **3-60** before being added to a grid and plunge frozen. Cryo-EM data were acquired at the Leicester Institute of Structural and Chemical Biology.

Refining the electron density map and building a model resulted in 2.64 Å structure of tetrameric TRPA1 with 4 lipid molecules, 1 Ca^{2+} ion, 1 N-acetylglucosamine, 1 N,N'-diacetylchitobiose (a dimer of β(1,4) linked N-acetylglucosamine) and 1 molecule of **3-60** located in a lipid binding pocket between W711 and N855 per TRPA1 monomer. The electron density observed for **3-60** enabled its

modelling in two different conformations, giving new insight into the binding mode of theophylline-based TRPA1 antagonists (fig. 3-23).

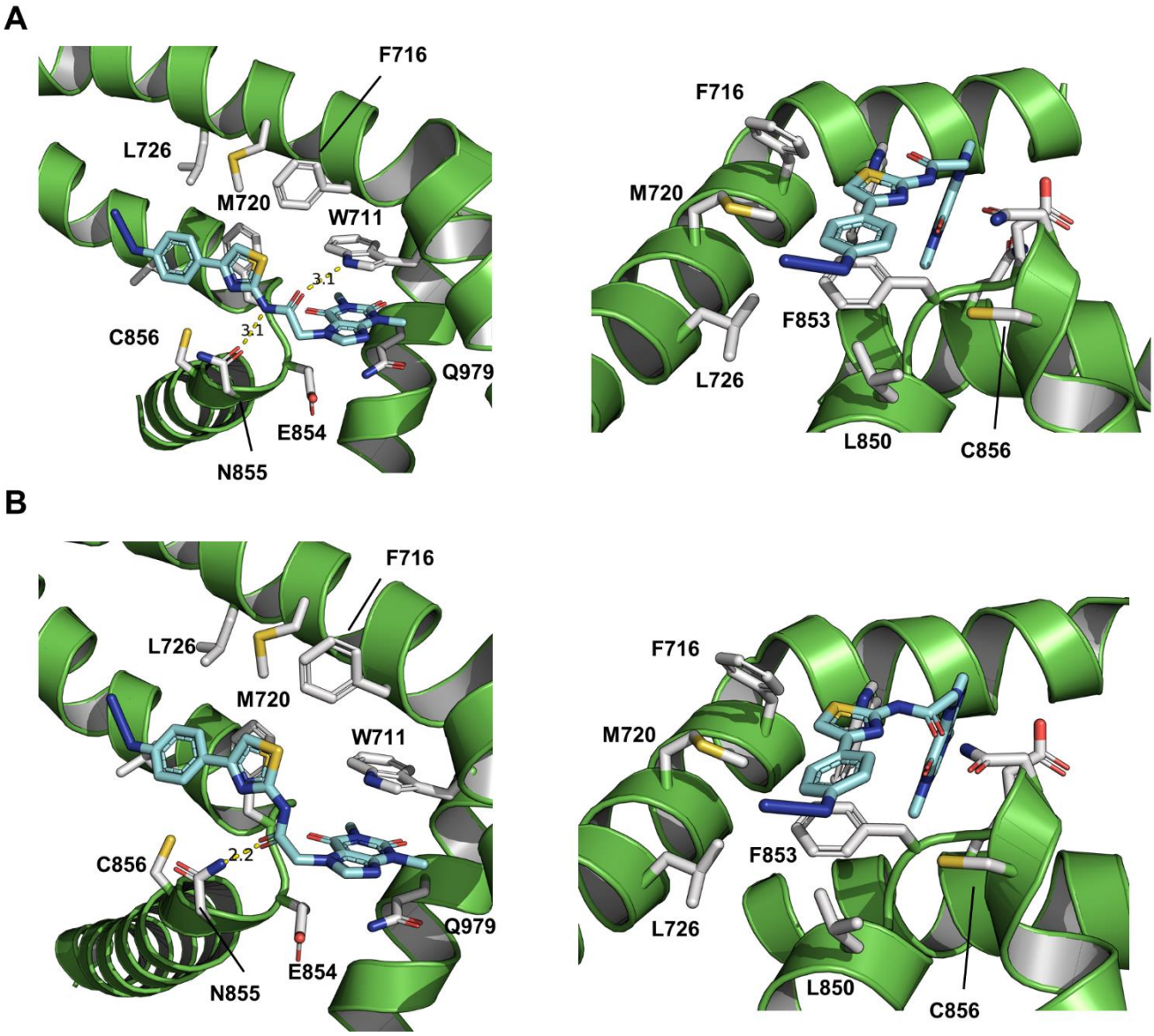


Figure 3-23: Views from the cryo-EM structure of 3-60 bound to hTRPA1 with 3-60 modelled in differing conformations. A) Two hydrogen bonds form containing the amide bond of 3-60, one with W711 and the other with N855. **B)** One hydrogen bond forms between the 3-60 carbonyl oxygen atom and the sidechain NH of N855 at a shorter distance than those formed in A.

In both modelled conformations of **3-60**, the amide group of **3-60** interacts closely with the sidechain amide group of N855 *via* hydrogen bonding. This explains why **HC-030031** has a reduced binding affinity to the N855S mutant of TRPA1, partly confirming the work by Gupta *et al.*[10] W711 is

positioned to form a π-stacking interaction with the theophylline ring in both conformations. The theophylline ring also interacts closely with the sidechain amide group of Q979 of the TRP-like box, presumably through π-interactions, stabilising the channel in the closed state. F853 forms a π-stacking interaction with the phenyl/thiazole moiety of **3-60**. The hydrophobic character of the pocket is contributed to by the sidechains of L723 and L850.

The cryo-EM structure of **3-60** bound to hTRPA1 helps to explain the different binding affinities observed by the calcium fluorescence assay outlined previously. The *meta*-phenyl azide **3-61** likely shows the poorest binding due to a steric clash between the azide and the sidechains of L723 or L850 (see fig. 2-3). Additionally, if the phenyl ring rotated to allow the azide into the pocket, interactions between the phenyl ring and F853 would be lost. The reason for the poor binding of the BGSP- series of compounds is also apparently clear. The alkyl linker of **BGSP-001** and **BGSP-002** is directed further into the lipid bilayer, mimicking the binding of the fatty acid chain of phospholipids in a highly lipophilic environment. The charged sulforhodamine and polar biotin tag of **BGSP-001** and **BGSP-002** respectively ablate binding to hTRPA1 due to their incompatibility with the highly lipophilic environment. The loss of efficacy of the probes by introducing a lipophilic alkyne tag (**BGSP-003**) can be explained by the pentynyl group being able to interact with the lipid bilayer, allowing it to be displaced on the channel transitioning to an open state, as described previously. **3-64** and **4-35** show good binding to hTRPA1 on account of the 8-azido group on the theophylline ring strengthening the π-stack with W711.

3-60 clearly binds in the **POPC** binding site, as described above. In the cryo-EM structure of the C621S hTRPA1 mutant in the absence of an agonist[31] each of these residues lining the **3-60** binding site play a key role in binding **POPC**. In the case of N855, the sidechain amide nitrogen atom is placed approximately 4 Å from the phosphonate oxygen atoms of **POPC**, allowing the formation of a hydrogen bonding interaction. W711 is approximately 4.4 Å from the trimethyl-ammonium ion of

POPC, allowing a cation-π interaction. This is supplemented by an ionic interaction with the negatively charged sidechain of E854 which is 4.2 Å away from the **POPC** trimethyl-ammonium ion. W711 also forms an interaction with an ester bond on the glycerol portion of **POPC**, an interaction which is replicated between the tryptophan ring NH and the oxygen atom of the amide group in **3-60** (fig. 3-24).

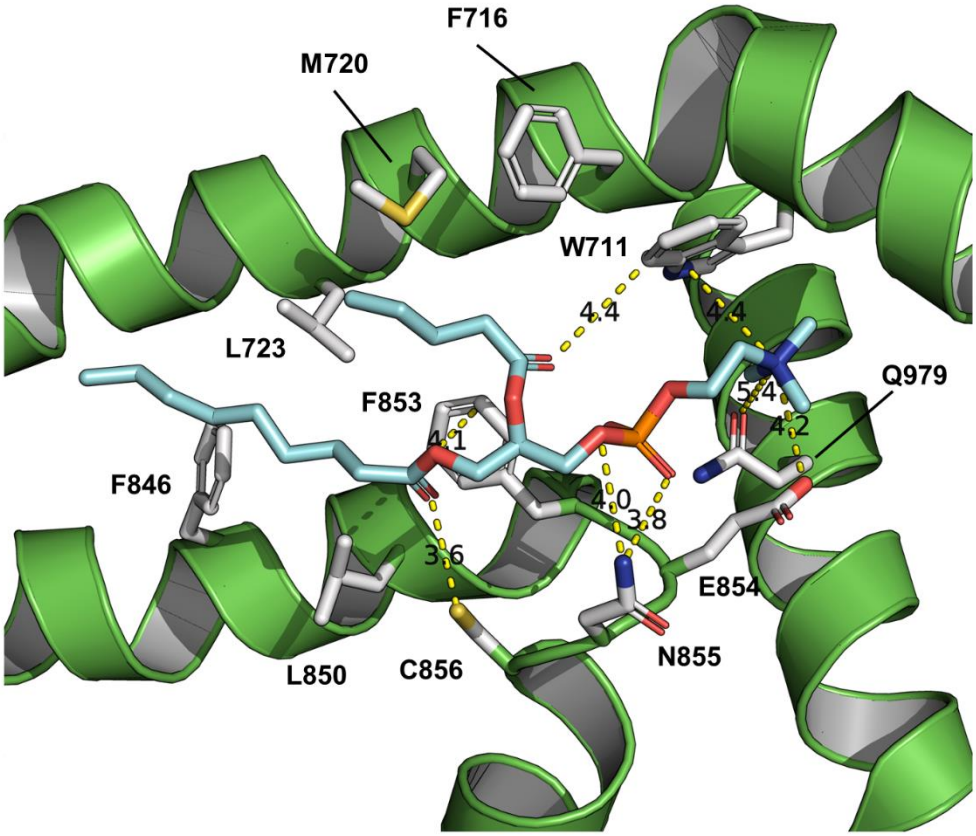


Figure 3-24: Views from the cryo-EM structure of the POPC binding site of hTRPA1 (PDB accession code: 6PQQ). Key interactions are highlighted (yellow dashes with distance in Å labelled) between POPC (blue) and TRPA1 (grey sidechains, green backbone). Residue identity is labelled for clarity.

Overlaying the C621S hTRPA1 mutant structure, containing **POPC**, with the **3-60**-hTRPA1 structure, it is clear that the theophylline-7-acetic acid moiety of **3-60** mimics the phosphorylcholine headgroup (**84**) of **POPC**. The quaternary ammonium of **POPC** overlaps with the N(9) atom of **3-60**, suggesting binding of **3-60** could be improved by the introduction of a positive charge on the theophylline ring by methylation of the N(9) atom to give **85**. The formation of a quaternary ammonium salt of **3-60** could also improve the solubility of the theophylline-based ligands, giving a salt that is less hygroscopic and prone to decomposition than the sodium salts used thus far (fig. 3-25).

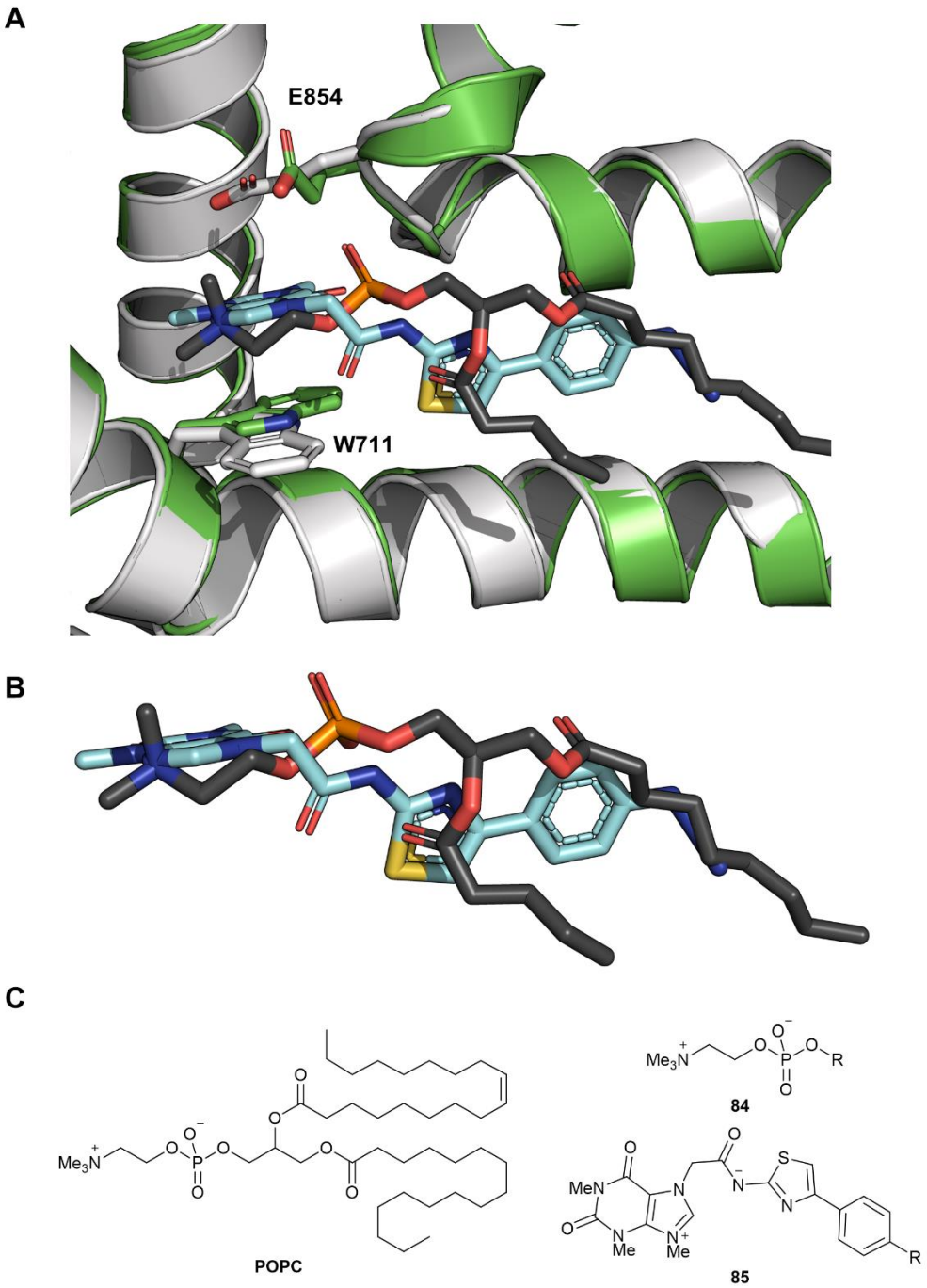


Figure 3-25: Improving the binding affinity of theophylline based hTRPA1 antagonists by comparison of 3-60 and POPC binding modes. A) views from the C621S hTRPA1 mutant (grey, PDB accession code 6PQQ) and 3-60-hTRPA1 (green with blue ligand) cryo-EM structures. **B) 3-60** overlaid with **POPC** (as in A) with hTRPA1 omitted for clarity. **C)** theophylline-7-acetic acid moiety of **3-60** mimics **84** of **POPC**, thus the zwitterion **85** is proposed as potential improvement on **3-60**.

The W711 – $RNMe_3^+$ – E854 interaction could be replicated without the formation of a quaternary ammonium salt by introducing electron deficient character on the theophylline ring through the use

of electron withdrawing substituents. The presence of C856 in the binding site could also enable the design of a covalent inhibitor of TRPA1, though attention would have to be paid to the high reactivity of C621, which may cause TRPA1 activation if C856 is not selectively modified by the covalent inhibitor.

The recent publication of a cryo-EM structure of a related series of compounds bound to hTRPA1 allows for a fragment-based approach to drug discovery.[37] This series of compounds, based on a hypoxanthine instead of a theophylline scaffold, bind to hTRPA1 *via* an interaction with W711 while the rest of the ligand is directed towards the space between the S4-S5 linker and TRP-like helix. Overlaying the structures of hTRPA1 with **3-60** bound and **86** bound, it is clear that both ligands can incorporate elements of the other one to improve binding (fig. 3-26 a,b). There is a slight difference in positioning of the xanthine ring between the two structures, although if the hypoxanthine moiety of **86** were to flip by 180 degrees or were replaced with the theophylline moiety from **3-60**, the rings may overlay more closely (fig. 3-26 c). Additionally, a flexible portion of the S4-S5 linker is displaced in the structure with **86** bound to make space for **86**. Residues, such as F853 and N855, which are part of this flexible loop form important interactions with **3-60**. F853 moves upwards by 1.4Å in the structure with **86**, and N855 moves up and directs away from the position occupied in the structure with **3-60** (fig. 3-26 d).

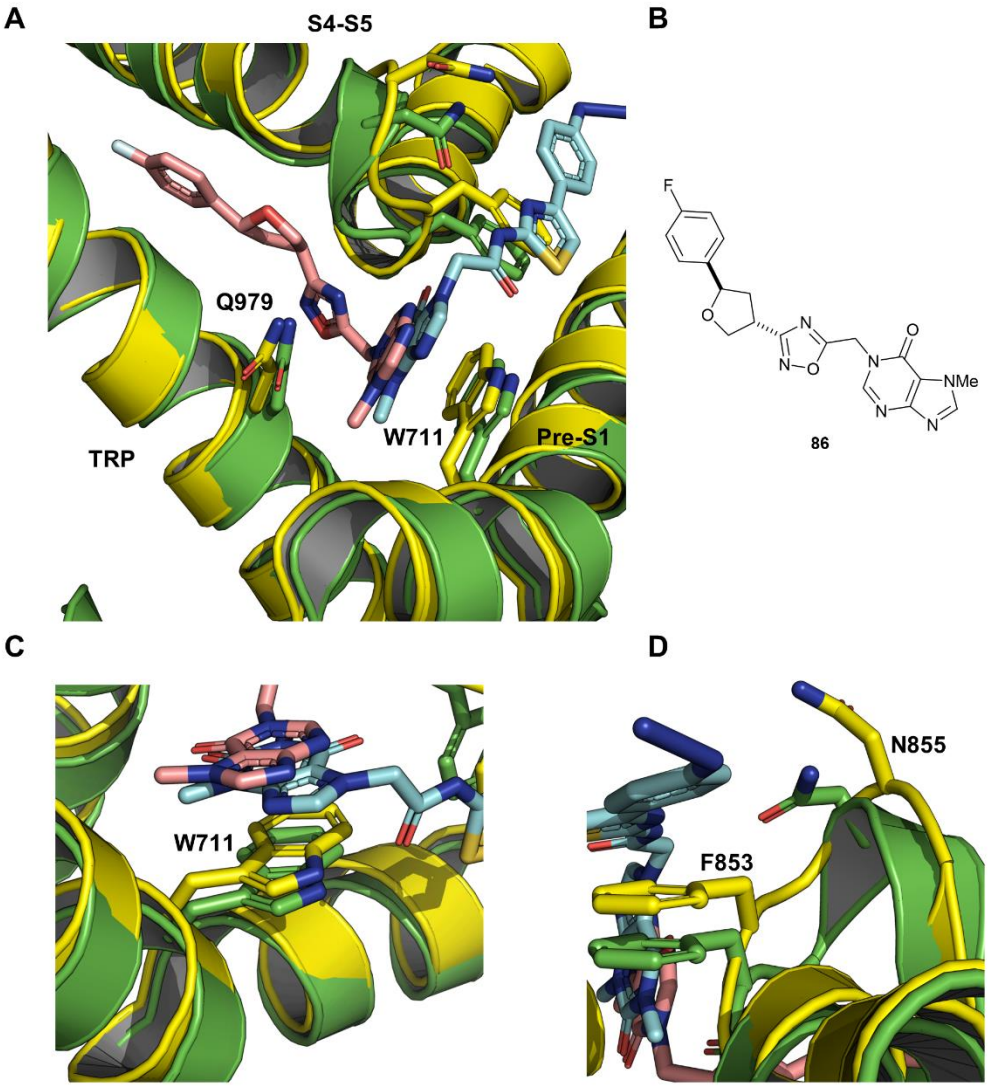


Figure 3-26: Comparison of ligand binding between 3-60 and a recently reported hypoxanthine series. A) views from the cryo-EM structure with **3-60** (green with blue ligand) and **86** bound (yellow with pink ligand, PDB accession code 7JUP). Residues and helices are labelled for clarity. **B)** structure of **86**. **C)** both **86** and **3-60** interact with W711 whilst having slightly different positioning. **D)** the flexible loop of the S4-S5 linker is displaced when binding **86** from its conformation in the cryo-EM structure with **3-60** bound.

6. Conclusions and Further Work

In this chapter a series of photoactivatable aryl azides that bind TRPA1 were designed and synthesised. It was found that the phenyl azide ligands could be synthesised from aryl bromides *via* a three-step process involving Miyaura borylation, Chan-Lam azidonation and amide coupling. The theophyllinyl azide ligands could be synthesised by bromination of theophylline followed by azidonation and subsequent amide coupling (see §2.2). A medium-throughput LCMS based solubility

assay was implemented to evaluate the effects of formulation on ligand solubility (see §3). The efficacies of the synthesised ligands were tested by a calcium fluorescence cell-based assay at DSTL. It was found that compound **3-60** was the most efficient binder of TRPA1 (as judged by having the lowest IC_{50} value in the assay, see fig. 3-22) and was taken forward for cryo-EM studies in collaboration with the Carpenter group at the SGC. A cryo-EM structure of **3-60** bound to TRPA1 was resolved by Dr. Mariana Grieben (MG) and Dr. Ashley Pike (AP) which gave new insight into the mode of binding of theophylline-based ligands to TRPA1 (see fig. 3-23).

A number of changes to the structure of **3-60**, **HC-030031** and **GRC-17536** can be proposed as a result of the structure acquired by MG and AP, and also through comparison of the structure with a recently published structure with a hypoxanthine series of compounds.[37] In particular, replacement of aminothiazole moiety, along with the formation of a zwitterionic species by alkylation of N(9) of the theophylline ring, could further improve the limited and challenging solubility of these compounds. The synthesis of a molecule combining the features of **86** with those of **HC-030031** at the N(7) position of the hypoxanthine ring would of be great interest due to possibly enhancing binding to hTRPA1. Photo-crosslinking **3-60**, **3-64** or **4-35** with hTRPA1 in the presence of UV light could be used verify the binding mode observed in the cryo-EM structure at room temperature in a native lipid environment (*i.e*. the cell membrane).

Combining binding and solubility optimisation, further pharmaceutical properties could be optimised and a lead compound brought to *in-vivo* studies. The development of a new lead molecule would be of great benefit to human health, given the role of TRPA1 in pain and inflammation, and could comprise a new non-opioid class of medication for pain relief.

7. Chapter III Bibliography

(1) Bayley, H.; Knowles, J. R. Photoaffinity Labeling. *Methods in Enzymology* **1977**, *46* (C), 69–114. https://doi.org/10.1016/S0076-6879(77)46012-9.

(2) Woll, K. A.; Skinner, K. A.; Gianti, E.; Bhanu, N. v.; Garcia, B. A.; Carnevale, V.; Eckenhoff, R. G.; Gaudet, R. Sites Contributing to TRPA1 Activation by the Anesthetic Propofol Identified by Photoaffinity Labeling. *Biophysical Journal* **2017**. https://doi.org/10.1016/j.bpj.2017.08.040.

(3) Berger, A. B.; Vitorino, P. M.; Bogyo, M. Activity-Based Protein Profiling: Applications to Biomarker Discovery, in Vivo Imaging and Drug Discovery. *American Journal of PharmacoGenomics*. Springer August 21, 2004, pp 371–381. https://doi.org/10.2165/00129785-200404060-00004.

(4) Wang, S.; Tian, Y.; Wang, M.; Wang, M.; Sun, G. B.; Sun, X. B. Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. *Frontiers in Pharmacology*. Frontiers Media S.A. April 9, 2018. https://doi.org/10.3389/fphar.2018.00353.

(5) Takaya, J.; Mio, K.; Shiraishi, T.; Kurokawa, T.; Otsuka, S.; Mori, Y.; Uesugi, M. A Potent and Site-Selective Agonist of TRPA1. *Journal of the American Chemical Society* **2015**, *137* (50), 15859–15864. https://doi.org/10.1021/jacs.5b10162.

(6) Gamal El-Din, T. M.; Lenaeus, M. J.; Zheng, N.; Catterall, W. A. Fenestrations Control Resting-State Block of a Voltage-Gated Sodium Channel. *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115* (51), 13111–13116. https://doi.org/10.1073/pnas.1814928115.

(7) Paulsen, C. E.; Armache, J. P.; Gao, Y.; Cheng, Y.; Julius, D. Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms. *Nature* **2015**, *520* (7548), 511–517. https://doi.org/10.1038/nature14367.

(8) Skerratt, S. Recent Progress in the Discovery and Development of TRPA1 Modulators. In *Progress in Medicinal Chemistry*; Elsevier B.V., 2017; Vol. 56, pp 81–115. https://doi.org/10.1016/bs.pmch.2016.11.003.

(9) Weigele, M.; Chong, J.; Larsen, G.; Hayward, N.; Ripka, A.; Ng, H.; Moran, M.; Fanger, C.; del Camino, D.; Adams, S. Methods and Compositions for Treating Disorders. US20180133170A1, June 30, 2017.

(10) Gupta, R.; Saito, S.; Mori, Y.; Itoh, S. G.; Okumura, H.; Tominaga, M. Structural Basis of TRPA1 Inhibition by HC-030031 Utilizing Species-Specific Differences. *Scientific Reports* **2016**, *6* (October), 37460. https://doi.org/10.1038/srep37460.

(11) Lenz, T.; Fischer, J. J.; Dreger, M. Probing Small Molecule-Protein Interactions: A New Perspective for Functional Proteomics. *Journal of Proteomics*. Elsevier December 10, 2011, pp 100–115. https://doi.org/10.1016/j.jprot.2011.07.017.

(12) Li, Z.; Silber, B. M.; Rao, S.; Gever, J. R.; Bryant, C.; Gallardo-Godoy, A.; Dolghih, E.; Widjaja, K.; Elepano, M.; Jacobson, M. P.; Prusiner, S. B.; Renslo, A. R. 2-Aminothiazoles with Improved Pharmacotherapeutic Properties for Treatment of Prion Disease. *ChemMedChem* **2013**, *8* (5), 847–857. https://doi.org/10.1002/cmdc.201300007.

(13) Patel, N. A.; Surti, S. C.; Patel, R. G.; Patel, M. P. Synthesis, Characterization, and Biological Activity of Some New Benzoic Acid and Thiazoloacridine Derivatives. *Phosphorus, Sulfur, and Silicon and the Related Elements* **2008**, *183* (9), 2191–2203. https://doi.org/10.1080/10426500701852687.

(14) Bruneau, A.; Roche, M.; Alami, M.; Messaoudi, S. 2-Aminobiphenyl Palladacycles: The "Most Powerful" Precatalysts in C-C and C-Heteroatom Cross-Couplings. *ACS Catalysis* **2015**, *5* (2), 1386–1396. https://doi.org/10.1021/cs502011x.

(15) Bordwell, F. G. Bordwell pKa Table https://organicchemistrydata.org/hansreich/resources/pka/ (accessed Aug 20, 2021).

(16) Grimes, K. D.; Gupte, A.; Aldrich, C. C. Copper(II)-Catalyzed Conversion of Aryl/Heteroaryl Boronic Acids, Boronates, and Trifluoroborates into the Corresponding Azides: Substrate Scope and Limitations. *Synthesis* **2010**, *2010* (9), 1441–1448. https://doi.org/10.1055/s-0029-1218683.
(17) Siméon, F. G.; Wendahl, M. T.; Pike, V. W. Efficient and Regioselective Halogenations of 2-Amino-1,3-Thiazoles with Copper Salts. *Journal of Organic Chemistry* **2009**, *74* (6), 2578–2580. https://doi.org/10.1021/jo802799c.
(18) Kende, A. S.; Liu, K. The Facile Fragmentation of Trifluoroacetyl Groups to Nitriles. *Tetrahedron Letters* **1995**, *36* (23), 4035–4038. https://doi.org/10.1016/0040-4039(95)00710-T.
(19) Ingle, J. D. Jr.; Crouch, S. R. Spectrochemical Analysis. 1988.
(20) Wang, L.; Asher, S. A. Refractive-Index Matching Avoids Local Field Corrections and Scattering Bias in Solid-State Na 2SO 4 Ultraviolet Raman Cross-Section Measurements. *Applied Spectroscopy* **2012**, *66* (2), 157–162. https://doi.org/10.1366/11-06468.
(21) Bergström, C. A. S.; Luthman, K.; Artursson, P. Accuracy of Calculated PH-Dependent Aqueous Drug Solubility. *European Journal of Pharmaceutical Sciences* **2004**, *22* (5), 387–398. https://doi.org/10.1016/j.ejps.2004.04.006.
(22) Adnani, N.; Michel, C. R.; Bugni, T. S. Universal Quantification of Structurally Diverse Natural Products Using an Evaporative Light Scattering Detector. *Journal of Natural Products* **2012**, *75* (4), 802–806. https://doi.org/10.1021/np300034c.
(23) Sou, T.; Bergström, C. A. S. Automated Assays for Thermodynamic (Equilibrium) Solubility Determination. *Drug Discovery Today: Technologies* **2018**, *27*, 11–19.
(24) Maes, E.; Nikolaev, V.; Patek, M. Kinetic Solubility: Measurement and Data Processing. *Icagen* **2017**, 1–14.
(25) Chemicalize - Instant Cheminformatics Solutions https://chemicalize.com/app/calculation/2-(1%2C3-dimethyl-2%2C6-dioxo-1%2C2%2C3%2C6-tetrahydro-7H-purin-7-yl)-N-(4-(p-tolyl)thiazol-2-yl)acetamide (accessed Mar 7, 2021).
(26) Macpherson, L. J.; Dubin, A. E.; Evans, M. J.; Marr, F.; Schultz, P. G.; Cravatt, B. F.; Patapoutian, A. Noxious Compounds Activate TRPA1 Ion Channels through Covalent Modification of Cysteines. *Nature* **2007**, *445* (February), 541–545. https://doi.org/10.1038/nature05544.
(27) Bandell, M.; Story, G. M.; Hwang, S. W.; Viswanath, V.; Eid, S. R.; Petrus, M. J.; Earley, T. J.; Patapoutian, A. Noxious Cold Ion Channel TRPA1 Is Activated by Pungent Compounds and Bradykinin. *Neuron* **2004**, *41* (6), 849–857. https://doi.org/10.1016/S0896-6273(04)00150-3.
(28) Lindsay, C. D.; Green, C.; Bird, M.; Jones, J. T. A.; Riches, J. R.; McKee, K. K.; Sandford, M. S.; Wakefield, D. A.; Timperley, C. M. Potency of Irritation by Benzylidenemalononitriles in Humans Correlates with TRPA1 Ion Channel Activation. *Royal Society Open Science* **2015**, *2* (1), 140160. https://doi.org/10.1098/rsos.140160.
(29) Eid, S. R.; Crown, E. D.; Moore, E. L.; Liang, H. A.; Choong, K. C.; Dima, S.; Henze, D. A.; Kane, S. A.; Urban, M. O. HC-030031, a TRPA1 Selective Antagonist, Attenuates Inflammatory- and Neuropathy-Induced Mechanical Hypersensitivity. *Molecular Pain* **2008**, *4*, 48. https://doi.org/10.1186/1744-8069-4-48.
(30) Gee, K. R.; Brown, K. A.; Chen, W. N. U.; Bishop-Stewart, J.; Gray, D.; Johnson, I. Chemical and Physiological Characterization of Fluo-4 Ca2+-Indicator Dyes. *Cell Calcium* **2000**, *27* (2), 97–106. https://doi.org/10.1054/ceca.1999.0095.
(31) Suo, Y.; Wang, Z.; Zubcevic, L.; Hsu, A. L.; He, Q.; Borgnia, M. J.; Ji, R. R.; Lee, S. Y. Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. *Neuron* **2020**, *105* (5), 882-894.e5. https://doi.org/10.1016/j.neuron.2019.11.023.
(32) Zhao, J.; Lin King, J. v.; Paulsen, C. E.; Cheng, Y.; Julius, D. Irritant-Evoked Activation and Calcium Modulation of the TRPA1 Receptor. *Nature* **2020**, *585* (7823), 141–145. https://doi.org/10.1038/s41586-020-2480-9.

(33) Eid, S. R.; Crown, E. D.; Moore, E. L.; Liang, H. A.; Choong, K. C.; Dima, S.; Henze, D. A.; Kane, S. A.; Urban, M. O. HC-030031, a TRPA1 Selective Antagonist, Attenuates Inflammatory- and Neuropathy-Induced Mechanical Hypersensitivity. *Molecular Pain* **2008**, *4*, 48. https://doi.org/10.1186/1744-8069-4-48.
(34) McNamara, C. R.; Mandel-Brehm, J.; Bautista, D. M.; Siemens, J.; Deranian, K. L.; Zhao, M.; Hayward, N. J.; Chong, J. A.; Julius, D.; Moran, M. M.; Fanger, C. M. TRPA1 Mediates Formalin-Induced Pain. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104* (33), 13525–13530. https://doi.org/10.1073/pnas.0705924104.
(35) Invitrogen. Bac-to-Bac® Baculovirus Expression System. An Efficient Site-Specific Transposition System to Generate Baculovirus for High-Level Expression of Recombinant Proteins. *Computer programs in biomedicine* **2010**, *18* (1–2), 1–78.
(36) Hauer, F.; Gerle, C.; Fischer, N.; Oshima, A.; Shinzawa-Itoh, K.; Shimada, S.; Yokoyama, K.; Fujiyoshi, Y.; Stark, H. GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. *Structure* **2015**, *23* (9), 1769–1775. https://doi.org/10.1016/j.str.2015.06.029.
(37) Terrett, J. A.; Chen, H.; Shore, D. G.; Villemure, E.; Larouche-Gauthier, R.; Déry, M.; Beaumier, F.; Constantineau-Forget, L.; Grand-Maître, C.; Lépissier, L.; Ciblat, S.; Sturino, C.; Chen, Y.; Hu, B.; Lu, A.; Wang, Y.; Cridland, A. P.; Ward, S. I.; Hackos, D. H.; Reese, R. M.; Shields, S. D.; Chen, J.; Balestrini, A.; Riol-Blanco, L.; Lee, W. P.; Liu, J.; Suto, E.; Wu, X.; Zhang, J.; Ly, J. Q.; La, H.; Johnson, K.; Baumgardner, M.; Chou, K. J.; Rohou, A.; Rougé, L.; Safina, B. S.; Magnuson, S.; Volgraf, M. Tetrahydrofuran-Based Transient Receptor Potential Ankyrin 1 (TRPA1) Antagonists: Ligand-Based Discovery, Activity in a Rodent Asthma Model, and Mechanism-of-Action via Cryogenic Electron Microscopy. *Journal of Medicinal Chemistry* **2021**, *64* (7), 3843–3869. https://doi.org/10.1021/acs.jmedchem.0c02023.

Chapter IV:

Investigating TRPA1 as a Substrate for the HIF Hydroxylases

1. Previous Work

1.1 Factor Inhibiting Hypoxia Induced Factor (FIH) Substrates

There is good evidence that FIH catalyses the hydroxylation of many cytoplasmic proteins including HIF,[1] however, the number of reported FIH substrates that are membrane proteins is low. The Notch family (Notch 1-3) are membrane proteins which are hydroxylated by FIH on their cytoplasmic C-terminal ankyrin repeat domain.[2,3] Hydroxylation of Notch seems to have little effect on the Notch/CSL pathway or to modulate the binding of other known Notch binding proteins. There is evidence that Notch binds to FIH more tightly than HIF-1α, suggesting that Notch may serve as a competitive inhibitor of FIH mediated hydroxylation of HIF-1α.[3] Another reported membrane-associated substrate of FIH is Rabankyrin-5,[2] which has an important role in regulating endocytosis.[4] Rabankyrin-5 is found in both the cytosolic and membrane fraction of cell lysates,[5] however, it is unclear whether membrane-associated Rabankyrin-5 is hydroxylated by FIH or whether the cytosolic component of Rabankyrin-5 is the FIH substrate.

The activities of both TRPA1 and TRPV3 are reported to be modulated by the HIF hydroxylases.[6] TRPA1 is reported to be hydroxylated on P394 by PHD2.[7] TRPA1 sensitivity is lowest under normoxia (20% O_2) and rises to a maximum under "physiological normoxia" (2-8% O_2) suggesting that proline hydroxylation dampens TRPA1 activity.[7] It was also found that the P394A mutant of TRPA1 had increased sensitivity to agonists and cool temperatures, further lending weight to hydroxylation of P394 as inhibiting channel function.[8] However, it should be noted that subsequent work by Cockman *et al.* has cast doubt on whether several reported substrates of PHDs are in fact genuine.[9] In particular, Cockman *et al.* failed to reproduce the hydroxylation of TRPA1 P394 by PHD2 both using peptides and recombinant TRPA1 produced by *in-vitro* transcription/translation (IVTT),[9] although the conditions used

for the hydroxylation assay and methods of recombinant PHD2 production vary considerably between the two studies (fig. 4-1). Cockman *et al.*[9] also used the catalytic domain of PHD2 in addition to full length PHD2 and found no activity. Furthermore, Takahashi *et al.*[7] reported that in-vitro hydroxylation of TRPA1 P394 by isolated PHD2 was inhibited by DMOG. This is surprising given that DMOG is an inactive prodrug which is cleaved by cellular esterases to give the active inhibitor *N*-oxalylglycine (NOG). NOG shows broad activity against 2-oxoglutarate (2OG) dependent hydroxylases, including FIH and PHD2.[10]

Study	Cockman *et al.*[9]	Takahashi *et al.*[7]
Hydroxylation observed?	No	Yes
PHD2 Construct	Full Length	Full Length
PHD2 Expression Organism	Insect (H5)	Human (HEK293T)
PHD2 Tag	C-terminal Flag-His_6	N-terminal Flag
Assay Buffer	50 µM FAS, 300 µM 2OG, 300 µM ASC, 50 mM Tris pH 7.5	50 mM CsCl, 50 mM CsOH, 50 mM aspartic acid, 3 mM ASC, 100 µM $FeCl_2$, 300 µM 2OG, 1 mM DTT, 0.3 mg ml^{-1} catalase, 10 mM Hepes pH 7.4
PHD2 concentration	2 µM	2 µM
Peptide concentration	50 µM	40 µM

Figure 4-1: Comparison of experimental procedures for the hydroxylation of the TRPA1(386-405) peptide as reported by *Cockman et al.*[9] and *Takahashi et al.*[7] which give contrasting results. Abbreviations: FAS = $NH_4[Fe(H_2O)_6](SO_4)_2 \cdot 6\,H_2O$, 2OG = 2-oxoglutarate, ASC = sodium ascorbate, DTT = dithioreitol.

TRPV3 is also reported to be a substrate of FIH, undergoing hydroxylation on N242.[11] Similarly to the reported, but controversial,[9] hydroxylation on P394 of TRPA1 by PHD2, hydroxylated TRPV3 shows decreased sensitivity to agonists under normoxia and conditions when 2OG dependent enzymes, such as the HIF hydroxylases, are inhibited by DMOG. TRPA1 ankyrin repeats align well with the consensus sequence for FIH,[11] and similar channel activities

are observed for TRPV3[11] and TRPA1[7] under changing oxygen concentrations and the presence of DMOG. Therefore, it was thus proposed that TRPA1 is a substrate for FIH.

Work in the group by Dr. R. Hopkinson and Dr. T. Leissing identified a peptide comprised of TRPA1 residues 313-339 which was efficiently hydroxylated by FIH. A crystal structure of FIH in complex with TRPA1(313-339), zinc and NOG was obtained showing N336 of TRPA1 present in the active site of FIH (fig. 4-2).[12]

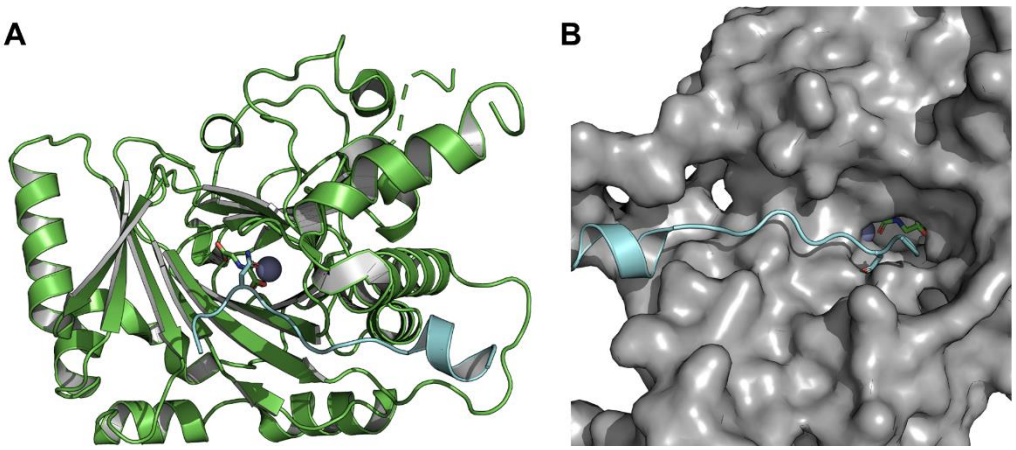


Figure 4-2: Views from the crystal structure of FIH (green/grey) in complex with TRPA1(313-339). A) Cartoon model. **B)** Surface model. (PDB accession code: 6HC8)

1.2 Aims

The work by Dr. T. Leissing and Dr. R. Hopkinson was focused on peptides binding to site 1 of FIH (*i.e.* residues of peptides on the N-terminal side of the hydroxylation site).[12] It was therefore hypothesised that peptides which contained additional residues to the C-terminal side of the hydroxylation site would be able to bind both site 1 and site 2 of FIH. It was hypothesised that these peptides would be better substrates for FIH and potentially produce better crystal structures than obtained previously. The aims of the work described in this chapter were to:

- Design and evaluate the TRPA1 ankyrin loop peptides as FIH substrates

- Investigate TRPA1 hydroxylation mediated by the HIF-hydroxylases at the protein level
- Investigate the functional consequences of TRPA1 hydroxylation

2. Screening TRPA1 Ankyrins as Substrates for FIH

TRPA1 has 16 putative ankyrin repeats which are 30-34 residue secondary structures comprised of a helix-turn-helix motif followed by a flexible loop.[13] FIH mediated hydroxylation of asparagine residues in the ankyrin repeat domain of proteins typically occurs on the flexible loop of an ankyrin.[3] In order for the ankyrin loop peptides to bind to site 1 and site 2 of FIH, 28-mer peptides were constructed based on the 28 residues following the 15th residue of the assigned ankyrin repeats of TRPA1.[14] Residues 374 – 403 of TRPA1 were not included in the analysis due to not appearing to be an ankyrin by sequence analysis. Alignment of the 28mers with the Schofield group consensus ankryin (1CA)[15] and the Mosavi[16] consensus ankyrin sequences showed that the peptides initiate at the start of the 2nd α-helix and typically terminate in the first portion of the 1st α-helix in the next ankyrin repeat (fig. 4-3).

