Affinity-guided chemical probes for the study of protein interactions

Hester Annie Beard

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Chemistry

April, 2018

Intellectual property and publication statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Portions of the work presented in Chapter 2 were published as a preprint article in October 2017: 'The small molecule KHS101 induces bioenergetics dysfunction in glioblastoma cells through inhibition of mitochondrial HSPD1', bioRxiv 205203, 2017; https://doi.org/10.1101/205203; Polson, E.S., Kuchler, V.B., Abbosh, C., Ross, E.M., Mathew, R.K., Beard, H.A., Chuntharpursat-Bon, E., Williams, J., Da Silva, B., Shao, H., Patel, A., Davies, A.J., Droop, A., Griffiths, H.B.S., Chumas, P., Short, S.C., Lorger, M., Gestwicki, J., Roberts, L.D., Bon, R.S., Allison, S.J., Zhu, S., Markowetz, F., Wurdak, H. The contributions of the authors to the work described in this chapter *published in the preprint* were as follows: HAB (the candidate) performed all synthetic chemistry, ESP and VBK performed the phenotypic and chaperone activity assays, SZ performed the chemical proteomics and *in vitro* pull-down experiments, ESP, VBK, RKM, and HW prepared the manuscript and figures, RSB, JG, SJA, SZ, FM, and HW provided project leadership. The contributions of researchers that did not contribute to work detailed in the preprint are indicated throughout Chapter 2.

Chapter 3 refers to work from a research article that was published in September 2016: ('Economical and scalable synthesis of 6-amino-2-cyanobenzothiazole'; Hauser, J.R., Beard, H.A., Bayana, M.E, Jolley, K.E., Warriner, S.L., Bon, R.S, *Beilstein J Org Chem*., 2016, **12**, 2019-2025). HAB (the candidate) and JRH performed initial synthesis studies, JRH synthesised the final compounds, JRH performed calorimetry experiments with assistance from MEB and KEJ, SLW and RSB provided supervision, JRH prepared a draft which was edited into its final form by RSB.

Acknowledgements

Firstly, I would like to thank my primary supervisor Dr Robin Bon for giving me the opportunity to work on this project and for all the support and guidance he has given throughout my PhD. I would also like to thank my co-supervisor Prof Andy Wilson for the incredibly useful ideas and suggestions for the project. The detailed feedback and advice on written work and presentations you have both given throughout the years has helped me enormously with my confidence, and I am very grateful.

The work described in this thesis would not have been possible without help from the technical staff in the School of Chemistry and the Faculty of Biological Sciences. Particularly, Martin Huscroft for assistance with HPLC purification, Simon Barrett for NMR experiments, Dr Iain Manfield for SPR and Dr Rachel George for performing the tandem mass spectrometry experiments.

I would also like to thank Zoe Arnott for helping with the numerous protein gels, Dr Katherine Horner for advice on fluorescence anisotropy assays and Dr Fruzsina Hobor for teaching me how to express and purify protein. I must also thank Justine Raymond for helping with the synthesis of KHS101 intermediates in Chapter 2 (and for being a lovely student) and David Klebl for his contributions to the biophysical analysis of the KHS101-target interaction.

A special thank you must be given to Aisling, my ‘partner in crime’. I have honestly made a best friend for life and am really going to miss you! Thank you Jake for all your useful science suggestions, funny office conversations and interesting lab music. Thanks also to the rest of the Bon and Wilson group members (past and present).

Thank you Mum and Dad, for all the love and support you continue to give me. Finally to Sam, I could not have done this without you.

Abstract

Chemical methods that allow for the targeted labelling of a specific protein within a complex biological environment can enable valuable information regarding the structure and function of proteins to be gained. This thesis explores two different projects where affinity-guided chemical probes were used to study the interactions of proteins, both with small molecules (Chapter 2) and interacting protein partners (Chapter 3). Firstly, chemical labelling methods based on a recognition unit for the protein of interest are reviewed in Chapter 1. Then, Chapter 2 describes how a combination of chemical tools (including photoaffinity, biotinylated and fluorescent probes) were used to study the interaction of a small molecule inducer of human glioblastoma cell death and its relevant target. This work resulted in the identification of HSPD1 as a potential therapeutic target for the treatment of glioblastoma. Chapter 3 details the development of a method for traceless labelling of B-cell lymphoma 2 (BCL-2) family proteins, using a ruthenium-bipyridyl modified peptide. Myeloid cell leukaemia 1 (MCL-1) was rapidly and selectively labelled with fluorescent and biotinylated tags, *in vitro*, facilitated by the interaction with a peptide mimicking a binding partner. Overall, this thesis demonstrates how affinity-guided labelling of proteins can be used for understanding molecular mechanisms of disease and mapping protein interactomes.