The poor alignment of the TRPA1 ankyrin 15 loop peptide with the other ankyrin peptides is due to the residues from 604 onwards forming the electrophile binding domain (EBD) of TRPA1. However, this peptide was still included into the substrate screen due to the close proximity of the ankyrin loop to the EBD. Thus, hydroxylation of this ankyrin loop could potentially stabilise the EBD into a state that has an effect on electrophile binding and gating of the channel (fig. 4-3).

Peptide	Residues	Sequence
1CA		----------HLEVV-KLLLE-AGA-DVNAQ-DK----------
TRPA1 ankyrin 1	76-103	----------IELMEKITRD-SSLEVLHEM-DDYGNTPLH-------
" 2	111-138	----------IESV-KFLLS-RGA-NPNLR-NFNMMAPLHIA-----
" 3	144-171	----------NEVM-KVLLE-HRTIDVNLE-GENGNTAVII------
" 4	177-204	----------NSEAL-QILLK-KGA-KPCKS-NKWGCFPIHQ------
" 5	211-238	----------KECMEIILR-FGE-------EHGYSRQLHINFMNNG
" 6	252-279	----------LEMI-KMCLD-NGA-QIDPV-EKGRCTAIHFA-----
" 7	285-312	----------TEIV-KLMIS-SYSGSVDIVNTTDGCHETM-------
" 8	322-348	----------HELA-DYLIS-VGA-DINKI-DSEGRSPLIL------
" 9	355-382	----------WNIV-NLLLS-KGA-QVDIK-DNFGRNFLHLT-----
" 10	426-453	----------PGSV-NNLLG-FNV-SIHSK-SKDKKSPLHFA-----
" 11	459-486	----------INTCQRLLQDISDTRLLNEG-DLHGMTPL--------
" 12	495-522	----------DKVV-QLLLK-KGA-LF-LS-DHNGWTALHHAS----
" 13	527-554	----------TQTM-KVILD-TNLKCTDRL-DEDGNTALHF------
" 14	561-588	----------AKAV-ALLLS-HNA-DIVL--NKQQASFLHLAL----
" 15	593-620	----KEVVLTIIRSKRWDECLKIFSH-NSP-------GNK----------
Mosavi		DKNGRTPLHLAARNGHLEVV-KLLLE-AGA-DVNAK----------

Figure 4-3: Alignment of TRPA1 ankyrin loop peptides with the Schofield group consensus ankyrin (1CA)[15] and Mosavi *et al.*[16] consensus sequence using Clustal Omega reveals potential secondary structure of peptides.[33] The ankyrin loop 28mers begin with the 2nd helix of the ankyrin (H2) and terminate in the 1st helix of the next ankyrin (H1). A representation of the ankyrin repeat is shown aligned with the Mosavi consensus sequence.

Initial work focused on the solid phase peptide synthesis (SPPS) of the target ankyrin loop peptides. SPPS provides a convenient and automated way of synthesising peptides though the use of a resin on which amide coupling reactions can be performed. Peptides are synthesised from the C-terminus using *N*-Fmoc-protected amino acids with acid labile protecting groups on the side chain. Peptide synthesis is achieved through a repeated cycle of deprotecting the N-terminal Fmoc-group on the peptide with a mild base such as piperidine to give an amine which can then undergo an amide coupling reaction with an Fmoc-protected amino acid, extending the peptide (fig. 4-4).[17]

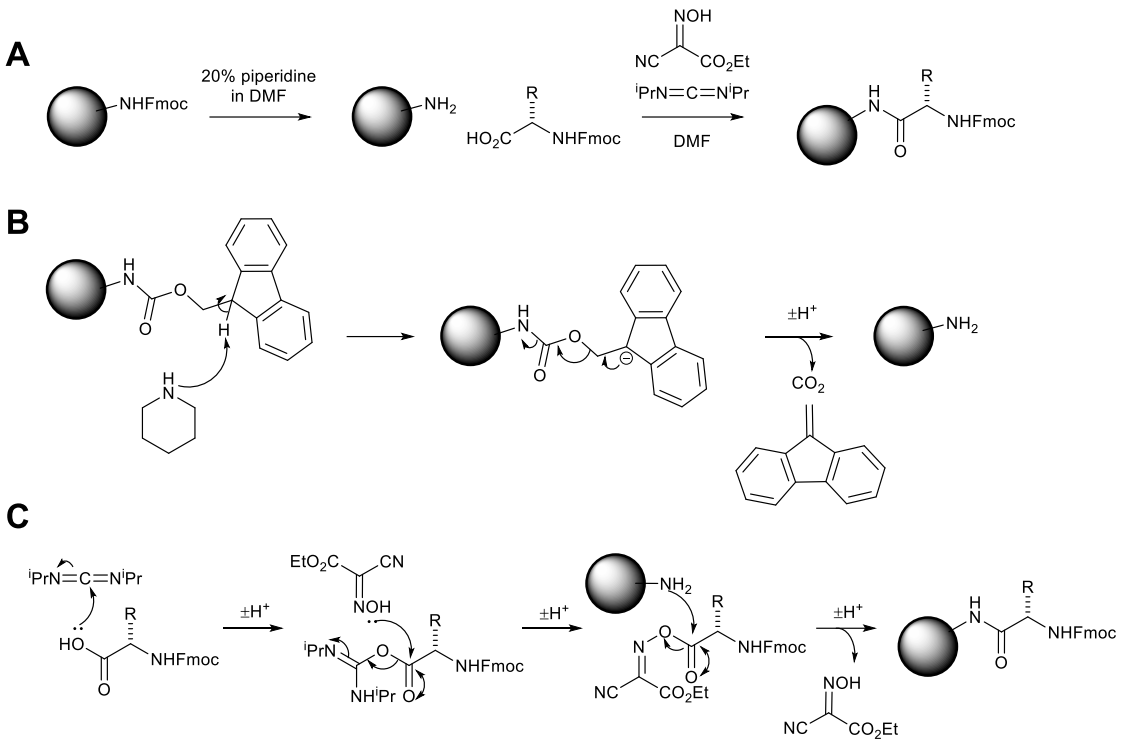


Figure 4-4: Outline of the transformations involved in SPPS. A) peptides are synthesised from the C-terminus by a repeated cycle of N-terminal Fmoc deprotection followed by amide coupling with a Fmoc-protected amino acid. **B)** Fmoc-deprotection is realised using a mild base such as piperidine via an E1 mechanism. **C)** Amide coupling is carried out using DIC with Oxyma Pure employed as a catalyst to supress side reactions such as racemisation of the active ester formed after reaction of the amino acid with DIC.

Most peptides underwent SPPS with success, however, the peptides for ankyrin loops 3, 6, 7 and 13 did not. Ankyrin loop 3 was thus ordered from a commercial supplier (GL Biochem), however, they also could not synthesise the peptide, thus a 26-mer peptides was synthesised containing TRPA1 residues 145 – 170. Interestingly, the 28-mer was synthesised previously by Dr. R. Hopkinson and was found to interact with FIH whilst not being hydroxylated. A crystal structure of FIH in complex with the ankyrin 3 peptide was obtained by Dr. T. Leissing; however, the complex gave rise to a hexagonal crystal morphology as opposed to the more commonly obtained tetragonal form and a lower resolution (3.04 Å) structure was obtained.[12]

The ankyrin loop peptides, along with 1CA and TRPV3(229-255) serving as positive controls,[18] were incubated with 10 μM FIH. The resulting mixtures were analysed by matrix-assisted laser desorption ionisation mass spectrometry (MALDI-MS). It was found that only the ankyrin 8 loop peptide, in addition to the positive controls, was hydroxylated efficiently at a final concentration of 10 μM FIH (fig. 4-5). Identification of the location of the hydroxylated residue was carried out using a pseudo-mass spectrometry coupled mass spectrometry (MS/MS) method developed by Bruker known as MALDI-LIFT. The MALDI-LIFT method causes fragmentation of the peptide, allowing for sequencing of the peptide.[19] A +16 Dalton shift was observed for the y13 ion in the hydroxylated peptide compared to the non-hydroxylated peptide while the y12 ion remains unchanged. This allowed for the hydroxylated peptide to be assigned as bearing a hydroxy-asparagine residue corresponding to N336 in TRPA1 (fig. 4-6).

Kinetic analysis of hydroxylation of ankyrin 8 in comparison to 1CA was undertaken with the help of Dr. A. Tumber using the RapidFire mass spectrometry platform (RFMS). RFMS uses a proprietary solid phase extraction (SPE) cartridge in place of the more typical column chromatography (*i.e.* liquid-chromatography coupled MS, LCMS) to separate an analyte of interest from buffer components that are incompatible with MS. The use of SPE greatly shortens analysis time, typically to seconds compared to minutes for LCMS samples, allowing for efficient reaction monitoring.[20] The ankyrin 8 loop peptide is a comparatively poor substrate compared to the 1CA peptide. This is shown by the low catalytic efficiency (k_{cat}/K_m) of the ankyrin 8 peptide (0.026 μM^{-1} min^{-1}) compared to 1CA (5.50 μM^{-1} min^{-1}) (fig. 4-7).

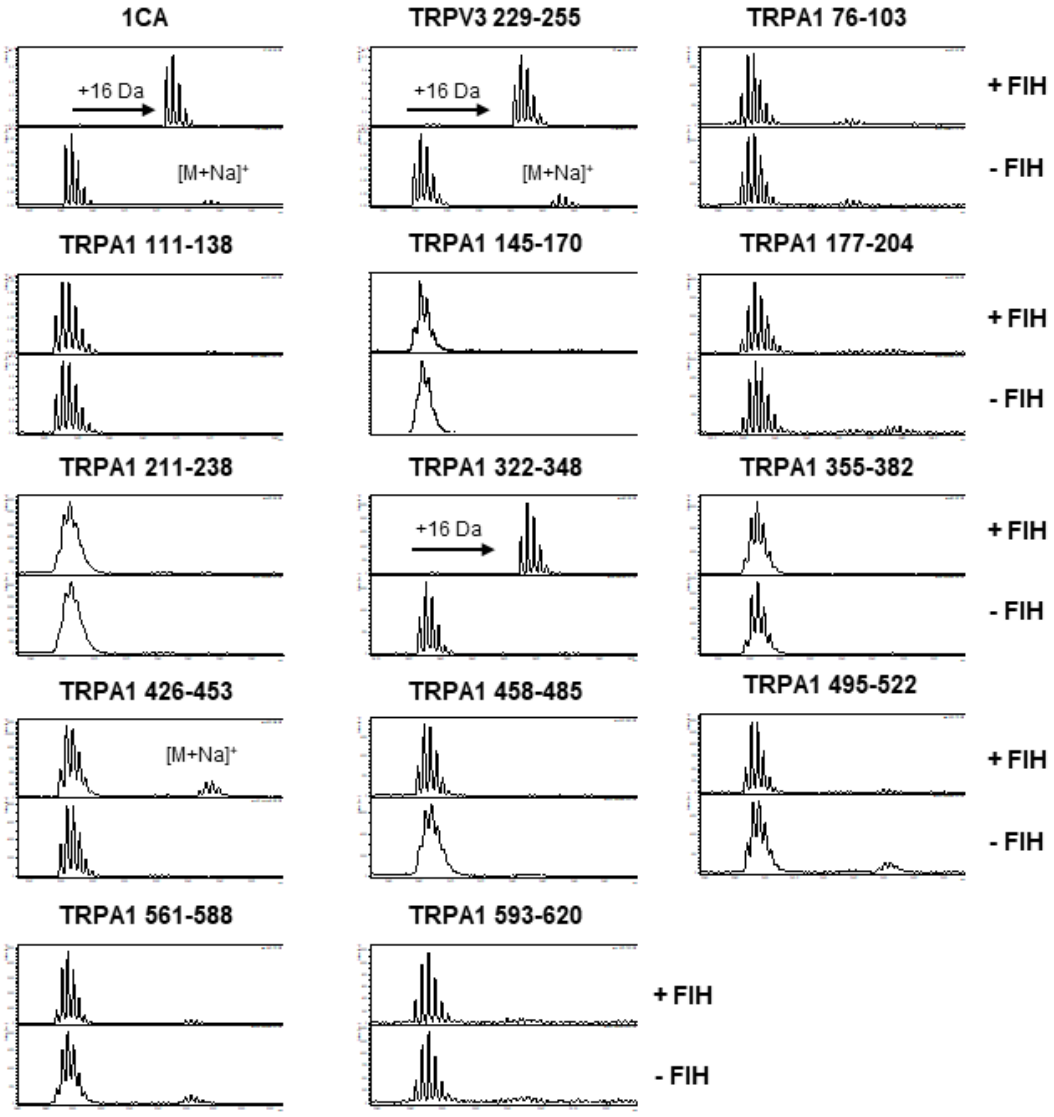


Figure 4-5: End point MALDI-MS spectra of peptides derived from TRPA1 and TRPV3 reacted with FIH. Assay conditions: 10 µM peptide, 20 µM ferric ammonium sulfate, 100 µM 2-oxoglutarate, 100 µM sodium ascorbate, 0 or 10 µM FIH, 50 mM Tris pH 7.5, 37 °C, 2 h. **NB.** TRPA1 ankyrin 8 (322-348) is the only TRPA1 derived peptide in this screen that undergoes hydroxylation by FIH.

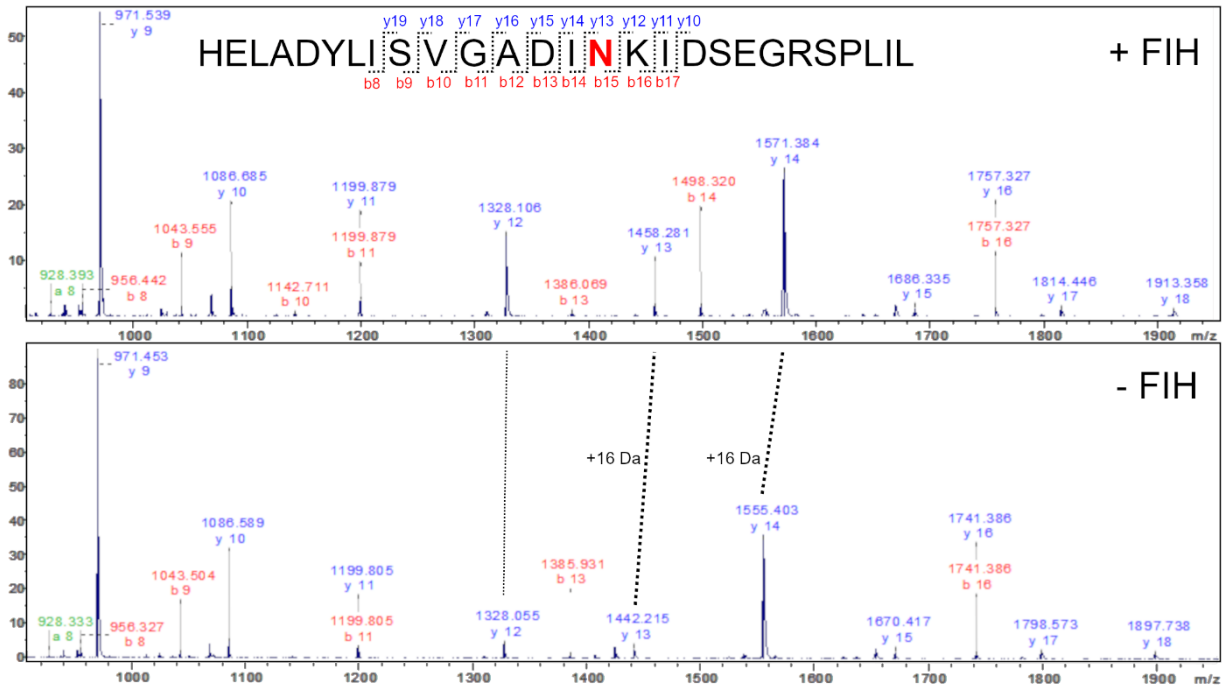


Figure 4-6: MALDI-LIFT spectra of ankyrin 8 loop peptide in different hydroxylation states. 10 mM peptide, 20 µM ferric ammonium sulfate, 100 µM 2-oxoglutarate, 100 µM sodium ascorbate, 0 or 10 µM FIH, 50 mM Tris pH 7.5, 37 °C, 2 h.

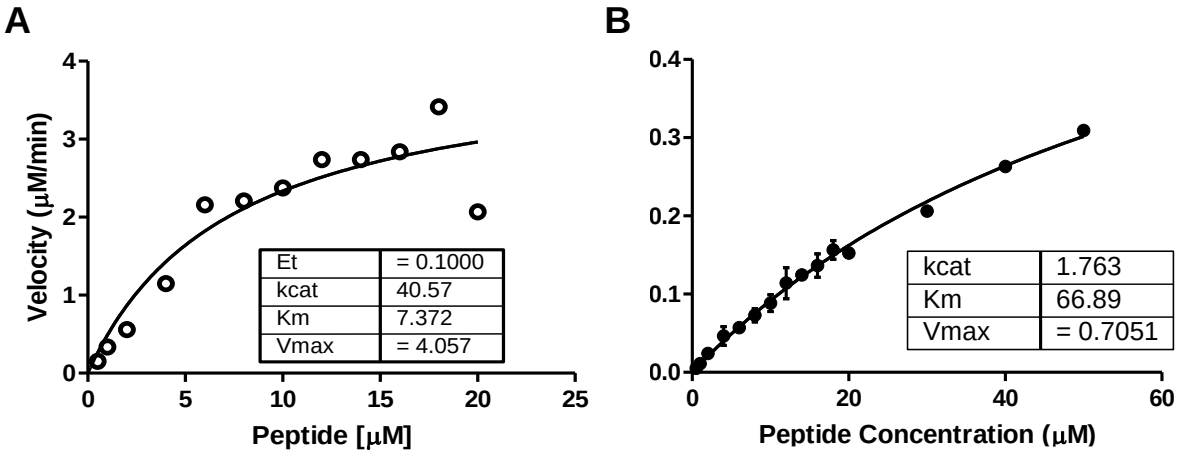


Figure 4-7: Comparative kinetics of hydroxylation of ankyrin peptides. A) Schofield group consensus (1CA). Assay conditions: 0.1 µM FIH, 100 µM L-ascorbate, 20 µM ferric ammonium sulfate, 20 µM 2-oxoglutarate; **B)** TRPA1 ankyrin 8 loop peptide. Assay conditions: 0.4 µM FIH, 100 µM L-ascorbate, 20 µM ferric ammonium sulfate, 20 µM 2-oxoglutarate. Error bars represent SEM for n=2 experiments performed in triplicate.

A possible reason for the TRPA1 ankyrin 8 loop peptide being hydroxylated by FIH, while the other TRPA1 ankyrin loop peptides were not, is that the ankyrin 8 loop peptide closely resembles the FIH consensus sequence (fig. 4-8 a). The reason for TRPA1 ankyrin 3 not undergoing FIH-mediated hydroxylation is not immediately clear. Like the TRPA1 ankyrin 8 peptide, the TRPA1 ankyrin 3 peptide closely resembles the FIH consensus sequence (fig. 4-8 a) and has been shown to interact with FIH *via* soaking the peptide into FIH crystals to give a structure of TRPA1 ankyrin 3 in complex with FIH. In the structure, N159 of TRPA1 is clearly visible inside the active site of FIH in a comparable geometry to that of TRPV3 ankyrin 1 which is readily hydroxylated by FIH (fig. 4-5 and fig. 4-8 b). Notably, the TRPA1 ankyrin 3 peptide N-terminus is unstructured which is unlike other substrates where the N-terminus of the peptide forms a helix (fig. 4-8 c). The alignment of the TRPA1 ankyrin 3 peptide with the 1CA and Mosavi consensus ankyrins shows the presence of an additional residue in the ankyrin loop between the conserved leucine in the second helix of the ankyrin and the target asparagine residue (see fig. 4-3). In order for this additional residue to be accommodated into the FIH consensus sequence, the secondary structure of TRPA1 ankyrin peptide must be disrupted. Although the crystal structure shows that this unfolding may occur, the activation energy involved in unfolding the peptide may present a kinetic barrier to hydroxylation. This would prevent hydroxylation from being observed, while a prolonged interaction, *via* soaking crystals of FIH with the peptide, allows for the thermodynamically favourable interaction to be observed *via* crystallography.

A

Peptide	Sequence
HIF-1α (788-818)	DESGLPQLTSYDCEVNAPIQGSRNLLQGEEL
HIF-2α (832-862)	ESYLLPELTRYDCEVNVPVLGSSTLLQGGDL
Notch 1 (1940-1970)	RSDAAKRLLEASADANIQDNMGRTPLHAAVS
TRPV3 (229-255)	--DIAALLIAAGADVNAHAKGAFFNPKYQ--
1CA	HLEVVKLLLEAGADVNAQDK----------
TRPA1 Ank 8 (322-348)	-HELADYLISVGADINKIDSEGRSPLIL---
TRPA1 Ank 3 (145-170) *	-EVMKVLLEHRTIDVNLEGENGNTAVI----
Consensus Sequence	-------L----$\pi\frac{D}{E}\phi$N----------

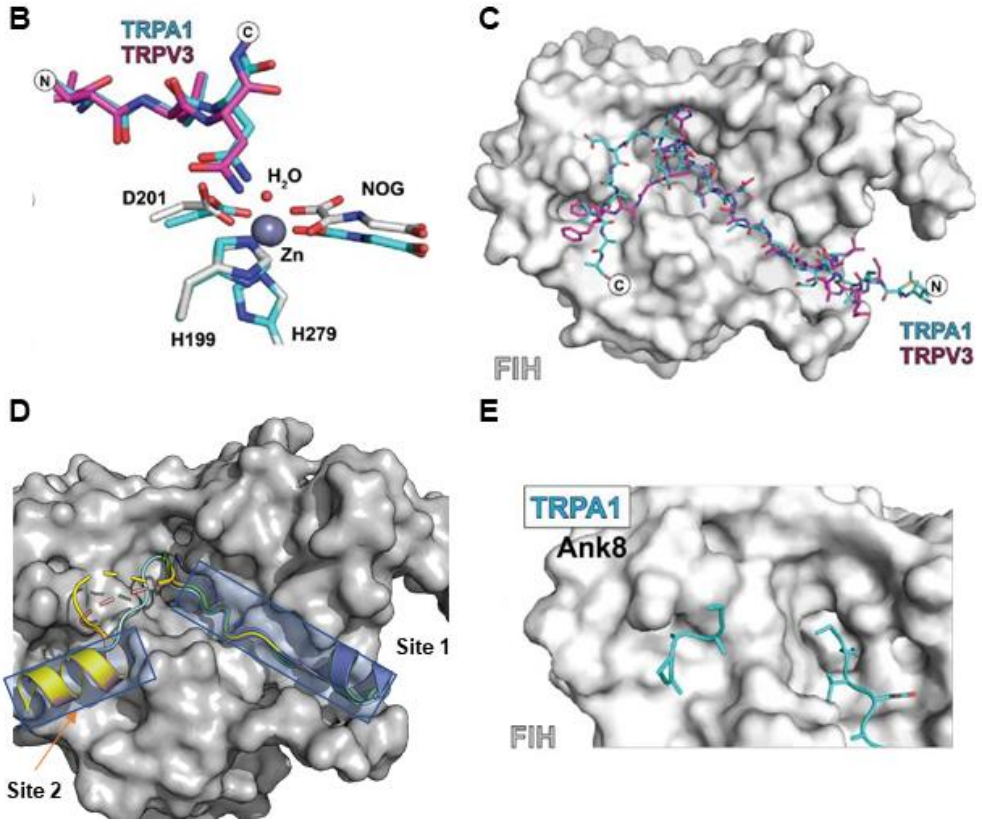


Figure 4-8: Comparison of different FIH substrates reveals subtle differences resulting in differences in observed substrate turnover. Figures B, C & E are reproduced from Dr. T. Leissing's DPhil thesis.[12] **A)** Alignment of FIH substrates with the FIH consensus sequence where π represents a small, uncharged residue and φ represents a small, hydrophobic residue. * TRPA1 ankyrin 3 peptide is included for comparison but is not hydroxylated by FIH. **B)** N159 of the TRPA1 ankyrin 3 peptide (blue) occupies the active site of FIH in a similar manner to N242 of the TRPV3 ankyrin 1 peptide (purple). **C)** The TRPA1 ankyrin 3 peptide (blue) shows disordered in comparison to the TRPV3 ankyrin 1 peptide (purple) at its N-terminus. **D)** Binding to site 1 and site 2 of FIH is conserved whilst a flexible loop bridges the two sites for HIF-1α (yellow, 1H2K) and HIF-2α (pink, 1H2M). Site 1 binding is also conserved for 1CA (purple, 4B7E) and TRPV3 (blue, 6H9J). (PDB accession codes given in brackets). **E)** Site 2 binding of the TRPA1 ankyrin 8 peptide is poorly resolved suggesting a degree of flexibility.

The TRPA1 ankyrin 8 peptide shows a similar binding mode to the HIF-1α and HIF-2α C-terminal transactivation domain (CTAD) peptides for which the N-terminus of the peptide binds tightly to site 1 of FIH, while the C-terminus binds to site 2 of FIH (fig. 4-8 d). However, the HIF-α CTAD peptides form a flexible loop between site 1 and site 2 which has not been resolved in previously obtained crystal structures.[21] The TRPA1 ankyrin 8 peptide density observed between site 1 and site 2 also cannot be modelled. Additionally, the identity of the TRPA1 residues in site 2 of FIH could not be unambiguously established, suggesting there is a degree in flexibility in binding (fig. 4-8 e). This helps to explain why the TRPA1 ankyrin 8 peptide is a comparatively poor substrate for FIH.

3. Investigating the Hydroxylation State of Recombinant Human TRPA1

It was proposed that analysing isolated TRPA1 for hydroxylation could be of use in unravelling whether TRPA1 is a substrate for PHD2 or FIH. The two sites could potentially be discriminated through cleaving the protein into smaller peptides with a protease such as trypsin. The resulting peptide mixture could be analysed by mass spectrometry and subsequent MS/MS to sequence a particular peptide of interest to unambiguously assign a post-translational modification.

Using hTRPA1 produced recombinantly from Sf9 insect cells by Dr. Mariana Grieben (see Chapter III for further information), a range of techniques were applied to determine the potential hydroxylation state of the recombinant protein. The buffer used to solubilise TRPA1 contained a high concentration of the detergent lauryl maltose neopentyl glycol (LMNG). In general, detergents have an adverse effect on the quality of mass spectrum[22] and in this case, prevented the formation of matrix crystals, preventing the acquisition of MALDI-MS spectra.

In order to remove the LMNG and denature TRPA1, making the protease cleavage sites more exposed, a chloroform-methanol extraction was performed. The protein pellet was then resuspended in 50 mM Tris pH 7.5 and 8 M urea and heated to 60 °C to further denature the protein. The sample was then cooled to rt and the disulphide bridges reduced with 5 mM dithioreitol (DTT) at 37 °C for 30 minutes. Alkylation of cysteine residues, preventing the disulphide bridges from forming, was achieved using 14 mM iodoacetamide at rt for 30 min. A mixture of trypsin and Lys-C was added to the mixture as TRPA1 was considered to be a proteolytically resistant protein. It was hoped that digestion of TRPA1 under denaturing conditions (6 M urea) with Lys-C would leave the trypsin cleavage sites more accessible and hence increase the coverage of the protein.[23]

After digesting the protein for 3 h at 37 °C, the mixture was diluted with deionised water to give a final urea concentration of 1 M. Diluting the mixture and lowering the urea concentration reverses the inhibition of trypsin by urea, allowing for further digestion of TRPA1. The digestion mixture was incubated at 37 °C for 16 h before 5 % formic acid was added and the sample concentrated *in vacuo*. The sample was diluted with 0.1% aq. CF_3CO_2H (TFA) before urea and other contaminants were removed through the use of a C18-ZipTip. The peptide mixture was eluted directly onto a MALDI-MS target using matrix buffer (50% aq. MeCN, 10 mg/mL α-cyano-4-hydroxycinnamic acid, 0.1% TFA). Acquisition of the MALDI-MS spectrum showed putative coverage of both the PHD2 and FIH hydroxylation sites (fig. 4-9).

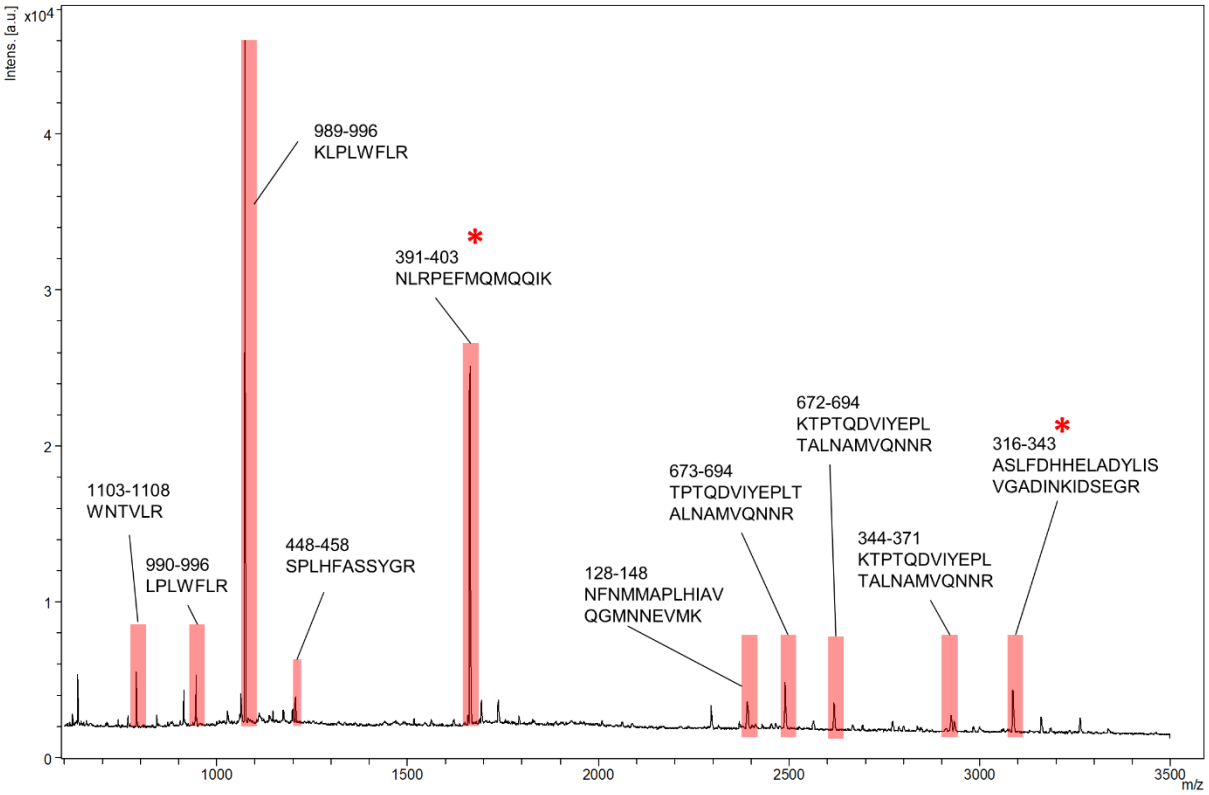


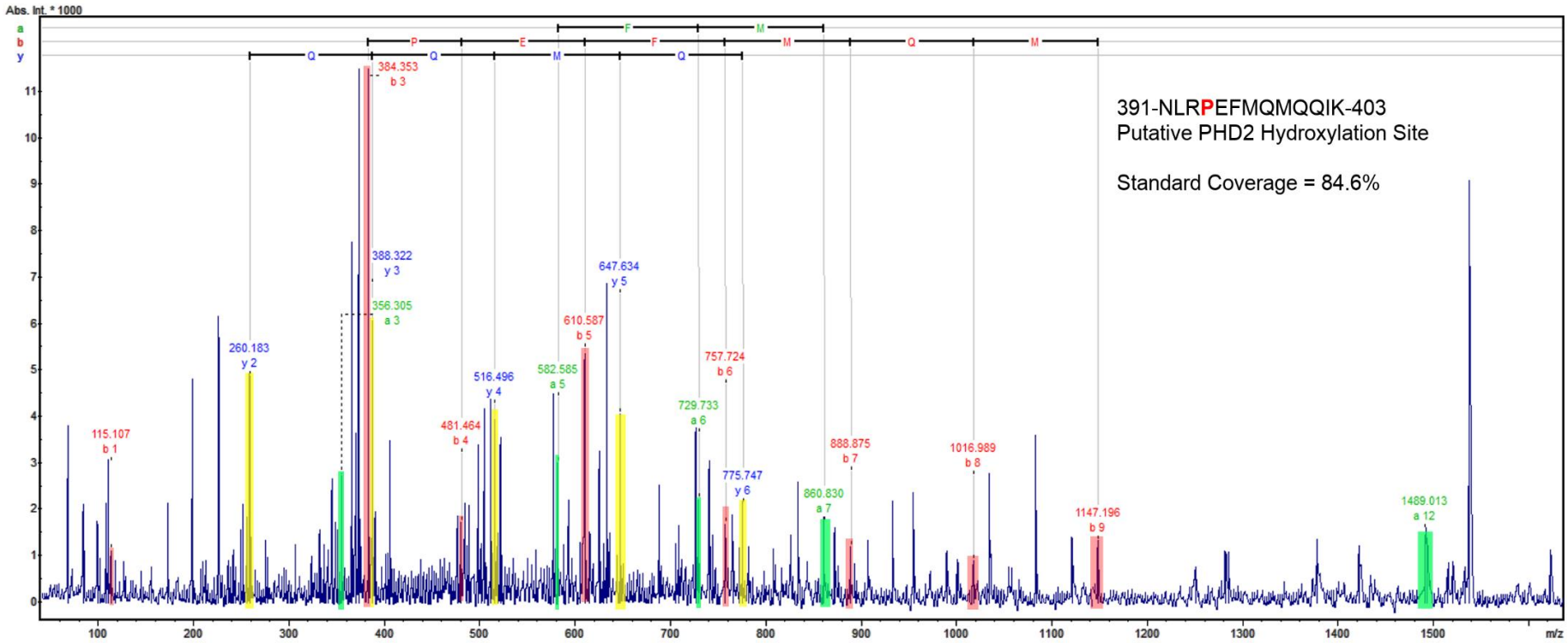
Figure 4-9: MALDI-MS spectrum with putative assignment of the peptides obtained from the trypsin/Lys-C digest of recombinant hTRPA1. * = peptide containing residue that may be hydroxylated by a HIF-hydroxylase.

Utilising the MALDI-LIFT method, it was possible to select a narrow mass range (± 10 Da) of the ions present in the MALDI-MS spectrum for MS/MS. This enabled the sequencing of the peptide that produced the peak at 1663 Da which was putatively assigned as containing non-hydroxylated P394 from the PHD2 hydroxylation site. Analysis of the MS/MS in biotools (Bruker) confirmed the assignment of the non-hydroxylated peptide with a standardised coverage of 84.6% (fig. 4-10).

The assignment of non-hydroxylated P394 is consistent with the results published by Cockman *et al.* that TRPA1 is not a substrate for hydroxylation by PHD2.[9] Unfortunately, the peptide corresponding to the FIH hydroxylation site could not be sequenced in the same way as the PHD2 hydroxylation site. Satisfactory MS/MS spectra could not be obtained *via* the MALDI-

LIFT method on account of the low intensity of the parent ion in the MALDI-MS spectrum. Thus, attempts were made to increase the amount of the peptide which was putatively assigned as containing N336 so that its identity could be verified. Two approaches were proposed: increasing the catalytic efficiency of the trysin/Lys-C mixture and attempting to separate the peptides from one another.

It has been reported that trypsin has improved activity in 80% aqueous acetonitrile (MeCN).[24] Additionally, it was hoped that the use of an organic solvent such as MeCN would help disrupt hydrophobic interactions in TRPA1 making the protease cleavage sites more accessible and thus increasing the efficiency of the cleavage. To separate the peptides from one another, increasing the signal to noise ratio in the MALDI-MS spectrum, the C18-ZipTip was eluted using different concentrations of aqueous MeCN, mimicking the reverse phase column chromatography aspect of an LCMS system.



	N	L	R	P	E	F	M	Q	M	Q	Q	I	K	Asn	Leu	Arg	Pro	Glu	Phe	Met	Gln	Met	Gln	Gln	Ile	Lys
Ion	1	2	3	4	5	6	7	8	9	1	1	1	1	1	2	3	4	5	6	7	8	9	10	11	12	13
a	N	L	R	P	E	F	M	Q	M	Q	Q	I	K	87.055	200.139	356.240	453.293	582.336	729.404	860.445	988.503	1119.544	1247.602	1375.661	1488.745	1616.840
b	N	L	R	P	E	F	M	Q	M	Q	Q	I	K	115.050	228.134	384.235	481.288	610.331	757.399	888.440	1016.498	1147.539	1275.597	1403.656	1516.740	1644.835
y	N	L	R	P	E	F	M	Q	M	Q	Q	I	K	147.113	260.197	388.255	516.314	647.355	775.413	906.454	1053.522	1182.565	1279.617	1435.718	1548.803	1662.845
	1	1	1	1	9	8	7	6	5	4	3	2	1	Lys	Ile	Gln	Gln	Met	Gln	Met	Phe	Glu	Pro	Arg	Leu	Asn

Figure 4-10: Assigned MALDI-LIFT spectra of peptide containing P394 which is reported to be hydroxylated by PHD2. MS/MS assigned using Bruker Biotools.

Each fraction was concentrated *in vacuo* before being resuspended in deionised water and co-spotted on a MALDI-MS target with matrix buffer. It was found that this method increased the coverage of hTRPA1 but not the intensity of the peptide containing N336 (fig. 4-11). Thus, it was not possible to unambiguously assign the identity of the peptide which was putatively assigned as the FIH hydroxylation site.

LCMS or LCMS/MS is employed in an attempt to better separate peptides from one another and the background, increasing the signal to noise ratio. LCMS also does not suffer from issues with matrix crystallisation which can occur with MALDI-MS. Given these advantages, it may have been possible to sequence the FIH hydroxylation site in addition to the PHD2 site. However, delays in sample processing on the departmental LCMS/MS system and subsequent delays due to the SARS-COV-2 pandemic meant that this system was unavailable for use.

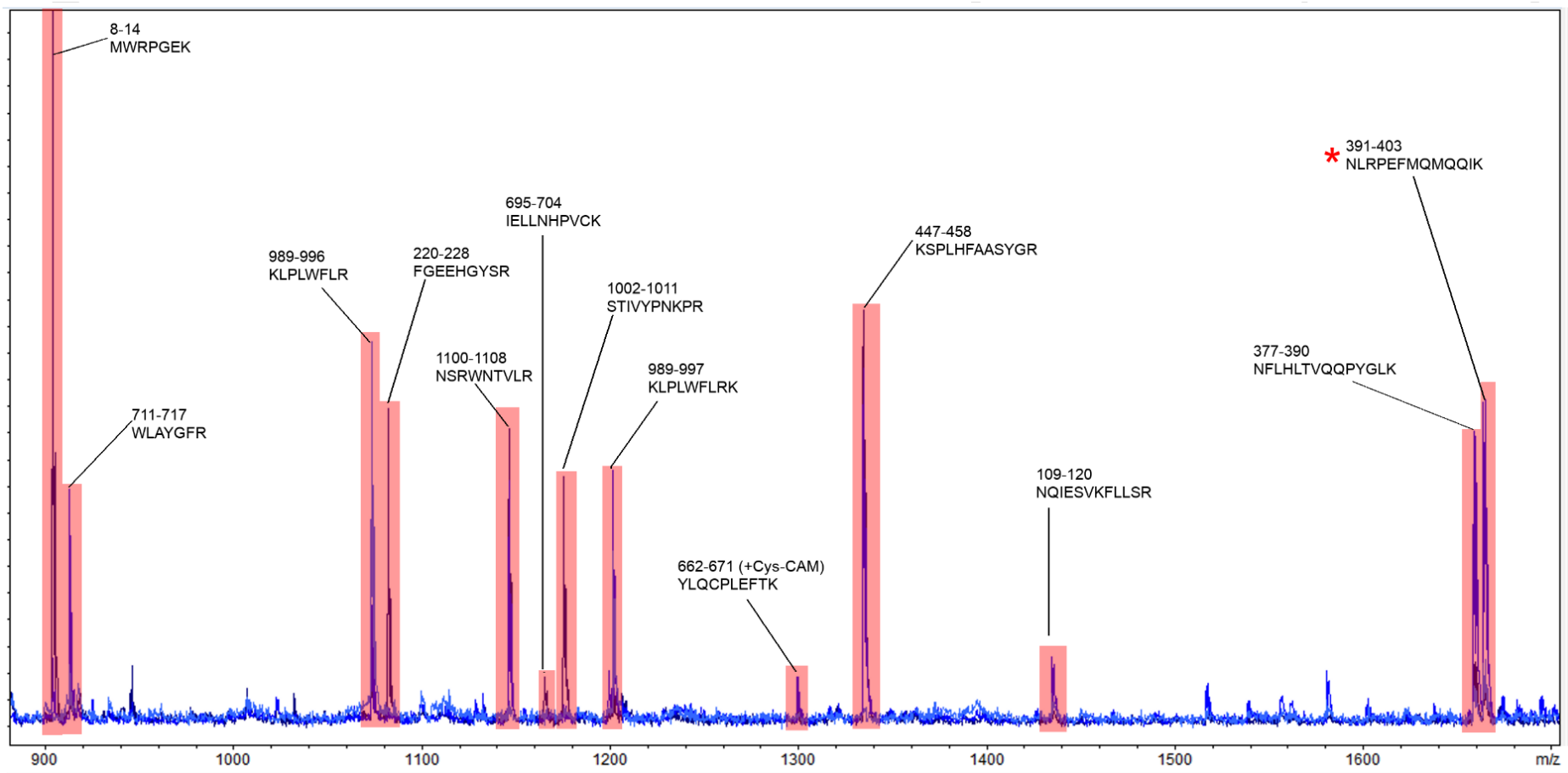


Figure 4-11: MALDI-MS spectrum of peptides arising from the digestion of TRPA1 with a mixture of trypsin and Lys-C in 80% aq. MeCN with putative assignment. * = peptide containing residue that may be hydroxylated by a HIF-hydroxylase.

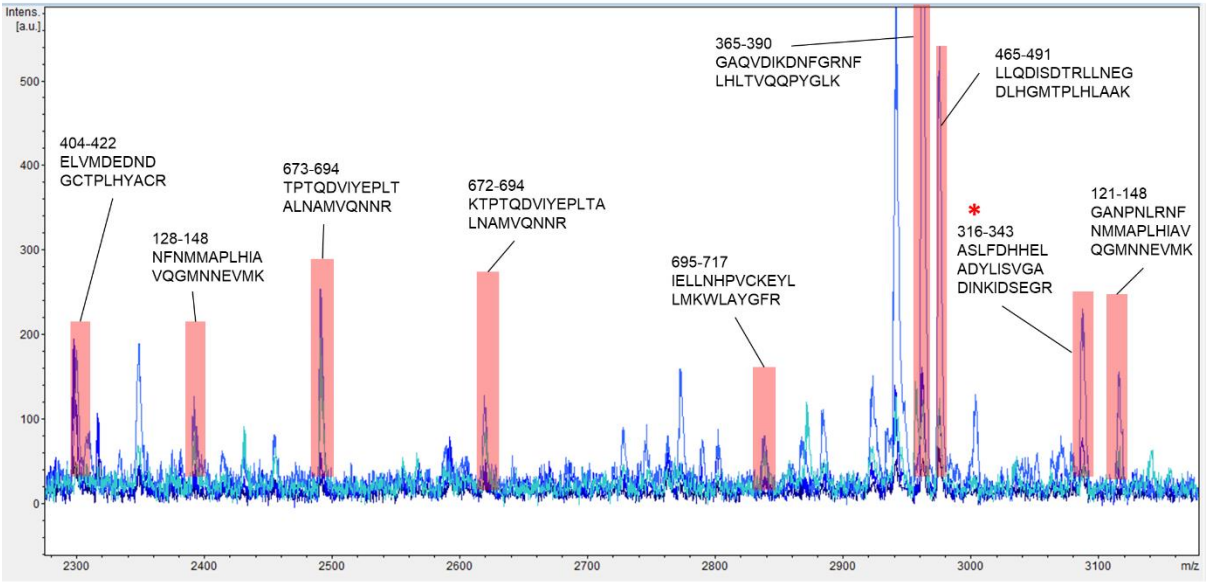


Figure 4-11 (continued): MALDI-MS spectrum of peptides arising from the digestion of TRPA1 with a mixture of trypsin and Lys-C in 80% aq. MeCN with putative assignment. * = peptide containing residue that may be hydroxylated by a HIF-hydroxylase.

4. Investigating the Effect of FIH Knockout on TRPA1 Sensitivity

It was envisaged that the sensitivity of TRPA1 to agonists in different hydroxylation states could be investigated through the expression of TRPA1 in an FIH knockout (KO) cell line. It was hoped that dose-response curves for several agonists could be recorded and the resulting EC_{50} could then be compared to those in the literature. It was envisaged that the intracellular calcium fluorescence assays performed at DSTL (see chapter III) could easily be implemented using other fluorescence plate readers such as the Pherastar FSX at the University of Oxford.

4.1 Subcloning

A plasmid encoding wild type (WT) hTRPA1 with the pcDNA3.1 backbone was a kind gift from Dr. Yihua Wang (University of Southampton). Neomycin, or G418, would typically be used as a selection antibiotic in the culture medium to select for cells that have been transfected

with the pcDNA3.1 plasmid on account of the *neo* gene which confers resistance to these antibiotics. However, HEK293T, which are transformed with a plasmid bearing the SV40 large T antigen (which improves plasmid copy numbers and hence protein expression) and *neo* genes, are naturally resistant to neomycin/G418. Colonies of FIH knock out (KO) HEK293T (a kind gift from Dr. Ya-Min Tian, Nuffield Department of Medicine, University of Oxford) expressing hTRPA1 could therefore not be selected using neomycin/G418. The pcDNA5/TO plasmid (a kind gift from Dr. Tom Durrant, Department of Chemistry, University of Oxford) contains a gene conferring resistance to the antibiotic hygromycin B (HB) which HEK293T are susceptible to. Therefore, the construction of a TRPA1-pcDNA5 plasmid would allow for transfection and subsequent selection of FIH KO HEK293T expressing TRPA1. Therefore, subcloning the TRPA1 gene into an empty pcDNA5 vector was attempted.

Polymerase chain reaction (PCR) primers were synthesised corresponding to the T7 promoter region (forward) and bovine growth hormone region (reverse) at the 5′- and 3′- end of the TRPA1 gene respectively (see Chapter VI §3.2 for sequences). PCR was carried out using Q5 polymerase which yielded a band at approximately 3500 base pairs, corresponding to the full length TRPA1 insert (fig. 4-12). The band was excised and purified using a GeneJet gel extract kit (Thermo Scientific). The TRPA1 insert was then digested with XhoI and HindIII before being purified with a GeneJet PCR purification kit (Thermo Scientific). The pcDNA5/TO plasmid was also digested using XhoI and HindIII to generate a linearised plasmid with complimentary overhangs to the insert. The insert and linearised plasmid were incubated in 3:1, 5:1 and 10:1 molar ratios at rt before being treated with T4 ligase to generate a circular plasmid. The reaction mixture was then used to transform XL-10 Gold *E. coli*. Selection of colonies was achieved through the use of the selection antibiotic ampicillin. Individual colonies were picked and a small amount used to inoculate 2-YT media and a small amount used to carry out PCR with primers corresponding to the BamhI and XhoI sites, thus confirming the

presence of the TRPA1 insert (fig. 4-12). After growing the individual colonies, the plasmids were extracted using a GeneJet plasmid purification kit (Thermo Scientific) and sent for sequencing to confirm the integrity of the TRPA1 gene.

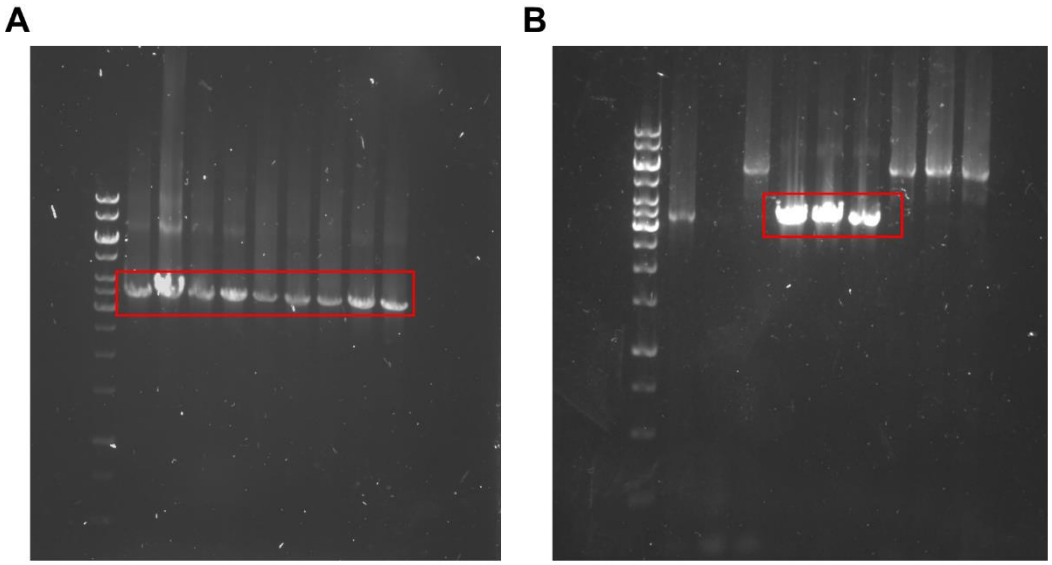


Figure 4-12: Subcloning of the TRPA1 gene into the pcDNA5/TO vector. A) 0.5% agarose gel showing the products of amplification of the TRPA1 gene from the TRPA1-pcDNA3 plasmid. Band contained in the red square was excised and extracted. **B)** 0.5% agarose gel showing the products of bacterial colony PCR. Bands at the approximate number of base pairs of the TRPA1 gene (red square) were taken to be a positive result for the presence of the TRPA1 gene.

4.2 Transfection Optimisation

Having generated a TRPA1-pcDNA5 plasmid which confers resistance to the antibiotic hygromycin B (HB), the FIH KO HEK293T cell line was tested for susceptibility to HB. The cell line was incubated with differing concentrations of HB for 10 days with fresh media containing HB added every 48 h. After 10 days the cell viability was assayed using CellTiter-Glo (Promega) which showed that the most effective concentration of HB for killing FIH KO HEK293T was 300 µg/mL (fig. 4-13).

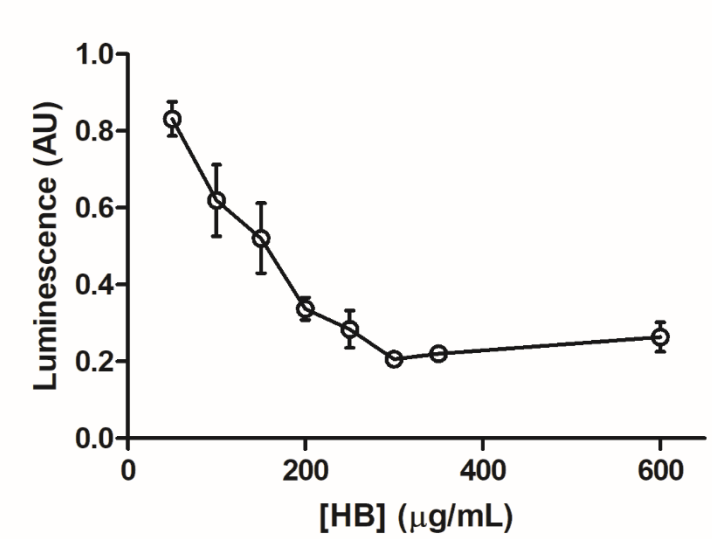


Figure 4-13: Kill curve of FIH KO HEK293T cells with different concentrations of Hygromycin B (HB). Error bars represent SD of triplicate measurements.

In order to optimise transfection conditions using the TRPA1-pcDNA5 plasmid and the transfection reagent Fugene-HD (Promega), FIH KO HEK293T cells were seeded in 96-plates at 5 x 10^3 cells per well in Dulbecco's modified Eagle's medium (DMEM) supplemented with 2 mM L-glutamine and 10% fetal bovine serum (FBS). The cells were allowed to grow for 48 h at 37 °C before being transfected with different concentrations of transfection reagent and plasmid. After a further 24 h incubation at 37 °C, the media was replaced with selective media containing 300 µg/mL HB and cell viability was measured after 10 days, with the selective media being replaced every 48 h.

In addition to testing viability, the transfected cells were assayed for the ability to respond to the known TRPA1 agonist cinnamaldehyde.[25] After 24 h incubation with the transfection mix, the media was aspirated and Fluo-4 direct assay kit buffer (Invitrogen) was added. The cells were incubated with the dye for 30 min at 37 °C before a 1:1 mixture of DMEM and Earle's salts supplemented with 20 mM Hepes pH 7.4 were added. The cells were then loaded into a

plate reader and the background fluorescence intensity measured. 250 µM cinnamaldehyde was then applied to the cells using an electronic pipette and the change in fluorescence intensity measured. Combining the viability and fluorescence intensity data allowed for the determination of the optimum composition of the transfection mixture (fig. 4-14).

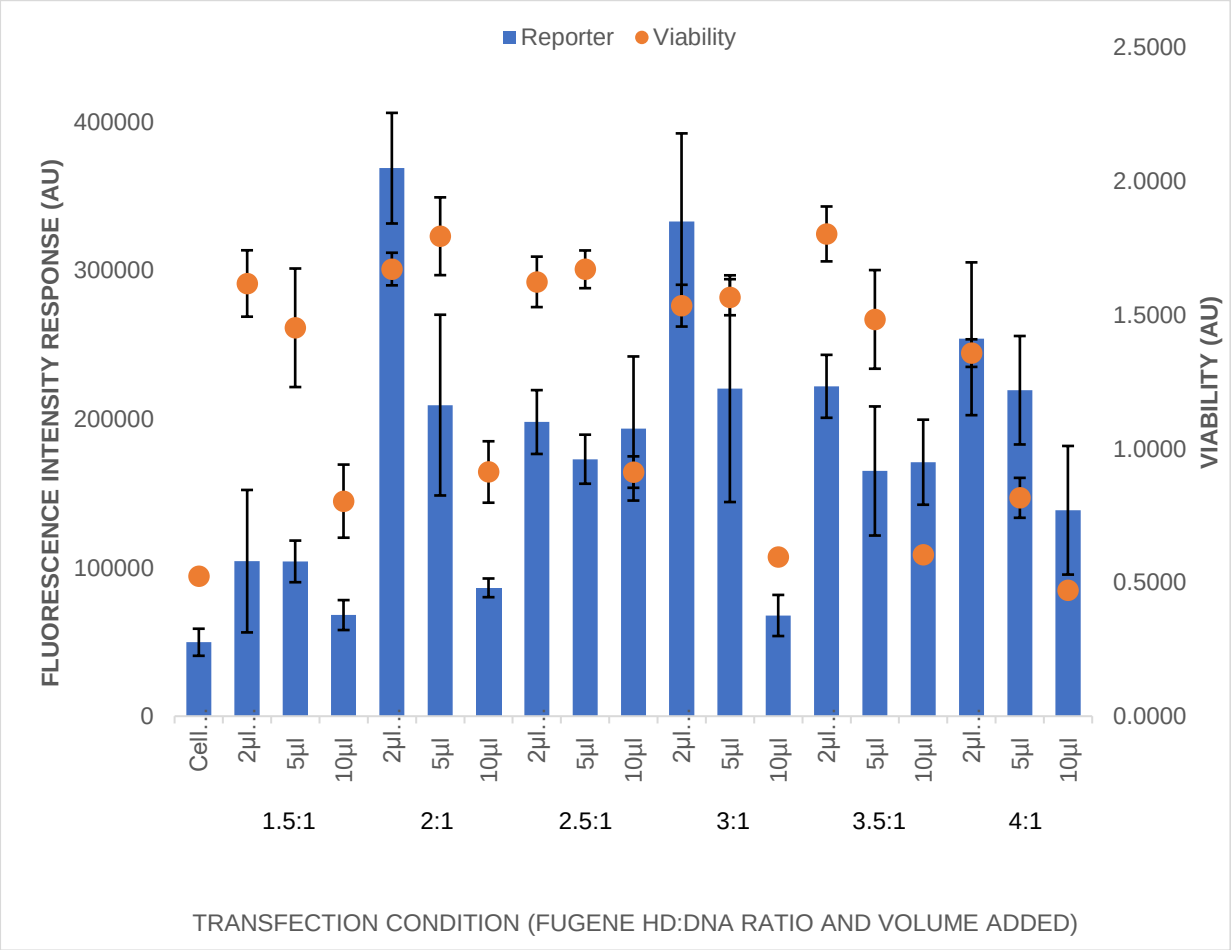


Figure 4-14: Optimisation of transfection mixture composition *via* viability in selective media (orange circle) and fluorescence intensity changes on application of 250 µM cinnamaldehyde (blue rectange). Error bars represent SD of triplicate measurement.

4.3 Stable Cell Line Generation

To try and minimise errors in experiments due to differences in levels of recombinant protein production and to ensure that TRPA1 is properly regulated *via* PTMs or other mechanisms, FIH KO HEK293T cells stably expressing the TRPA1 gene were sought. An initial pool of cells in a T75 flask was transfected with 360 µL (*i.e.* 2µL per 100 µL medium) of a 2:1 mixture

of fugene HD to the TRPA1-pcDNA5 plasmid which was previously determined to be the optimum conditions. The cells were incubated with the transfection mixture for 24 h at 37 °C before the media was replaced selective media containing 300 µg/mL HB. The selective media was replaced every 48 h for 10 days before the HB resistant cells were harvested using trypsin. Separately, conditioned media was collected from non-transfected FIH KO HEK293T cells and diluted with fresh media containing a final concentration of 150 µg/mL HB. The harvested cells were diluted with the conditioned media mix and seeded into 96-well plates at a rate of 0.5 cells per well. The plates were then incubated at 37 °C for 2 weeks before being checked for clonal cell growth. Investigation of the wells using a microscope revealed several wells that contained monoclonal cells (fig. 4-15).

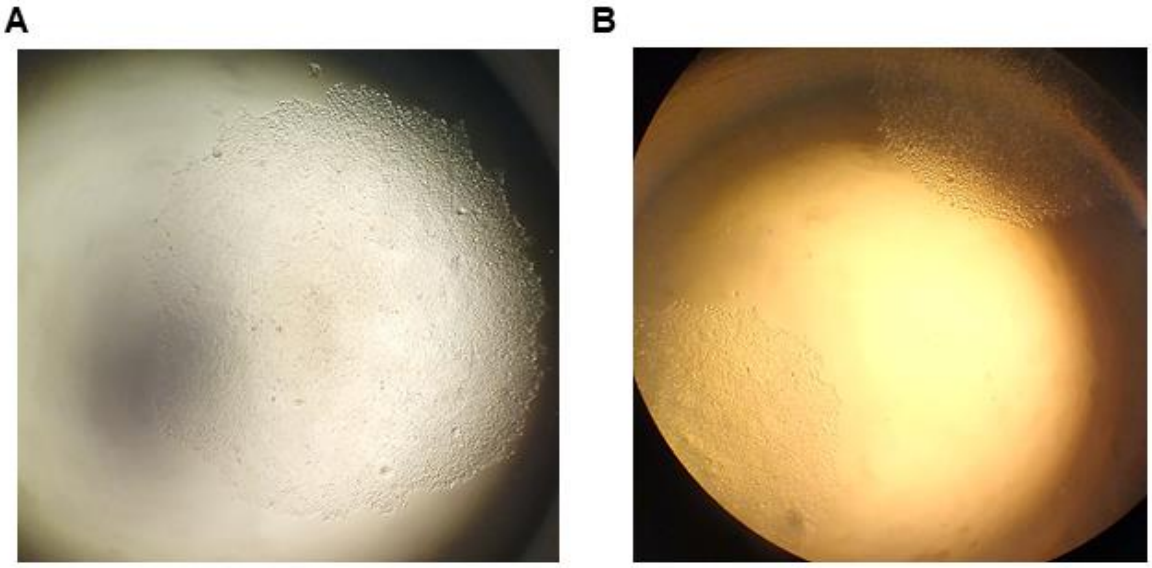


Figure 4-15: Microscopy images of individual wells from plate generating stable cell lines. A) A representative well containing monoclonal growth. **B)** A well showing polyclonal growth (*ie* two or more colonies).

The medium from wells containing monoclonal cells was aspirated and the cells detached with trypsin before being seeded into T25 flasks. The monoclonal cells were allowed to reach confluence before being passaged and a small amount cryo-preserved in vapor-phase liquid nitrogen storage. This process was repeated for 5 different wells to give 5 different monoclonal

cell lines which had integrated the TRPA1-pcDNA5 plasmid into their genome. Each cell line was maintained in complete medium containing 150 µg/mL HB. Each cell line was tested for a response to the TRPA1 agonists cinnamaldehyde and allyl-isothiocyanate using the Fluo-4 based assay described above. However, none of the isolated cell lines elicited a robust response upon treatment with the agonists. Plasmids must be linearised before incorporation into the genome of the host cell,[26] therefore it may be possible that linearisation of the TRPA1-pcDNA5 vector occurred within the TRPA1 gene insert, preventing the expression of functional TRPA1.

5. Conclusions and Further Work

The flexible loop of TRPA1 ankyrin 8 has been shown to be a substrate for isolated recombinant FIH using the peptide TRPA1(322-348) (see fig. 4-5). Sequencing of the hydroxylated peptide using MALDI-LIFT to generate MS/MS unambiguously shows that hydroxylation of the peptide occurs on N336 (see fig. 4-6). This supports previous work by Dr. T. Leissing and Dr. R. Hopkinson showing the FIH-mediated hydroxylation of the peptide TRPA1(313-339) and a crystal structure of the same peptide bound to FIH (see fig. 4-8).[12] Kinetic analysis of the hydroxylation of the TRPA1(322-348) peptide gives a K_m value of 66.9 µM (see fig. 4-7). This value in a similar range to the K_m value for HIF-1α C-terminal transactivation domain (CTAD) of 50 – 100 µM.[27,28] However, comparing the catalytic efficiency (k_{cat}/K_m) for the two peptides gives an approximately 16-fold difference, with TRPA1(322-348) (0.026 µM min^{-1}) appearing to be a worse substrate than HIF-1α CTAD (0.42 ± 0.08 µM min^{-1}).[29] Other substrates such as Notch-1 show an inverse correlation between peptide length and K_m so caution should be taken when comparing K_m values for peptide substrates.[30] Further investigation is needed using the isolated TRPA1 ankyrin repeat domain to discern hydroxylation and kinetic parameters at the protein level.

Attempts to examine the hydroxylation state of TRPA1 which was recombinantly produced in Sf9 cells yielded an apparently interesting result in that the protein appeared not to be hydroxylated by endogenous levels of PHD2 or FIH (see §3). This finding appears to contrast with the claim that TRPA1 is a substrate of PHD2[7] while appearing to support the claim that it is not.[9] The finding also appears to conflict with the evidence showing that the loop of TRPA1 ankyrin 8 is a substrate for recombinant FIH. However, the level of TRPA1 production in Sf9 cells may be too high, or the production of the HIF hydroxylases may be too low or non-existent for substantial hydroxylation to occur. There may additionally be differences in selectivity between the HIF hydroxylases produced in Sf9 cells compared to those in humans. Little is known about the presence or absence of the HIF hydroxylases in *Spodoptera frugiperda*, the insect from which Sf9 cells originate.[31] Other invertebrates such as *Drosophila melanogaster* and *Caenorhabditis elegans* express an orthologue of the PHDs known as Hypoxia-inducible factor prolyl hydroxylase (EGL-9) which hydroxylates HIF-1, although little is known about its function outside of HIF regulation.[32] However, the work described in this section has shown it is possible to obtain coverage of the two putative TRPA1 hydroxylation sites.

6. Chapter IV Bibliography

(1) Mahon, P. C.; Hirota, K.; Semenza, G. L. FIH-1: A Novel Protein That Interacts with HIF-1α and VHL to Mediate Repression of HIF-1 Transcriptional Activity. *Genes and Development* **2001**, *15* (20), 2675–2686. https://doi.org/10.1101/gad.924501.

(2) Cockman, M. E.; Webb, J. D.; Kramer, H. B.; Kessler, B. M.; Ratcliffe, P. J. Proteomics-Based Identification of Novel Factor Inhibiting Hypoxia-Inducible Factor (FIH) Substrates Indicates Widespread Asparaginyl Hydroxylation of Ankyrin Repeat Domain-Containing Proteins. *Molecular and Cellular Proteomics* **2009**, *8* (3), 535–546. https://doi.org/10.1074/mcp.M800340-MCP200.

(3) Coleman, M. L.; McDonough, M. A.; Hewitson, K. S.; Coles, C.; Mecinović, J.; Edelmann, M.; Cook, K. M.; Cockman, M. E.; Lancaster, D. E.; Kessler, B. M.; Oldham, N. J.; Ratcliffe, P. J.; Schofield, C. J. Asparaginyl Hydroxylation of the Notch Ankyrin Repeat Domain by Factor Inhibiting Hypoxia-Inducible Factor. *Journal of Biological Chemistry* **2007**, *282* (33), 24027–24038. https://doi.org/10.1074/jbc.M704102200.

(4) Schnatwinkel, C.; Christoforidis, S.; Lindsay, M. R.; Uttenweiler-Joseph, S.; Wilm, M.; Parton, R. G.; Zerial, M. The Rab5 Effector Rabankyrin-5 Regulates and Coordinates Different Endocytic Mechanisms. *PLoS Biology* **2004**, *2* (9), e261. https://doi.org/10.1371/journal.pbio.0020261.

(5) Kuriyama, H.; Asakawa, S.; Minoshima, S.; Maruyama, H.; Ishii, N.; Ito, K.; Gejyo, F.; Arakawa, M.; Shimizu, N.; Kuwano, R. Characterization and Chromosomal Mapping of a Novel Human Gene, ANKHZN. *Gene* **2000**, *253* (2), 151–160. https://doi.org/10.1016/S0378-1119(00)00247-X.

(6) Nagarajan, Y.; Rychkov, G. Y.; Peet, D. J. Modulation of TRP Channel Activity by Hydroxylation and Its Therapeutic Potential. *Pharmaceuticals*. MDPI AG June 1, 2017, p 35. https://doi.org/10.3390/ph10020035.

(7) Takahashi, N.; Kuwaki, T.; Kiyonaka, S.; Numata, T.; Kozai, D.; Mizuno, Y.; Yamamoto, S.; Naito, S.; Knevels, E.; Carmeliet, P.; Oga, T.; Kaneko, S.; Suga, S.; Nokami, T.; Yoshida, J.; Mori, Y. TRPA1 Underlies a Sensing Mechanism for O2. *Nature Chemical Biology* **2011**, *7* (10), 701–711. https://doi.org/10.1038/nchembio.640.

(8) Miyake, T.; Nakamura, S.; Zhao, M.; So, K.; Inoue, K.; Numata, T.; Takahashi, N.; Shirakawa, H.; Mori, Y.; Nakagawa, T.; Kaneko, S. Cold Sensitivity of TRPA1 Is Unveiled by the Prolyl Hydroxylation Blockade-Induced Sensitization to ROS. *Nature Communications* **2016**, *7* (1), 12840. https://doi.org/10.1038/ncomms12840.

(9) Cockman, M. E.; Lippl, K.; Tian, Y. M.; Pegg, H. B.; Figg, W. D.; Abboud, M. I.; Heilig, R.; Fischer, R.; Myllyharju, J.; Schofield, C. J.; Ratcliffe, P. J. Lack of Activity of Recombinant HIF Prolyl Hydroxylases (PHDs) on Reported Non-HIF Substrates. *eLife* **2019**, *8*. https://doi.org/10.7554/eLife.46490.

(10) Jaakkola, P.; Mole, D. R.; Tian, Y. M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; von Kriegsheim, A.; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J. Targeting of HIF-α to the von Hippel-Lindau Ubiquitylation Complex by O2-Regulated Prolyl Hydroxylation. *Science* **2001**, *292* (5516), 468–472. https://doi.org/10.1126/science.1059796.

(11) Karttunen, S.; Duffield, M.; Scrimgeour, N. R.; Squires, L.; Lim, W. L.; Dallas, M. L.; Scragg, J. L.; Chicher, J.; Dave, K. A.; Whitelaw, M. L.; Peers, C.; Gorman, J. J.; Gleadle, J. M.; Rychkov, G. Y.; Peet, D. J. Oxygen-Dependent Hydroxylation by FIH Regulates the TRPV3 Ion Channel. *Journal of cell science* **2015**, *128* (2), 225–231. https://doi.org/10.1242/jcs.158451.

(12) Leissing, T. Studies on FIH Interactions, DPhil Thesis, 2017, Department of Chemistry, University of Oxford.

(13) Li, J.; Mahajan, A.; Tsai, M. D. Ankyrin Repeat: A Unique Motif Mediating Protein-Protein Interactions. *Biochemistry*. Biochemistry December 26, 2006, pp 15168–15178. https://doi.org/10.1021/bi062188q.

(14) Paulsen, C. E.; Armache, J. P.; Gao, Y.; Cheng, Y.; Julius, D. Structure of the TRPA1 Ion Channel Suggests Regulatory Mechanisms. *Nature* **2015**, *520* (7548), 511–517. https://doi.org/10.1038/nature14367.

(15) Kelly, L.; McDonough, M. a; Coleman, M. L.; Ratcliffe, P. J.; Schofield, C. J. Asparagine Beta-Hydroxylation Stabilizes the Ankyrin Repeat Domain Fold. *Molecular bioSystems* **2009**, *5* (1), 52–58. https://doi.org/10.1039/b815271c.

(16) Mosavi, L. K.; Minor, D. L.; Peng, Z. yu. Consensus-Derived Structural Determinants of the Ankyrin Repeat Motif. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99* (25), 16029–16034. https://doi.org/10.1073/pnas.252537899.

(17) Fields, G. B. Introduction to Peptide Synthesis. *Current Protocols in Protein Science* **2001**, *26* (1). https://doi.org/10.1002/0471140864.ps1801s26.

(18) Karttunen, S.; Duffield, M.; Scrimgeour, N. R.; Squires, L.; Lim, W. L.; Dallas, M. L.; Scragg, J. L.; Chicher, J.; Dave, K. A.; Whitelaw, M. L.; Peers, C.; Gorman, J. J.; Gleadle, J. M.; Rychkov, G. Y.; Peet, D. J. Oxygen-Dependent Hydroxylation by FIH Regulates the TRPV3 Ion Channel. *Journal of Cell Science* **2015**, *128* (2), 225–231. https://doi.org/10.1242/jcs.158451.

(19) Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A. A Novel MALDI LIFT-TOF/TOF Mass Spectrometer for Proteomics. *Analytical and Bioanalytical Chemistry* **2003**, *376* (7), 952–965. https://doi.org/10.1007/s00216-003-2057-0.

(20) Campuzano, I. D. G.; San Miguel, T.; Rowe, T.; Onea, D.; Cee, V. J.; Arvedson, T.; McCarter, J. D. High-Throughput Mass Spectrometric Analysis of Covalent Protein-Inhibitor Adducts for the Discovery of Irreversible Inhibitors. *Journal of Biomolecular Screening* **2016**, *21* (2), 136–144. https://doi.org/10.1177/1087057115621288.

(21) Elkins, J. M.; Hewitson, K. S.; McNeill, L. A.; Seibel, J. F.; Schlemminger, I.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J. Structure of Factor-Inhibiting Hypoxia-Inducible Factor (HIF) Reveals Mechanism of Oxidative Modification of HIF-1α. *Journal of Biological Chemistry* **2003**, *278* (3), 1802–1806. https://doi.org/10.1074/jbc.C200644200.

(22) Rey, M.; Mrázek, H.; Pompach, P.; Novák, P.; Pelosi, L.; Brandolin, G.; Forest, E.; Havlíček, V.; Man, P. Effective Removal of Nonionic Detergents in Protein Mass Spectrometry, Hydrogen/Deuterium Exchange, and Proteomics. *Analytical Chemistry* **2010**, *82* (12), 5107–5116. https://doi.org/10.1021/ac100171m.

(23) Saveliev, S.; Bratz, M.; Zubarev, R.; Szapacs, M.; Budamgunta, H.; Urh, M. Trypsin/Lys-C Protease Mix for Enhanced Protein Mass Spectrometry Analysis. *Nature Methods* **2013**, *10* (11), i–ii. https://doi.org/10.1038/nmeth.f.371.

(24) Strader, M. B.; Tabb, D. L.; Hervey, W. J.; Pan, C.; Hurst, G. B. Efficient and Specific Trypsin Digestion of Microgram to Nanogram Quantities of Proteins in Organic-Aqueous Solvent Systems. *Analytical Chemistry* **2006**, *78* (1), 125–134. https://doi.org/10.1021/ac051348l.

(25) Macpherson, L. J.; Dubin, A. E.; Evans, M. J.; Marr, F.; Schultz, P. G.; Cravatt, B. F.; Patapoutian, A. Noxious Compounds Activate TRPA1 Ion Channels through Covalent Modification of Cysteines. *Nature* **2007**, *445* (7127), 541–545. https://doi.org/10.1038/nature05544.

(26) Stuchbury, G.; Münch, G. Optimizing the Generation of Stable Neuronal Cell Lines via Pre-Transfection Restriction Enzyme Digestion of Plasmid DNA. *Cytotechnology* **2010**, *62* (3), 189–194. https://doi.org/10.1007/s10616-010-9273-1.

(27) Tarhonskaya, H.; Hardy, A. P.; Howe, E. A.; Loik, N. D.; Kramer, H. B.; McCullagh, J. S. O.; Schofield, C. J.; Flashman, E. Kinetic Investigations of the Role of Factor Inhibiting Hypoxia-Inducible Factor (FIH) as an Oxygen Sensor. *Journal of Biological Chemistry* **2015**, *290* (32), 19726–19742. https://doi.org/10.1074/jbc.M115.653014.

(28) Wilkins, S. E.; Hyvärinen, J.; Chicher, J.; Gorman, J. J.; Peet, D. J.; Bilton, R. L.; Koivunen, P. Differences in Hydroxylation and Binding of Notch and HIF-1α Demonstrate Substrate Selectivity for Factor Inhibiting HIF-1 (FIH-1). *International Journal of Biochemistry and Cell Biology* **2009**, *41* (7), 1563–1571. https://doi.org/10.1016/j.biocel.2009.01.005.

(29) Hangasky, J. A.; Taabazuing, C. Y.; Martin, C. B.; Eron, S. J.; Knapp, M. J. The Facial Triad in the α-Ketoglutarate Dependent Oxygenase FIH: A Role for Sterics in Linking Substrate Binding to O2 Activation. *Journal of Inorganic Biochemistry* **2017**, *166*, 26–33. https://doi.org/10.1016/j.jinorgbio.2016.10.007.

(30) Cockman, M. E.; Webb, J. D.; Ratcliffe, P. J. FIH-Dependent Asparaginyl Hydroxylation of Ankyrin Repeat Domain-Containing Proteins. In *Annals of the New York Academy of Sciences*; Blackwell Publishing Inc., 2009; Vol. 1177, pp 9–18. https://doi.org/10.1111/j.1749-6632.2009.05042.x.

(31) Vaughn, J. L.; Goodwin, R. H.; Tompkins, G. J.; McCawley, P. The Establishment of Two Cell Lines from the Insect Spodoptera Frugiperda (Lepidoptera; Noctuidae). *In Vitro* **1977**, *13* (4), 213–217. https://doi.org/10.1007/BF02615077.

(32) Wang, L.; Cui, S.; Ma, L.; Kong, L.; Geng, X. Current Advances in the Novel Functions of Hypoxia-Inducible Factor and Prolyl Hydroxylase in Invertebrates. *Insect Molecular Biology* **2015**, *24* (6), 634–648. https://doi.org/10.1111/imb.12189.

(33) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Molecular Systems Biology* **2011**, *7* (1), 539. https://doi.org/10.1038/msb.2011.75.

Chapter V:
Summary & Conclusions

Nociception is the neural process by which damaging or potentially damaging (noxious) stimuli are encoded to produce the psychological state of pain. Neurons and proteins which are sensitive towards noxious stimuli are known as nociceptors and nocisensors.[1] Transient receptor potential (TRP) channels form a superfamily of ion channels which are responsible for sensing an array of stimuli.[2,3] Of the TRP channels, TRPA1, TRPV1 and TRPM8 are the most physiologically relevant to human nociception.[4] TRPA1 is a thermosensitive, calcium permeable ion channel that is present in small diameter neurons in humans and other animals.[5,6] These nerves have free endings in most organs, skin, lungs and trachea.[7–9] TRPA1 is activated by a wide variety of ligands, most notably lipophilic electrophiles including α,β-unsaturated aldehydes such as cinnamaldehyde and 4-hydroxynonenal, and the tear gases **CR** and **CS** (see Chapter I fig. 1-3).[10–12] TRPA1 is also activated by small, hydrophilic electrophiles such as allyl isothiocyanate (**AITC**), iodoacetamide (**IAA**) formaldehyde (HCHO), furthermore, knock-out of the TRPA1 gene in mice causes a complete loss in sensitivity to HCHO.[13–15]

It is thought that TRPA1 activation is the result of covalent modification of C621 in the electrophile binding domain (EBD) on the channel's N-terminal cytosolic domain.[13,16] However, at the outset of this project, the reason for the high reactivity of C621 towards electrophiles and the mechanism by which electrophile binding causes TRPA1 to transition to an open state were unclear. The work described in Chapter II explores the reaction between cysteine containing peptides and formaldehyde (HCHO) via NMR. It was found that cysteine reacts to form *S*-hydroxymethylcysteine (supporting previous work in the group on the reaction between amino acids and HCHO)[17] and that this functional group has a characteristic chemical shift in the ^{13}C-NMR of the peptides tested (see Chapter II §2). It was also found that HCHO reacts with peptides containing a N-terminal proline to form a 5,5-bicyclic structure with a methylene group bridging the proline amine and amide nitrogen atoms (see Chapter II fig. 2-

11). It was also found that glutathione (GSH) acts as a "sink" for electrophiles such as HCHO and is able to remove HCHO from peptide-HCHO adducts (see Chapter II fig. 2-13 and fig. 2-16).

In order to model the reaction between electrophiles and TRPA1, a synthetic construct (BK3) containing the EBD of TRPA1 supported by ankyrin repeats to improve protein solubility and mimic the native environment of the EBD was designed and expressed from *E. coli* (see Chapter II §4). Evaluation of BK3 as a model for the EBD of TRPA1 showed it incorporated one equivalent of the electrophiles **CS** and **IAA** in the presence of a large excess of the electrophiles (see Chapter II fig. 2-22). Attempts to determine the location of the residue undergoing covalent modification were unsuccessful. Although **CR** binding to BK3 was not observed by denaturing mass spectrometry, binding was observed by NMR (see Chapter II fig. 2-23). Attempts to determine the structure of BK3 by crystallography were abandoned as a result of the publication of high resolution cryo-electron microscopy (cryo-EM) structures of TRPA1 in 2020 showing a change in conformation of the EBD and pore region upon electrophile binding to C621.[15,18] This allowed for the authors of the studies to propose why C621 is reactive towards electrophiles and how its modification leads to TRPA1 opening (see Chapter I §2.3).[15,18]

TRPA1 has gathered attention as a potential target for pain-relieving medications.[19–24] **GRC-17536**, a TRPA1 antagonist, has undergone phase IIa clinical trials for painful diabetic neuropathy. However, the binding mode of **GRC-17536**, and the related compound **HC-030031**, to TRPA1 is unknown. The work in Chapter III describes the synthesis of a library of photolabile compounds based on **GRC-17536** and **HC-030031** (see Chapter III §2.2).

Additionally, small molecule probes for TRPA1 with a pulldown (biotin), fluorescent (sulforhodamine B) and alkyne tag were synthesised. The synthesised compounds showed poor solubility in aqueous buffers, thus a medium-throughput solubility assay was implemented using liquid chromatography coupled mass spectrometry (LCMS). It was found that the formation of a sodium salt of the compounds improved their solubility, allowing their use in activity assays (see Chapter III fig. 3-17). Each of the compounds were then evaluated for their ability to inhibit cinnamaldehyde evoked TRPA1 activity using a Ca^{2+} cellular fluorescence assay at the Defence, Science and Technology Laboratory (DSTL) in Porton Down as part of an industrial placement. It was found that compound **3-60** had the lowest IC_{50} value (2.17 – 3.09 µM) of the compounds evaluated (see Chapter III fig. 3-21 and fig. 3-22).

3-60 was taken forward for cryo-EM studies by Dr. Mariana Grieben and Dr. Ashley Pike at the Structural Genomics Consortium (SGC, University of Oxford). The resulting cryo-EM structure of **3-60** bound to human TRPA1 showed that **3-60** occupies a lipid binding site between the pre-S1 helix and the S4-S5 linker in the transmembrane region of TRPA1 (see Chapter III §5). It was found that the theophylline ring of 3-60 interacts with the sidechain of W711 via π-stacking interactions and the amide group of **3-60** with the sidechain of N855 via hydrogen bonding (see Chapter III fig.3-23). As a result of this structure, a number of modifications to **3-60** and **GRC-17536** were proposed which may strengthen these interactions and improve the binding, efficacy and solubility of this class of compound (see Chapter III fig. 3-25).

TRPA1 displays a "U"-shaped sensitivity towards agonists under changing O_2 concentration with minimum sensitivity at normoxia (*i.e.* 20% O_2).[25] It is proposed that this is the result of

hydroxylation of P394 by PHD2,[25] however, this is disputed by work showing TRPA1 is not a substrate for the PHDs.[26] The work in Chapter IV builds on previous work in the group showing that a peptide derived from TRPA1 ankyrin 8 is a substrate for hydroxylation by FIH. Ankyrin loop peptides were designed using the amino acid sequence of TRPA1 such that site 1 and site 2 of FIH would be bound to the peptide and screened against FIH for activity. It was found that only the TRPA1 ankyrin 8 loop peptide (containing TRPA1 residues 322-348) was hydroxylated by FIH (see Chapter IV fig. 4-5). Fragmentation mass spectrometry was performed on the hydroxylated TRPA1 ankyrin 8 loop peptide and it was found that the asparagine corresponding to N336 was hydroxylated by FIH. Kinetic analysis of the hydroxylation of this peptide (kindly performed by Dr. A. Tumber) showed that it is a comparatively poor substrate for FIH when compared to HIF-1α peptides[27–29] and 1CA consensus ankyrin as shown by its low catalytic efficiency (k_{cat}/K_m) of 0.026 μM^{-1} min^{-1} (see Chapter IV fig. 4-6 and fig. 4-7).

Recombinant hTRPA1 expressed from Sf9 cells was examined for evidence of hydroxylation by the HIF hydroxylases *via* digestion of the protein and sequencing of the resulting peptides by MS/MS (see Chapter IV §3). Using this approach, no evidence was accrued that P394 is hydroxylated, in accordance with studies showing that isolated TRPA1 is not a substrate for recombinant PHD2.[26] N855 was putatively assigned as not being hydroxylated because a low intensity peak in the mass spectrum was observed which corresponds to the mass of a peptide containing non-hydroxylated N855. Attempts to increase the intensity of this peptide and confirm its identity through sequencing were unsuccessful. In order to further investigate the effect of FIH-mediated hydroxylation on TRPA1 sensitivity, attempts were made to create FIH-knockout cell lines which stably express the TRPA1 gene (see Chapter IV §4). It was found that the monoclonal cell lines that were generated were completely insensitive to the TRPA1

agonist cinnamaldehyde, suggesting that linearisation of the TRPA1-pcDNA5/TO plasmid occurred within the TRPA1 gene before incorporation into the cell line genome, preventing the expression of TRPA1. Further work investigating the hydroxylation state of TRPA1 produced from human cell lines and treatment of the isolated ankyrin repeat domain of TRPA1 with FIH is required to confirm that TRPA1 is substrate of FIH at the protein level.