Table of contents

List of abbreviations

Ac	Acyl
AGD	Affinity-Guided 4-(Dimethylamino)pyridine chemistry
APS	Ammonium Persulfate
BAD	BCl-2 Associated Death promoter protein
BAK	BCL-2 homologous Antagonist Killer protein
BCL-2	B-Cell Lymphoma-2 anti-apoptotic protein
BCL-X_L	B-Cell Lymphoma 'extra-large' anti-apoptotic protein
BH	BCL-2 Homologous region
BID	BH3 Interacting Domain death agonist
BIM	BCL-2 Interacting Mediator of cell death
Boc	*tert*-Butyloxycarbonyl
BP	Benzophenone
bpy	2,2'-Bipyridine
CD	Circular Dichroism
COSY	Correlation Spectroscopy
DEPT	Distortionless Enhancement by Polarisation Transfer
DIPEA	Di*iso*propylethylamine
DMAP	4-(Dimethylamino)pyridine
DMF	*N,N*-dimethylformamide
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EDCI	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDT	1,2-Ethanedithiol
ESI	Electrospray Ionisation
FA	Fluorescence Anisotropy
FITC	Fluorescein Isothiocyanate
Fmoc	9-Fluorenylmethoxycarbonyl
GBM	Glioblastoma Multiforme
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HCTU	O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
*h*DM2	Human Double-Minute 2
HIF-1α	Hypoxia-Inducible Factor 1 alpha

HMBC	Heteronuclear Multiple Bond Correlation
HMQC	Heteronuclear Multiple Quantum Coherence
HOBt	1-Hydroxybenzotriazole
HPLC	High Pressure Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HSPD1	Heat Shock Protein family D member 1 (HSP60)
HSPE1	Heat Shock Protein family E member 1 (HSP10)
IC_{50}	Half maximal inhibitory concentration
IR	Infrared (spectroscopy)
ITC	Isothermal Titration Calorimetry
K_d	Dissociation Constant
LC-MS	Liquid Chromatography Mass Spectrometry
LDAI	Ligand-Directed Acyl Imidazole chemistry
LDT	Ligand-Directed Tosyl chemistry
MCL-1	Myeloid Cell Leukaemia 1
m.p.	Melting Point
Ms	Methanesulfonyl
NHS	*N*-Hydroxysuccinimide
NMM	*N*-Methylmorpholine
NMP	*N*-Methylpyrrolidinone
NMR	Nuclear Magnetic Resonance
NOXA	Phorbol-12-myristate-13-acetate-induced protein 1
NPC	Neural Progenitor Cell
OEG	Oligo Ethylene Glycol
p300	E1A binding Protein 300
p53	Tumour protein 53
PDB	Protein Data Bank
PPI	Protein-Protein Interaction
PUMA	p53 Upregulated Modulator of Apoptosis
PyBOP	Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
RTA	Radical Trapping Agent
SPPS	Solid Phase Peptide Synthesis
SPR	Surface Plasmon Resonance
^{t}Bu	*tertiary*-Butyl
TTR	Tag Transfer Reagent

TAMRA	Tetramethylrhodamine
TCEP	*tris*-2-Carboxyethylphosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIPS	Tri*iso*propylsilane
TLC	Thin Layer Chromatography
TMS	Tetramethylsilane
TOF	Time of Flight
Trt	Trityl
UV	Ultraviolet
WT	Wild Type

Amino acid	**Three-letter code**	**One-letter code**
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1 – Chemical methods for protein labelling

1.1 Importance of chemical labelling and its uses

In order to elucidate the structure and function of specific proteins, it is important that they can be studied within their native environment, *i.e.* living cells. It is thus important that techniques are developed that modify native proteins at the endogenous level, without genetic manipulations.[1] Despite their uses, current biological techniques such as the use of green fluorescent protein (GFP, used in the visualisation of proteins)[2] are inadequate for the precise study of protein function in living systems, due to large protein tags which can interfere with the structure, function, localisation and dynamics of the protein under study.[3]

A variety of methods utilising chemistry-based protein modification have been established; these show promise over molecular-biology based techniques due to smaller probe sizes which are less likely to perturb the function of the target protein. In addition, a variety of labels can be attached to the protein of interest, such as affinity tags, fluorescent dyes, NMR and MRI probes. When used *in vitro*, these modifications can be attached to the protein through nucleophilic amino acid residues on the surface of the protein, for example the thiol group of cysteine.[4] Cysteine is generally chosen because of its comparatively low level of occurrence on protein surfaces, in addition to its higher nucleophilicity. If not naturally occurring in the protein, cysteine can be introduced into the protein's amino acid chain by site-directed mutagenesis.[5] Labels are commonly incorporated *via* Michael additions of thiols (*e.g.* Cys) with maleimides,[6] or the reaction of amines (*e.g.* Lys and protein *N*- termini) with isocyanates, isothiocyanates or activated esters.[7] Although these reagents allow efficient labelling of proteins *in vitro*, they do not allow target-specific labelling within a complex cellular environment.