Overall, the studies presented in this thesis have provided chemical insights into the reaction of cysteine containing peptides with electrophiles and biochemical insights into ligand binding to TRPA1 to modulate its activity. Additionally, evidence at the peptide level has shown that TRPA1 may be a substrate of FIH-mediated hydroxylation. The key findings presented in this thesis are:

- HCHO reacts on cysteine in peptides to give the *S*-hydroxymethyl adduct. GSH is able to remove HCHO from this adduct to give HMG and the unmodified peptide (see Chapter II §2-3).
- Synthesis of a series of photolabile TRPA1 antagonists which show efficacy in inhibiting cinnamaldehyde evoked TRPA1 activity (see Chapter III §2 and 4).
- A cryo-electron microscopy structure of **3-60** bound to human TRPA1 (see Chapter III §5).
- A peptide derived from the TRPA1 ankyrin 8 loop is a substrate for FIH (see Chapter IV §2).

Chapter V Bibliography:

(1) Dubin, A. E.; Patapoutian, A. Nociceptors: The Sensors of the Pain Pathway. *Journal of Clinical Investigation* **2010**, *120* (11), 3760–3772. https://doi.org/10.1172/JCI42843.

(2) Jardín, I.; López, J. J.; Diez, R.; Sánchez-Collado, J.; Cantonero, C.; Albarrán, L.; Woodard, G. E.; Redondo, P. C.; Salido, G. M.; Smani, T.; Rosado, J. A. TRPs in Pain Sensation. *Frontiers in Physiology* **2017**, *8* (JUN), 392. https://doi.org/10.3389/fphys.2017.00392.

(3) Clapham, D. E. TRP Channels as Cellular Sensors. *Nature* **2003**, *426* (6966), 517–524. https://doi.org/10.1038/nature02196.

(4) Julius, D. TRP Channels and Pain. *Annual Review of Cell and Developmental Biology* **2013**, *29*, 355–384. https://doi.org/10.1146/annurev-cellbio-101011-155833.

(5) Kobayashi, K.; Fukuoka, T.; Obata, K.; Yamanaka, H.; Dai, Y.; Tokunaga, A.; Noguchi, K. Distinct Expression of TRPM8, TRPA1, and TRPV1 MRNAs in Rat Primary Afferent Neurons with Aδ/C-Fibers and Colocalization with Trk Receptors. *Journal of Comparative Neurology* **2005**, *493* (4), 596–606. https://doi.org/10.1002/cne.20794.

(6) Bautista, D. M.; Jordt, S. E.; Nikai, T.; Tsuruda, P. R.; Read, A. J.; Poblete, J.; Yamoah, E. N.; Basbaum, A. I.; Julius, D. TRPA1 Mediates the Inflammatory Actions of Environmental Irritants and Proalgesic Agents. *Cell* **2006**, *124* (6), 1269–1282. https://doi.org/10.1016/j.cell.2006.02.023.

(7) Gerhold, K. A.; Bautista, D. M. TRPA1: Irritant Detector of the Airways. *The Journal of Physiology* **2008**, *586* (14), 3303–3303. https://doi.org/10.1113/jphysiol.2008.158030.

(8) Cattaruzza, F.; Spreadbury, I.; Miranda-Morales, M.; Grady, E. F.; Vanner, S.; Bunnett, N. W. Transient Receptor Potential Ankyrin-1 Has a Major Role in Mediating Visceral Pain in Mice. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **2010**, *298* (1), G81–G91. https://doi.org/10.1152/ajpgi.00221.2009.

(9) Vandewauw, I.; Owsianik, G.; Voets, T. Systematic and Quantitative MRNA Expression Analysis of TRP Channel Genes at the Single Trigeminal and Dorsal Root Ganglion Level in Mouse. *BMC Neuroscience* **2013**, *14* (1), 21. https://doi.org/10.1186/1471-2202-14-21.

(10) Lindsay, C. D.; Green, C.; Bird, M.; Jones, J. T. A.; Riches, J. R.; McKee, K. K.; Sandford, M. S.; Wakefield, D. A.; Timperley, C. M. Potency of Irritation by Benzylidenemalononitriles in Humans Correlates with TRPA1 Ion Channel Activation. *Royal Society Open Science* **2015**, *2* (1), 140160. https://doi.org/10.1098/rsos.140160.

(11) Taylor-Clark, T. E.; McAlexander, M. A.; Nassenstein, C.; Sheardown, S. A.; Wilson, S.; Thornton, J.; Carr, M. J.; Undem, B. J. Relative Contributions of TRPA1 and TRPV1 Channels in the Activation of Vagal Bronchopulmonary C-Fibres by the Endogenous Autacoid 4-Oxononenal. *The Journal of Physiology* **2008**, *586* (14), 3447–3459. https://doi.org/10.1113/jphysiol.2008.153585.

(12) Taylor-Clark, T. E.; Undem, B. J.; MacGlashan, D. W.; Ghatta, S.; Carr, M. J.; McAlexander, M. A. Prostaglandin-Induced Activation of Nociceptive Neurons via Direct Interaction with Transient Receptor Potential A1 (TRPA1). *Molecular Pharmacology* **2007**, *73* (2), 274–281. https://doi.org/10.1124/mol.107.040832.

(13) Macpherson, L. J.; Dubin, A. E.; Evans, M. J.; Marr, F.; Schultz, P. G.; Cravatt, B. F.; Patapoutian, A. Noxious Compounds Activate TRPA1 Ion Channels through Covalent Modification of Cysteines. *Nature* **2007**, *445* (7127), 541–545. https://doi.org/10.1038/nature05544.

(14) McNamara, C. R.; Mandel-Brehm, J.; Bautista, D. M.; Siemens, J.; Deranian, K. L.; Zhao, M.; Hayward, N. J.; Chong, J. A.; Julius, D.; Moran, M. M.; Fanger, C. M. TRPA1 Mediates Formalin-Induced Pain. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104* (33), 13525–13530. https://doi.org/10.1073/pnas.0705924104.

(15) Zhao, J.; Lin King, J. v.; Paulsen, C. E.; Cheng, Y.; Julius, D. Irritant-Evoked Activation and Calcium Modulation of the TRPA1 Receptor. *Nature* **2020**, *585* (7823), 141–145. https://doi.org/10.1038/s41586-020-2480-9.

(16) Bahia, P. K.; Parks, T. A.; Stanford, K. R.; Mitchell, D. A.; Varma, S.; Stevens, S. M.; Taylor-Clark, T. E. The Exceptionally High Reactivity of Cys 621 Is Critical for Electrophilic Activation of the Sensory Nerve Ion Channel TRPA1. *Journal of General Physiology* **2016**, *147* (6), 451–465. https://doi.org/10.1085/jgp.201611581.

(17) Kamps, J. J. A. G.; Hopkinson, R. J.; Schofield, C. J.; Claridge, T. D. W. How Formaldehyde Reacts with Amino Acids. *Communications Chemistry* **2019**, *2* (1). https://doi.org/10.1038/s42004-019-0224-2.

(18) Suo, Y.; Wang, Z.; Zubcevic, L.; Hsu, A. L.; He, Q.; Borgnia, M. J.; Ji, R. R.; Lee, S. Y. Structural Insights into Electrophile Irritant Sensing by the Human TRPA1 Channel. *Neuron* **2020**, *105* (5), 882-894.e5. https://doi.org/10.1016/j.neuron.2019.11.023.

(19) Mukhopadhyay, I.; Kulkarni, A.; Aranake, S.; Karnik, P.; Shetty, M.; Thorat, S.; Ghosh, I.; Wale, D.; Bhosale, V.; Khairatkar-Joshi, N. Transient Receptor Potential Ankyrin 1 Receptor Activation In Vitro and In Vivo by Pro-Tussive Agents: GRC 17536 as a Promising Anti-Tussive Therapeutic. *PLoS ONE* **2014**, *9* (5), e97005. https://doi.org/10.1371/journal.pone.0097005.

(20) Eid, S. R.; Crown, E. D.; Moore, E. L.; Liang, H. A.; Choong, K. C.; Dima, S.; Henze, D. A.; Kane, S. A.; Urban, M. O. HC-030031, a TRPA1 Selective Antagonist, Attenuates Inflammatory- and Neuropathy-Induced Mechanical Hypersensitivity. *Molecular Pain* **2008**, *4*, 48. https://doi.org/10.1186/1744-8069-4-48.

(21) Pryde, D. C.; Marron, B. E.; West, C. W.; Reister, S.; Amato, G.; Yoger, K.; Antonio, B.; Padilla, K.; Cox, P. J.; Turner, J.; Warmus, J. S.; Swain, N. A.; Omoto, K.; Mahoney, J. H.; Gerlach, A. C. Discovery of a Series of Indazole TRPA1 Antagonists. *ACS Medicinal Chemistry Letters* **2017**, *8* (6), 666–671. https://doi.org/10.1021/acsmedchemlett.7b00140.

(22) Skerratt, S. Recent Progress in the Discovery and Development of TRPA1 Modulators. In *Progress in Medicinal Chemistry*; Elsevier B.V., 2017; Vol. 56, pp 81–115. https://doi.org/10.1016/bs.pmch.2016.11.003.

(23) Rooney, L.; Vidal, A.; D'Souza, A.-M.; Devereux, N.; Masick, B.; Boissel, V.; West, R.; Head, V.; Stringer, R.; Lao, J.; Petrus, M. J.; Patapoutian, A.; Nash, M.; Stoakley, N.; Panesar, M.; Verkuyl, J. M.; Schumacher, A. M.; Petrassi, H. M.; Tully, D. C. Discovery, Optimization, and Biological Evaluation of 5-(2-(Trifluoromethyl)Phenyl)Indazoles as a Novel Class of Transient Receptor Potential A1 (TRPA1) Antagonists. *Journal of Medicinal Chemistry* **2014**, *57* (12), 5129–5140. https://doi.org/10.1021/jm401986p.

(24) Terrett, J. A.; Chen, H.; Shore, D. G.; Villemure, E.; Larouche-Gauthier, R.; Déry, M.; Beaumier, F.; Constantineau-Forget, L.; Grand-Maître, C.; Lépissier, L.; Ciblat, S.; Sturino, C.; Chen, Y.; Hu, B.; Lu, A.; Wang, Y.; Cridland, A. P.; Ward, S. I.; Hackos, D. H.; Reese, R. M.; Shields, S. D.; Chen, J.; Balestrini, A.; Riol-Blanco, L.; Lee, W. P.; Liu, J.; Suto, E.; Wu, X.; Zhang, J.; Ly, J. Q.; La, H.; Johnson, K.; Baumgardner, M.; Chou, K. J.; Rohou, A.; Rougé, L.; Safina, B. S.; Magnuson, S.; Volgraf, M. Tetrahydrofuran-Based Transient Receptor Potential Ankyrin 1 (TRPA1) Antagonists: Ligand-Based Discovery, Activity in a Rodent Asthma Model, and Mechanism-of-Action via Cryogenic Electron Microscopy. *Journal of Medicinal Chemistry* **2021**, *64* (7), 3843–3869. https://doi.org/10.1021/acs.jmedchem.0c02023.

(25) Takahashi, N.; Kuwaki, T.; Kiyonaka, S.; Numata, T.; Kozai, D.; Mizuno, Y.; Yamamoto, S.; Naito, S.; Knevels, E.; Carmeliet, P.; Oga, T.; Kaneko, S.; Suga, S.; Nokami, T.; Yoshida, J.; Mori, Y. TRPA1 Underlies a Sensing Mechanism for O2. *Nature Chemical Biology* **2011**, *7* (10), 701–711. https://doi.org/10.1038/nchembio.640.

(26) Cockman, M. E.; Lippl, K.; Tian, Y. M.; Pegg, H. B.; Figg, W. D.; Abboud, M. I.; Heilig, R.; Fischer, R.; Myllyharju, J.; Schofield, C. J.; Ratcliffe, P. J. Lack of Activity of Recombinant HIF Prolyl Hydroxylases (PHDs) on Reported Non-HIF Substrates. *eLife* **2019**, *8*. https://doi.org/10.7554/eLife.46490.

(27) Hangasky, J. A.; Taabazuing, C. Y.; Martin, C. B.; Eron, S. J.; Knapp, M. J. The Facial Triad in the α-Ketoglutarate Dependent Oxygenase FIH: A Role for Sterics in Linking Substrate Binding to O2 Activation. *Journal of Inorganic Biochemistry* **2017**, *166*, 26–33. https://doi.org/10.1016/j.jinorgbio.2016.10.007.

(28) Tarhonskaya, H.; Hardy, A. P.; Howe, E. A.; Loik, N. D.; Kramer, H. B.; McCullagh, J. S. O.; Schofield, C. J.; Flashman, E. Kinetic Investigations of the Role of Factor Inhibiting Hypoxia-Inducible Factor (FIH) as an Oxygen Sensor. *Journal of Biological Chemistry* **2015**, *290* (32), 19726–19742. https://doi.org/10.1074/jbc.M115.653014.

(29) Wilkins, S. E.; Hyvärinen, J.; Chicher, J.; Gorman, J. J.; Peet, D. J.; Bilton, R. L.; Koivunen, P. Differences in Hydroxylation and Binding of Notch and HIF-1α Demonstrate Substrate Selectivity for Factor Inhibiting HIF-1 (FIH-1). *International Journal of Biochemistry and Cell Biology* **2009**, *41* (7), 1563–1571. https://doi.org/10.1016/j.biocel.2009.01.005.

Chapter VI:
Materials & Methods

1. Organic Synthesis

1.1 General Procedures

All reactions involving moisture-sensitive reagents or anhydrous solvents were carried out under a nitrogen or argon atmosphere using standard vacuum line techniques and glassware that was dried in an oven and cooled under nitrogen before use. Anhydrous solvents were dispensed from a MBraun MB SPS 5 Benchtop system or used as supplied. Water was purified using an Elix® UV-10 or MilliQ (MQ H_2O) system. All other reagents were used as supplied (analytical or HPLC grade) without prior purification. Organic layers were dried over Na_2SO_4 unless specified.

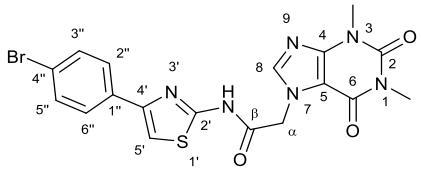
Thin layer chromatography was performed on aluminium plates coated with 60 F254 silica. Plates were visualised using UV light (254 & 365 nm) and 1% aq. $KMnO_4$. Flash column chromatography was performed using a Biotage Isolera automated flash column chromatography platform using Biotage KP-SNAP-Sil columns. Reverse phase flash column chromatography (RP-FCC) was performed using the same system utilising Biotage RP-C18-Sil columns. Semi-preparative HPLC was performed using a Shimadzu semi-preparative HPLC system utilising an ACE 5 C18 (250 x 21.2 mm) column utilising the following solvents: MQ H_2O + 0.1% CF_3CO_2H (v/v, solvent A), MeCN + 0.1% CF_3CO_2H (v/v, solvent B). Strong Cation Exchange – Solid Phase Extraction (SCX-SPE) was carried out using an Isolute Flash SCX-2 tube.

IR spectra were recorded using a Bruker Tensor 27 FT-IR spectrometer equipped with an attenuated total reflection (ATR) crystal in the solid state or as a thin film. Wavenumbers (ν_{max}) are reported in cm^{-1}. NMR spectra were recorded using Bruker Avance spectrometers in the deuterated solvent stated as an internal deuterium lock. Spectra were recorded at rt unless otherwise stated. Assignments given correspond with the numbering system as drawn, where assignments are not given it was not possible to assign the resonances due to overlapping signals or ambiguity in assignment. The multiplicity of each signal is indicated by: s (singlet), br. s (broad singlet), d (doublet), t (triplet), q (quartet), p (pentet), app. p. (apparent pentet), sept (septet), dd (doublet of doublets), ddd (doublet of doublets of doublets) or m (multiplet).

Low-resolution mass spectra were recorded using an Agilent Technologies 1260 Infinity LC-MS system fitted with a 6120 Quadrupole mass spectrometer. LC-MS was performed using the same system utilising a Merc Chromolith C18 (100 x 4.2 mm) column. Accurate mass measurements were run using a Thermo Exactive High-Resolution Orbitrap FTMS.

1.2 Procedures and Characterisation Data

_N_-(4′-(4′′-Bromophenyl)thiazol-2′-yl)-α-(theophyllin-7-yl)acetamide 26

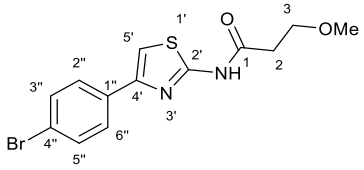


Step 1: $SOCl_2$ (541 µL, 7.42 mmol) was added to a solution of theophylline-7-acetic acid (1.82 g, 7.62 mmol) in anhydrous DMF (60 mL) at rt. The resulting solution was allowed to stir for 1 h at rt and was used without further modification.

Step 2: pyridine (5.00 mL, 62.0 mmol) and **25** (510 mg, 2.00 mmol) were added sequentially to the crude solution from step 1 at rt and the resulting solution was allowed to stir for 16 h at rt. The solution was diluted with CH_2Cl_2 (200 mL), then washed sequentially with 1M aq. HCl (3 x 200 mL), a saturated solution of $NaHCO_3$ (3 x 200 mL), and brine (5 x 300 mL). The solid precipitate was collected by filtration to give **26** as an off-white solid which was used without further purification. Additionally, the organic layer was dried and concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 50% EtOAc in *c*-hexane to 25% MeOH in EtOAc) gave **26** as an off-white solid (combined yield: 760 mg, 80%); ν_{max} (ATR, cm^{-1}) 1650 (NC=O), 1638 (NC=O); δ_H (400 MHz, DMSO-d_6): 3.18 (3H, s, N(1)*Me*), 3.46 (3H, s, N(3)*Me*), 5.35 (2H, s, C(α)H_2), 7.60 – 7.68 (2H, m, C(3′′)*H*, C(5′′)*H*), 7.73 (1H, s, C(5′)*H*), 7.81 – 7.90 (2H, m, C(2′′)*H*, C(6′′)*H*), 8.09 (1H, s, C(8)*H*), 12.79 (1H, s, C(2′)N*H*); δ_C (101 MHz, DMSO-d_6): 27.44 (N(1)*Me*), 29.47 (N(3)*Me*), 48.19 (*C*(α)), 106.39 (*C*(5)), 109.24 (*C*(5′)), 120.94 (*C*(4′′)), 127.68 (*C*(2′′), *C*(6′′)), 131.69 (*C*(3′′), *C*(5′′)),

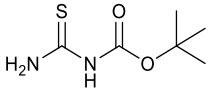
133.31 (*C*(1″)), 143.66 (*C*(8)), 147.78 (*C*(4′)), 148.01 (*C*(4)), 150.99 (*C*(2)), 154.47 (*C*(6)), 157.60 (*C*(2′)), 165.70 (C(β)); *m/z* (API-ES$^+$) 475 ([M(^{79}Br)+H]$^+$, 100%); HRMS (ESI$^+$) $C_{18}H_{16}O_3N_6{}^{79}BrS^+$ calculated 475.0183; found 475.0188.

_N_-(4′-(4″-Bromophenyl)thiazol-2′-yl)-3-methoxypropanamide 32



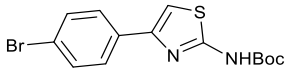
(1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxide hexafluorophosphate (HATU, 570 mg, 1.50 mmol) was added to a solution of **25** (255 mg, 1.00 mmol), N,N-diisopropylethylamine (DIPEA, 261 µL, 1.50 mmol) and 3-methoxypropanoic acid (142 µL, 1.50 mmol) in DMF (3.0 mL) at rt and the resulting solution was heated to 100 °C for 96 h. The reaction mixture was allowed to cool, diluted with EtOAc (40 mL), then washed sequentially with a saturated solution of $NaHCO_3$ (40 mL), H_2O (3 x 40 mL), brine (40 mL), dried and concentrated *in vacuo*, to give **32** as a brown solid (289 mg, 85%); ν_{max} (ATR, cm^{-1}) 1681 (NC=O); δ_H (400 MHz, DMSO-d_6) 2.69 (2H, t, *J* 6.0, C(2)H_2), 3.24 (3H, s, O*Me*), 3.64 (2H, t, *J* 6.0, C(3)H_2), 7.62 (2H, m, C(2″)*H*, C(6″)*H*), 7.69 (1H, s, C(5′)*H*), 7.84 (2H, m, C(3″)*H*, C(5″)*H*), 12.29 (1H, s, C(2′)N*H*); δ_C (101 MHz, DMSO-d_6): 35.61 (*C*(2)), 57.93 (O*Me*), 67.63 (*C*(3)), 108.76 (*C*(5′)), 120.81 (*C*(4″)), 127.66 (*C*(3″), *C*(5″)), 131.65 (*C*(2″), (*C*(6″)), 133.51 (*C*(1″)), 147.55 (*C*(4′)), 157.97 (*C*(2′)), 169.72 (*C*(1)); *m/z* 341 ([M(^{79}Br)+H]$^+$, 50%), 343 ([M(^{81}Br)+H]$^+$, 50%); HRMS (ESI$^+$) $C_{13}H_{14}O_2N_2{}^{79}BrS^+$ calculated 340.9954; found 340.9955.

_N_-Butyloxycarbonyl-thiourea 34 [1]



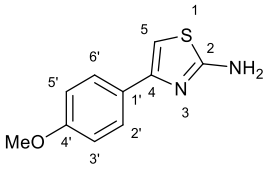
Following a literature procedure,[1] NaH (60% dispersion in mineral oil, 184 mg, 7.67 mmol) was added to a solution of thiourea (152 mg, 2.00 mmol) in anhydrous THF (15 mL) at 0 °C and the resulting suspension was allowed to warm to rt and was stirred for 1 h. The reaction mixture was diluted with CH_2Cl_2 (25 mL), then the layers were separated. The organic layer was washed sequentially with a saturated solution of $NaHCO_3$ (25 mL), water (25 mL), then dried, concentrated *in vacuo*, then triturated with *n*-heptane to give **34** as an off white solid (249 mg, 71%) which was used without further purification; δ_H (400 MHz, $CDCl_3$) 1.61 (9H, s); *m/z* (API-ES$^+$) 176 ([M+H]$^+$, 100%); all other data were in accordance with the literature values.[1]

_N_-Butyloxycarbonyl-2-amino-4-(4′-bromophenyl)-thiazole 35 [1]



Following a literature procedure,[1] **34** (249 mg, 1.41 mmol) was added to a solution of 2,4′-dibromoacetophenone (435 mg, 1.57 mmol) in acetone (20 mL) at rt and the resulting solution was allowed to stir for 16 h. The resulting precipitate was filtered and washed with acetone. Purification by flash column chromatography (eluent: 1% MeOH in CH_2Cl_2) gave **35** as a white solid (167 mg, 33%); δ_H (400 MHz, DMSO-d_6): 1.49 (9H, s), 7.56 – 7.67 (3H, m), 7.78 – 7.85 (2H, m); *m/z* 355 ([M(^{79}Br)+H]$^+$, 50%), 357 ([M(^{81}Br)+H]$^+$, 50%); all other data were in accordance with the literature values.[1]

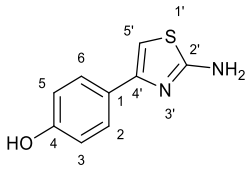
4-(4′-Methoxyphenyl)thiazol-2-amine 43 [2]



Following a literature procedure,[2] thiourea (3.21 g, 42.2 mmol) was added to a solution of **42** (9.30 g, 40.6 mmol) in EtOH (80 mL) at rt and the resulting mixture was heated to reflux for 2 h. The reaction mixture was allowed to cool to rt and was concentrated *in vacuo*. The resulting white precipitate was

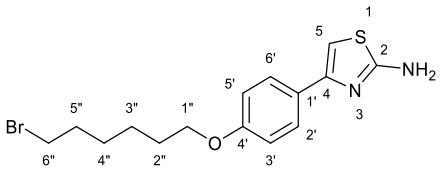
suspended in H_2O/saturated $NaHCO_3$ (30/70 v/v, 50 mL) and stirred at rt for 1 h. The precipitate was collected by filtration, dried and concentrated *in vacuo* to give **43** as a yellow solid (9.60 g, apparent quant.) which was used without further purification; ν_{max} (ATR, cm^{-1}) 3437 (C(2)N-H), 1234 (C(4′)-O); δ_H (400 MHz, DMSO-d_6) 3.77 (3H, O*Me*), 6.88 (1H, s, C(5)*H*), 6.92 – 6.98 (2H, m, C(2′)*H*, C(6′)*H*), 7.58 (2H, br. s, NH_2), 7.65 – 7.74 (2H, m, C(3′)*H*, C(5′)*H*); δ_C (101 MHz, DMSO-d_6) 55.16 (O*Me*), 99.71 (*C*(5)), 113.98 (*C*(2′), *C*(6′)), 125.92 (*C*(4)), 126.98 (*C*(3′), *C*(5′)), 146.51 (*C*(1′)), 158.98 (*C*(4′)), 168.71 (*C*(2)); *m/z* (API-ES+) 207 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{10}H_{11}ON_2S^+$ calculated 207.0587; found 207.0587.

4-(2′-Aminothiazol-4′-yl)phenol 44 [2]



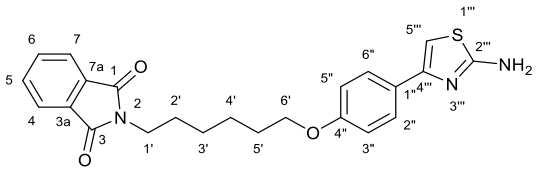
Following a literature procedure,[2] BBr_3 (72.3 mL, 72.3 mmol, 1.0 M in CH_2Cl_2) was added dropwise to a solution of **43** (5.00 g, 24.2 mmol) in anhydrous CH_2Cl_2 (130 mL) at 0 °C. The resulting solution was allowed to warm to rt, then stirred for 48 h. A saturated solution of $NaHCO_3$ (50 mL) was added, then the layers separated. The organic layer was washed with brine (100 mL), then dried and concentrated *in vacuo* to give **44** as a brown solid (3.76 g, 81%) which was used without further purification; ν_{max} (ATR, cm^{-1}) 3448 (C(2′)N-H), 1263 (C(4)-O); δ_H (400 MHz, DMSO-d_6) 6.65 – 6.78 (3H, m, C(5′)*H*, C(3)*H*, C(5)*H*), 6.94 (2H, br. s, NH_2), 7.52 – 7.65 (2H, m, C(2)*H*, C(6)*H*), 9.44 (1H, s, O*H*); δ_C (101 MHz, DMSO-d_6) 98.40 (*C*(5′), 115.13 (*C*(3), *C*(5)), 126.39 (*C*(1)), 126.89 (*C*(2), *C*(6)), 150.12 (*C*(4′)), 156.73 (*C*(4)), 167.94 (*C*(2′)); *m/z* (API-ES^+) 193 ($[M+H]^+$); HRMS (ESI^+) $C_9H_9ON_2S^+$ calculated 193.0430; found 193.0431.

4-(4′-((6″-Bromohexyl)oxy)phenyl)thiazol-2-amine 45



1,6-Dibromohexane (1.60 mL, 10.4 mmol) was added to a suspension of **44** (1.00 g, 5.20 mmol) and K_2CO_3 (2.54 g, 18.4 mmol) in acetone (10 mL) at rt and the resulting mixture was heated to 70 °C for 24 h. The reaction mixture was allowed to cool, concentrated *in vacuo*, diluted with water (100 mL), then extracted with EtOAc (4 x 75 mL). The combined organic layers were washed with brine (100 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (30% EtOAc in *c*-hexane) gave **45** as a yellow solid (1.14 g, 62%); ν_{max} (ATR, cm^{-1}) 3438 (C(2)N-H), 2943 (C-H), 1255 (C(4′)-O); δ_H (400 MHz, DMSO-d_6) 1.44 (4H, app. p, *J* 4.0, C(3″)H_2, C(4″)H_2), 1.71 (2H, ddd, *J* 10.0, 6.5, 4.0, C(2″)H_2), 1.76 – 1.88 (2H, m, C(5″)H_2), 3.54 (2H, t, *J* 7.0, C(6″)H_2), 3.96 (2H, t, *J* 6.5, C(1″)H_2), 6.81 (1H, s, C(5)*H*), 6.86 – 6.95 (2H, m, C(3′)*H*, C(5′)*H*), 6.99 (2H, br. s, NH_2), 7.65 – 7.74 (2H, m, C(2′)*H*, C(6′)*H*); δ_C (101 MHz, DMSO-d_6) 24.69 (*C*(3″) or *C*(4″)), 27.30 (*C*(3″) or *C*(4″)), 28.55 (*C*(2″)), 32.19 (*C*(5″)), 35.13 (*C*(6″)), 67.29 (*C*(1″)), 99.25 (*C*(5)), 114.29 (*C*(3′), *C*(5′)), 126.80 (*C*(2′), *C*(6′)), 127.69 (*C*(1′)), 149.62 (*C*(4)), 157.92 (*C*(4′)), 168.04 (*C*(2)); *m/z* (API-ES^+) 355 ([M(^{79}Br)+H]$^+$, 50%), 357 ([M(^{81}Br)+H]$^+$, 50%); HRMS (ESI^+) $C_{15}H_{20}ON_2{}^{79}BrS^+$ calculated 355.0474; found 355.0475.

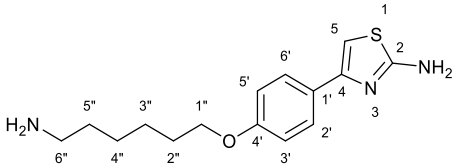
2-(6′-(4″-(2‴-Aminothiazol-4‴-yl)phenoxy)hexyl)isoindoline-1,3-dione 50



Potassium phthalimide (1.55 g, 8.38 mmol) was added to a solution of **45** (2.48 g, 6.98 mmol) in anhydrous DMF (55 mL) and the resulting mixture was heated to 70 °C for 8 h. The reaction mixture was allowed to cool, then diluted with EtOAc (200 mL) and H_2O (200 mL), then the layers were

separated. The organic layer was washed with H_2O (5 x 100 mL), brine (100 mL), then dried and concentrated *in vacuo* to give **50** as yellow solid (2.51 g, 85%) which was used without further purification; ν_{max} (ATR, cm^{-1}) 3432 (C(2‴)N-H), 2947 (C-H), 1703 (NC=O), 1252 (C(4″)-O); δ_H (400 MHz, DMSO-d_6) 1.26 - 1.39 (2H, m, C(3′)H_2), 1.40 – 1.50 (2H, m, C(4′)H_2), 1.62 (2H, app. quin, *J* 7.0, C(2′)H_2), 1.69 (2H, p, *J* 6.5, C(5′)H_2), 3.58 (2H, t, *J* 7.0, C(1′)H_2), 3.94 (2H, t, *J* 6.5, C(6′)H_2), 6.81 (1H, s, C(5‴)*H*), 6.88 (2H, m, C(3″)*H*, C(5″)*H*), 7.00 (2H, br. s, NH_2), 7.68 (2H, m, C(2″)*H*, C(6″)*H*), 7.79 - 7.90 (4H, m, C(4)*H*, C(5)*H*, C(6)*H*, C(7)*H*); δ_C (101 MHz, DMSO-d_6) 25.15 (*C*(4′)), 26.01 (*C*(3′)), 27.87 (*C*(2′)), 28.51 (*C*(5′)), 37.34 (*C*(1′)), 67.35 (*C*(6′)), 99.24 (*C*(5‴)), 114.27 (*C*(3′), *C*(5′)), 123.00 (*C*(4), *C*(7) or *C*(5), *C*(6)), 126.80 (*C*(2″), *C*(6″)), 127.77 (*C*(1″)), 131.61 (*C*(3a), *C*(7a)), 134.38 (*C*(4), *C*(7) or *C*(5), *C*(6)), 149.78 (*C*(4‴)), 157.93 (*C*(4″)), 167.98 (*C*(1), *C*(3)), 168.05 (*C*(2‴)); *m/z* (API-ES^+) 422 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{23}H_{24}O_3N_3S^+$ calculated 422.1533; found 422.1523.

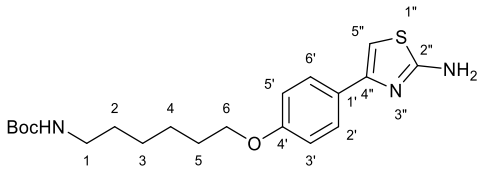
4-(4′-((6″-Aminohexyl)oxy)phenyl)thiazol-2-amine 51



Hydrazine monohydrate (1.44 mL, 29.8 mmol) was added to a solution of **50** (2.18 g, 5.17 mmol) in a mixture of MeOH and CH_2Cl_2 (50 mL, 1:1 v/v) at rt. The resulting solution was stirred for 16 h at rt before the precipitate was filtered off. The filtrate was diluted with $CHCl_3$ (100 mL), then washed sequentially with a saturated solution of $NaHCO_3$ (100 mL), brine (100 mL), then dried and concentrated *in vacuo* to give **51** as an off-white solid (1.21 g, 80%) which was used without further purification; ν_{max} (ATR, cm^{-1}) 3264 – 3132 (br, NH_2), 2981 (C-H), 1541 (N-H), 1249 (C(4′)-O); δ_H (400 MHz, DMSO-d_6) 1.27 - 1.47 (6H, m, C(4″)H_2, C(5″)H_2, C(3″)H_2), 1.64 - 1.78 (2H, m, C(2″)H_2), 2.51 - 2.55 (2H, obsc., C(6″)H_2), 3.25 (2H, br. s, C(6″)NH_2), 3.96 (2H, t, *J* 6.5, C(1″)H_2), 6.80 (1H,

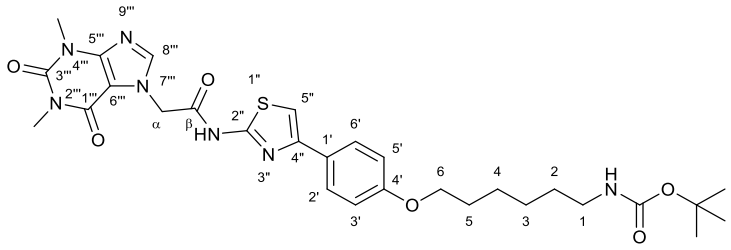
s, Ar*H*), 6.87 - 6.93 (2H, m, C(3′)*H,* C(5′)*H*), 6.96 (2H, br. s, C(2)N*H*$_2$), 7.66 - 7.74 (2H, m, C(2′)*H,* C(6′)*H*); δ_C (101 MHz, DMSO-*d*$_6$) 25.48 (*C*(3″)), 26.19 (*C*(4″) or *C*(5″)), 28.76 (*C*(2″)H$_2$), 33.25 (*C*(4″)H$_2$ or *C*(5″)H$_2$), 41.59 (*C*(6″)), 67.38 (*C*(1″)), 99.23 (*C*(5)), 113.80 (*C*(1′)), 114.30 (*C*(3′), *C*(5′)), 127.71 (*C*(2′), *C*(6′)), 149.72 (*C*(4)), 157.96 (*C*(4′)), 168.04 (*C*(2)); *m/z* (API-ES$^+$) 147 ([M+2H]$^{2+}$, 100%), 292 ([M+H]$^+$, 34%); HRMS (ESI$^+$) $C_{15}H_{22}ON_3S^+$ calculated 292.1478; found 292.1474.

***tert*-Butyl (6-(4′-(2″-aminothiazol-4″-yl)phenoxy)hexyl)carbamate 52**



Boc$_2$O (1.01 g, 4.65 mmol) in MeOH (5.0 mL) was added to a stirred solution of **51** (1.21 g, 4.15 mmol) in MeOH (5.0 mL) at rt and the resultant solution was stirred at rt for 90 min. The reaction mixture was then concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 50 – 75% EtOAc in *c*-hexane) gave **52** as a colourless oil (1.28 g, 79%); ν_{max} (ATR, cm^{-1}) 1702 (NH(C=O)O), 1245 (C(4′)-O); δ_H (400 MHz, DMSO-*d*$_6$) 1.30 (2H, m, C(3)*H*$_2$), 1.36 – 1.44 (13H, m, C*Me*$_3$, C(2)*H*$_2$, C(4)*H*$_2$), 1.64 – 1.73 (2H, m, C(5)*H*$_2$), 2.91 (2H, q, *J* 6.5, C(1)*H*$_2$), 3.94 (2H, t, *J* 6.4, C(6)*H*$_2$, 6.79 (1H, s, C(5″)*H*), 6.85 – 6.93 (2H, m, C(5′)*H*, C(3′)*H*), 7.01 (2H, br. s, N*H*$_2$), 7.64 – 7.73 (2H, m, C(2′)*H*, C(6′)*H*); δ_C (101 MHz, DMSO-*d*$_6$) 25.28 (*C*(4) or *C*(2)), 26.08 (*C*(3)), 28.29 (C*Me*$_3$), 28.71 (*C*(5)), 29.46 (*C*(2) or *C*(4)), 39.81 (*C*(1)), 67.38 (*C*(6)), 77.33 (*C*Me$_3$), 99.25 (*C*(5″)), 114.30 (*C*(3′), *C*(5′)), 126.83 (*C*(2′), *C*(6′)), 127.66 (*C*(1′)), 149.62 (*C*(4″)), 155.62 (N*C*O$_2$^{t}Bu), 157.99 (*C*(4′)), 168.11 (*C*(2″)); *m/z* (API-ES$^+$) 392 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{20}H_{30}O_3N_3{}^{32}S^+$ calculated 392.2002; found 392.2003.

***tert*-Butyl (6-(4′-(2″-(α-(theophyllin-7‴-yl)acetamido)thiazol-4″-yl)phenoxy)hexyl)carbamate 53**

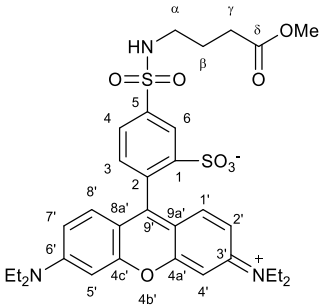


Step 1: $SOCl_2$ (887 µL, 12.2 mmol) was added to a solution of theophylline-7-acetic acid (2.97 g, 12.5 mmol) in anhydrous DMF (100 mL) at rt and then resulting solution was allowed to stir for 1 h at rt.

Step 2: Anhydrous pyridine (3.98 mL, 49.2 mmol) and a solution of **52** (1.28 g, 3.28 mmol) in anhydrous DMF (8.0 mL) were added sequentially to the solution from step 1 at rt and the resulting mixture was allowed to stir for 16 h at rt. The solution was diluted with CH_2Cl_2 (200 mL), then washed sequentially with 1M aq. HCl (3 x 200 mL), a saturated solution of $NaHCO_3$ (3 x 200 mL), brine (5 x 300 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 5% MeOH in $CHCl_3$) gave **53** as a yellow solid (473 mg, 30%); ν_{max} (ATR, cm^{-1}) 1698 (NH(C=O)O), 1648 (NC=O), 1252 (C(4′)-O); δ_H (400 MHz, DMSO-d_6) 1.26 - 1.47 (15H, m, C(CH_3)$_3$, C(2)H_2, C(3)H_2, C(4)H_2), 1.72 (2H, dt, *J* 14.2, 6.9, C(5)H_2), 2.92 (2H, q, *J* 6.6, C(1)H_2), 3.19 (3H, s, N(2‴)*Me*), 3.46 (3H, s, N(4‴)*Me*), 4.00 (2H, t, *J* 6.6, C(6)H_2), 5.34 (2H, s, C(α)H_2), 6.96 - 7.02 (2H, m, C(3′)*H*, C(5′)*H*) 7.50 (1H, s, C(5″)*H*), 7.79 - 7.85 (2H, m, C(2′)H, C(6′)H), 8.08 (1H, s, C(6‴)*H*), 12.75 (1H, s, C(2″)N*H*); δ_C (101 MHz, DMSO-d_6) 25.24 (*C*(2) or *C*(3) or *C*(4)), 26.04 (*C*(2) or *C*(3) or *C*(4)), 27.47 (N(2‴)*Me*), 28.28 (CMe_3), 28.65 (*C*(4)), 29.43 (N(4‴)*Me*), 29.50 (*C*(2) or *C*(3) or *C*(4)), 39.52 (obsc., *C*(1)), 48.20 (*C*(α)), 67.45 (*C*(6)), 77.30 (*C*Me$_3$), 106.30 (*C*(5″)), 106.42 (*C*(6‴)), 114.61 (*C*(3′), *C*(5′)), 126.83 (*C*(1′)), 127.01 (*C*(2′), *C*(6′)), 143.69 (*C*(6‴)), 148.03 (*C*(5‴)), 148.93 (*C*(4″)), 151.02 (*C*(3‴)), 154.49 (*C*(1‴)), 155.59 (N*C*O_2^{t}Bu), 157.24 (*C*(2″)), 158.49

(*C*(4′)), 165.54 (*C*(β)); m/z (API-ES$^+$) 612 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{29}H_{38}O_6N_7S^+$ calculated 612.2599; found 612.2585.

2-(6′-(Diethylamino)-3′-(diethyliminio)-3H-xanthen-9′-yl)-5-(*N*-(δ-methoxy-δ-oxobutyl)sulfamoyl)benzenesulfonate 55

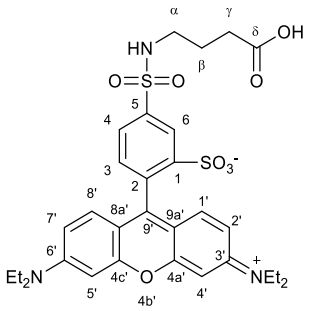


Methyl 4-aminobutyric acid hydrochloride salt (154 mg, 1.00 mmol) and DMAP (24.4 mg, 200 μmol) were added sequentially to a solution of **54** (577 mg, 1.00 mmol) in anhydrous CH_2Cl_2 (20 mL) at rt. DIPEA (523 μL, 3.00 mmol) was added dropwise to the reaction mixture at rt and the resulting solution was stirred further at rt for 16 h then concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 1 – 5% MeOH in $CHCl_3$) gave **55** as a purple solid (185 mg, 28%); ν_{max} (ATR, cm^{-1}) 1646 (C=N), 1182 (NS=O); δ_H (400 MHz, DMSO-*d*$_6$): 1.21 (12H, t, *J* 7.0, NCH_2CH_3), 1.68 (2H, p, *J* 7.0, C(β)H_2), 2.36 (2H, t, *J* 7.0, C(γ)H_2), 2.90 (2H, t, *J* 7.0, C(α)H_2), 3.59 (3H, s, O*Me*), 3.60 – 3.70 (8H, m, NCH_2CH_3), 6.91 – 7.08 (6H, m, C(1′)*H*, C(2′)*H*, C(4′)*H*, C(5′)*H*, C(7′)*H*, C(8′)*H*), 7.47 (1H, d, *J* 8.0, C(3)*H*), 7.92 (1H, dd, *J* 8.0, 2.0, C(4)*H*), 8.41 (1H, d, *J* 2.0, C(6)*H*); δ_C (101 MHz, DMSO-*d*$_6$): 12.44 (NCH_2CH_3), 24.55 (*C*(β)), 30.24 (*C*(γ)), 41.82 (*C*(α)), 45.23 (NCH_2CH_3), 51.27 (O*Me*), 95.36 (*), 113.46 (*), 113.60 (*), 125.63 (*C*(6)), 126.44 (*C*(4)), 130.60 (*C*(3)), 132.68 (*), 132.97 (*C*(2)), 141.57 (*C*(1)), 148.07 (*C*(5)), 155.00 (*C*(3′), *C*(6′)), 157.09 (*), 157.50 (*C*(9′)), 172.88 (*C*(δ));

* = C(1′), C(8′) or C(2′), C(7′) or C(4′), C(5′) or C(4a′), C(4c′) or C(8a′), C(9a′)

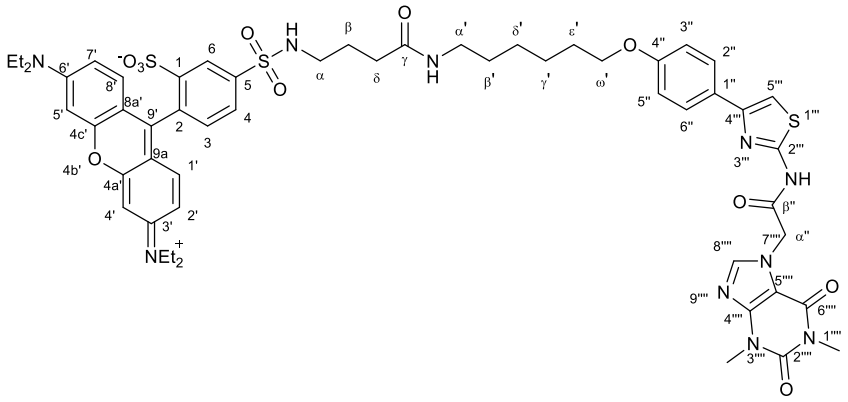
m/z (API-ES$^+$) 658 ([M+H]$^+$, 100%), 330 ([M+2H]$^{2+}$, 96%); HRMS (ESI$^+$) $C_{32}H_{40}O_8N_3S_2^+$ calculated 658.2251; found 658.2247.

2-(6′-(Diethylamino)-3′-(diethyliminio)-3H-xanthen-9′-yl)-5-(*N*-(δ-hydroxy-δ-oxobutyl) sulfamoyl)benzenesulfonate 56



1M aq. NaOH (821 µL, 821 µmol) was added to a solution of **55** (180 mg, 274 µmol) in a MeOH/THF mixture (2:1 v/v, 24 mL) at rt and the resulting solution was stirred at rt for 16 h, then concentrated *in vacuo*. The residue was suspended in 1M aq. HCl (15 mL), then extracted with a mixture of $CHCl_3$ and MeOH (3 x 25 mL, 4:1 v/v), dried and concentrated *in vacuo* to give **56** as a purple solid which was used without further purification (177 mg, apparent quant.); ν_{max} (ATR, cm^{-1}) 1731 (C(δ)=O), 1646 (C=N), 1181 (NS=O); δ_H (400 MHz, DMSO-d_6) 1.21 (12H, t, *J* 7.0, NCH_2CH_3), 1.56 – 1.74 (2H, m, C(β)H_2), 2.24 (2H, t, *J* 7.5, C(γ)H_2), 2.90 (2 H, q, *J* 6.5, C(α)H_2), 3.60 – 3.72 (8H, q, *J* 7.0, NCH_2Me), 6.94 (2H, d, *J* 2.0, C(4′)*H*, C(5′)*H*), 6.98 (2H, d, *J* 9.5, C(1′)*H*, C(8′)*H*), 7.04 (2H, dd, *J* 9.5, 2.0, C(2′)*H*, C(7′)*H*), 7.47 (1H, d, *J* 8.0, C(3)*H*)), 7.87 – 8.03 (2H, m, C(4)H, C(α)N*H*), 8.41 (1H, d, *J* 2.0, C(6)*H*); δ_C (101 MHz, DMSO-d_6) 12.48 (NCH_2Me), 24.63 (*C*(β)), 30.62 (*C*(γ)), 41.97 (*C*(α)), 45.26 (NCH_2Me), 95.38 (*C*(4′), *C*(5′)), 113.48 (*C*(2′), *C*(7′)), 113.65 (*C*(8a′), *C*(9a′)), 125.64 (*C*(6′)), 126.51 (*C*(4)), 130.65 (*C*(3)), 132.71 (*C*(1′), *C*(8′)), 133.02 (*C*(2)), 141.56 (*C*(1)), 148.04 (*C*(5)), 155.02 (*C*(3′), *C*(6′)), 157.11 (*C*(4a′), *C*(4c′)), 157.47 (*C*(9′)), 173.98 (*C*(δ)); *m/z* (ESI^-) 642 ($[M-H]^-$, 100%); HRMS (ESI^-) $C_{31}H_{36}O_8N_3S_2^-$ calculated 642.1950; found 642.1955.

2-(6′-(Diethylamino)-3′-(diethyliminio)-3H-xanthen-9′-yl)-5-(*N*-(δ-((ω′-(4′′-(2′′′-(α′′-(theophyllin-7′′′′-yl)acetamido)thiazol-4′′′-yl)phenoxy)hexyl)amino)-δ-oxobutyl)sulfamoyl) benzenesulfonate BGSP-001

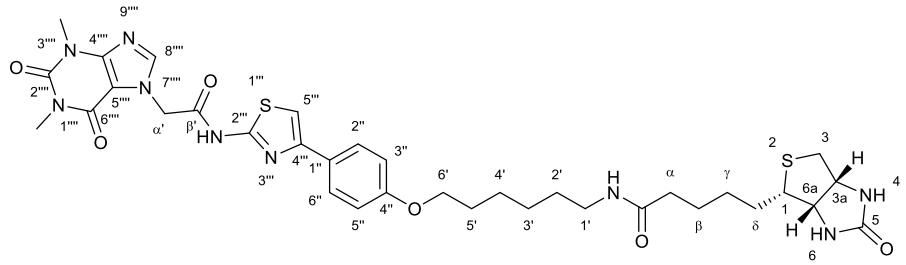


Step 1: TFA (0.50 mL) was added to a solution of **53** (30.0 mg, 49.0 µmol) in $CHCl_3$ (1.5 mL) at rt and the resulting solution was stirred for 1 h at rt. The reaction mixture was then concentrated in vacuo. The concentrate was then diluted with CH_2Cl_2 (5 mL) and concentrated *in vacuo*. This process was repeated twice to give a pale yellow oil.

Step 2: Oxyma Pure (6.97 mg, 49.0 µmol) and DIC (40.0 µL, 49.0 µmol) were added to a solution of theophylline-7-acetic acid (31.6 mg, 49.0 µmol) in DMF (1.0 mL) at 0 °C and the resulting solution was stirred at 0 °C for 20 min. The solution was warmed to rt, the crude residue from step 1 in DMF (1.0 mL) was added to the reaction mixture and the resulting mixture was stirred at rt for 16 h. The reaction mixture was then concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 5 – 10% MeOH in $CHCl_3$) gave **BGSP-001** as a purple solid (7.03 mg, 13%); δ_H (600 MHz, DMSO-d_6): 1.16 (12H, t, *J* 7.5, NCH_2CH_3), 1.25 (2H, dd, *J* 9.0, 6.0, C(δ′)H_2), 1.28 – 1.38 (4H, m, C(β′)H_2, C(γ′)H_2), 1.57 – 1.65 (4H, m, C(ε′)H_2, C(β)H_2), 2.06 (2H, t, *J* 8.0, C(δ)H_2), 2.86 (2H, q, *J* 7.0, C(α)H_2), 3.00 (2H, q, *J* 6.5, C(α′)H_2), 3.16 (3H, s, N(1′′′′)*Me*), 3.44 (3H, s, N(3′′′′)*Me*), 3.56 (8H, q, *J* 7.5, NCH_2CH_3), 3.89 (2H, obsc., C(ω′)H_2), 5.29 (2H, s, C(α′′)H_2), 6.81 – 6.85 (2H, d, *J* 2.2, C(4′)*H*, C(6′)*H*), 6.88 (2H, m, C(3′′)*H*, C(5′′)*H*), 6.92 – 7.03 (4H, m, C(1′)*H*, C(2′)*H*, C(7′)*H*, C(8′)*H*), 7.36 (1H, s, C(5′′′)*H*), 7.42 (1H, d, *J* 8.0, C(3)*H*), 7.70 – 7.75 (2H, m, C(2′′)*H*, C(6′′)*H*), 7.83

(1H, t, *J* 5.5, C(α′)N*H*), 7.90 – 7.96 (2H, m, C(α)N*H*, C(4)*H*), 8.04 (1H, s, C(8′′′′)*H*), 8.38 (1H, d, *J* 2.0, C(6)*H*); δ_C (151 MHz, DMSO-d_6): 12.88 (NCH$_2$*C*H$_3$), 25.64 (*C*(γ′)), 25.90 (*C*(β)), 26.52 (*C*(δ′)), 28.05 (N(1′′′′)*Me*), 28.99 (*C*(ε′)), 29.34 (*C*(β′)), 30.10 (N(3′′′′)*Me*), 33.06 (*C*(δ)), 38.99 (*C*(α′)), 42.76 (*C*(α′)), 45.83 (N*C*H$_2$Me), 48.77 (*C*(α′′)), 67.97 (*C*(ω′)), 95.92 (*C*(4′), *C*(5′)), 106.80 (*C*(5′′′′)), 107.03 (*C*(5′′′)), 113.81 (*C*(4a′), *C*(4c′) or *C*(8a′), *C*(9a′)), 114.17 (C(1′), (C(8′) or C(2′), C(7′)), 115.09 (*C*(3′′), *C*(5′′)), 126.23 (*C*(6)), 127.19 (*C*(4)), 127.40 (*C*(1′′)), 127.55 (*C*(2′′), *C*(6′′)), 131.34 (*C*(3)), 132.92 (*C*(1′), *C*(8′) or *C*(2′), *C*(7′)), 133.64 (*C*(2)), 142.08 (*C*(1)), 144.17 (*C*(8′′′′), 147.77 (*C*(5)), 148.50 (*C*(4′′′′), 149.42 (*C*(4′′′), 151.70 (*C*(2′′′′)), 155.05 (*C*(6′′′′)), 155.52 (C(3′), C(5′)), 157.23 (*C*(9′)), 157.58 (*C*(4a′), *C*(4c′) or *C*(8a′), *C*(9a′)), 157.77 (C(2′′′), 158.92 (*C*(4′′)), 166.07 (*C*(β′′)), 172.50 (*C*(γ)); *m/z* (API-ES$^+$) 569 ([M+2H]$^{2+}$, 100%); HRMS (ESI$^+$) $C_{55}H_{65}O_{11}N_{10}S_3^+$ calculated 1137.3991; found 1137.3986.

_N_-(6′-(4′′-(2′′′-(α′-(Theophyllin-7′′′′-yl)acetamido)thiazol-4′′′-yl)phenoxy)hexyl) biotin amide BGSP-002



Step 1: TFA (0.50 mL) was added to a solution of **53** (30.0 mg, 49.0 µmol) in $CHCl_3$ (1.5 mL) at rt and the resulting solution was stirred for 1 h at rt. The reaction mixture was then concentrated in vacuo. The concentrate was then diluted with CH_2Cl_2 (5 mL) and concentrated *in vacuo*. This process was repeated twice to give a pale yellow oil.

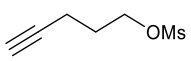
Step 2: A solution of biotin-*N*-hydroxysuccinimide ester (30.6 mg, 89.8 µmol) and Et_3N (34.2 µL, 245 µL) in DMF (1.0 mL) was added to a solution of the residue from step 1 in DMF (1.0 mL) at rt and the resulting solution was stirred at rt for 16 h. The reaction mixture diluted with $CHCl_3$ (25 mL),

then washed sequentially with 1M aq. HCl (3 x 10 mL), brine (10 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 7 – 11% MeOH in $CHCl_3$) gave **BGSP-002** as a white solid (16.1 mg, 45%); δ_H (600 MHz, DMSO-d_6): 1.07 – 1.58 (11H, m, C(2′)H_2, C(3′)H_2, C(4′)H_2, C(β)H_2, C(γ)H_2, C(δ)H_AH$_B$), 1.57 – 1.63 (1H, m, C(δ)H$_A$$H_B$), 1.72 (2H, p, *J* 6.5, C(5′)H_2), 2.04 (2H, t, *J* 7.5, C(1′)H_2 or C(α)H_2), 2.57 (1H, d, *J* 12.5, C(3)H_AH$_B$), 2.81 (1H, dd, *J* 12.5, 5.0, C(3)H$_A$$H_B$), 3.03 (2H, q, *J* 6.5, C(1′)H_2 or C(α)H_2), 3.06 – 3.11 (1H, m, C(1)*H*), 3.19 (3H, s, N(1‴′)*Me*), 3.46 (3H, s, N(3‴′)*Me*), 3.99 (2H, t, *J* 6.5, C(6′)H_2), 4.12 (1H, ddd, *J* 7.5, 4.5, 2.0, C(6a)*H*), 4.29 (1H, ddt, *J* 7.5, 5.0, 1.0, C(3a)*H*), 5.34 (2H, s, C(α′)H_2), 6.34 (1H, s, N(4)*H*), 6.41 (1H, s, N(6)*H*), 6.94 – 7.01 (2H, m, C(3″)*H*, C(5″)*H*), 7.47 (1H, s, C(5‴′)*H*), 7.73 (1H, t, *J* 5.7, C(1′)N*H*), 7.77 – 7.86 (2H, m, C(2″)*H*, C(6″)*H*), 8.09 (1H, s, C(8‴′)*H*), 12.71 (1H, s, C(2‴′)N*H*); δ_C (151 MHz, DMSO-d_6): 25.25 (*), 25.34 (*), 26.16 (*), 27.45 (N(1‴′)*Me*), 28.04 (*), 28.21 (*C*(δ)), 28.63 (*), 29.11 (*C*(5′)) , 29.48 (N(3‴′)*Me*), 35.22 (*C*(1′) or *C*(α)), 38.30 (*C*(1′) or *C*(α)), 40.06 (*C*(3)), 48.19 (*C*(α′)), 55.43 (*C*(1)), 59.18 (*C*(3a)), 61.04 (*C*(6a)), 67.45 (*C*(6′)), 106.28 (*C*(5‴′)), 106.40 (*C*(5‴)), 114.61 (*C*(3″), *C*(5″)), 126.83 (*C*(1″)), 127.00 (*C*(2″), *C*(6″)), 143.67 (*C*(8‴′)), 148.01 (*C*(4‴′)), 148.91 (*C*(4‴)), 151.00 (*C*(2‴′)), 154.47 (*C*(6‴′)), 157.24 (*C*(2‴)), 158.48 (*C*(4″)), 162.67 (*C*(5)), 165.52 (*C*(β′)), 171.78 (C(α)*C*ONH);

* = *C*(2′), *C*(3′), *C*(4′), *C*(β) or *C*(γ);

m/z (API-ES$^+$) 370 ([M+2H]$^{2+}$, 100%), 738 ([M+H]$^+$, 10%); HRMS (ESI$^+$) $C_{34}H_{44}O_6N_9S_2^+$ calculated 738.2851; found 738.2849.

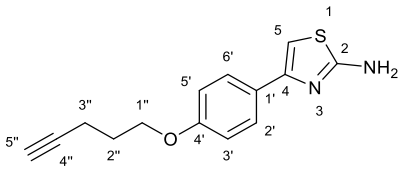
Pent-4-yn-1-yl methanesulfonate 58[3]



Following a literature procedure,[3] Et_3N (544 µL, 3.90 mmol) and MsCl (279 µL, 3.60 mmol) were added to a solution of pent-4-yne-1-ol (279 µL, 3.00 mmol) in CH_2Cl_2 (2.0 mL) at 0 °C. The reaction

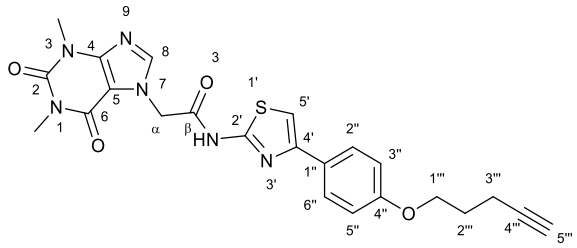
mixture was allowed to warm to rt and was stirred further for 3 h before H_2O (15 mL) and CH_2Cl_2 (15 mL) were added. The layers were separated and the organic layer was washed with a saturated solution of $NaHCO_3$ (15 mL) then brine (15 mL), dried and concentrated *in vacuo* to give **58** as a pale yellow oil which was used without further purification (429 mg, 88%); δ_H (200 MHz, $CDCl_3$) 1.84 – 2.07 (3H, m), 2.27 – 2.46 (2H, m), 3.03 (3H, s), 4.36 (2H, t, *J* 6.0); *m/z* (API-ES$^+$) 163 ([M+H]$^+$, 100%); all other data were in accordance with the literature values.[3]

4-(4′-(Pent-4′′-yn-1-yloxy)phenyl)thiazol-2-amine 59



58 (428mg, 2.64 mmol) was added to a suspension of K_2CO_3 (274 mg, 1.98 mmol) and **44** (254 mg, 1.32 mmol) in DMF (3.0 mL) at rt. The resulting suspension was heated to 40 °C for 88 h before being cooled to rt, diluted with EtOAc (40 mL), washed sequentially with brine (40 mL), 1M aq. NaOH (2 x 40 mL), brine (40 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 15 – 25% EtOAc in *c*-hexane) gave **59** as a white solid (96.1 mg, 28%); ν_{max} (ATR, cm^{-1}) 1242 (C(4′)-O); δ_H (400 MHz, $CDCl_3$): 1.94 – 2.07 (3H, m, C(2′′)H_2, C(5′′)*H*), 2.42 (2H, td, *J* 7.0, 2.5, C(3′′)H_2), 4.09 (2H, t, *J* 6.0, C(1′′)H_2), 6.58 (1H, s, C(5)*H*), 6.86 – 6.95 (2H, m, C(3′)*H*, C(5′)*H*), 7.65 – 7.74 (2H, m, C(2′)*H*, C(6′)*H*); δ_C (101 MHz, $CDCl_3$): 15.34 (*C*(3′′)), 28.34 (*C*(2′′)), 66.33 (*C*(1′′)), 69.02 (*C*(5′′)), 83.63 (*C*(4′′)), 101.23 (*C*(5)), 114.69 (*C*(3′), *C*(5′)), 127.41 (*C*(2′), *C*(6′)), 127.85 (C(1′)), 151.24 (*C*(4)), 158.75 (*C*(4′)), 167.22 (*C*(2)); *m/z* (API-ES$^+$) 259 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{14}H_{15}N_2S^+$ calculated 259.0900; found 259.0900.

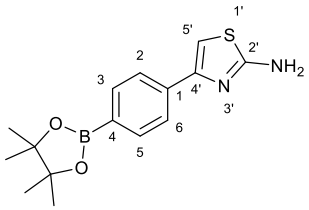
α-(Theophyllin-7-yl)-*N*-(4′-(4″-(pent-4‴-yn-1‴-yloxy)phenyl)thiazol-2′-yl)acetamide BGSP-003



Step 1: $SOCl_2$ (157 µL, 2.15 mmol) was added to a solution of theophylline-7-acetic acid (527 mg, 2.21 mmol) in anhydrous DMF (10 mL) at rt and the resulting solution was stirred for 1 h at rt.

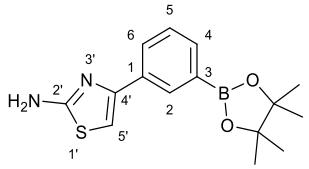
Step 2: Anhydrous pyridine (704 µL, 8.71 mmol) and a solution of **59** (150 mg, 581 µmol) in anhydrous DMF (2.0 mL) were added sequentially to the solution from step 1 at rt and the resulting solution was allowed to stir at rt for 16 h. The reaction mixture was diluted with $CHCl_3$ (50 mL), then washed sequentially with 0.1 M HCl (3 x 50 mL), 10% aq. NH_4OH (3 x 50 mL), brine (5 x 50 mL), then dried and concentrated *in vacuo*. The crude residue was washed with 20% aq. MeCN (50 mL) then Et_2O (50 mL). Purification by flash column chromatography (gradient elution: 40 – 80% acetone in *c*-hexane) gave **BGSP-003** as a white solid (44.2 mg, 16%); δ_H (600 MHz, DMSO-d_6): 1.91 (2H, q, *J* 6.5, C(2‴)H_2), 2.35 (2H, td, *J* 7.0, 2.5, C(3‴)H_2), 2.82 (1H, t, *J* 2.5, C(5‴)*H*), 3.19 (3H, s, N(1)*Me*), 3.46 (3H, s, N(3)*Me*), 4.08 (2H, t, *J* 6.5, C(1‴)H_2), 5.34 (2H, s, C(α)H_2), 6.96 – 7.03 (2H, m, C(3″)*H*, C(5″)*H*), 7.49 (1H, s, C(5′)*H*), 7.75 – 7.87 (2H, m, C(2″)*H*, C(6″)*H*), 8.09 (1H, s, C(8)*H*); δ_C (151 MHz, DMSO-d_6): 14.46 (*C*(3‴)), 27.44 (N(1)*Me*), 27.70 (*C*(2‴)), 29.47 (N(3)*Me*), 48.18 (*C*(α)), 66.00 (*C*(1‴)), 71.63 (*C*(5‴)), 83.64 (*C*(4‴)), 106.41 (*C*(5′), *C*(5)), 114.65 (*C*(3″), *C*(5″)), 127.04 (*C*(1″), *C*(2″), *C*(6″)), 143.68 (*C*(8)), 148.02 (*C*(4)), 148.86 (*C*(4′)) , 151.01 (*C*(2)), 154.48 (*C*(6)), 157.25 (*C*(2)), 158.28 (*C*(4″)), 165.53 (*C*(β)); m/z 479 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{23}H_{23}O_4N_6S^+$ calculated 479.1496; found 479.1497.

(4-(2′-Aminothiazol-4′-yl)phenyl)boronic acid pinacol ester *p*-60



$Pd(dppf)Cl_2 \cdot CH_2Cl_2$ complex (123 mg, 150 µmol) was added to a stirred solution of **25** (765 mg, 3.00 mmol), bis(pinacolato)diboron (952 mg, 3.75 mmol) and KOAc (883 mg, 9.00 mmol) in 1,4-dioxane (30 mL) at rt. Ar gas was bubbled through the reaction mixture until it turned dark brown and the resulting mixture was heated to 100 °C for 16 h before being cooled to rt. The reaction mixture was then filtered through Celite® (eluent: EtOAc). The filtrate was washed sequentially with H_2O (50 mL), brine (50 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 20 – 25% EtOAc in *c*-hexane) gave ***p*-60** as an off-white solid (653 mg, 72%); ν_{max} (ATR, cm^{-1}) 3312 (N-H); δ_H (400 MHz, DMSO-d_6): 1.30 (12H, s, CMe_2), 7.07 (2H, br. s, NH_2), 7.10 (1H, s, C(5′)*H*), 7.62 – 7.70 (2H, m, C(3)*H*, C(5)*H*), 7.77 – 7.84 (2H, m, C(2)*H*, C(5)*H*); δ_C (101 MHz, DMSO-d_6): 24.70 (CMe_2), 83.57 (*C*Me_2), 102.88 (*C*(5′)), 124.87 (*C*(3), *C*(5)), 127.51 (*C*(2), *C*(6)), 134.68 (*C*(4)), 137.50 (*C*(1)), 149.54 (*C*(4′)), 168.16 (*C*(2′)); *m/z* (API-ES^+) 303 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{15}H_{20}O_2N_2{}^{10}BS^+$ calculated 302.1369; found 302.1369.

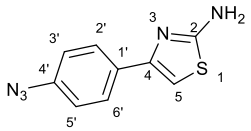
(3-(2′-Aminothiazol-4′-yl)phenyl)boronic acid pinacol ester *m*-60



$Pd(dppf)Cl_2 \cdot CH_2Cl_2$ complex (41.8 mg, 51.2 µmol) was added to a stirred solution of 2-amino-4-(3′-bromophenyl)thiazole (261 mg, 1.02 mmol), bis(pinacolato)diboron (325 mg, 1.28 mmol) and KOAc (301 mg, 3.07 mmol) in 1,4-dioxane (10 mL) at rt. Ar gas was bubbled through the reaction mixture until it turned dark brown and the resulting mixture was heated to 100 °C for 16 h before being cooled to rt. The reaction mixture was then filtered through Celite® (eluent: EtOAc). The filtrate was washed

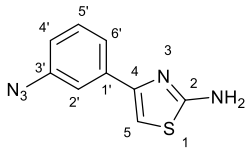
sequentially with H_2O (20 mL), brine (20 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 20 – 30% EtOAc in *c*-hexane) gave ***m*-60** as an off-white solid (230 mg, 75%); ν_{max} (ATR, cm^{-1}) 3414 (N-H); δ_H (400 MHz, DMSO-d_6): 1.30 (12H, s, CMe_2), 7.02 (1H, s, C(5′)*H*), 7.07 (2H, br. s, C(2′)NH_2), 7.37 (1H, t, *J* 7.5, C(5)*H*), 7.54 (1H, dt, *J* 7.5, 1.5, C(4)*H*), 7.90 (1H, ddd, *J* 7.5, 2.0, 1.5, C(6)*H*), 8.17 (1H, t, *J* 1.5, C(2)*H*); δ_C (101 MHz, DMSO-d_6): 24.71 (CMe_2), 83.66 (*C*Me_2), 101.54 (*C*(5′)), 127.99 (*C*(5)), 128.26 (*C*(6)), 131.93 (*C*(4)), 133.11 (*C*(2), *C*(3)), 134.25 (*C*(1)), 149.53 (*C*(4′)), 168.20 (*C*(2′)); *m/z* (API-ES^+) 303 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{15}H_{20}O_2N_2{}^{10}BS^+$ calculated 302.1369; found 302.1369.

4-(4′-Azidophenyl)thiazol-2-amine *p*-61



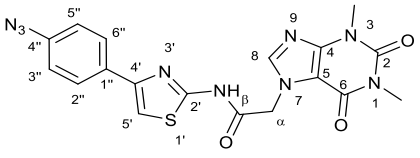
NaN_3 (699 mg, 10.8 mmol) and $Cu(OAc)_2$ (78.4 mg, 430 μmol) were added to a solution of ***p*-60** (650 mg, 2.15 mmol) in MeOH (50 mL) at rt. Compressed air was bubbled through the reaction mixture and the resulting suspension was heated to 50 °C for 5 h before being cooled to rt and filtered through Celite® (eluent: EtOAc). The filtrate was washed sequentially with H_2O (2 x 100 mL), brine (100 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 20% EtOAc in *c*-hexane) gave ***p*-61** as a yellow solid (301 mg, 64%); ν_{max} (ATR, cm^{-1}) 2122 (N_3); δ_H (400 MHz, DMSO-d_6): 6.99 (1H, s, C(5)*H*), 7.06 (2H, br. s, C(2)NH_2), 7.08 – 7.15 (2H, m, C(3′)*H*, C(5′)*H*), 7.78 – 7.87 (2H, m, C(2′)*H*, C(6′)*H*); δ_C (101 MHz, DMSO-d_6): 101.39 (*C*(5)), 119.19 (*C*(3′), *C*(5′)), 127.10 (*C*(2′), *C*(6′)), 132.03 (*C*(1′)), 137.88 (*C*(4′)), 148.96 (*C*(4)), 168.24 (*C*(2)); *m/z* (API-ES^+) 218 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_9H_8N_5S^+$ calculated 218.0495; found 218.0497.

4-(3′-Azidophenyl)thiazol-2-amine *m*-61



NaN_3 (553 mg, 8.50 mmol) and $Cu(OAc)_2$ (61.8 mg, 340 µmol) were added to a solution of ***m*-60** (514 mg, 1.70 mmol) in MeOH (40 mL). Compressed air was bubbled through the reaction mixture and the resulting suspension was heated to 50 °C for 48 h before being cooled to rt and filtered through Celite® (eluent: EtOAc). The filtrate was washed sequentially with H_2O (2 x 100 mL), brine (100 mL), then dried and concentrated *in vacuo*. The crude residue was resuspended in Et_2O (5 mL), then diethanolamine (200 µL) was added at rt and the reaction mixture was allowed to stir for 30 min at rt before being filtered and the filtrate concentrated *in vacuo*. Purification by flash column chromatography (eluent 25% EtOAc in *c*-hexane) gave ***m*-61** as an orange solid (116 mg, 31%); ν_{max} (ATR, cm^{-1}) 2115 (N_3); δ_H (400 MHz, DMSO-*d*$_6$): 6.99 (1H, ddd, *J* 8.0, 2.5, 1.0, C(4′)*H*), 7.11 (2H, br. s, C(2)N*H*$_2$), 7.13 (1H, s, C(5)*H*), 7.39 (1H, t, *J* 8.0, C(5′)*H*), 7.53 (1H, dd, *J* 2.5, 1.0, C(2′)*H*), 7.62 (1H, dt, *J* 8.0, 1.0, C(6′)*H*); δ_C (101 MHz, DMSO-*d*$_6$): 102.86 (*C*(5)), 115.87 (*C*(2′)), 117.81 (*C*(4′)), 122.22 (*C*(6′)), 130.12 (*C*(5′)), 136.64 (*C*(1′)), 139.60 (*C*(3′)), 148.67 (*C*(4)), 168.27 (*C*(2)); *m/z* (API-ES$^+$) 218 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_9H_8N_5S^+$ calculated 218.0495; found 218.0496.

***N*-(4′-(4′′-Azidophenyl)thiazol-2′-yl)-α-(theophyllin-7-yl)acetamide 3-60**

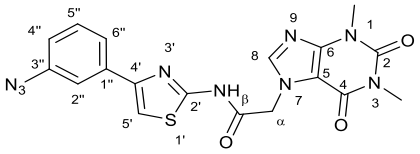


Step 1: $SOCl_2$ (249 µL, 3.42 mmol) was added to a solution of theophylline-7-acetic acid (836 mg, 3.51 mmol) in anhydrous DMF (15 mL) at rt and the resulting solution was allowed to stir for 1 h.

Step 2: Anhydrous pyridine (1.12 mL, 13.8 mmol) and a solution of ***p*-61** (200 mg, 921 µmol) in anhydrous DMF (2 mL) were added to the crude solution from step 1 at rt and the resulting solution

was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (50 mL), then washed sequentially with H_2O (3 x 50 mL), brine (50 mL), then dried and concentrated *in vacuo*. Purification by HPLC (sample loading: DMSO, gradient elution: 2 - 98% solvent B in solvent A) gave **3-60** as an off-white solid (32.4 mg, 8%); ν_{max} (ATR, cm^{-1}) 2122 (N_3), 1645 (NC=O); δ_H (600 MHz, DMSO-d_6): 3.18 (3H, s, N(1)*Me*), 3.46 (3H, s, N(3)*Me*), 5.34 (2H, s, C(α)H_2), 7.16 – 7.22 (2H, m, C(3″)*H*, C(5″)*H*), 7.64 – 7.67 (1H, s, C(5′)*H*), 7.92 – 7.97 (2H, m, C(2″)*H*, C(6″)*H*), 8.09 (1H, s, C(8)*H*), 12.77 (1H, s, C(2′)N*H*); δ_C (151 MHz, DMSO-d_6): 27.46 (N(1)*Me*), 29.49 (N(3)*Me*), 48.20 (*C*(α)), 106.41 (*C*(5)), 108.27 (*C*(5′)), 119.51 (*C*(3″), *C*(5″)) 127.27 (*C*(2″), *C*(6″)), 131.17 (*C*(1″)), 138.73 (*C*(4″)), 143.67 (*C*(8)), 148.02 (*C*(4′)), 148.14 (*C*(4)), 151.01 (*C*(2)), 154.48 (*C*(6)), 157.49 (*C*(2′)), 165.66 (*C*(β)); *m/z* (API-ES^+) 438 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{18}H_{16}O_3N_9S^+$ calculated 438.1091; found 438.1097.

_N_-(4′-(3″-Azidophenyl)thiazol-2′-yl)-α-(theophyllin-7-yl)acetamide 3-61

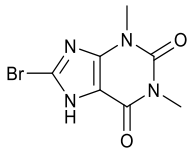


Step 1: $SOCl_2$ (113 µL, 1.83 mmol) was added to a solution of theophylline-7-acetic acid (447 mg, 1.88 mmol) in anhydrous DMF (7.0 mL) at rt and the resulting solution was allowed to stir for 1 h.

Step 2: Anhydrous pyridine (598 µL, 7.39 mmol) and a solution of ***m*-61** (107 mg, 493 µmol) in anhydrous DMF (5.0 mL) were added sequentially to the crude solution from step 1 at rt and the resulting solution was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (50 mL), then washed sequentially with H_2O (3 x 50 mL), brine (50 mL), then dried and concentrated *in vacuo*. Purification by HPLC (sample loading: DMSO, gradient elution: 2 - 98% solvent B in solvent A) gave **3-61** as an off-white solid (8.85 mg, 4%); ν_{max} (ATR, cm^{-1}) 2113 (N_3), 1645 (NC=O), 1697 (NC=O); δ_H (600 MHz, DMSO-d_6): 3.19 (3H, s, N(1)*Me*), 3.46 (3H, s, N(3)*Me*), 5.35 (2H, s, C(α)H_2),

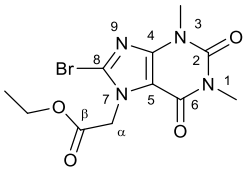
7.08 (1H, ddd, *J* 8.0, 2.5, 1.0, C(4″)*H*), 7.48 (1H, t, *J* 8.0, C(5′)*H*), 7.65 (1H, t, *J* 2.0, C(2″)*H*), 7.74 (1H, dt, *J* 8.0, 1.0, (C(6″)*H*), 7.80 (1H, s, C(5′)*H*), 8.09 (1H, s, C(8)*H*), 12.80 (1H, s, C(2′)N*H*); δ_C (151 MHz, DMSO-d_6): 27.46 (N(1)*Me*), 29.49 (N(3)*Me*), 48.20 (*C*(α)), 106.41 (*C*(5)), 109.69 (*C*(5′)), 115.97 (*C*(2″)), 118.57 (*C*(4″)), 122.40 (*C*(6″)), 130.48 (*C*(5′)), 135.83 (*C*(1″)), 139.96 (*C*(3″)), 143.67 (*C*(8)), 147.83 (*C*(4′)), 148.03 (*C*(4)), 151.01 (*C*(2)), 154.49 (*C*(6)), 157.54 (*C*(2′)), 165.73 (*C*(β)); m/z (API-ES$^+$) 438 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{18}H_{16}O_3N_9S^+$ calculated 438.1091; found 438.1087.

8-Bromotheophylline 65[4]



Following a literature procedure,[4] Br_2 (563 µL, 11.0 mmol) was added to a solution of theophylline (1.80 g, 10.0 mmol) in an AcOH/H_2O mixture (25 mL, 3:2 v/v) at rt. The resulting solution was heated to 50 °C for 4 h before being cooled to rt. The resulting suspension was collected by filtration, washed with water (2 x 100 mL), then dried *in vacuo* to give **65** as a white solid (2.10 g, 81%) which was used without further purification; δ_H (400 MHz, DMSO-d_6) 3.20 (3H, s, N*Me*), 3.35 (1H, br. s, N*H*), 3.37 (3H, s, N*Me*); δ_C (101 MHz, DMSO-d_6) 27.77, 29.84, 109.10, 124.29, 147.86, 150.83, 153.37; *m/z* (API-ES$^+$) 259 ([M(^{79}Br)+H]$^+$, 50%), 261 ([M(^{81}Br)+H]$^+$, 50%).

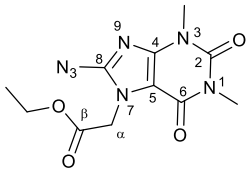
Ethyl 8-bromo-theophylline-7-acetate 66[5]



Following a literature procedure,[5] K_2CO_3 (5.29 g, 38.3 mmol) and ethyl bromoacetate (2.18 mL, 18.4 mmol) were added sequentially to a solution of **65** (3.97 g, 15.3 mmol) in DMF (75 mL) at rt and the

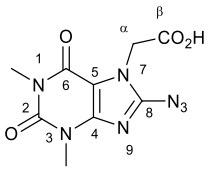
resulting suspension was heated to 70 °C for 4 h. The reaction mixture was cooled to rt, then poured over ice water. The resulting precipitate was collected by filtration, washed with H_2O (2 x 100 mL), then dried *in vacuo* to give **66** as a white solid (4.58 g, 87%) which was used without further purification; δ_H (400 MHz, DMSO-d_6): 1.22 (3H, t, *J* 7.1, *Me*CH_2O), 3.20 (3H, s, N(1)*Me*), 3.41 (3H, s, N(3)*Me*), 4.20 (2H, q, *J* 7.1, MeCH_2O), 5.17 (2H, s, C(α)H_2); δ_C (101 MHz, DMSO-d6): 13.97 (*Me*CH_2O), 27.59 (N(1)*Me*), 29.58 (N(3)*Me*), 47.52 (*C*(α)), 61.86 (Me*C*H_2O), 108.31 (*C*(5)), 129.12 (*C*(8)), 147.47 (*C*(4)), 150.59 (*C*(2)), 153.68 (*C*(6)), 166.59 (*C*(β)); *m/z* (API-ES$^+$) 345 ([M(^{79}Br)+H]+, 50%), 347 ([M(^{81}Br)+H]+, 50%).

Ethyl α-(8-Azido-theophyllin-7-yl)acetate 67



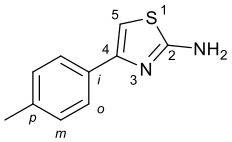
NaN_3 (3.25 g, 50.0 mmol) was added to a solution of **66** (3.45 g, 10.0 mmol) in a DMF/H_2O mixture (33 mL, 10:1 v/v) at rt and the resulting suspension was heated to 50 °C for 2 h. The reaction was then cooled to rt, diluted with $CHCl_3$/MeOH (4:1), then washed sequentially with H_2O (5 x 100 mL) and brine (100 mL), then dried and concentrated *in vacuo* to give **67** as a pale yellow solid (2.92 g, 95%) which was used without further purification; vmax (ATR, cm^{-1}) 2163 (N_3), 1749 (C=O), 1695, 1651 (NC=O); δ_H (500 MHz, DMSO-d_6): 1.21 (3H, t, *J* 7.0, OCH_2*Me*), 3.19 (3H, s, N(1)*Me*), 3.41 (3H, s, N(3)*Me*), 4.17 (2H, q, *J* 7.0, OCH_2Me), 4.92 (2H, s, C(α)H_2); δ_C (126 MHz, DMSO-d_6): 13.97 (OCH2*Me*), 27.49 (N(1)*Me*), 29.67 (N(3)*Me*), 45.07 (*C*(α)), 61.74 (O*C*H_2Me), 105.13 (*C*(5)), 145.68 (*C*(8)), 146.65 (*C*(4)), 150.71 (*C*(2)), 153.84 (*C*(6)), 166.80 (*C*(β)); *m/z* (API-ES$^+$) 308 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{11}H_{13}O_4N_7Na^+$ calculated 330.0921; found 330.0921.

2-(8-Azido-theophyllin-7-yl)acetic acid 68



NaOH (1M aq, 976 µL, 976 µmol) was added to a solution of the **67** (100 mg, 325 µmol) in a THF/MeOH mixture (6 mL, 1:2 v/v) at rt and the resulting solution was stirred for 30 min at rt. The solution was acidified to pH <3 with HCl (1M aq.) then concentrated *in vacuo*. The mixture was extracted with a mixture of $CHCl_3$ and MeOH (3 x 10 mL, 4:1 v/v), then the resulting organic layers were washed with brine (10 mL), dried and concentrated *in vacuo* to give **68** as an off-white solid (86.7 mg, 95%) which was used without further purification; ν_{max} (ATR, cm^{-1}) 2990 (br., OH), 2156 (N_3), 1745 (C=O), 1696, 1632 (NC=O); δ_H (400 MHz, DMSO-d_6): 3.19 (3H, s, N(1)*Me*), 3.40 (3H, s, N(3)*Me*), 4.81 (2H, s, C(α)H_2); δ_C (101 MHz, DMSO-d_6): 27.45 (N(1)*Me*), 29.63 (N(3)*Me*), 45.07 (*C*(α)), 105.17 (*C*(5)), 145.60 (*C*(8)), 146.56 (*C*(4)), 150.67 (*C*(2)), 153.83 (*C*(6)), 168.05 (C(β)); *m/z* (API-ES^+) 280 ($[M+H]^+$, 100%); HRMS (ESI^-) $C_9H_8O_4N_7^-$ calculated 278.0643; found 278.0640.

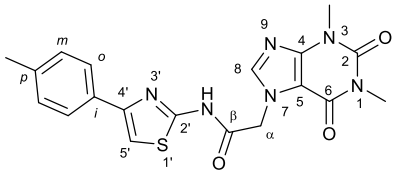
4-(*p*-Tolyl)-thiazol-2-amine 70[6]



I_2 (5.08 g, 20.0 mmol) was added to a solution of thiourea (3.04 g, 40.0 mmol) and 4-methylacetophenone (2.68 mL, 20.0 mmol) in EtOH (25 mL) and the resulting solution was refluxed for 24 h. The reaction mixture was allowed to cool, diluted with EtOAc (150 mL), then washed sequentially with a saturated solution of sodium thiosulfate (2 x 100 mL), a saturated solution of $NaHCO_3$ (2 x 100 mL), brine (100 mL), then dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 15% EtOAc in *c*-hexane) gave **70** as a pale yellow solid (2.47 g, 65%); δ_H (400 MHz, DMSO): 2.29 (3H, s, *p-Me*), 6.90 (1H, s, C(5)*H*), 7.01 (2H, br. s, NH_2), 7.09 – 7.24 (2H, m, C(*m*-)*H*, C(*m*-)*H*), 7.58 – 7.74 (2H, m, C(*o*-)*H*, C(*o*-)*H*); δ_C (101 MHz, DMSO-d_6): 20.84 (*p-Me*), 100.63 (*C*(5)), 125.54 (*C*(o-)), 129.09 (*C*(*m*-)), 132.31 (*C*(*i*-)), 136.45 (*C*(*p*-)), 149.93 (*C*(4)),

168.17 (*C*(2)); *m/z* (API-ES+) 191 ([M+H]$^+$, 100%); All other data were in accordance with the literature values.[6]

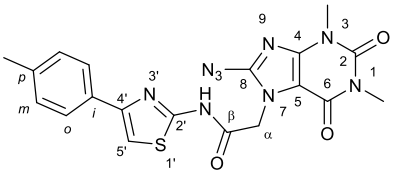
_N_-(4′-(_p_-Tolyl)thiazol-2′-yl)-α-(theopyllin-7-yl)-acetamide 3-55



Step 1: $SOCl_2$ (542 µL, 7.42 mmol) was added to a solution of theophylline-7-acetic acid (1.82 g, 7.62 mmol) in anhydrous DMF (30 mL) at rt and the resulting solution was allowed to stir for 1 h.