The development of bioorthogonal chemistry is of paramount importance to this area; this exploits chemical reactions which do not affect biological processes, while the biological environment equally does not impede the progress of the reaction.[8] Selective labelling of a protein of interest within a crude biological system can be enabled with the use of two reactive groups that are non-native to the system and only have the capacity to react with each other. Pioneered by the Bertozzi group with the introduction of the modified Staudinger ligation,[9] a variety of other bioorthogonal reactions have been developed, including the copper catalysed alkyne-

azide cycloaddition (CuAAC),[10,11] strain-promoted alkyne-azide cycloadditions (SPAAC)[12,13] and inverse electron-demand Diels-Alder reactions.[14] However, incorporating non-natural reactive handles into proteins often requires genetic manipulations; as a result these proteins are no longer expressed at endogenous levels. Furthermore, the physiological conditions within the cell are majorly altered.[4] Non-natural handles can also be incorporated into proteins through metabolic incorporation of molecular building blocks such as non-natural amino acids,[15] sugars[16] and lipids,[17] although this may not always be completely selective to a particular protein or protein class.

Using a ligand or recognition unit that binds specifically to a protein of interest to direct labelling onto native proteins, presents unique advantages over the previously described techniques. Many of these chemistries have been used to identify molecular targets of small molecules to study their mechanism of action, in addition to studying interactomes of proteins. These methods will be explored in the rest of this chapter, together with a discussion of the advantages and limitations of each approach.

1.2 Affinity labelling

Affinity labelling[18] allows the study of endogenously expressed proteins through their modification with a reactive group attached to a recognition unit (*Figure 1*, a). The ligand component of the affinity probe imparts selectivity for a protein of interest, allowing targeted labelling in whole proteomes of cells. Coupled with mass spectrometry techniques, structural information about the protein can be obtained, such as the amino acid composition around the ligand binding site.[19] Following complexation of the ligand with the target protein, an amino acid residue on the surface of the protein reacts with the reactive probe moiety, forming a covalent bond (*Figure 1*, c). Probes containing reactive groups that are activated upon light irradiation, such as diazirines,[20] benzophenones[21] or aryl azides[22] are termed “photoaffinity probes”; these have been used extensively in the identification of protein targets and binding sites, in conjunction with proteomics analyses.[23,24,25]