Step 2: Anhydrous pyridine (2.34 mL, 30.0 mmol) followed by **70** (381 mg, 2.00 mmol) were added to the crude solution from step 1 and the resulting solution was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (150 mL), then washed sequentially with H_2O (3 x 100 mL), brine (100 mL), then dried and concentrated *in vacuo*. Purification by RP-FCC (sample loading: DMSO, gradient elution: 2 - 60% solvent B in solvent A) gave **3-55** as a yellow solid (357 mg, 44%); ν_{max} (ATR, cm^{-1}) 1702, 1649 (NC=O); δ_H (500 MHz, DMSO-d_6): 2.33 (3H, s, *p-Me*), 3.18 (3H, s, N(1)*Me*), 3.46 (3H, s, N(3)*Me*), 5.34 (2H, s, C(α)H_2), 7.24 (2H, d, *J* 8.0, C(*m*-)*H*), 7.57 (1H, s, C(5′)*H*), 7.71 – 7.86 (2H, m, C(*o*-)*H*), 8.09 (1H, s, C(8)*H*), 12.75 (1H, s. N*H*); δ_C (126 MHz, DMSO-d_6): 20.82 (*p-Me*), 27.47 (N(1)*Me*), 29.51 (N(3)*Me*), 48.22 (C(α)H_2, 106.43 (*C*(5)), 107.56 (*C*(5′)), 125.64 (*C*(*o*-)), 129.35 (*C*(*m*-)), 131.49 (*C*(*i*-)), 137.25 (*C*(*p*-)), 143.71 (*C*(8)), 148.04 (*C*(4)), 149.09 (*C*(4′)), 151.03 (C(2)), 154.51 (*C*(6)), 157.33 (*C*(2′)), 165.62 (*C*(β)); *m/z* (API-ES$^+$) 411 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{19}H_{19}O_3N_6S^+$ calculated 411.1234; found 411.1234.

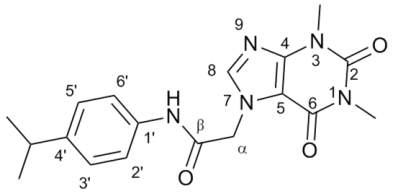
***N*-(4′-(*p*-Tolyl)thiazol-2′-yl)-α-(8-azido-theophyllin-7-yl)-acetamide 3-64**



Step 1: $SOCl_2$ (129 µL, 1.77 mmol) was added to a solution of **68** (506 mg, 1.81 mmol) in anhydrous DMF (7.0 mL) at rt and the resulting solution was allowed to stir for 1 h.

Step 2: Anhydrous pyridine (578 µL, 7.14 mmol) followed by **70** (90.6 mg, 476 µmol) were added to the crude solution from step 1 at rt and the resulting solution was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (50 mL), then washed sequentially with H_2O (3 x 50 mL), brine (50 mL), then dried and concentrated *in vacuo*. Purification by RP-FCC (sample loading: DMSO, gradient elution: 2 - 60% solvent B in solvent A) gave **3-64** as a brown solid (26.0 mg, 12%); ν_{max} (ATR, cm^{-1}) 2154 (N_3), 1699, 1647 (NC=O); δ_H (400 MHz, DMSO-d_6): 2.33 (3H, s, C(*p*-)*Me*), 3.18 (3H, s, N(1)*Me*), 3.43 (3H, s, N(3)*Me*), 5.08 (2H, s, C(α)H_2), 7.24 (2H, m, C(*m*-)H, C(*m*-)H), 7.59 (1H, s, C(5′)*H*), 7.69 – 7.89 (2H, m, C(*o*-)*H*, C(*o*-)*H*), 12.76 (1H, s, N*H*); δ_C (101 MHz, DMSO-d_6): 20.81 (C(*p*-)*Me*, 27.48 (N(1)*Me*), 29.66 (N(3)*Me*), 46.09 (*C*(α)), 105.33 (*C*(5)), 107.67 (*C*(5′)), 125.63 (*C*(*o*-), *C*(*o*-)), 129.34 (*C*(*m*-), *C*(*m*-)), 131.44 (*C*(*i*-)), 137.26 (*C*(*p*-)), 146.08 (*C*(8)), 146.67 (*C*(4)), 149.10 (*C*(4′)), 150.71 (*C*(2)), 153.86 (*C*(6)), 157.16 (*C*(2′)), 164.64 (*C*(β)); *m/z* (API-ES^+) 452 ($[M+H]^+$, 100%); HRMS (ESI^+) $C_{19}H_{18}N_9O_3S^+$ calculated 452.1248; found 452.1247.

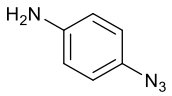
***N*-(4′-Isopropylphenyl)-α-(theophyllin-7-yl)-acetamide HC-030031**



Step 1: $SOCl_2$ (271 µL, 3.71 mmol) was added to a solution of theophylline-7-acetic acid (908 mg, 3.81 mmol) in anhydrous DMF (15 mL) at rt and the resulting solution was allowed to stir for 1 h.

Step 2: Anhydrous pyridine (1.21 mL, 15.0 mmol) and isopropylaniline (137 µL, 1.00 mmol) were added sequentially to the crude solution from step 1 and the resulting solution was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (50 mL), then washed sequentially with H_2O (3 x 50 mL), brine (50 mL), then dried and concentrated *in vacuo*. Purification by RP-FCC (sample loading: DMSO, gradient elution: 2 - 48% solvent B in solvent A) gave **HC-030031** as an off-white solid (213 mg, 60%); ν_{max} (ATR, cm^{-1}) 1701, 1652 (NC=O); δ_H (400 MHz, DMSO-*d*$_6$): 1.17 (6H, d, *J* 6.9, CH*Me*$_2$), 2.83 (1H, p, *J* 6.9, C*H*Me$_2$), 3.19 (3H, s, N(1)*Me*), 3.45 (3H, s, N(3)*Me*), 5.18 (2H, s, C(α)*H*$_2$), 7.18 (2H, d, *J* 8.5, C(3′)*H*, C(5′)*H*), 7.47 (2H, d, *J* 8.5, C(2′)*H*, C(6′)*H*), 8.06 (1H, s, C(8)*H*), 10.32 (1H, s, N*H*); δ_C (101 MHz, DMSO): 23.93 (CH*Me*$_2$), 27.46 (N(1)*Me*), 29.48 (N(3)*Me*), 32.85 (*C*HMe$_2$), 48.68 (*C*(α)), 106.47 (*C*(5)), 119.15 (*C*(2′), *C*(6′)), 126.57 (*C*(3′), *C*(5′)), 136.36 (*C*(1′)), 143.64 (*C*(6′)), 143.79 (*C*(8)), 147.96 (*C*(4)), 151.03 (*C*(2), 154.48 (*C*(6)), 164.62 (*C*(β)); *m/z* (ESI^+) 356 ([M+H]+, 53%), 378 ($[M+Na]^+$, 100%); HRMS (ESI^+) $C_{18}H_{22}N_5O_3^+$ calculated 356.1717; found 356.1719.

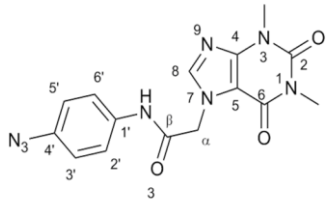
4-Azido-aniline 63 [7]



According to a literature procedure,[7] CuI (74.7 mg, 392 µmol) was added to a solution of 4-bromoaniline (750 mg, 4.36 mmol), NaN_3 (510 mg, 7.85 mmol), sodium ascorbate (43.2 mg, 218 µmol) and N,N′-dimethyl-ethyldiamine (64 µL, 610 µmol) in an EtOH/H_2O mixture (7:3 v/v, 8 mL). The resultant mixture was degassed with Ar for 15 min and then heated to 90 °C for 40 min. The mixture was diluted with H_2O (50 mL), then extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with brine (50 mL), then dried and concentrated *in vacuo* to give **63** as a

brown solid which was used without further purification (335 mg, 57%); δ_H (400 MHz, $CDCl_3$): 3.62 (2H, br. s, N*H*$_2$), 6.63 – 6.73 (2H, m), 6.80 – 6.94 (2H, m); δ_C (101 MHz, $CDCl_3$): 116.30, 120.00, 130.21, 143.77; *m/z* (API-ES$^+$) 135 ([M+H]$^+$, 100%); All other data were in accordance with the literature values.[7]

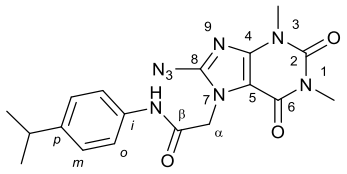
_N_-(4′-Azidophenyl)-α-(theophyllin-7-yl)-acetamide 4-33



Step 1: $SOCl_2$ (80 µL, 1.11 mmol) was added to a solution of theophylline-7-acetic acid (271 mg, 1.14 mmol) in anhydrous DMF (5 mL) at rt and the resulting solution was allowed to stir for 1 h.

Step 2: Anhydrous pyridine (362 µL, 4.47 mmol) followed by a solution of **63** (40 mg, 298 µmol) in anhydrous DMF (5 mL) were added to the crude solution from step 1 and the resulting solution was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (50 mL), then washed sequentially with H_2O (3 x 50 mL), 1M HCl (3 x 50 mL), a saturated solution of $NaHCO_3$ (3 x 50 mL), brine (50 mL), then dried and concentrated *in vacuo*. Purification by RP-FCC (sample loading: DMSO, gradient elution: 2 - 40% solvent B in solvent A) gave **4-33** as an off-white solid (55 mg, 52%); ν_{max} (ATR, cm^{-1}) 2122 (N_3), 1700, 1653 (NC=O); δ_H (500 MHz, DMSO-*d*$_6$): 3.19 (3H, s, N(1)*Me*), 3.45 (3H, s, N(3)*Me*), 5.19 (2H, s, C(α)*H*$_2$), 7.09 (2H, d, *J* 9.0, C(2′)*H*, C(6′)*H*), 7.61 (2H, d, *J* 9.0, C(3′)*H*, C(5′)*H*), 8.06 (1H, s, C(8)*H*), 10.50 (1H, s, N*H*); δ_C (126 MHz, DMSO-*d*$_6$): 27.48 (N(1)*Me*), 29.50 (N(3)*Me*), 48.70 (*C*(α)), 106.47 (*C*(5)), 119.61 (*C*(2′), *C*(6′)), 120.59 (*C*(3′), *C*(5′)), 134.22 (*C*(4′)), 135.85 (*C*(1′)), 143.79 (*C*(8)), 147.97 (*C*(4)), 151.04 (*C*(2)), 154.51 (*C*(6)), 164.86 (*C*(β)); *m/z* (API-ES$^+$) 355 ([M+H]$^+$, 100%); HRMS (ESI$^+$) $C_{15}H_{14}N_8O_3Na^+$ calculated 377.1081; found 377.1081.

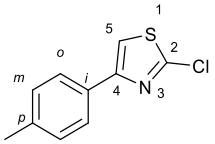
2-(8-Azido-theophyllin-7-yl)-*N*-(*p*-isopropylphenyl)acetamide 4-35



Step 1: $SOCl_2$ (284 µL, 3.90 mmol) was added to a solution of **68** (1.12 g, 4.00 mmol) in anhydrous DMF (25 mL) at rt and the resulting solution was allowed to stir for 1 h.

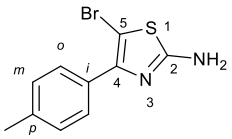
Step 2: Anhydrous pyridine (1.27 mL, 15.8 mmol) followed by isopropylaniline (143 µL, 1.05 mmol) were added to the crude solution from step 1 and the resulting solution was allowed to stir for 16 h at rt. The reaction mixture was diluted with CH_2Cl_2 (100 mL), then washed sequentially with H_2O (3 x 100 mL), 1M HCl (3 x 100 mL), a saturated solution of $NaHCO_3$ (3 x 100 mL), brine (100 mL), then dried and concentrated *in vacuo*. Purification by RP-FCC (sample loading: DMSO, gradient elution: 2 - 50% solvent B in solvent A) gave **4-35** as an off-white solid (23 mg, 6%); ν_{max} (ATR, cm^{-1}) 2154 (N_3), 1704, 1657 (NC=O); δ_H (500 MHz, DMSO-*d*$_6$): 1.17 (6H, d, *J* 7.0, (CH*Me*$_2$), 2.83 (1H, hept, *J* 7.0, C*H*Me$_2$), 3.19 (3H, s, N(1)*Me*), 3.43 (3H, s, N(3)*Me*), 4.92 (2H, s, C(α)*H*$_2$), 7.18 (2H, d, *J* 8.5, C(*m*-)*H*), 7.45 (2H, d, *J* 8.5, C(*o*-)*H*), 10.32 (1H, s, N*H*CO); δ_C (126 MHz, DMSO-d_6): 23.95 (CH*Me*$_2$), 27.49 (N(1)*Me*), 29.67 (N(3)*Me*), 32.88 (*C*HMe$_2$), 46.61 (*C*(α)), 105.48 (*C*(5)), 119.21 (*C*(*o*-)), 126.61 (*C*(*m*-)), 136.18 (*C*(*i*-)), 143.83 (*C*(*p*-)), 146.11 (*C*(8)), 146.60 (*C*(4)), 150.74 (*C*(2)), 153.90 (*C*(6)), 163.62 (*C*(β)); *m/z* (ESI^+) 419 ($[M+Na]^+$, 100 %); HRMS (ESI^+) $C_{18}H_{20}N_8O_3Na^+$ calculated 419.1551; found 419.1549.

2-Chloro-4-(*p*-tolyl)thiazole 72



tert-Butyl nitrite (59.3 µL, 788 µmol) was added to a stirred solution of **70** (100 mg, 526 µmol) and CuCl (78.1 mg, 788 µmol) in MeCN (5 mL) at rt and the resulting solution was stirred for 4 h at rt. The reaction mixture was concentrated *in vacuo*. The crude residue was dissolved in EtOAc (50 mL), washed with brine (50 mL), dried, filtered through Celite® (eluent EtOAc) and concentrated *in vacuo*. Purification by flash column chromatography (gradient elution: 3 – 6% Et_2O in *c*-hexane) gave **72** as an off-white solid (42.9 mg, 39%); δ_H (400 MHz, DMSO-*d*$_6$): 2.33 (3H, s, C*H*$_3$), 7.10 – 7.36 (2H, m, C(*m*-)*H*), 7.65 – 7.87 (2H, m, C(*o*-)*H*), 8.03 (1H, s, C(5)*H*); δ_C (101 MHz, DMSO-*d*$_6$): 20.80 (*C*H$_3$), 116.11 (*C*(5)), 125.73 (*C*(*o*-)), 129.45 (*C*(*m*-)), 130.35 (*C*(*i*-)), 138.06 (*C*(*p*-)), 150.42 (*C*(4)), 152.96 (*C*(2)); m/z (API-ES$^+$) 210 ([M(^{35}Cl)+H]$^+$, 100%), 212 ([M(^{37}Cl)+H]$^+$, 37%); All other data were in accordance with the literature values.[8]

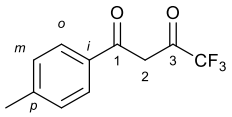
5-Bromo-4-(*p*-tolyl)thiazol-2-amine 74



Copper (II) bromide (500 mg, 2.24 mmol) was added to a stirred solution of **70** (381 mg, 2.00 mmol) in MeCN (8 mL) at rt and the resulting mixture was stirred at rt for 48 h. The reaction mixture was diluted with EtOAc (100 mL), washed with saturated solution of NH_4Cl (2 x 50 mL), dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 15 – 20% EtOAc in *c*-hexane) gave **74** as a pink solid (395 mg, 73%); δ_H (400 MHz, DMSO-*d*$_6$): 2.32 (3H, s, C*H*$_3$), 7.19 – 7.25 (2H, m, C(*m*-)*H*), 7.27 (2H, s, N*H*$_2$), 7.66 – 7.76 (2H, m, C(*o*-)*H*); δ_C (101 MHz, DMSO-*d*$_6$): 20.83 (*C*H$_3$), 86.37 (*C*(5)), 127.76 (*C*(o-)), 128.68 (*C*(*m*-)), 130.97 (*C*(*i*-)), 137.25 (C(*p*-)), 147.17

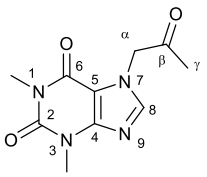
(*C*(4)), 166.82 (*C*(2)); *m/z* (API-ES$^+$) 269 ([M(^{79}Br)+H]$^+$, 100%), 271 ([M(^{81}Br)+H]$^+$, 100%); All other data were in accordance with the literature values.[9]

4,4,4-Trifluoro-1-(*p*-tolyl)butane-1,3-dione 77[10]



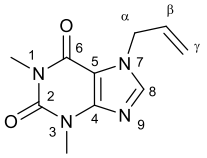
60% NaH in mineral oil (1.04 g, 26.0 mmol) was added to a stirred solution of ethyl trifluoroacetate (2.57 mL, 24.0 mmol) in anhydrous THF (40 mL) at rt. The resulting suspension was cooled to 0 °C before 4-methylacetophenone (2.65 mL, 20 mmol) was added dropwise at 0 °C. The reaction mixture was then heated to reflux for 2 h before being cooled to 0 °C. The mixture was acidified with 1 M aq. HCl and the reaction mixture stirred further for 15 min. The reaction mixture was then neutralised with a saturated solution of $NaHCO_3$, and partitioned between H_2O (100 mL) and EtOAc (100 mL). The aqueous layer was extracted with EtOAc (2 x50 mL). The combined organic layers were washed with brine (50 mL), dried and concentrated *in vacuo*. Purification by flash column chromatography (eluent 1 – 5% EtOAc in *c*-hexane) gave **77** as an orange solid (1.97 g, 42%); δ_H (400 MHz, DMSO-*d*$_6$) 2.30 (3H, s, C*H*$_3$), 5.91 (2H, s, C(2)*H*$_2$), 7.12 – 7.26 (2H, m, C(*m*-)*H*), 7.61 – 7.78 (2H, m, C(o-)*H*); δ_c (400 MHz, DMSO-*d*$_6$) 20.90 (*C*H$_3$), 86.84 (*C*(2)), 119.43 (q, *J* 2.9, *C*F$_3$), 126.66 (C(o-)), 128.71 (C(m-), 138.97 (C(i-)), 139.90 (C(p-)), 168.85 (q, *J* 0.28, *C*OCF$_3$), 185.63 (*C*(1)O); *m/z* (API-ES$^+$) 231 ([M+H]$^+$, 100%); all other data were in accordance with the literature values.[10]

7-(β-Oxopropyl)-theophylline 80[11]



According to a literature procedure,[11] chloroacetone (438 µL, 5.50 mmol) was added to a stirred suspension of theophylline (901 mg, 5.00 mmol) and K_2CO_3 (691 mg, 5.00 mmol) in anhydrous DMF (9 mL) at rt. The resulting suspension was heated to 120 °C for 1 h before being cooled to rt. H_2O (100 mL) was added, the resulting mixture neutralised with 0.1 M aq. HCl, then extracted with $CHCl_3$ (3 x 100 mL). The combined organic layers were dried and concentrated *in vacuo*. Recrystalisation of the crude mixture from EtOH gave **80** as an off-white solid (662 mg, 56%); δ_H (400 MHz, DMSO-d_6): 2.22 (3H, s, C(γ)H_3), 3.18 (3H, s, N(1)*Me*), 3.43 (3H, s, N(3)*Me*), 5.24 (2H, s, C(β)H_2), 7.94 (1H, s, C(8)*H*); δ_C (101 MHz, DMSO-d_6): 26.88 (*C*(γ)), 27.43 (N(1)*Me*), 29.47 (N(3)*Me*), 54.92 (*C*(α)), 106.29 (*C*(5)), 143.05 (*C*(8)), 147.90 (*C*(4)), 151.02 (*C*(2)), 154.45 (*C*(6)), 201.19 (C(β)); *m/z* (API-ES$^+$) 237 ([M+H]+, 100%); all other data were in accordance with the literature values.[11]

7-Allyl-theophylline 81[12]



According to a literature procedure,[12] allyl bromide (480 µL, 5.55 mmol) was added to a stirred suspension of theophylline (901 mg, 5.00 mmol) and K_2CO_3 (781 mg, 5.55 mmol) in anhydrous DMF (9 mL) at rt and the resulting mixture was heated to 35 °C for 4 h. The reaction mixture was cooled to rt, diluted with H_2O (100 mL), then extracted with CHCl3 (3 x 50 mL). The combined organic layers were dried and concentrated *in vacuo* to give **81** as an off-white solid which was used without further purification (1.07 g, 97%); δ_H (400 MHz, DMSO-d_6): 3.21 (3H, s, N(1)*Me*), 3.42 (3H, s,

N(3)*Me*), 4.89 (2H, dt, *J* 1.5, 5.5, C(α)H_2), 5.08 (1H, dq, *J* 1.5, 17.0, C(γ)H_AH_B), 5.19 (1H, dq, *J* 1.5, 10.5, C(γ)H_AH_B), 6.05 (1H, ddt, *J* 5.5, 10.5, 17.0, C(β)*H*), 8.07 (1H, s, C(8)*H*); δ_C (101 MHz, DMSO-d_6): 27.50 (N(1)*Me*), 29.42 (N(3)*Me*), 48.01 (*C*(α)), 105.90 (*C*(5)), 117.73 (*C*(γ)), 133.64 (*C*(β)), 142.24 (*C*(8)), 148.27 (*C*(4)), 151.01 (*C*(2)), 154.24 (*C*(6)); *m/z* (API-ES$^+$) 221 ([M+H]$^+$, 100%); all other data were in accordance with the literature values.[12]

2. Solid Phase Peptide Synthesis (SPPS)

SPPS was performed using a Liberty Blue (CEM inc.) automated peptide synthesis platform using standard FMOC-chemistry and rink-amide MHBA resin (Merck). Solutions of Fmoc-protected amino acids (0.2 M in DMF), Oxyma Pure (1.0 M in DMF with 8% v/v DIPEA) and diisoproylcarbodiimide (DIC, 0.5 M in DMF) were freshly prepared for each SPPS run. 20% piperidine in DMF was used to remove N-terminal FMOC-groups. Peptides bound to resin were washed with CH_2Cl_2 and dried *in vacuo*, N-terminal acetylation was performed using Ac_2O (3 mL) at rt for 3 h. Global deprotection and peptide-resin cleavage was realised using a mixture of CF_3CO_2H, H_2O, $HSiMe_3$, and 1,3-dimethoxybenzene (92.5:2.5:2.5:2.5 v/v/v/v). Typically the resin was incubated with the cleavage mixture (5 mL) for 3 h at rt.[13]

In the case of longer peptides (>10mer), the resin was filtered off, washed with CH_2Cl_2 and the filtrate diluted with Et_2O and cooled to 4 °C for 16 h. The crude peptide was collected by centrifugation (Beckman Coulter, SX-4400, 4000 rpm, 5 min) and the supernatant discarded. The crude peptide was diluted with H_2O and lyophilised.

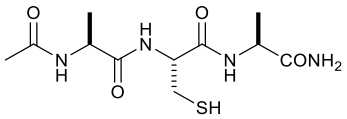
In the case of shorter peptides (<10mer), the resin was filtered, washed with CH_2Cl_2 and the filtrate concentrated *in vacuo*.

Purification was carried out by semi-preparative HPLC using a Phenomenex Gemini-NX 5 C18 (250 x 30 mm) or Waters Sunfire (150 x 10 mm) column on a Shimadzu HPLC system (see §1.1), or by reverse phase column chromatography utilising a Biotage system (see §1.1) and KP-SNAP-C18

columns. Peptide identity was confirmed using LCMS (see §1.1) and MALDI-TOF (see §7) for longer peptides. Identity of shorter peptides was confirmed by LCMS and NMR.

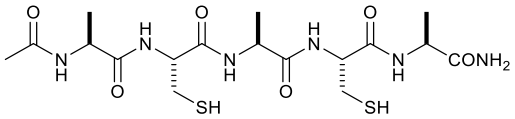
2.1 Peptide Characterisation Data

Acetyl-Alanine-Cysteine-Alanine (Ac-ACA-NH_2)



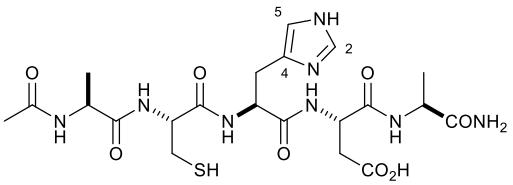
m/z (API-ES^+) 305 ($[M+H]^+$, 100%); δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.33 (3H, d, *J* 7.3, Ala-*Me*), 1.36 (3H, d, *J* 7.3, Ala-*Me*), 1.97 (3H, s, NCO*Me*), 2.88 (2H, m, Cys-C(β)H_2), 4.24 (2H, app ddt, *J* 3.2, 7.3, 10.0, Ala-C(α)*H*), 4.42 - 4.47 (1H, m, Cys-C(α)*H*); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.34 (Ala-*Me*), 16.78 (Ala-*Me*), 21.64 (NCO*Me*), 25.16 (Cys-*C*(β)), 49.60 (Ala-*C*(α)), 49.85 (Ala-*C*(α)), 55.36 (Cys-*C*(α)), 171.56 (Cys-*C*O), 174.25 (NCO*Me*), 175.71 (Ala-*C*O), 177.60 (Ala-*C*$CONH_2$).

Acetyl-Alanine-Cysteine-Alanine-Cysteine-Alanine (Ac-ACACA-NH_2)



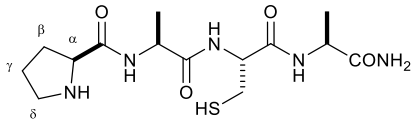
m/z (API-ES^+) 479 ($[M+H]^+$, 100%); δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.31 (3H, d, *J* 7.4, Ala-*Me*), 1.35 (6H, d, *J* 7.2, Ala-*Me,* Ala-*Me*), 1.88 (3H, s, NCO*Me*), 2.83 – 2.90 (4H, m, Cys-C(β)H_2, Cys-C(β)H_2), 4.18 – 4.29 (3H, m, Ala-C(α)*H*, Ala-C(α)*H*, Ala-C(α)*H*), 4.42 – 4.48 (2H, m, Cys-C(α)*H*, Cys-C(α)*H*); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.24 (Ala-*Me*), 16.55 (Ala-*Me*), 16.80 (Ala-*Me*), 21.64 (NCO*Me*), 25.12 (Cys-*C*(β)), 25.20 (Cys-*C*(β)), 49.62 (Ala-*C*(α)), 49.83 (Ala-*C*(α)), 50.04 (Ala-*C*(α)), 53.20 (Cys-*C*(α)), 53.49 (Cys-*C*(α)), 171.16 (Cys-*C*O), 171.74 (Cys-*C*O), 174.32 (NCO*Me*), 175.64 (Ala-*C*O), 175.75 (Ala-*C*O), 177.60 (Ala-*C*$CONH_2$).

Acetyl-Alanine-Cysteine-Histidine-Aspartate-Alanine (Ac-ACHDA-NH$_2$)



m/z (API-ES$^+$) 557 ([M+H]$^+$, 100%); δ_H (600 MHz, 50 mM Tris-*d$_{11}$* in D_2O, pH 7.5): 1.29 (3H, d, *J* 7.4, Ala-*Me*), 1.35 (3H, d, *J* 7.4, Ala-*Me*), 1.96 (3H, s, NCO*Me*), 2.72 (1H, dd, *J* 7.8, 16.3, Asp-C(β)*H$_A$*H$_B$), 2.79 – 2.86 (3H, m, Cys-C(β)*H$_2$*, Asp-C(β)H$_A$*H$_B$*), 3.14 (1H, dd, *J* 7.7, 15.3, His-C(β)*H$_A$*H$_B$), 3.25 (1H, dd, *J* 6.0, 15.3, His-C(β)H$_A$*H$_B$*), 4.14 – 4.26 (2H, m, Ala-C(α)*H*, Ala-C(α)*H*), 4.35 – 4.40 (1H, m, Cys-C(α)*H*), 4.57 – 4.72 (2H, obscured, Asp-C(α)*H*, His-C(α)*H*), 7.25 (1H, s, His-C(5)*H*), 8.57 (1H, s, His-C(2)*H*); δ_C (151 MHz, 50 mM Tris-*d$_{11}$* in D_2O, pH 7.5): 16.43 (Ala-*Me*), 16.64 (Ala-*Me*), 21.63 (NCO*Me*), 24.95 (Cys-*C*(β)), 26.08 (His-*C*(β)), 36.38 (Asp-*C*(β)), 49.61 (Ala-*C*(α)), 49.80 (Ala-*C*(α)), 50.57 (Asp-*C*(α)), 52.53 (His-*C*(α)), 55.57 (Cys-*C*(α)), 117.54 (His-*C*(5)), 128.25 (His-*C*(4)), 133.31 (His-*C*(2)), 171.09 (His-*C*O), 172.00 (Asp-*C*O), 173.27 (Cys-CO) 174.19 (N*C*OMe) 174.86 (Asp-*C*(γ)O), 175.46 (Ala-*C*O), 177.57 (Ala-*C*ONH$_2$).

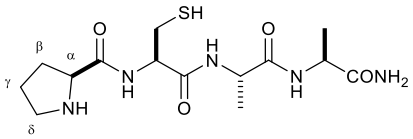
Proline-Alanine-Cysteine-Alanine (PACA-NH$_2$)



m/z (API-ES$^+$) 360 ([M+H]$^+$, 100%); δ_H (600 MHz, 50 mM Tris-*d$_{11}$* in D_2O, pH 7.5): 1.31 – 1.41 (6H, m, Ala-*Me*, Ala-*Me*), 1.93 – 2.14 (3H, m, Pro-C(γ)*H$_2$*, Pro-C(β)*H$_A$*H$_B$), 2.33 – 2.55 (1H, m, Pro-C(β)H$_A$*H$_B$*), 2.79 – 2.99 (2H, m, Cys-C(β)*H$_2$*), 3.31 – 3.56 (2H, m, Pro-C(δ)*H$_2$*), 4.20 – 4.38 (3H, m, Ala-C(α)*H,* Ala-C(α)*H*, Pro-C(α)*H*), 4.42 – 4.46 (1H, m, Cys-C(α)*H*); δ_C (151 MHz, 50 mM Tris-*d$_{11}$* in D_2O, pH 7.5): 16.24 (Ala-*Me*), 16.74 (Ala-*Me*), 23.73 (Pro-*C*(γ)), 25.25 (Cys-*C*(β)), 29.76 (Pro-

C(β)), 46.58 (Pro-*C*(δ)), 49.46 (Ala-*C*(α)), 49.93 (Ala-*C*(α)), 55.39 (Cys-*C*(α)), 59.51 (Pro-*C*(α)), 169.52 (Pro-*C*O), 171.32 (Cys-*C*O), 174.63 (Ala-*C*O), 177.39 (Ala-*C*$CONH_2$).

Proline-Cysteine-Alanine-Alanine (PCAA-NH_2)



m/z (API-ES^+) 360 ($[M+H]^+$, 100%); δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.30 – 1.40 (6H, m, Ala-*Me*, Ala-*Me*), 1.95 – 2.06 (3H, m, Pro-C(γ)H_2, Pro-C(β)H_AH_B), 2.33 – 2.51 (1H, m, Pro-C(β)H_AH_B), 2.84 – 2.92 (2H, m, Cys-C(β)H_2), 3.29 – 3.44 (2H, m, Pro-C(δ)H_2), 4.22 (1H, q, *J* 7.0, Ala-C(α)*H*), 4.27 (1H, q, *J* 7.0, Ala-C(α)*H*), 4.34 – 4.40 (1H, m, Cys-C(α)*H)* 4.41 – 4.50, (1H, m, Pro-C(α)*H*); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.42 (Ala-*Me*), 16.71 (Ala-*Me*), 23.92 (Pro-*C*(γ)), 25.43 (Cys-*C*(β)), 29.84 (Pro-*C*(β)), 46.61 (Pro-C(δ)), 49.44 (Ala-*C*(α)), 49.79 (Ala-*C*(α)), 56.07 (Cys-*C*(α)), 59.63 (Pro-*C*(α)), 169.31 (Pro-*C*O), 171.60 (Cys-*C*O), 174.57 (Ala-*C*O), 177.71 (Ala-*C*$CONH_2$).

3. Cloning, Recombinant Protein Expression and Purification

3.1 General Procedures

DNA sequences encoding for recombinant protein constructs were designed, synthesised (Gene Art), and amplified using PCR. The DNA sequence of the construct was confirmed using Sanger sequencing (Source Bioscience). Heat shock transformation of *E. coli* was performed by incubation of 20 µL of the relevant cell line with 1 µL plasmid DNA on ice for 10 min. The mixture was then incubated at 42 °C for 30 seconds before being incubated on ice for 20 min. 250 µL SOC media was added, then the culture was incubated at 37 °C for 45 min.[14] The culture was then spread over an agar plate supplemented with a selection antibiotic (100 µg/mL ampicillin or 50 µg/mL kanamycin) and incubated at 37 °C for 16 h. A single colony was then selected and inoculated into 2-YT (100 mL) supplemented with a selection antibiotic at 37 °C for 16 h. This culture was then split evenly between

8 x 1 L 2-YT media supplemented with a selection antibiotic and grown to OD_{600} ~ 0.6 at 37 °C before recombinant protein production was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG).[15,16] Cells were harvested by centrifugation (Beckman Coulter Avanti J-26s XP, JLA-10-500, 8000 rpm, 10 min, 4 °C). The cell pellet was then stored at -80 °C until required. Cell lysis was carried out by sonication (Sonics VibraCell) with 2 second pulses followed by 2 second rests for a total of 15 min in 50 mM Tris pH 7.5 supplemented with 0.1 mM phenylmethylsulfonyl fluoride (Thermo Scientific) and 30 µg mL^{-1} DNase I (Roche) at 4 °C. The cell lysate was clarified by centrifugation (Beckman Coulter Avanti J-26s XP, JA-25.50, 20000 rpm, 30 min, 4 °C) and filtration through a 0.44 µm filter. Fast protein liquid chromatography (FPLC) was carried out using an AKTA platform (GE Healthcare). Nickel affinity chromatography was carried out using a 5 mL HisTrap (GE healthcare) eluting with a gradient of imidazole (50 mM to 250 mM) in 50 mM Tris pH 7.5. Proteins were concentrated using an Amicon® Ultra-15 centrifugal filter (Merck), flash frozen in liquid N_2 and stored at -80 °C until use.

3.2 Polymerase Chain Reaction[17]

Primers:

BK3-F:	*TACTTCCAATCCATGGGTAGCGATCTG*
BK3-R:	*TATCCACCTTTACTGTTACTGCAGGATTTCCGC*
T7F-pcDNA3:	*TAATACGACTCACTATAGGGAGA*
BGH-R:	*TAGAAGGCACAGTCGAGG*
BAMHI-F:	*GTACCGAGCTCGGATCC*
XHOI-R:	*GGGCCCTCTAGACTCGAG*

A master mixture of Q5 polymerase (New England Biolabs), 10 mM deoxynucleotide triphosphate mix, 10 µM forward primer, 10 µM reverse primer in Q5 buffer (New England Biolabs) was prepared and incubated on ice. A small amount of template DNA (<100 ng) was added to a PCR tube before master mix was added to give a final volume of 25 µL. In the case of bacterial colony PCR, a small

section of a colony on an agar plate was scraped using a pipette tip and mixed into 25 µL of MQ H_2O. 1 µL of the resulting mixture was added to a PCR tube containing 24 µL of master mixture.

The thermocycler was preheated to 98 °C and the PCR tubes added. The annealing temperature was calculated using an online tool from New England Biolabs (tmcalculator.neb.com). After an initial denaturing of the mixture at 98 °C for 5 min, the mixture was subjected to 40 cycles of denaturing at 98 °C for 10 s, annealing at the calculated temperature for 30 s, and polymerisation at 72 °C for 2 min. A final extension was carried out at 72 °C for 2 min at the end of the cycle. The samples were then cooled to 4 °C. Analysis of the PCR mixture was achieved through gel electrophoresis using either a 0.5 or 1% (w/v) agarose gel in TAE buffer. Agarose gels were pre-stained with SYBR safe (Invitrogen). Selected bands were excised using a scalpel and extracted using a GeneJet gel extraction kit (Thermo Scientific) according to the manufacturer's protocol.[18]

3.3 Ligation Independent Cloning[19]

50 µL of purified PCR amplification product was treated with LIC qualified T4 polymerase (New England Biolabs) in buffer 2 (New England Biolabs) supplemented with 2.5 mM deoxycytosine triphosphate, 5 mM dithioreitol and 0.1 mg mL^{-1} bovine serum albumin to give a give a final reaction volume of 100 µL. The reaction mixture was incubated at rt for 30 min before being heated to 65 °C for 10 min. The reaction mixture was then mixed with LIC pre-treated pNIC-BSA4 (a kind gift from Dr. G. Langley, University of Oxford) in a 3:1 ratio at rt for 10 min before the resulting mixture was used to transform chemically competent XL-10 gold *E. coli*.

3.4 Subcloning DNA Encoding for TRPA1 into the pcDNA5/TO Vector

The TRPA1 insert from the TRPA1-pcDNA3.1 vector (a kind gift from Dr. Y. Wang, University of Southampton) was amplified using PCR with the T7F-pcDNA3 and BGH-R primers (see §3.2). The

amplified insert and a pcDNA5/TO vector (a kind gift from Dr. T. Durrant, University of Oxford) were separately subjected to a double digestion protocol using 16 µL DNA construct, 1 µL XhoI (New England Biolabs/NEB), 1 µL HindIII (NEB) and 2 µL 10 x CutSmart buffer (NEB) at rt for 15 min. The mixture was then heated to 60 °C for 15 min before being cooled to rt. The DNA was purified using a GeneJet PCR purification kit (Thermo Scientific). T4 ligation was carried out at rt for 15 min using a 3:1, 5:1 and 10:1 molar ratio of the TRPA1 insert to pcDNA5/TO backbone with 1 µL T4 DNA ligase (NEB) in T4 DNA ligase buffer (NEB) with a final reaction mixture volume of 20 µL. The reaction mixture was then heated to 65 °C for 10 min before the mixture was used to transform chemically competent XL-10 gold *E. coli*.

3.5 Synthetic Ankyrin Protein (BK3)

Plasmid: pNIC-BSA4 (kanamycin resistance)[20]

DNA Insert Encoding for BK3:

ATGGGTAGCGATCTGGGTAAAAAACTGCTGGAAGCGGCGCGTGCGGGTCAGGATGATGAAG
TGCGTATCCTGATGGCGAACGGTGCGGATGTGGCGGCGAAAAAAATCTTCAGCCATAATAGC
CCTGGTAACAAATGTCCGATTACCGAAATGATTGAATATCTGCCGGAATGCATGAAAGTGCTGC
TGGATTTTTGTATGCTGCATAGCACCGAAGATAAAAGCTGTCGCGATTATTACATCGAGTACAA
CTTCAAATATCTGCAGTGTCCGCTGGAATTCACCAAAAAAACCCCGACACAGGATGTTATTTAT
GAACCGCTGACCGCACTGAATGCAATGGTTCAGAATGGTCACCTGGAAGTGGTGAAACTGCT
GCTGGAAGCGGGTGCGGATGTGGCGGCGAAAGATAAATTCGGTAAAACCGCGTTCGATATCA
GCATCGATAACGGTAACGAAGATCTGGCGGAAATCCTGCAGTAA

Protein Sequence for BK3 (as expressed using the PNIC-BSA4 vector):

MHHHHHHSSGVDLGTENLYFQSMGSDLGKKLLEAARAGQDDEVRILMANGADVAAKKIFSHNS
PGNKCPITEMIEYLPECMKVLLDFCMLHSTEDKSCRDYYIEYNFKYLQCPLEFTKKTPTQDVIYE
PLTALNAMVQNGHLEVVKLLLEAGADVAAKDKFGKTAFDISIDNGNEDLAEILQ

Recombinant Protein Production:

Recombinant BK3 protein was produced using transformed BL21-DE3 *E. coli* at 18 °C for 16 h in 2-YT media supplemented with 500 µM IPTG and 30 µg/mL kanamycin. Protein purification by nickel affinity chromatography was followed by size exclusion chromatography which was carried out using a Superdex S75-300 mL column (GE Healthcare) eluting with 50 mM Tris pH 7.5.

3.6 FIH:

A plasmid encoding FIH was a kind gift from Dr. R. J. Hopkinson. Recombinant protein production was carried out following a literature protocol.[21] In brief, FIH was produced utilising transformed BL21-DE3 *E. coli* at 37 °C for 4 h in 2-YT media supplemented with 500 µM IPTG and 30 µg/mL kanamycin. Protein purification by nickel affinity chromatography was followed by size exclusion chromatography which was carried out using a Superdex S75-300 mL column (GE Healthcare) eluting with 50 mM Tris pH 7.5.

4. Hydroxylation Assays Using FIH

Ankyrin loop peptides were dissolved in MQ H_2O to give a 20 mM stock solution. A portion of this solution was diluted to give a 100 µM working solution. 50 mM Tris pH 7.5 was freshly prepared by diluting 500 mM Tris pH 7.5 with MQ H_2O. Solutions of 2-oxoglutarate and ascorbate were freshly prepared by dissolving 2-oxoglutarate disodium salt and sodium ascorbate in 50 mM Tris pH 7.5 to give a 500 mM stock solution. Each of these stocks were diluted with 50 mM Tris pH 7.5 to give a 1 mM working solution. $NH_4[Fe(H_2O)_6](SO_4)_2 \cdot 6\ H_2O$ was dissolved in 20 mM HCl to give a stock solution of 100 mM which was diluted with MQ H_2O to give a 200 µM working solution. FIH stocks were retrieved from -80 °C storage and thawed using hand warmth before being placed on ice. A 100

µM working solution of FIH was prepared by diluting the thawed FIH stock with cold 50 mM Tris pH 7.5.