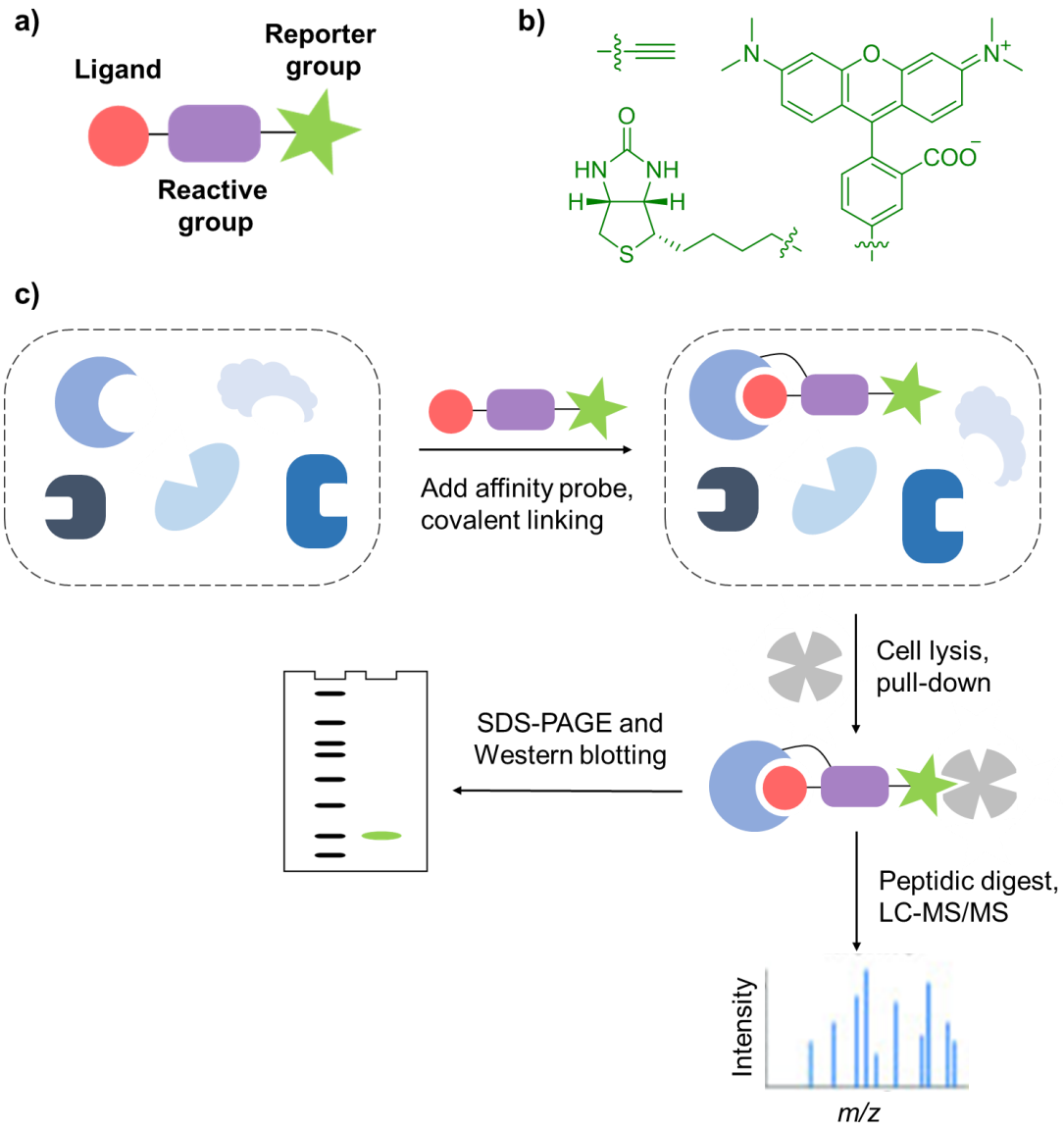


***Figure 1**: a) General structure of an affinity probe. b) Chemical structures of reporter groups, including an alkyne (used in a two-step labelling approach), a biotin and a fluorescent dye. c) Schematic illustration of an affinity labelling experiment using biotin as the reporter group. Affinity probe is used to specifically label a protein of interest within a cell. Labelled proteins are isolated using affinity purification and analysed using mass spectrometry and Western blotting.*

1.3 Photoaffinity labelling

1.3.1 Concepts in photoaffinity labelling

Photoaffinity labelling is a chemical crosslinking technique that uses light-activated chemical tools to form covalent bonds to target proteins, enabling the study of protein-ligand interactions. The photoaffinity probe typically comprises a ligand for the protein of interest, a photo-reactive group, and a reporter group (*Figure 2*, a). Upon irradiation with a specific wavelength of light, a reactive intermediate is formed, which rapidly reacts with functionalities in mutual proximity. Harnessing light in this way enables

transient and/or weak interactions to be trapped, and provides temporal control over the crosslinking reaction. Furthermore, samples can be left to equilibrate prior to irradiation, reducing the frequency of non-specific crosslinks. This powerful method can be used to identify unknown molecular targets of small molecules or drug compounds (*e.g.* targets of hit molecules identified through phenotypic screens), in addition to assisting structural elucidation of binding sites. Aryl azides, diazirines and benzophenones are the most frequently used photoaffinity groups;[23,26] these undergo photolysis upon irradiation to form nitrene, carbene and diradicals, respectively (*Figure 2*, b).

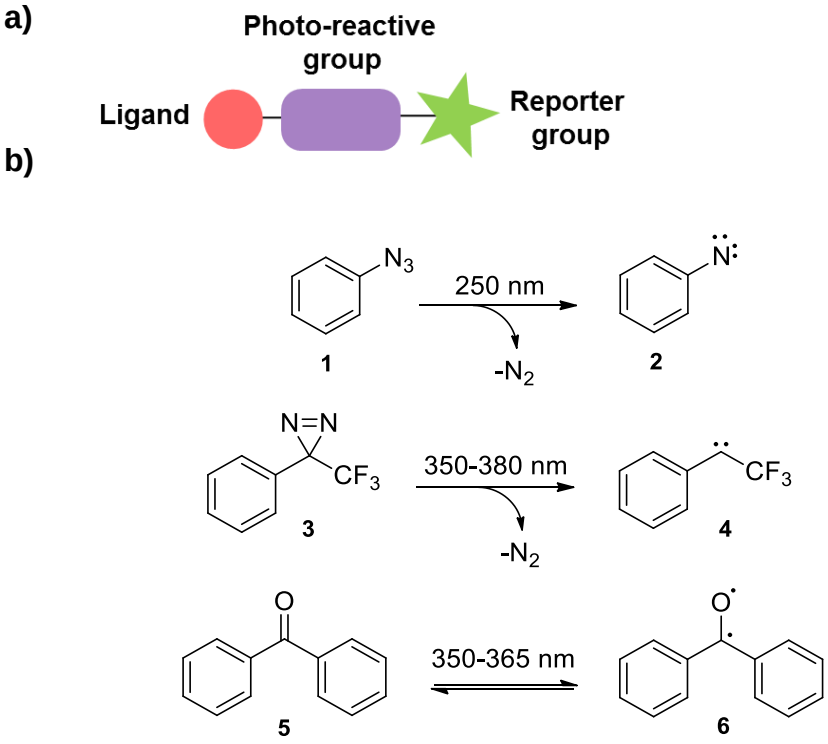


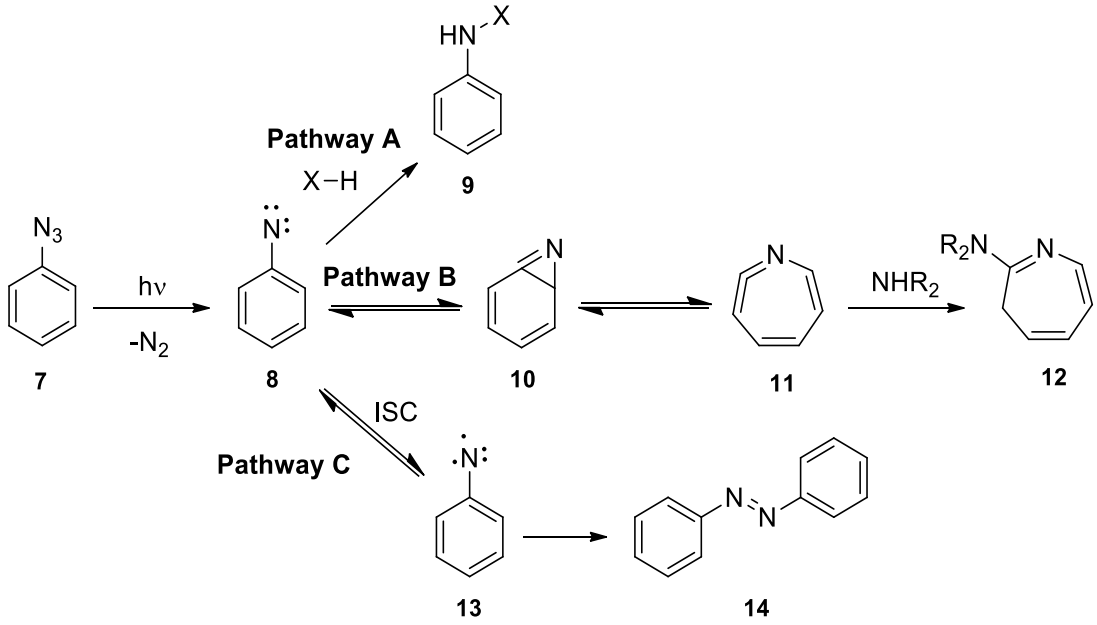
***Figure 2**: a) General structure of a photoaffinity probe, b) Chemical structures of the most commonly used photoaffinity groups and the reactive intermediates they form upon light irradiation.*

An ideal photoaffinity group should be activated at wavelengths that avoid damage to biomolecules, but that still produce highly reactive intermediates with the capacity to react and form stable adducts with many different bond types.[27] Surplus absorption of UV light by proteins, which absorb at 280 nm (due to Trp, Tyr and Phe residues) and 200 nm (peptide bonds) can result in degradative chemical transformations.[23] Furthermore, the photo-activatable group should be stable in the dark at a range of pHs, and not be excessively large that it interferes with protein-ligand binding.

1.3.2 Photo-induced crosslinking chemistries

1.3.2.1 Aryl azides

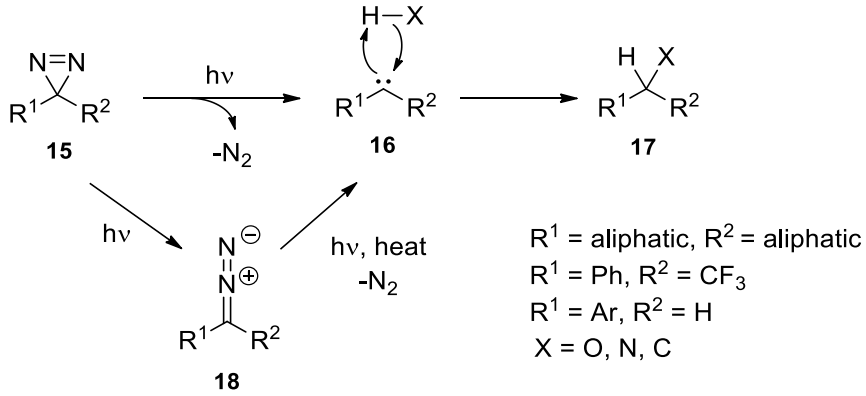
The synthetic tractability of aryl azides has led to the frequent use of these photoaffinity groups.[28,29,30] However, the shorter wavelengths required for excitation (~250 nm)[31] can damage proteins, limiting their practical utility. Upon photoirradiation, loss of molecular nitrogen results in the generation of a singlet nitrene reactive intermediate (**8**, *Scheme 1*). Although still indiscriminate in its reactivity, this species is less reactive than its carbon analogue, the carbene, and can rearrange to form undesired side-products resulting in decreased crosslinking yields.[32] For example, in addition to insertion into C-H and N-H bonds (*Scheme 1*, pathway A), ring expansion to dehydroazepine intermediate **11** *via* benzazirine **10**, permits reaction with nucleophiles such as primary amines (*Scheme 1*, pathway B).[27] However, substitution of the phenyl ring with fluorine (*e.g.* tetrafluorophenylazide), has been shown to reduce the rate of intramolecular ring expansion, allowing the singlet nitrene to undergo intermolecular crosslinking reactions.[33] While singlet nitrene **8** behaves as an electrophile, triplet state **13** formed through intersystem crossing (ISC) behaves as a diradical;[34] this species can dimerise in an additional unproductive side-reaction, affording azo **14** (*Scheme 1*, pathway C).