The 2 µL each of the working solutions (with the exception of FIH) were pipetted into wells in a transparent V-bottom 96-well plate (Greiner) and 50 mM Tris pH 7.5 was added to a final volume of 18 µL. The assay was initiated with the addition of 2 µL 100 µM FIH or 2 µL 50 mM Tris pH 7.5 in the case of the no enzyme control wells. The 96-well plate was then transferred to a preheated thermomixer and the plate was incubated at 37 °C for 2 h. The reaction was quenched by co-crystalisation of the sample on a MALDI-MS target with matrix buffer (see §7).

5. Nuclear Magnetic Resonance (NMR)

5.1 General Procedures

Spectra were recorded using a Bruker Avance III HD 600 MHz spectrometer fitted with a 5mm BB-^{19}F/^{1}H Prodigy N_2 cryoprobe or a Bruker Avance III 700 MHz spectrometer fitted with a 5mm ^{1}H(^{13}C/^{15}N) inverse cryoprobe. Concentrations were referenced to 1 mg/mL sodium 3-trimethylsilyl-2,2,3,3-tetradeuteropropionoate (TSP). Spectra were recorded at 298 K unless stated otherwise.

5.2 Reactions of Peptides with Formaldehyde

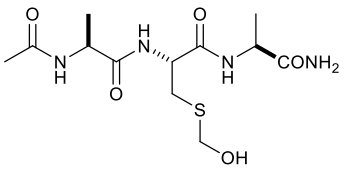
5.2.1 Procedures

Fresh solutions of formaldehyde in D_2O were prepared by heating paraformaldehyde (Sigma Aldrich) in D_2O in a sealed vial to give a solution of known concentration. The solution was then diluted further with D_2O to give a working solution of 80 mM HCHO. Experiments were typically performed using 3 mm tubes with a final volume of 160 µL. To determine the reaction of peptides with formaldehyde, formaldehyde in D_2O was added to a solution of the relevant peptide in 50 mM Tris-d_{11} in D_2O, the mixture briefly vortexed, pipetted into a 3 mm NMR tube and centrifuged using a

hand centrifuge to ensure sample homogeneity. For competition experiments between formaldehyde-peptide adducts and GSH, the peptide was pre-incubated with formaldehyde for 5 min at rt before 20 mM GSH in D_2O was added. Proton spectra were recorded using the excitation sculpting with perfect echo (ESPE) pulse sequence with 16 scans. Time courses were acquired using the `multi_zg` command.

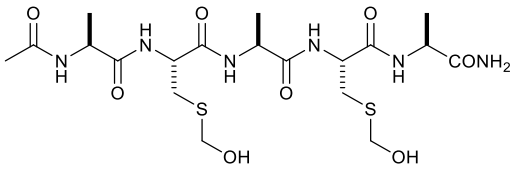
5.2.2 Characterisation Data

Ac-A(*S*-hydroxymethyl-C)A-NH$_2$



δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.32 (3H, d, *J* 7.2, Ala-*Me*), 1.35 (3H, d, *J* 7.3, Ala-*Me*), 1.97 (3H, s, NCO*Me*), 3.01 (1H, dd, *J* 7.4, 14.1, Cys-C(β)H_AH_B), 3.08 (1H, dd, *J* 5.5, 14.1, Cys-C(β)H_AH_B), 4.18 – 4.27 (2H, m, Ala-C(α)*H*, Ala-C(α)*H*), 4.49 – 4.60 (1H, m, Cys-C(α)*H*), 4.66 – 4.75 (2H, obscured, SCH_2OH); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.48 (Ala-*Me*), 16.72 (Ala-*Me*), 21.64 (NCO*Me*), 32.05 (Cys-*C*(β)), 49.73 (Ala-*C*(α)), 49.94 (Ala-*C*(α)), 53.22 (Cys-*C*(α)), 65.38 (S*C*H$_2$OH), 172.03 (Cys-*C*O), 174.33 (N*C*OMe), 175.58 (Ala-*C*O), 177.48 (Ala-*C*ONH$_2$).

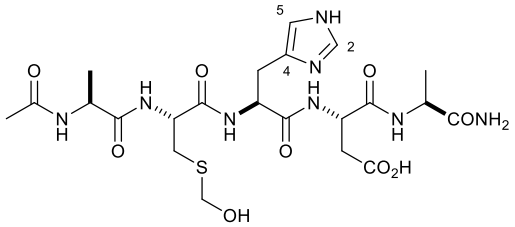
Ac-A(*S*-hydroxymethyl-C)A(*S*-hydroxymethyl-C)A-NH$_2$



δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.32 (3H, d, *J* 7.2, Ala-*Me*), 1.35 (6H, m, Ala-*Me*, Ala-*Me*), 1.97 (3H, s, NCO*Me*), 2.97 – 3.11 (4H, m, Cys-C(β)H_2, Cys-C(β)H_2), 4.19 – 4.30 (3H, m, Ala-C(α)*H*, Ala-C(α)*H*, Ala-C(α)*H*), 4.46 – 4.57 (2H, m, Cys-C(α)*H*, Cys-C(α)*H*), 4.68 – 4.78 (4H, obscured, SCH_2OH, SCH_2OH); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.22 (Ala-*Me*), 16.48

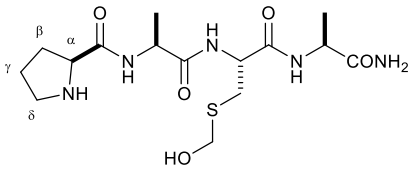
(Ala-*Me*), 16.84 (Ala-*Me*), 21.66 (NCO*Me*), 31.89 (Cys-*C*(β)), 31.95 (Cys-*C*(β)), 49.48 (Ala-*C*(α)), 49.75 (Ala-*C*(α)), 49.98 (Ala-*C*(α)), 53.14 (Cys-*C*(α)), 53.43 (Cys-*C*(α)), 65.36 (S*C*H$_2$OH), 65.41 (S*C*H$_2$OH), 171.84 (Cys-*C*O), 172.32 (Cys-*C*O), 174.33 (N*C*OMe), 175.58 (Ala-*C*O), 175.87 (Ala-*C*O), 177.48 (Ala-*C*ONH$_2$).

Ac-A(*S*-hydroxymethyl-C)HDA-NH$_2$



δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.29 (3H, d, *J* 7.3, Ala-*Me*), 1.35 (1H, d, *J* 7.3, Ala-*Me*), 1.96 (3H, s, NCO*Me*), 2.77 (1H, dd, *J* 7.4, 16.8, Asp-C(β)H_AH$_B$), 2.88 (1H, dd, *J* 5.5, 16.8, Asp-C(β)H$_A$$H_B$), 2.94 – 3.04 (2H, m, Cys-C(β)H$_2$), 3.13 (1H, dd, *J* 8.5, 15.3, His-C(β)H_AH$_B$), 3.26 (1H, dd, *J* 5.5, 15.3 , Asp-C(β)H$_A$$H_B$), 4.12 – 4.28 (2H, m, Ala-C(α)*H*, Ala-C(α)*H*), 4.40 – 4.52 (1H, m, Cys-C(α)*H*), 4.61 – 4.74 (4H, obscured, His-C(α)*H*, Asp-C(α)*H*, SCH_2OH), 7.25 (1H, s, His-C(5)*H*), 8.57 (1H, s, His-C(2)*H*); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.41 (Ala-*Me*), 16.64 (Ala-*Me*), 21.67 (NCO*Me*), 26.08 (His-*C*(β)), 32.03 (Cys-*C*(β)), 35.73 (Asp-*C*(β)), 49.70 (Ala-*C*(α)), 49.91 (Ala-*C*(α)), 50.32 (Asp-*C*(α)), 52.50 (Cys-*C*(α)), 65.38 (S*C*H$_2$OH), 117.31 (His-*C*(5)), 128.32 (His-*C*(4)), 133.38 (His-*C*(2)), 171.21 (His-*C*O), 171.86 (Asp-*C*O) 172.21 (Cys-*C*O), 174.22 (Asp-*C*O$_2$H), 174.36 (N*C*OMe), 175.43 (Ala-*C*O), 177.61 (Ala-*C*ONH$_2$).

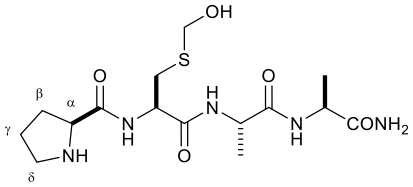
PA(*S*-hydroxymethyl-C)A-NH$_2$



δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.35 (6H, dd, *J* 4.5, 7.3, Ala-*Me*, Ala-*Me*), 1.89 – 2.20 (3H, m, Pro-C(γ)H$_2$, Pro-C(β)H_AH$_B$), 2.35 – 2.46 (1H, m, Pro-C(β)H$_A$$H_B$), 3.00 (1H, dd, *J* 7.5, 14.0,

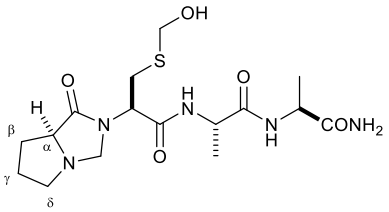
Cys-C(β)H_AH$_B$), 3.08 (1H, dd, *J* 5.9, 14.0, Cys-C(β)H$_A$$H_B$), 3.30 – 3.41 (2H, m, Pro-C(δ)H_2), 4.24 (1H, qd, *J* 5.5, 7.4, Ala-C(α)*H*), 4.28 – 4.38 (2H, m, Ala-C(α)*H,* Pro-C(α)*H*), 4.45 – 4.53 (1H, m, Cys-C(α)*H*), 4.65 – 4.74 (2H, obscured, SCH_2OH); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.41 (Ala-*Me*), 16.78 (Ala-*Me*), 23.72 (Pro-*C*(γ)), 29.78 (Pro-*C*(β)), 31.98 (Cys-*C*(β)), 46.62 (Pro-*C*(δ)), 49.64 (Ala-*C*(α)), 49.97 (Ala-*C*(α)), 53.32 (Cys-*C*(α)), 59.62 (Pro-*C*(α)), 65.42 (S*C*H$_2$OH), 169.49 (Pro-*C*O), 171.88 (Cys-*C*O), 174.68 (Ala-*C*O), 177.49 (Ala-*C*ONH$_2$).

P(*S*-hydroxymethyl-C)AA-NH$_2$



δ_H (600 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 1.30 – 1.38 (6H, m, Ala-*Me*, Ala-*Me*), 1.93 – 2.07 (3H, m, Pro-C(γ)H$_2$, Pro-C(β)H_AH$_B$), 2.34 – 2.49 (1H, m, Pro-C(β)H$_A$$H_B$), 2.98 (1H, dd, *J* 8.8, 13.9, Cys-C(β)H_AH$_B$), 3.09 (1H, dd, *J* 5.9, 13.9, Cys-C(β)H$_A$$H_B$), 3.29 – 3.47 (2H, m, Pro-C(δ)H_2), 4.17 – 4.30 (2H, m, Ala-C(α)*H*, Ala-C(α)*H*), 4.37 (1H, t, *J* 7.4, Pro-C(α)*H*), 4.50 – 4.61 (1H, m, Cys-C(α)*H*), 4.67 – 4.77 (2H, obscured, SCH_2OH); δ_C (151 MHz, 50 mM Tris-d_{11} in D_2O, pH 7.5): 16.39 (Ala-*Me*), 16.81 (Ala-*Me*), 23.84 (Pro-*C*(γ)), 29.68 (Pro-*C*(β)), 31.97 (Cys-*C*(β)), 46.65 (Pro-*C*(δ)), 49.49 (Ala-*C*(α)), 49.81 (Ala-*C*(α)), 53.68 (Cys-*C*(α)), 59.73 (Pro-*C*(α)), 65.43 (S*C*H$_2$OH), 169.97 (Pro-*C*O), 171.58 (Cys-*C*O), 174.41 (Ala-*C*O), 177.78 (Ala-*C*ONH$_2$).

(R)-*N*-((S)-1-(((S)-1-amino-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-((hydroxymethyl)thio)-2-((S)-1-oxotetrahydro-1H-pyrrolo[1,2-c]imidazol-2(3H)-yl)propanamide



δ_H (600 MHz, 50 mM Tris-*d*$_{11}$ in D_2O, pH 7.5): 1.34 (6H, d, *J* 7.3, Ala-*Me*, Ala-*Me*), 1.82 – 2.05 (2H, m, Pro-C(γ)*H*$_2$), 2.09 – 2.18 (1H, m, Pro-C(β)*H*$_A$H$_B$), 2.20 – 2.31 (1H, m, Pro-C(β)H$_A$*H*$_B$), 2.97 – 3.03 (1H, m, Cys-C(β)*H*$_A$H$_B$), 3.16 – 3.27 (2H, m, Cys-C(β)H$_A$*H*$_B$, Pro-C(δ)*H*$_A$H$_B$), 3.58 (1H, dt, *J* 6.6, 11.1, Pro-C(δ)*H*$_A$H$_B$), 4.23 (2H, dd, *J* 6.5, 13.1, Ala-C(α)*H*, Ala-C(α)*H*), 4.35 – 4.49 (1H, Pro-C(α)*H*), 4.62 – 4.99 (5H, obscured, Cys-C(α)*H*, NC*H*$_2$N, SC*H*$_2$OH); δ_C (151 MHz, 50 mM Tris-*d*$_{11}$ in D_2O, pH 7.5): 16.45 (Ala-*Me*), 16.81 (Ala-*Me*), 24.22 (Pro-*C*(γ)), 27.27 (Pro-*C*(β)), 29.03 (Cys-*C*(β)), 49.67 (Ala-*C*(α)), 49.89 (Ala-*C*(α)), 54.02 (Cys-*C*(α)), 57.08 (Pro-*C*(δ)), 64.57 (S*C*H$_2$OH), 65.35 (Pro-*C*(α)), 65.84 (N*C*H$_2$N), 169.20 (Cys-*C*O), 172.66 (Pro-*C*O), 174.45 (Ala-*C*O), 177.71 (Ala-*C*ONH$_2$).

5.3 Ligand Binding Experiments

CR was a kind from Dr. C. Lindsay, Defence Science & Technology Laboratory. A 0.5 mM solution of **CR** was prepared in 50 mM phosphate buffer pH 7.5 in a 90:10 (v/v) mixture of H_2O and D_2O using a 10 mM or 50 mM stock of **CR** in DMSO-d_6 to give a final concentration of 5% and 1% DMSO-d_6 respectively. For the heat denatured controls, 20 µL aliquots of BK3 were heated to 95 °C for 10 minutes before being cooled to rt. BK3 in 50 mM Tris pH 7.5 was added to the **CR** solution at rt to give a final volume of 160 µL, the mixture gently mixed by pipetting before being transferred to a 3 mm NMR tube and centrifuged briefly using a hand centrifuge to ensure sample homogeneity. Data were acquired using the CPMG-PROJECT sequence utilising 64 scans with water suppression achieved by pre-saturation.[22,23]

6. Digestion of TRPA1

Recombinant hTRPA1 produced from Sf9 cells was a kind gift from Dr. M. Grieben (University of Oxford). A 20 µl aliquot of TRPA1 was thawed using hand warmth and diluted with MQ H_2O (180 µL). MeOH (600 µL) and $CHCl_3$ (150 µL) were added and the solution was briefly vortexed. MQ H_2O (450 µL) was added and the mixture centrifuged (Sigma 1-4 microfuge, 14000 rpm) for 1 min

at rt. The upper layer was aspirated and MeOH (450 µL) added to the lower layer. The suspension was briefly vortexed before being centrifuged (Sigma 1-4 microfuge, 14000 rpm) for 2 min at rt. The supernatant was then discarded and the mixture resuspended in 150 µL 8M urea supplemented with 15 µL 500 mM Tris pH 7.5 and the resulting suspension was incubated at 60 °C for 10 minutes. The sample was then allowed to cool to rt before 9 µL of 1 M dithioreitol (5 mM final concentration) was added and the sample incubated at 37 °C for 30 min. 25 µL of 1 M iodoacetamide was added (14 mM final concentration) and the sample was incubated at rt in the dark for 30 min. 20 µg of Trypsin/Lys-C mix (Promega) was resuspended in 100 µL of 50 mM acetic acid. The mixture was then subjected to digestion by Trypsin/Lys-C under denaturing conditions or in 80% aq. MeCN.

6.1 Denaturing Conditions[24]

5 µL Trypsin/Lys-C mixture (2 µg mL^{-1}, Promega) was added to the protein mixture and the mixture incubated at 37 °C for 3 h. The mixture was then diluted with 1 mL of 50 mM Tris pH 7.5 and incubated for a further 16 h at 37 °C. The digest was then acidified using 50 µL formic acid and concentrated *in vacuo*. The peptide pellet was resuspended in 200 µL of 2% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v). ZipTips with 0.6 µL C18 resin (Millipore) were washed with MeCN (2 x 10 µL) and equilibrated with 2% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v, 2 x 10 µL). 10 µL of the digest was aspirated into the ZipTip and then dispensed back into the tube containing the digest. This process was repeated 10 times before the ZipTip was washed with 2% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v, 2 x 10 µL). 10 µL of maldi matrix (see §7) was aspirated into the ZipTip and eluted from the ZipTip into a clean Eppendorf tube. The resulting solution was then pipetted directly onto a maldi target for analysis by MALDI-MS.

6.2 Digestion with 80% Aqueous MeCN[25]

The protein mixture was diluted with 1 mL of 80% aq. MeCN supplemented with 50 mM Tris pH 7.5. 5 µL Trypsin/Lys-C mixture (2 µg mL^{-1}, Promega) was added to the protein mixture and the

mixture incubated at 37 °C for 16 h. The digest was then acidified using 50 µL formic acid and concentrated *in vacuo*. The peptide pellet was resuspended in 200 µL of 2% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v). ZipTips with 0.6 µL C18 resin (Millipore) were washed with MeCN (2 x 10 µL) and equilibrated with 2% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v, 2 x 10 µL). 10 µL of the digest was aspirated into the ZipTip and then dispensed back into the tube containing the digest. This process was repeated 10 times before the ZipTip was washed with 2% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v, 2 x 10 µL). The ZipTip was eluted sequentially with 10 µL of 10%, 20%, 30%, 40%, 50% and 70% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v) into separate clean Eppendorf tubes. The fractions were concentrated in vacuo before being resuspended in 10 µL of MQ H_2O. Each fraction was co-crystalised with maldi matrix on a maldi target and analysed by MALDI-MS (see §7) .

7. Mass Spectrometry

Low resolution, denaturing mass spectra of proteins were recorded on a Waters Micromass LCT Premier XE spectrometer utilising a ThermoFischer ProSwift RP-4E (1 x 50 mm) column. MALDI-MS spectra were recorded using a Bruker Autoflex Speed Maldi-TOF/TOF using a ground steel sample carrier. Samples were co-crystalised with maldi matrix containing 10 mg/mL α-cyano-4-hydroxycinnamic acid in 50% aq. MeCN supplemented with 0.1% CF_3CO_2H (v/v). Typically, 0.8 µL of the sample was added to the sample carrier followed by 0.8 µL of the matrix with mixing carried out by pipetting. The mixture was then allowed to dry at rt for 1 h before the sample carrier was loaded into the spectrometer. MALDI-MS were acquired in positive reflection mode.

8. Mammalian Cell Culture

8.1 General Procedures

FIH knock and wild type (WT) HEK293T cells were a kind gift from Dr. Y. Tien (Oxford). Cells were cultured in complete medium comprised of Dulbecco′s modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 2 mM L-glutamine. Cells were grown in humidified 5% CO_2 at 37 °C to approximately 90% confluence in T75 flasks before being passaged. Passaging was carried out by washing with phosphate buffered saline (PBS, 2 x 12 mL) followed by detachment with 0.05% Trypsin - 0.02% EDTA solution (4 mL, Sigma) at 37 °C for 5 minutes. Resuspended cells were added to 23 mL of complete medium, the flask washed with 8 mL of complete medium and centrifuged at 200 rcf for 5 minutes at rt. The resulting cell pellet was resuspended in 12 mL complete medium. 0.5 – 1.5 mL of this suspension was then used to seed a fresh T75 flask containing a total of 18 mL of complete medium. 20 µL of cell suspension was incubated with 20 µL 0.4% trypan blue for 2 min at rt before determining cell count and viability using a Countess II automated cell counter (Invitrogen).

8.2 TRPA1 Activity via Ca^{2+} Influx Assay[26]

The HEK293 hTRPA1 photoscreen cell line (Perkin Elmer) was cultured in accordance with the literature.[26] In brief, the cells were cultured with Eagle's minimum essential media with Earle's salts (EMEM) supplemented with fetal bovine serum (10% v/v), penicillin and streptomycin (100 units mL^{-1} and 100 µg mL^{-1}, respectively), L-glutamine (2 mM) and G418 (0.4 mg mL^{-1}). Cells were washed twice with Dulbecco′s phosphate buffered saline, detached using trypsin and harvested by centrifugation. Cells were plated at 1.25 x 10^4 viable cells per well in black poly-D-lysine coated 96-well cell culture plates (BD Biosciences) and were grown for 48 h prior to use. Cells at passage number 10-23 were used.

hTRPA1 cells were loaded with a Ca^{2+} fluorophore using a Calcium 5 assay kit (Molecular Devices Ltd.). The cell culture medium was aspirated and 60 µL of the proprietary loading buffer was added to each well. The cells were then incubated at 37 °C for 30 min in a humidified atmosphere of 5% CO_2 in air. Compound loading buffer (CLB) was prepared comprising a 1:1 (v/v) mixture of EMEM and Hank's balanced salt solution (Gibco) supplemented with 20 mmol HEPES pH 7.4.

For the antagonist dose-response curves, a 40 mM suspension of cinnamaldehyde (CNA) in CLB was prepared and further diluted to a 1 mM working stock solution in CLB. 200 µL of 1 mM CNA in CLB was added to each well of a transparent U-bottom 96-well plate (Corning). For the agonist dose-response curves, serial dilution in a transparent U-bottom 96-well plate (Corning) was carried out to give a fifth-log series with a maximum concentration of 625 µM. The plates containing agonists were loaded into a FlexStation 3 (Molecular Devices Ltd.).

Antagonists were dissolved in DMSO to give a 5 mM stock solution which was diluted with 1 mM NaOH to give a 0.5 mM working solution of the compounds in 1% (v/v) DMSO in 1 mM NaOH. A serial dilution series was prepared using a transparent U-bottom 96-well plate (Corning), diluting the working solution with CLB to give a quarter-log series from 250 µM to 445 nM. 30 µL of the resulting solution was added to the HEK hTRPA1 cells, which were then incubated at 37 °C for 10 min in a humidified atmosphere with 5% CO_2. The black 96-well plate containing the cells was then loaded into the FlexStation 3, in addition to the 96-well plate containing solutions of CNA in CLB.

Cellular fluorescence was measured at 1 second intervals before and after addition of 30 µL of CNA working solution (giving a final CNA concentration of 250 µM). Fluorescence from formation of the Calcium-5−Ca^{2+} complex was measured using excitation and emission wavelengths of 485 nm and 525 nm respectively. Baseline fluorescence data were collected for 15 s prior to addition of the CNA

solution and the response followed for a further 45 s. Data were collected in triplicate from at least 4 separate repeats.

Fluorescence emission data (fluorescence intensity versus time) were captured from the FlexStation 3 using SOFTMAX PRO software v5.4.1 (Molecular Devices Inc.). The area under the time–response curve was calculated for the entire recording and imported into PRISM v. 4.03 (GraphPad Software Inc.) for nonlinear curve fitting. The maximum and minimum response were normalised to 100% and 0% respectively. Normalised cell response was plotted against the logarithm of antagonist concentration and the data were fit using the inbuilt "log(inhibitor) vs. normalised response – variable slope" model.

9. Chapter VI Bibliography

(1) Schiavi, B.; Ahond, A.; Poupat, C.; Potier, P. Preparation of N-Tert-Butoxycarbonylthiourea Opens the Way to Protected 2-Aminothiazoles. *Synthetic Communications* **2002**, *32* (11), 1671–1674. https://doi.org/10.1081/SCC-120004257.

(2) Takaya, J.; Mio, K.; Shiraishi, T.; Kurokawa, T.; Otsuka, S.; Mori, Y.; Uesugi, M. A Potent and Site-Selective Agonist of TRPA1. *Journal of the American Chemical Society* **2015**, *137* (50), 15859–15864. https://doi.org/10.1021/jacs.5b10162.

(3) Day, F. A.; Launay, D. F. M.; Charlton, M. H.; Moffat, D. F. C. Inhibitors of Hsp90. WO2010043865A1, 2010.

(4) Chenard, B. L.; Gallaschun, R. J.; Kimball, S. D. Substituted Xanthines and Methods of Use Thereof. US 2014/0275528 A1, 2013.

(5) Ho, S. Y.; Blanchard, S. E.; Duraiswamy, A. J.; Jenefer, A.; Adsool, V. A. Purine Diones as WNT Pathway Modulators. WO2014189466, 2014.

(6) Shibasaki, K.; Togo, H. 2-Amino-4-Arylthiazoles through One-Pot Transformation of Alkylarenes with NBS and Thioureas. *European Journal of Organic Chemistry* **2019**, *2019* (14), 2520–2527. https://doi.org/10.1002/ejoc.201900100.

(7) Mariani, A.; Bartoli, A.; Atwal, M.; Lee, K. C.; Austin, C. A.; Rodriguez, R. Differential Targeting of Human Topoisomerase II Isoforms with Small Molecules. *Journal of Medicinal Chemistry* **2015**, *58* (11), 4851–4856. https://doi.org/10.1021/acs.jmedchem.5b00473.

(8) Merijanian, A.; Sharma, G. M.; Moushati, J.; Gunderson, K. Steric Effects of Ortho Substituents on Acid-Catalyzed Cyclization of Thiocyanatoacetophenones. *Journal of Organic Chemistry* **1986**, *51* (4), 543–545. https://doi.org/10.1021/jo00354a029.

(9) Ghabbour, H. A.; Kadi, A. A.; Eltahir, K. E. H.; Angawi, R. F.; El-Subbagh, H. I. Synthesis, Biological Evaluation and Molecular Docking Studies of Thiazole-Based Pyrrolidinones and Isoindolinediones as Anticonvulsant Agents. *Medicinal Chemistry Research* **2015**, *24* (8), 3194–3211. https://doi.org/10.1007/s00044-015-1371-3.

(10) Ahlström, M. M.; Ridderström, M.; Zamora, I.; Luthman, K. CYP2C9 Structure-Metabolism Relationships: Optimizing the Metabolic Stability of COX-2 Inhibitors. *Journal of Medicinal Chemistry* **2007**, *50* (18), 4444–4452. https://doi.org/10.1021/jm0705096.

(11) Cavallaro, R. A.; Filocamo, L.; Galuppi, A.; Galione, A.; Brufani, M.; Genazzani, A. A. Potentiation of CADPR-Induced Ca2+-Release by Methylxanthine Analogues. *Journal of Medicinal Chemistry* **1999**, *42* (14), 2527–2534. https://doi.org/10.1021/jm980469t.

(12) Daly, J. W.; Padgett, W. L.; Shamim, M. T. Analogues of Caffeine and Theophylline: Effect of Structural Alterations on Affinity at Adenosine Receptors. *Journal of Medicinal Chemistry* **1986**, *29* (7), 1305–1308. https://doi.org/10.1021/jm00157a035.

(13) *Liberty Blue™ User Guide Reagents and Concentrations*; 2016.

(14) Froger, A.; Hall, J. E. Transformation of Plasmid DNA into E. Coli Using the Heat Shock Method. *Journal of Visualized Experiments* **2007**, No. 6. https://doi.org/10.3791/253.

(15) Studier, F. W.; Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *Journal of Molecular Biology* **1986**, *189* (1), 113–130. https://doi.org/10.1016/0022-2836(86)90385-2.

(16) Mierendorf, R. C.; Morris, B. B.; Hammer, B.; Novy, R. E. Expression and Purification of Recombinant Proteins Using the PET System. In *Molecular Diagnosis of Infectious Diseases*; Humana Press, 2003; Vol. 13, pp 257–292. https://doi.org/10.1385/0-89603-485-2:257.

(17) PCR Using Q5® High-Fidelity DNA Polymerase (M0491) | NEB https://international.neb.com/protocols/2013/12/13/pcr-using-q5-high-fidelity-dna-polymerase-m0491 (accessed Jun 16, 2021).

(18) User Guide: GeneJET Gel Extraction Kit https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0012661_GeneJET_Gel_Extraction_UG.pdf&title=VXNlciBHdWlkZTogR2VuZUpFVCBHZWwgRXh0cmFjdGlvbiBLaXQ= (accessed Jun 16, 2021).

(19) Aslanidis, C.; de Jong, P. J. Ligation-Independent Cloning of PCR Products (LIC-PCR). *Nucleic Acids Research* **1990**, *18* (20), 6069–6074. https://doi.org/10.1093/nar/18.20.6069.

(20) Savitsky, P.; Bray, J.; Cooper, C. D. O.; Marsden, B. D.; Mahajan, P.; Burgess-Brown, N. A.; Gileadi, O. High-Throughput Production of Human Proteins for Crystallization: The SGC Experience. *Journal of Structural Biology* **2010**, *172* (1), 3–13. https://doi.org/10.1016/j.jsb.2010.06.008.

(21) Hewitson, K. S.; McNeill, L. A.; Riordan, M. v.; Tian, Y. M.; Bullock, A. N.; Welford, R. W.; Elkins, J. M.; Oldham, N. J.; Bhattacharya, S.; Gleadle, J. M.; Ratcliffe, P. J.; Pugh, C. W.; Schofield, C. J. Hypoxia-Inducible Factor (HIF) Asparagine Hydroxylase Is Identical to Factor Inhibiting HIF (FIH) and Is Related to the Cupin Structural Family. *Journal of Biological Chemistry* **2002**, *277* (29), 26351–26355. https://doi.org/10.1074/jbc.C200273200.

(22) Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. Spin Echo NMR Spectra without J Modulation. *Chemical Communications* **2012**, *48* (6), 811–813. https://doi.org/10.1039/c1cc16699a.

(23) Quevedo, C. E.; Cruz-Migoni, A.; Bery, N.; Miller, A.; Tanaka, T.; Petch, D.; Bataille, C. J. R.; Lee, L. Y. W.; Fallon, P. S.; Tulmin, H.; Ehebauer, M. T.; Fernandez-Fuentes, N.; Russell, A. J.; Carr, S. B.; Phillips, S. E. V.; Rabbitts, T. H. Small Molecule Inhibitors of RAS-Effector Protein Interactions Derived Using an Intracellular Antibody Fragment. *Nature Communications* **2018**, *9* (1), 1–12. https://doi.org/10.1038/s41467-018-05707-2.

(24) Saveliev, S.; Bratz, M.; Zubarev, R.; Szapacs, M.; Budamgunta, H.; Urh, M. Trypsin/Lys-C Protease Mix for Enhanced Protein Mass Spectrometry Analysis. *Nature Methods* **2013**, *10* (11), i–ii. https://doi.org/10.1038/nmeth.f.371.

(25) Strader, M. B.; Tabb, D. L.; Hervey, W. J.; Pan, C.; Hurst, G. B. Efficient and Specific Trypsin Digestion of Microgram to Nanogram Quantities of Proteins in Organic-Aqueous Solvent Systems. *Analytical Chemistry* **2006**, *78* (1), 125–134. https://doi.org/10.1021/ac051348l.

(26) Lindsay, C. D.; Green, C.; Bird, M.; Jones, J. T. A.; Riches, J. R.; McKee, K. K.; Sandford, M. S.; Wakefield, D. A.; Timperley, C. M. Potency of Irritation by Benzylidenemalononitriles in Humans Correlates with TRPA1 Ion Channel Activation. *Royal Society Open Science* **2015**, *2* (1), 140160. https://doi.org/10.1098/rsos.140160.