Scheme 1: *Reaction pathways of the singlet nitrene generated from aryl azide* ***7****. Pathway A: productive insertion into X-H bonds of protein. Pathways B and C are non-productive reactions.*

1.3.2.2 Diazirines

The high chemical stability and small size of diazirines (particularly the alkyl derivatives) are key advantages of these photoaffinity groups. The reactive species is efficiently generated at longer wavelengths compared with aryl azides, reducing the likelihood of damage to proteins. Upon irradiation with UV light (350 – 380 nm) an extremely reactive and indiscriminate singlet carbene intermediate **16** is formed, *via* irreversible loss of N_2 from either diazirine **15** or its diazo isomer **18** (*Scheme 2*). This reactive intermediate can insert into neighbouring C-H or heteroatom-H bonds through a concerted mechanism, involving abstraction of H by the non-bonding electrons of the carbene carbon, which is simultaneously attacked by the bonding electrons of the H-X bond. However, rearrangement of **15** to diazo **18** can result in undesired by-products *via* carbocations (produced by protonation at carbon to form the diazonium ion, followed by loss of N_2),[35] which are no longer indiscriminate in their reactivity.

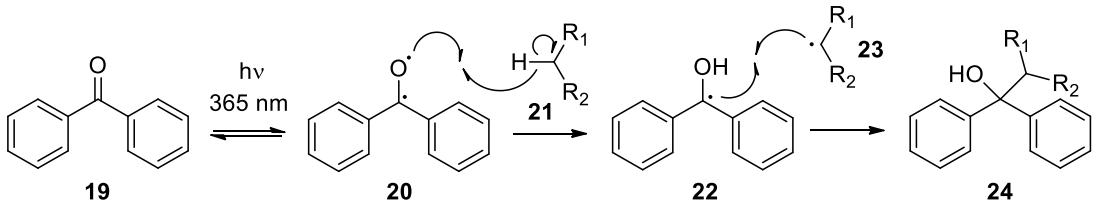


Scheme 2: *Mode of action of the diazirine photoaffinity group: the diazirine ring can be converted to its relatively stable linear diazo isomer **18**, which in turn can generate a carbene upon extrusion of molecular nitrogen from the excited state.*

Introduction of the 3-trifluoromethyl-3-aryldiazirine (TFMD) group by Brunner *et al.*[36] was found to stabilise diazo isomer **18**, reducing its propensity to form diazonium species. As a means to increase crosslinking yield, irradiation of diazo isomer **18** at 302 nm enables conversion to reactive carbene intermediate **16**.[37] It should be noted that the high reactivity of the carbene (which has a half-life in the nanosecond range) means that quenching by water is often observed, decreasing yields of crosslinked product.[38]

1.3.2.3 Benzophenones

Benzophenone (BP) was first introduced as an effective photoreactive group for labelling proteins in 1974 by Printz *et al.,*[39] where crosslinking to the α-carbon of glycine was reported. The mechanism by which this occurs involves the formation of a triplet ketyl biradical.[21] Absorption of a photon at ~350 nm by the chromophore of BP, causes the promotion of a non-bonding electron from the sp^2-like orbital of the carbonyl oxygen to an antibonding π* orbital of the carbonyl group (*Scheme 3*, producing **20**). It is this reactive intermediate which is able to react with functional groups on the protein. In this resultant triplet excited state, interaction with weak C-H σ-bonds (enabled due to the electron-deficient oxygen n-orbital) allows hydrogen abstraction to occur (*Scheme 3*, forming **22**). Alternatively, reaction with amines can involve electron transfer followed by proton abstraction. Recombination of the ketyl **22** and alkyl **23** radicals generates a new C-C bond, linking the BP group to the target protein (*Scheme 3*, **24**).



Scheme 3: *Mechanism of photo-induced crosslinking of benzophenone with a methylene group, via a sequential abstraction-recombination mechanism.*[40]

Advantages of using benzophenones in photo-crosslinking experiments include their improved chemical stability compared with other photo-activatable groups such as aryl azides and diazirines,[21] their ability to be activated at wavelengths that avoid damage to the protein,[39] and their preferential reactivity with 'unreactive' C-H bonds - even in the presence of bulk nucleophiles (reaction with water is reversible). Furthermore, they are relatively easy to prepare, with intermediates often being commercially available and reasonably priced.

On the other hand, steric hindrance attributed to the bulky, hydrophobic benzophenone group can interfere with protein-ligand interactions, reducing specificity for the intended protein target and resulting in non-specific labelling.[21] Furthermore, BP diradicals are reported to have preferential reactivity for methionine residues, compared with the indiscriminate reactivity ascribed to the carbene, which could result in mistaken binding-site identification.[41] Finally, a light-independent reaction of the aryl ketone with biological amines has been observed: the group of Schultz reported a covalent Schiff base adduct of a BP-containing cyclic peptide with

lysine residues on HIV protease.[42] This phenomenon should be kept in mind when analysing modifications on amino acid residues during peptide mapping experiments, for example.

1.4 Traceless affinity labelling

To enable studies into the function of the protein post-labelling, the original function of the protein must be restored. This point has been addressed in the development of "traceless affinity labelling" by the group of Hamachi.[43] Once the reactive group has been covalently linked to the target protein, the ligand component is removed, leaving the active site vacant and the protein often able to perform its native function. To date, eight different methods within this term have been developed, although the initial two chemistries[44,45] were only applicable for labelling proteins *in vitro*. The following six types of traceless labelling chemistry have been shown to work in cell lysates and live cells; ligand-directed tosyl (LDT) chemistry,[46,47,48] ligand-directed acyl imidazole (LDAI) chemistry,[49,50] ligand-directed (di)bromophenyl benzoate (LDBB) chemistry,[51] ligand-directed *N*-sulfonyl pyridone chemistry (LDSP),[52] affinity guided DMAP (AGD) catalysis[53,54,55] and affinity-guided oxime (AGOX) chemistry.[56] The first four reaction types are considered to be exchange/cleavage (*Figure 3*, a), while the latter two are termed catalyst tethering types (*Figure 3*, b).

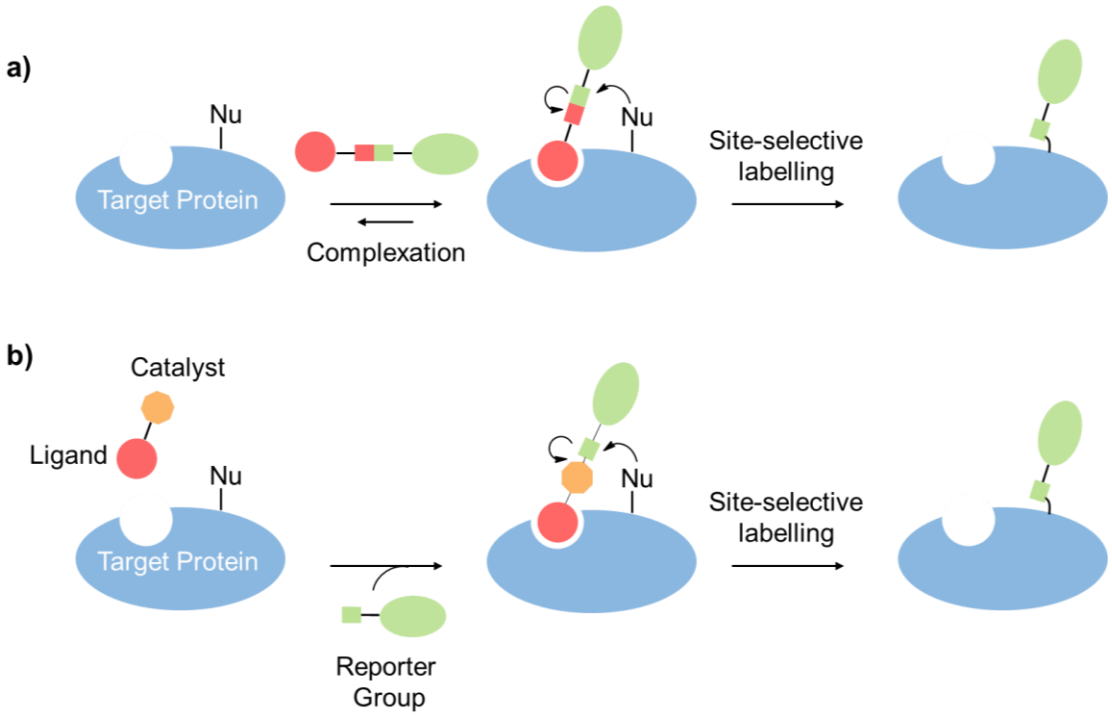


***Figure 3**: Schematic representation of traceless affinity labelling: a) exchange/cleavage reaction type (includes LDT, LDAI, LDBB and LDSP chemistries), b) catalyst tethering type (includes AGD and AGOX catalysis).*