Appendix I:
NMR Spectra of Selected Compounds

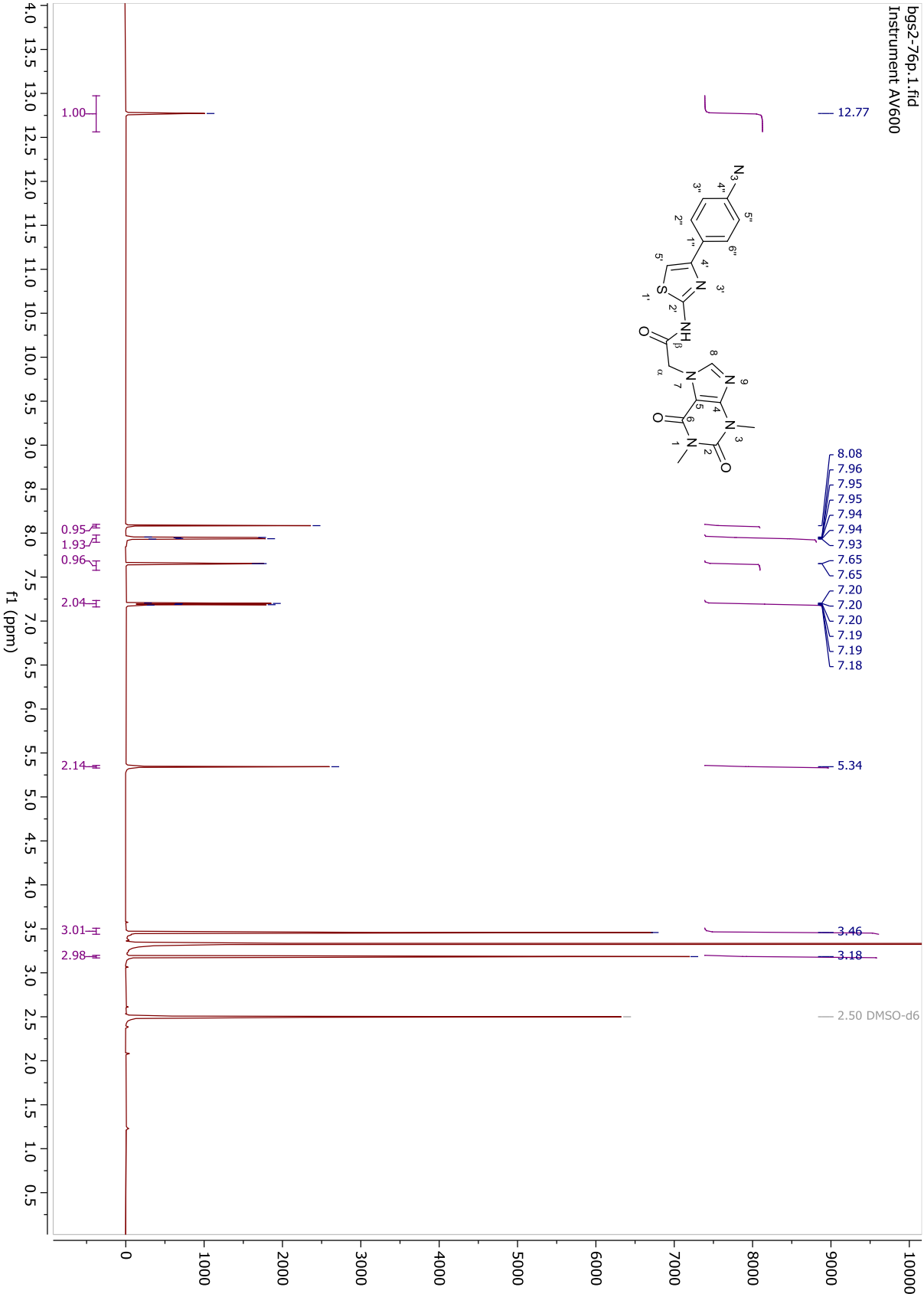


Figure S1: ^{1}H NMR spectrum of 3-60

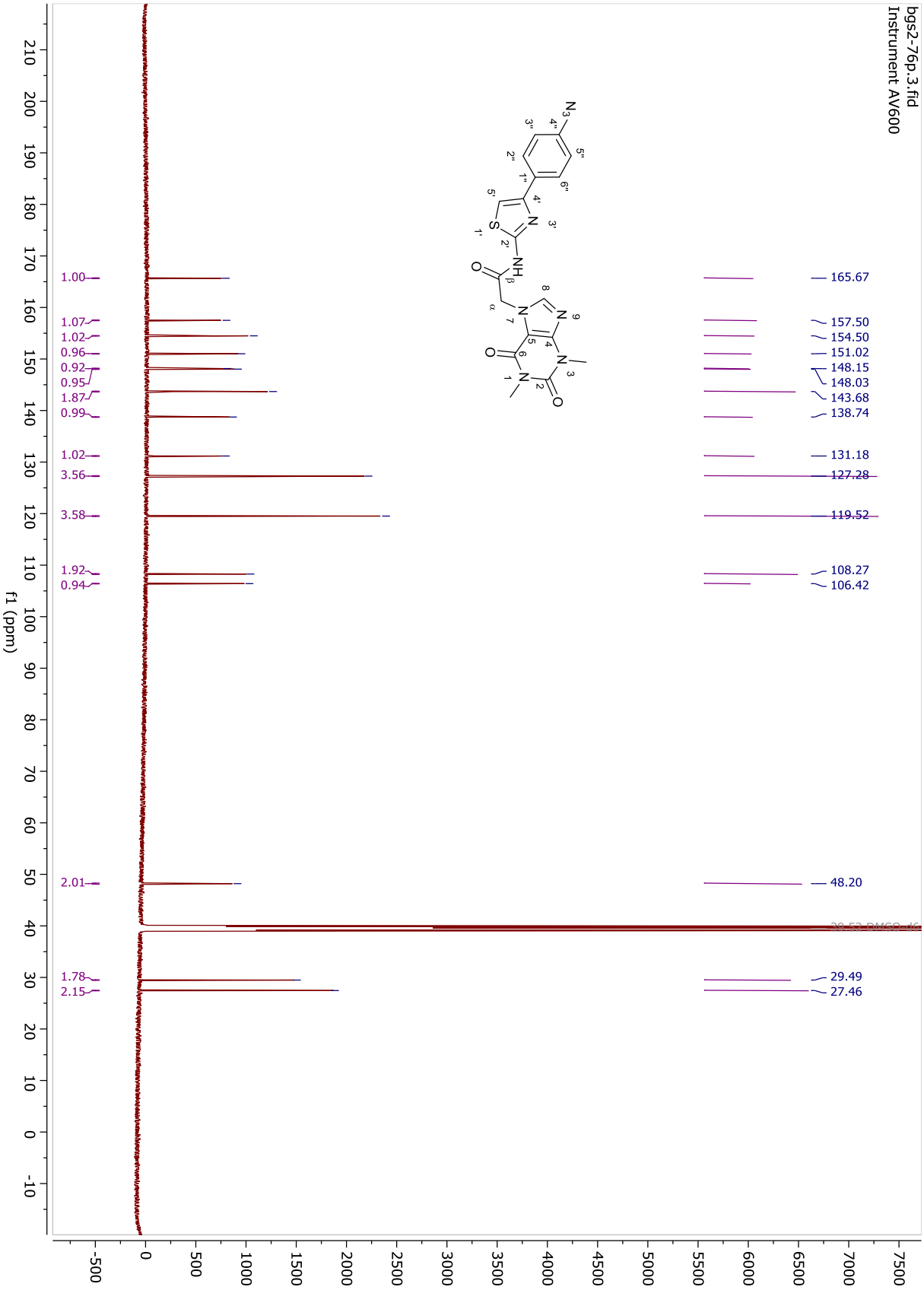


Figure S2: ^{13}C NMR spectrum of 3-60

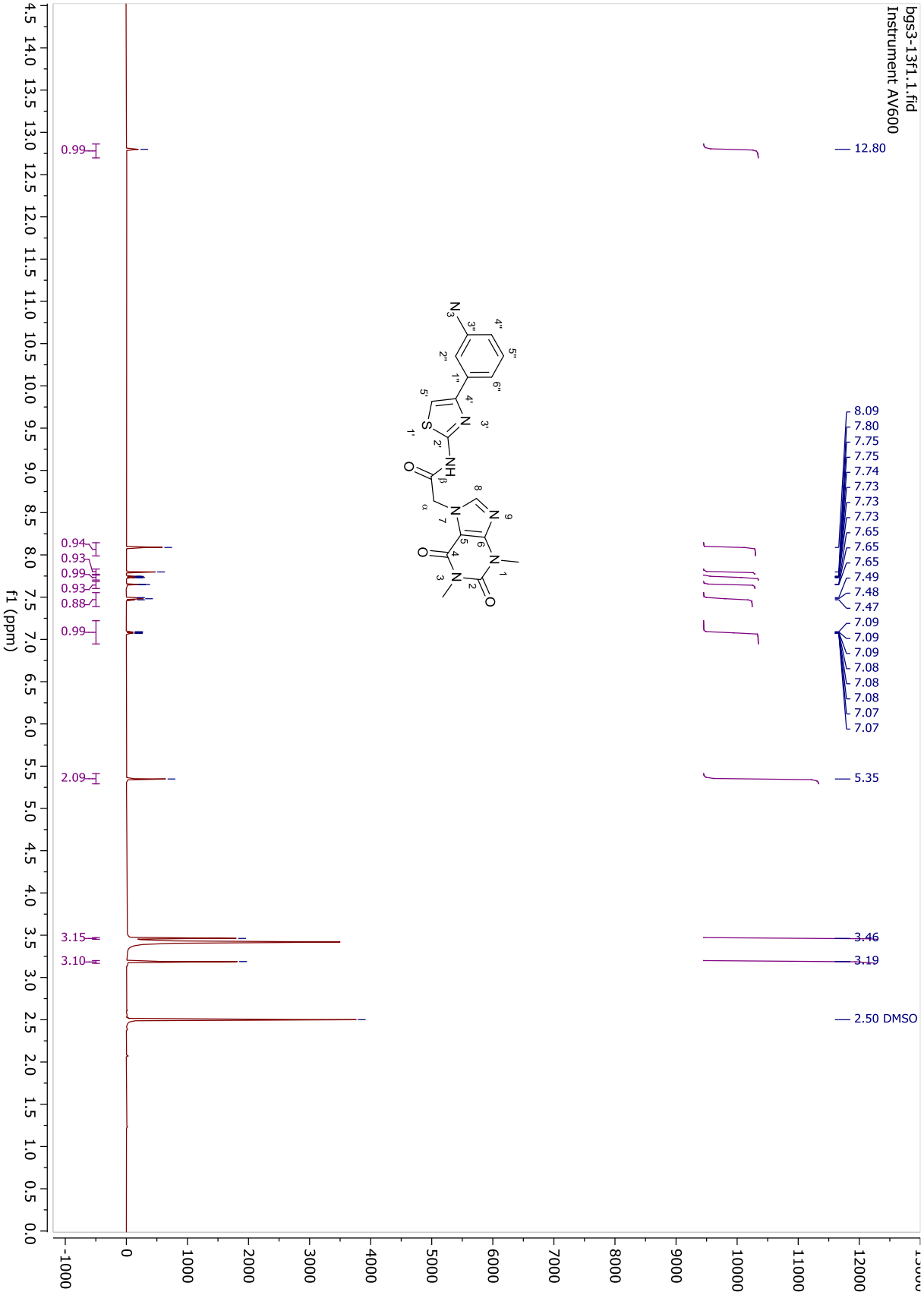


figure S3: ^{1}H NMR spectrum of 3-61

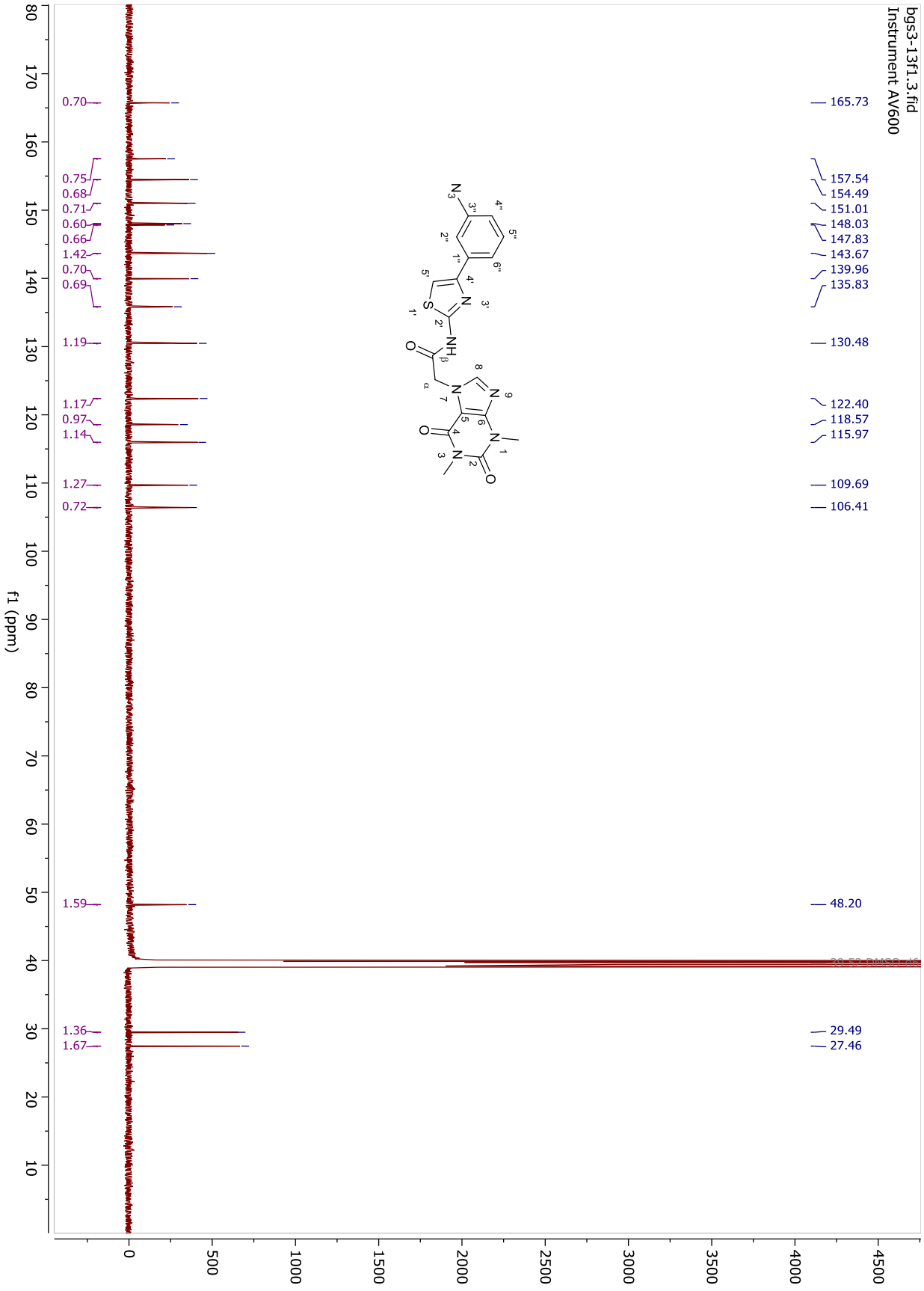


Figure S4: ^{13}C NMR spectrum of 3-61

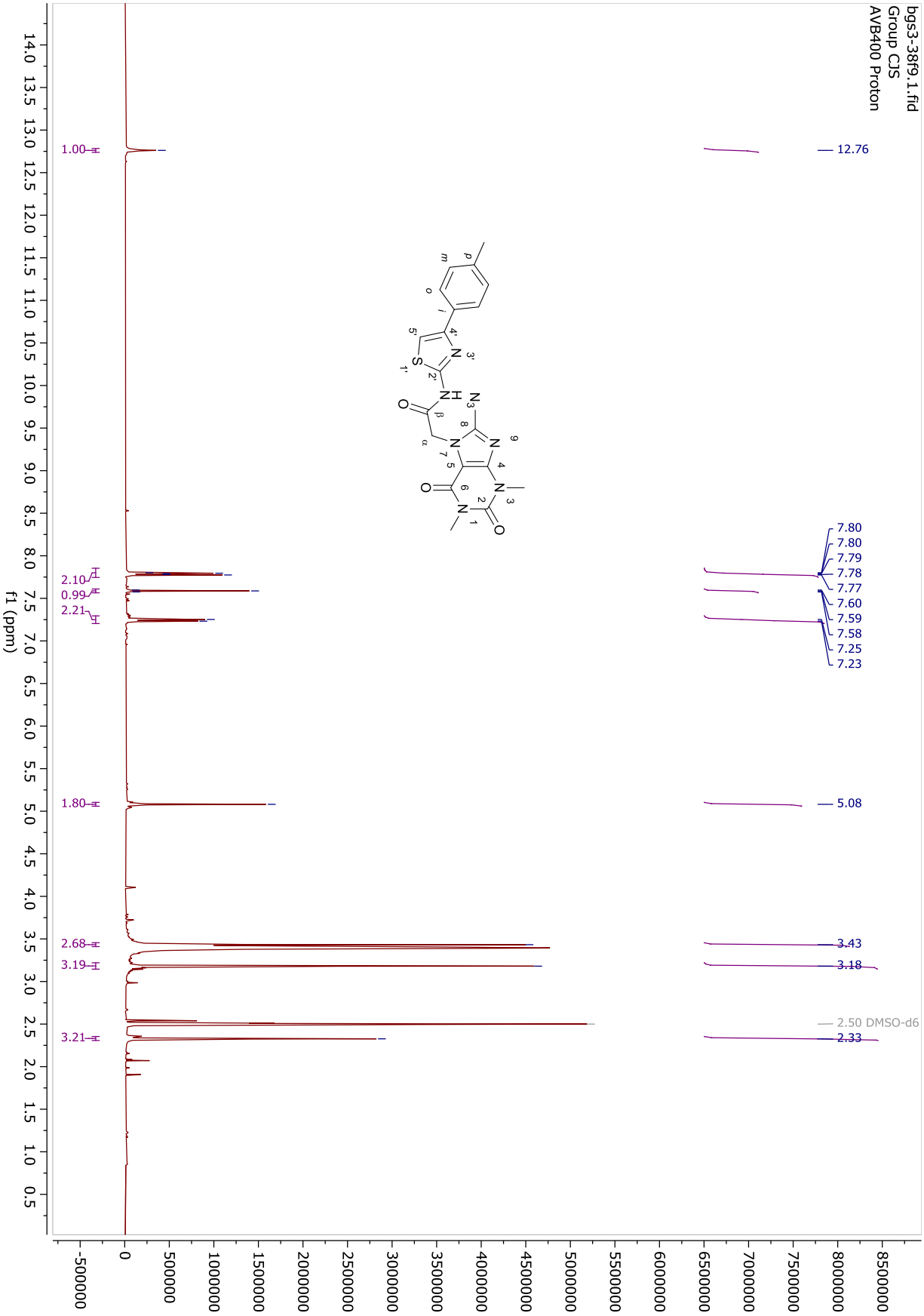


Figure S5: ^{1}H NMR spectrum of 3-64

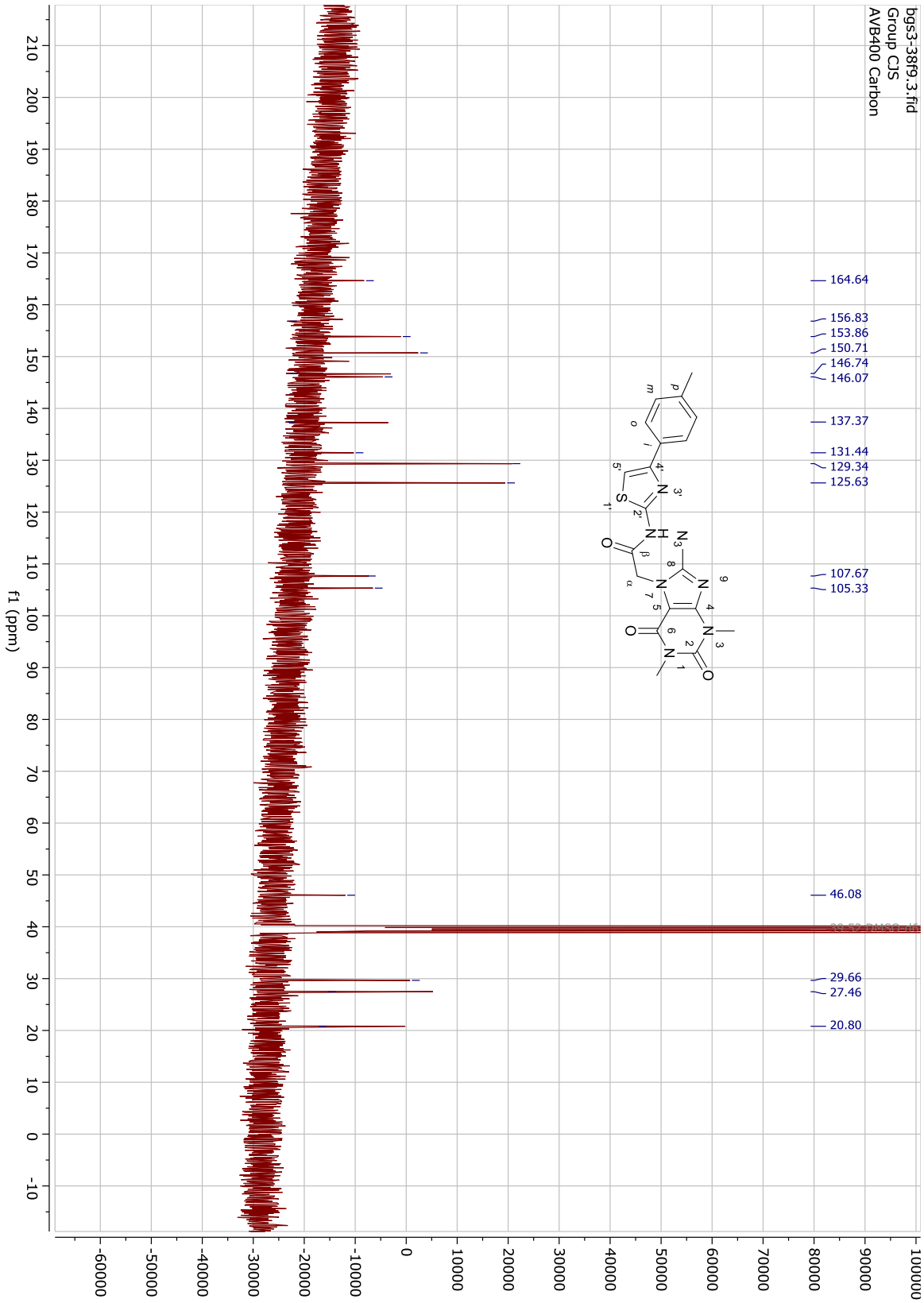


Figure S6: ^{13}C NMR of 3-64

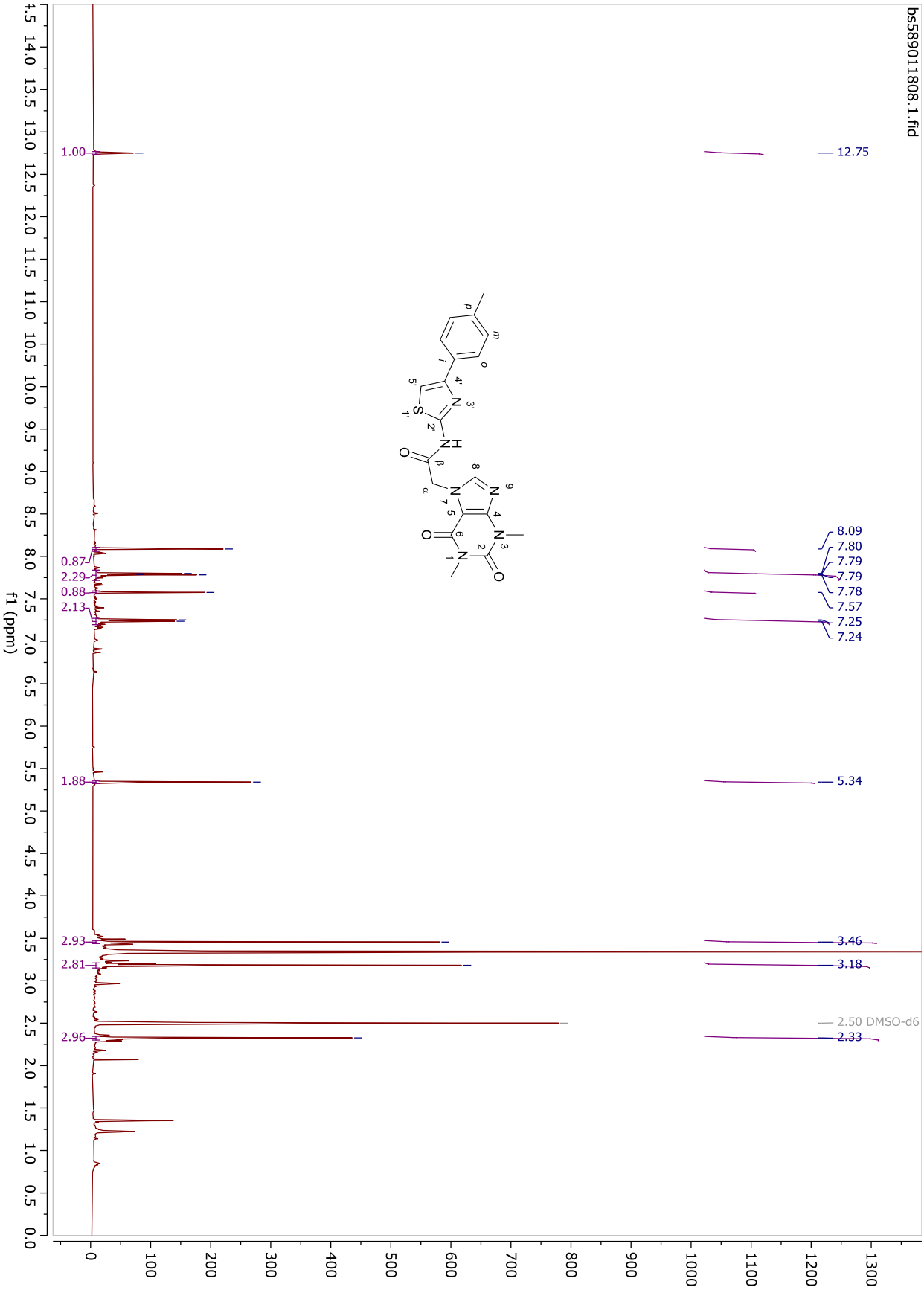


Figure S7: ^{1}H NMR spectrum of 3-55

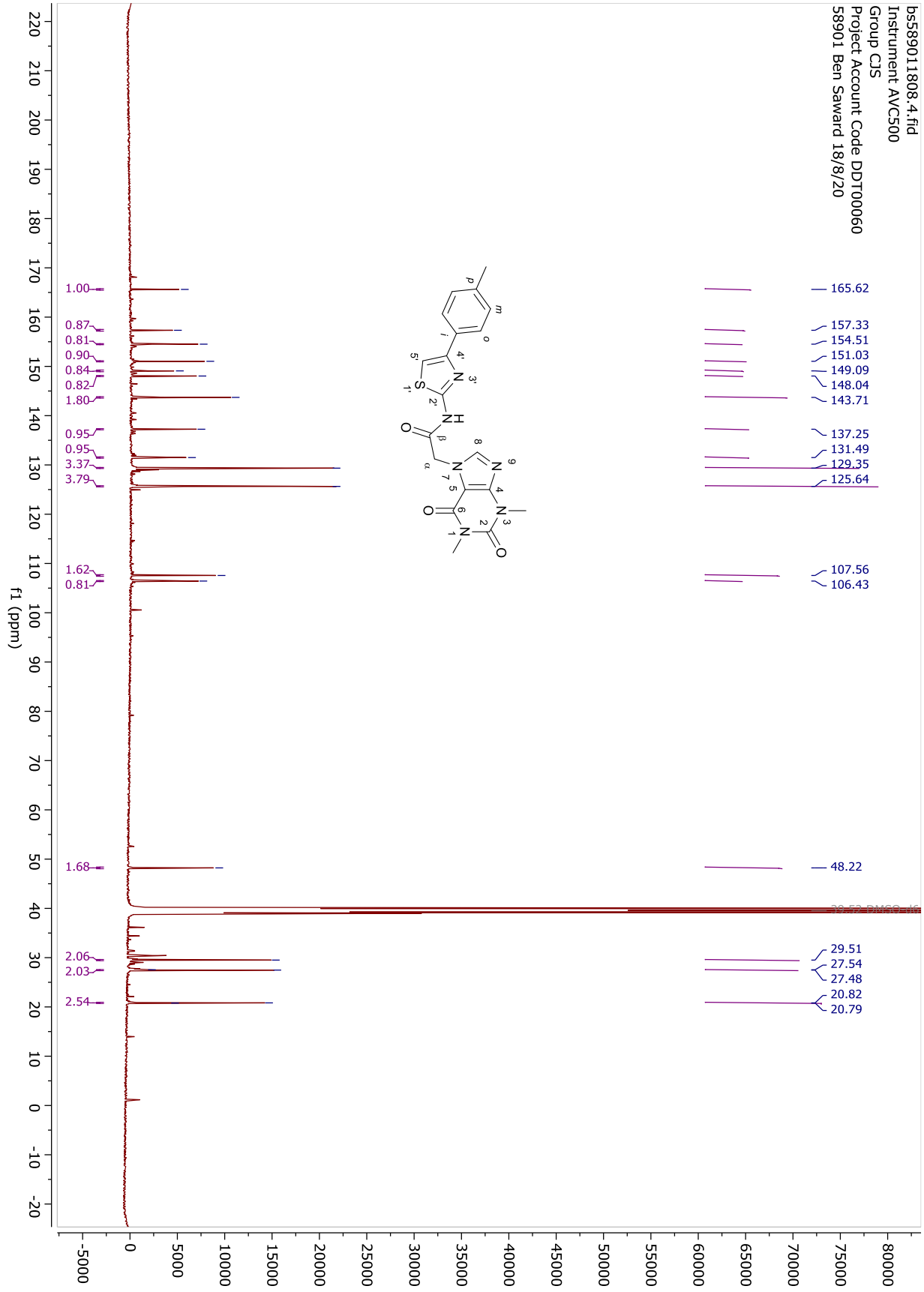


Figure S8: ^{13}C NMR spectrum of 3-55

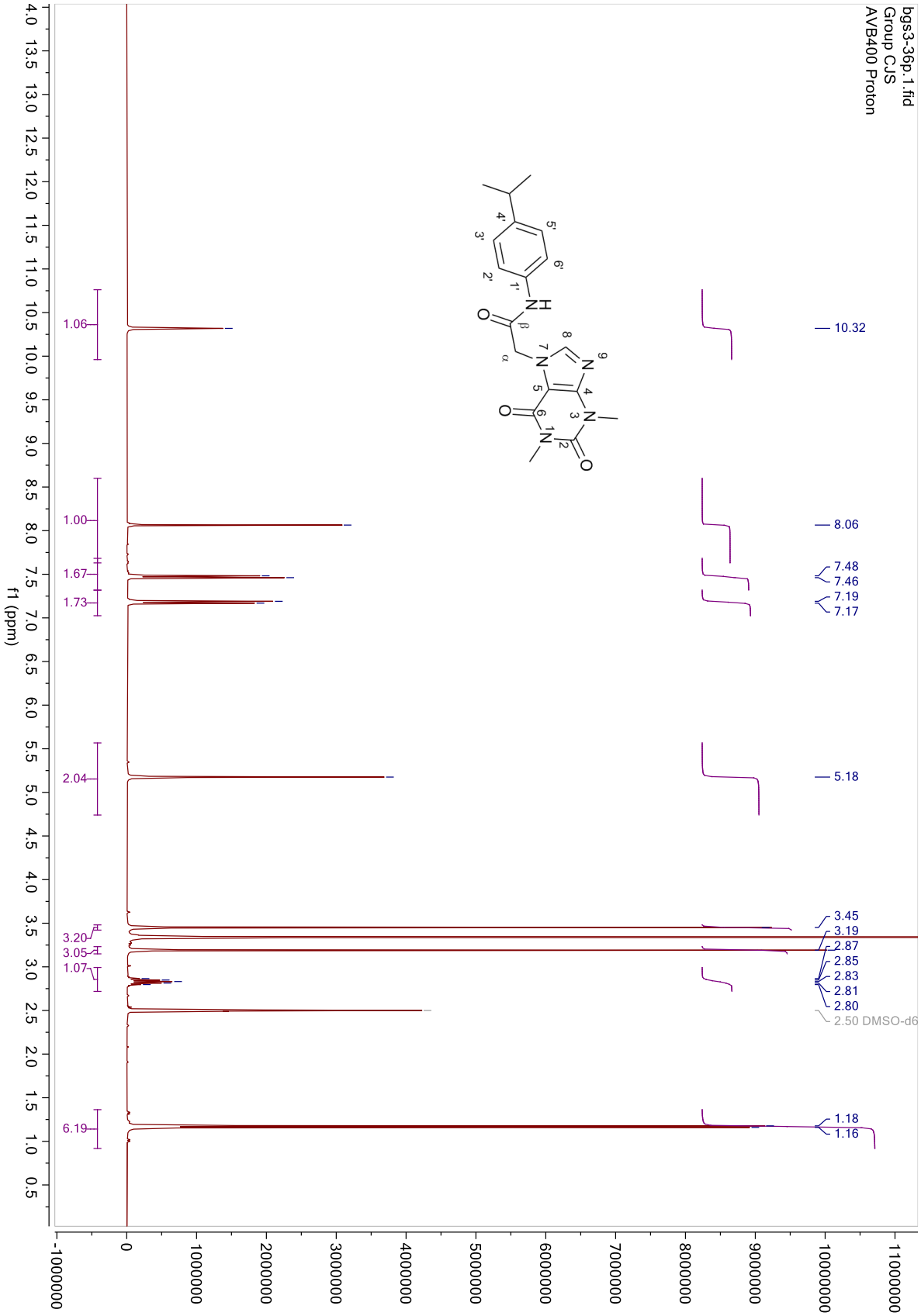


Figure S9: ^{1}H NMR spectrum of HC-030031

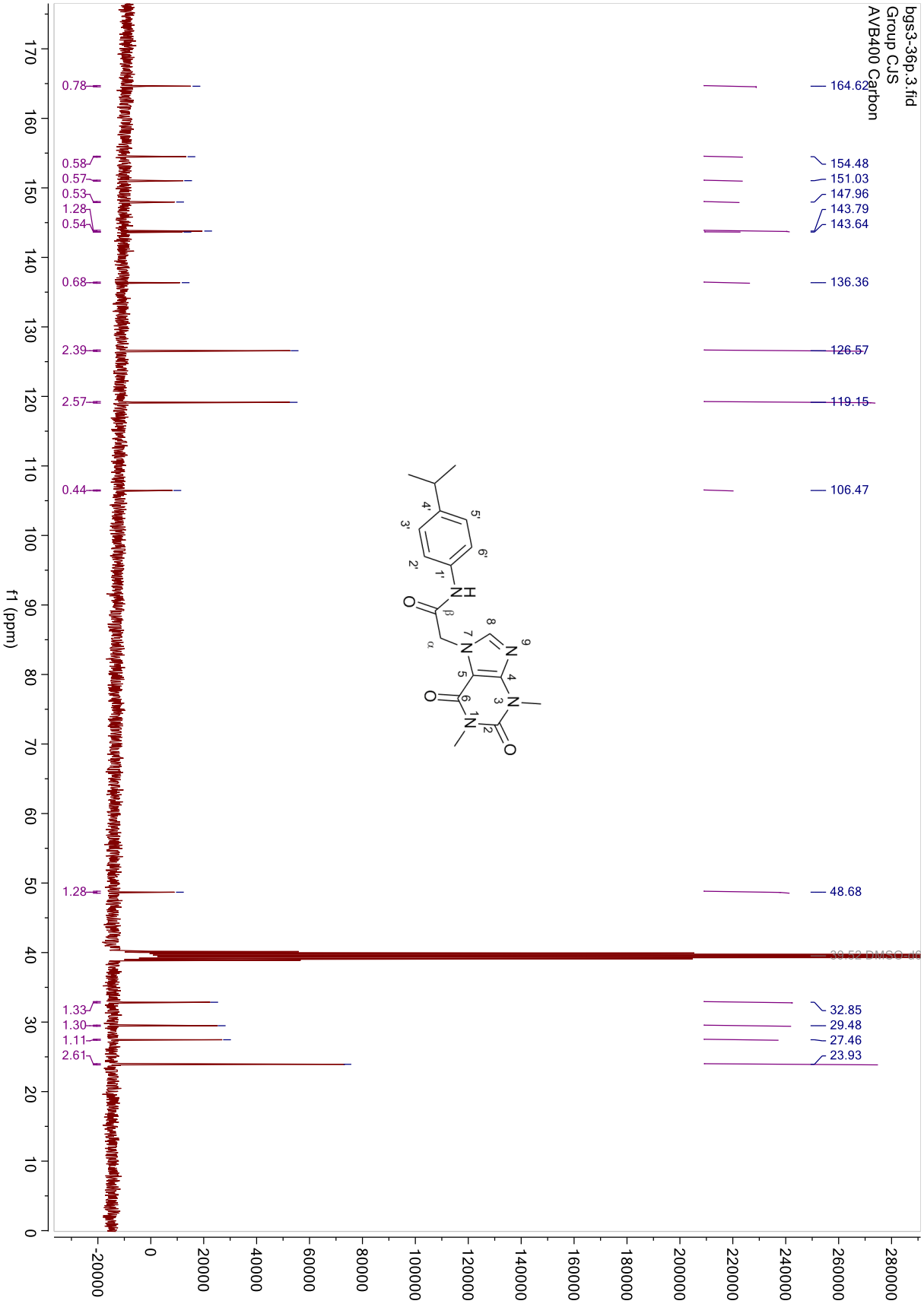


Figure S10: ^{13}C NMR spectrum of HC-030031

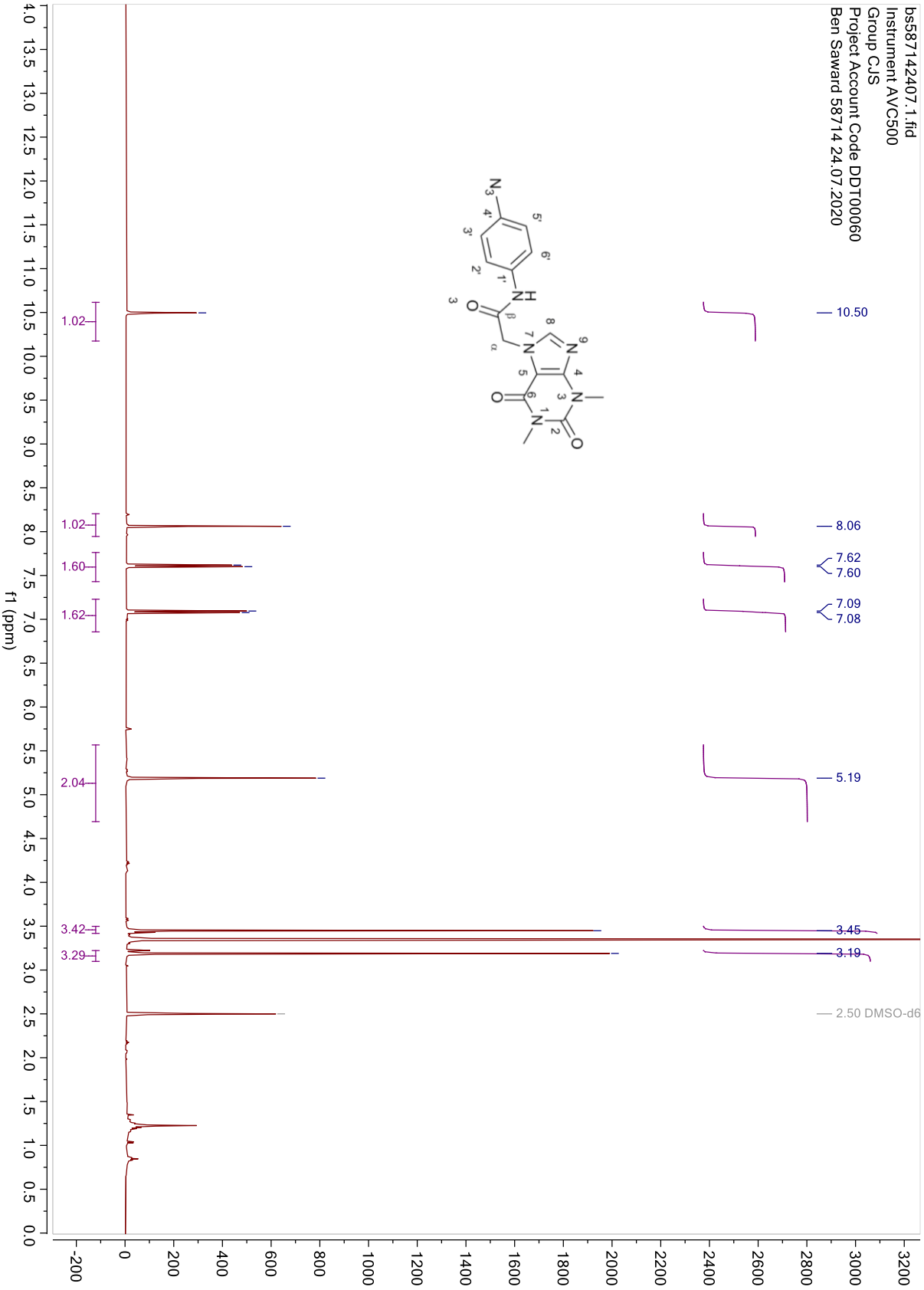


Figure S11: ^{1}H NMR spectrum of 4-33

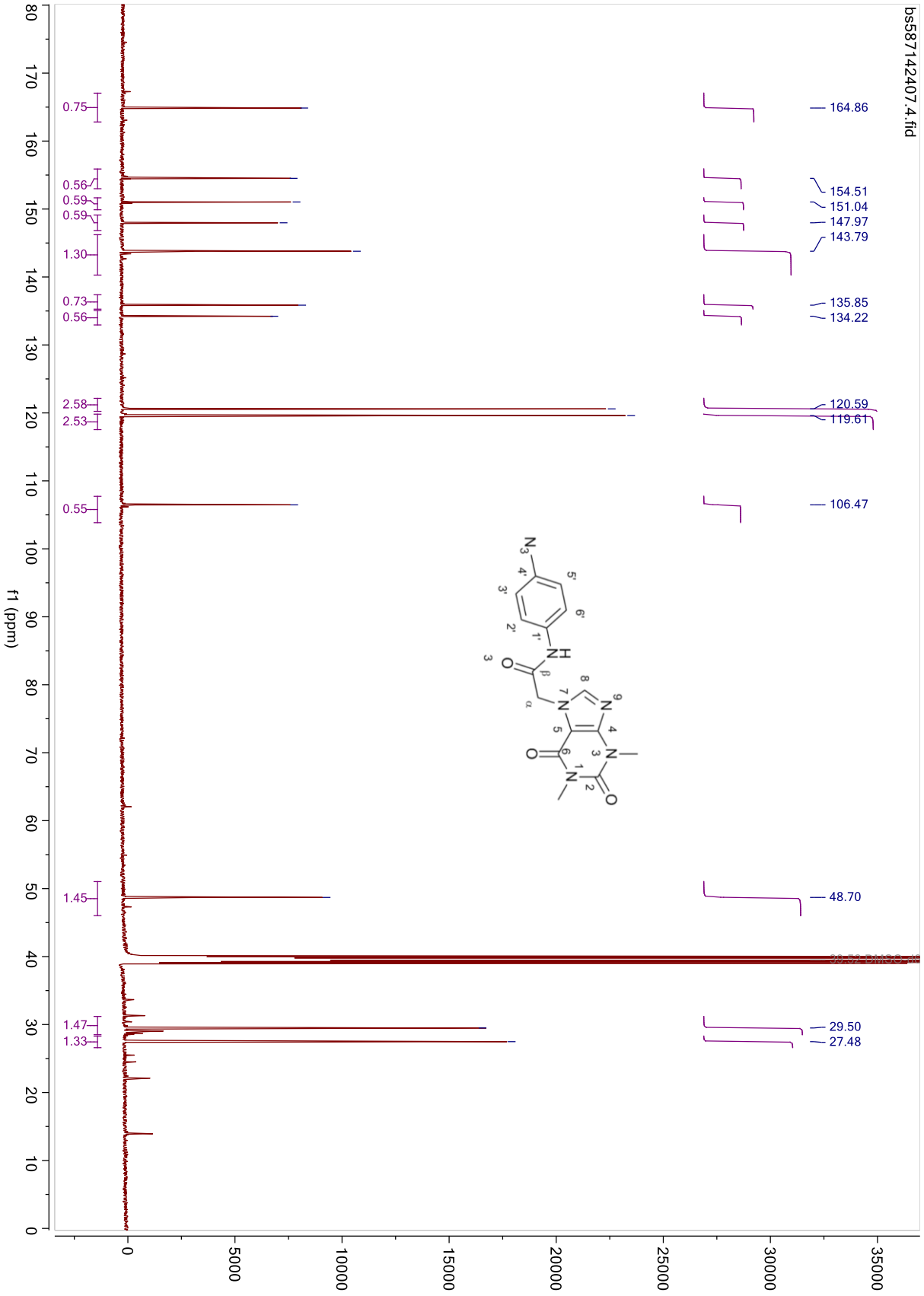


Figure S12: ^{13}C NMR spectrum of 4-33

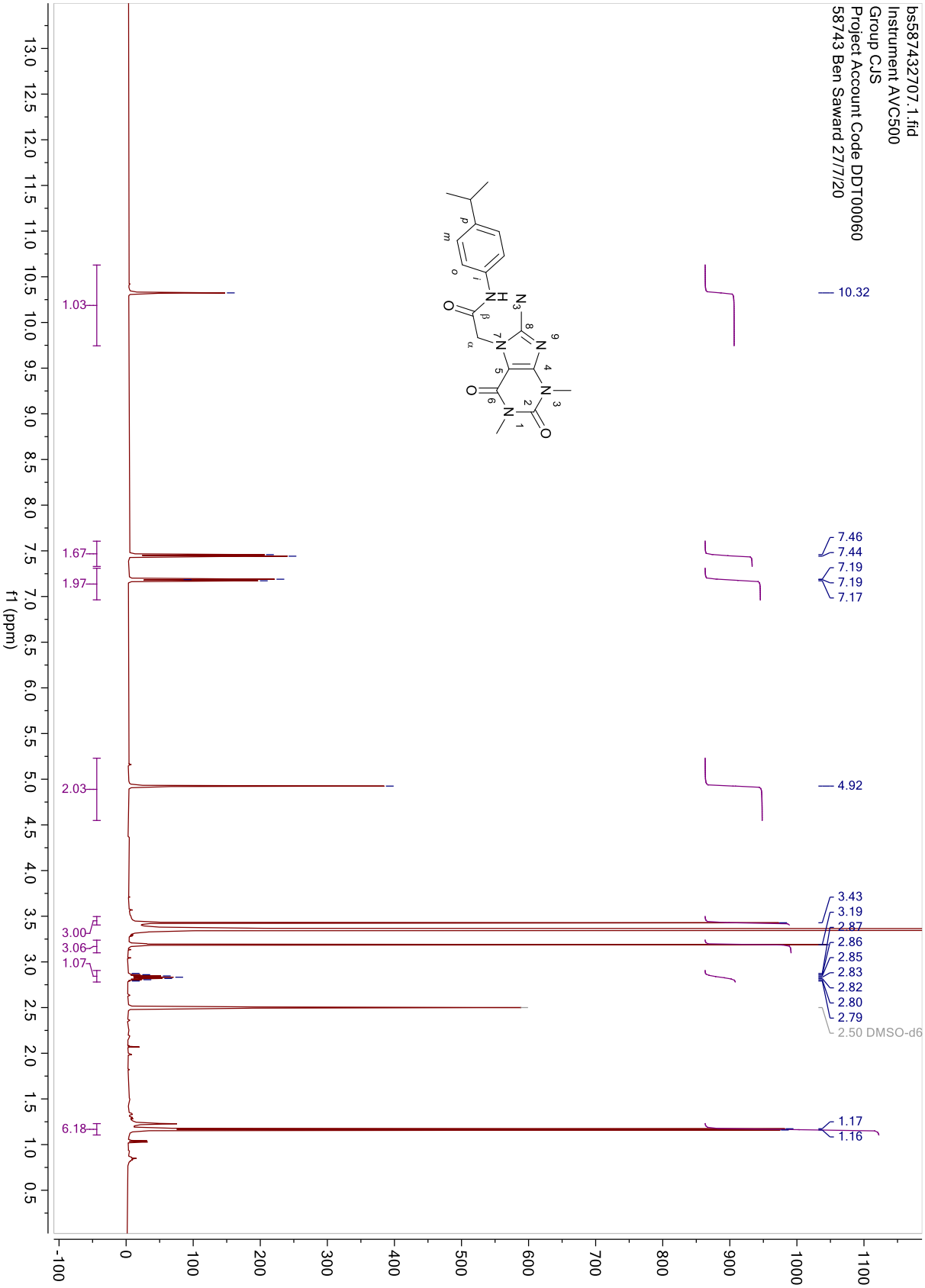


Figure S13: ^{1}H NMR spectrum of 4-35

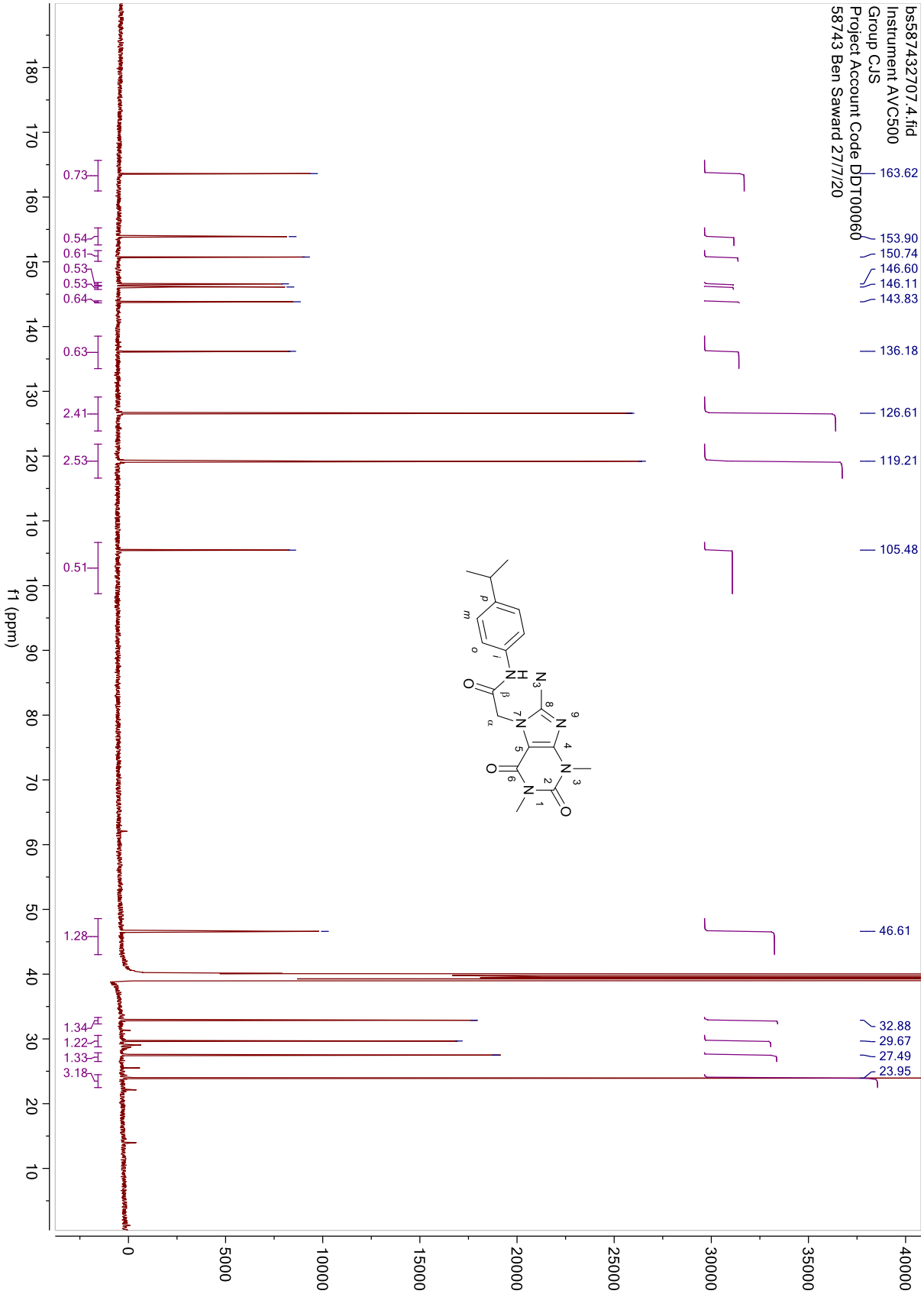


Figure S14: ^{13}C NMR spectrum of 4-35