1.4.1 Ligand-directed tosyl chemistry

The labelling reagent used in Ligand-Directed Tosyl (LDT) chemistry[46,47] comprises a ligand for the protein of interest and a reporter group, linked by an electrophilic aryl sulfonate ester (*Figure 4*). A nucleophilic amino acid residue adjacent to the ligand-binding pocket attacks the aryl sulfonate group in an S_N2 fashion, concomitantly releasing the ligand component. This process is driven by the "proximity effect" whereby the effective local concentration of the amino acid residue and reactive group is increased upon complexation. Analogous to enzymatic reactions, a bimolecular reaction is transformed into a pseudo-intramolecular reaction, resulting in enhanced reaction kinetics.[57] The reaction is made facile due to the aryl sulfonate ester moiety acting as an effective leaving group. As the ligand component is removed from the ligand-binding site post-labelling (and amino acids essential for protein function are protected from labelling by occupation of the 'active site') the labelled protein is often able to retain its native function.

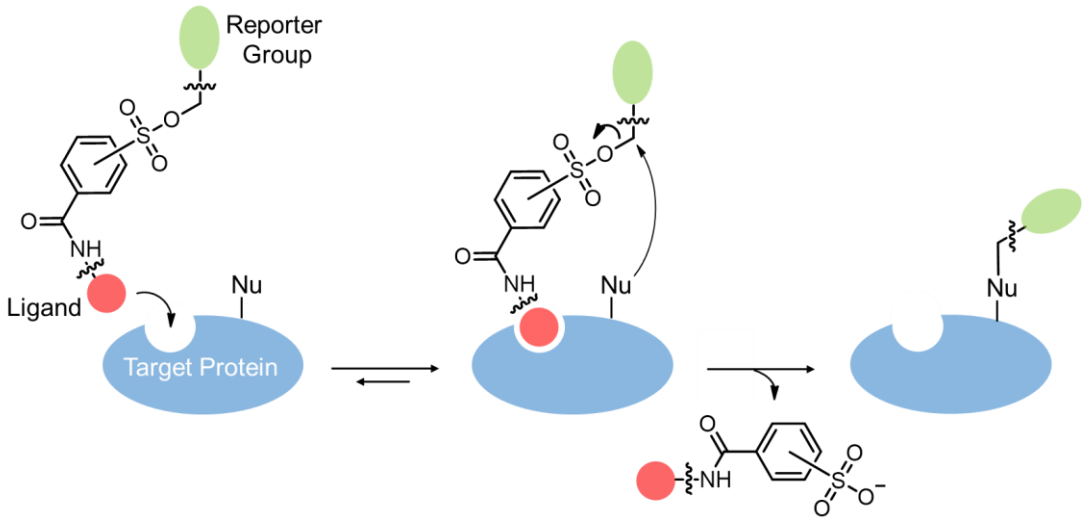


***Figure 4**: Schematic representing the mechanism by which target specific protein labelling is achieved using LDT chemistry. Attack of the electrophilic sulfonate ester by nucleophilic amino acids (such as His, Tyr and Glu residues) results in transfer of a reporter group onto the protein of interest and cleaving the ligand component of the reagent.*

In addition to identifying off-targets of the anticancer drug lapatinib in live NCI-N87 cells (a human gastric cancer cell line),[58] this technique has been shown to selectively and efficiently label carbonic anhydrase II (CAII) in live mice.[46] Furthermore, LDT chemistry has been used to construct a fluorescent biosensor for phosphotyrosine peptides in bacterial cell lysates,[47] using the *N*-terminal Src homology 2 (SH2) domain of the p85α subunit of human phosphatidylinositol-3-kinase (PI3K) as a scaffold. Peptide mapping analysis of labelled proteins revealed His, Tyr and Glu residues were modified. However, labelling of membrane proteins

using this chemistry has proved challenging, while LDT-mediated labelling generally suffers from slow reaction kinetics (labelling times typically take >10 hours).[57]

A particularly interesting application of this chemistry was demonstrated in the study of protein-protein interactions.[48] A bifunctional LDT reagent (*Figure 5*, a) containing tosyl ester and diazirine groups, designed to label FK506-binding protein (FKBP12) was used to capture interacting partners of FKBP12 upon UV irradiation (*Figure 5*, b). Evaluation of spacer length and rigidity found a piperazine linker to give the best labelling yields (compound **25** gave 60% conversion *in vitro* after 48 hrs, analysed by MALDI-TOF mass spectrometry). Labelling of FKBP12 using reagent **25**, followed by irradiation at 365 nm in a mixture with known interacting partner calcineurin, revealed bands corresponding to FKBP12-calcineurin crosslinked adducts *via* Western blot analysis. Although this experiment proved that the natural binding ability of labelled proteins was restored using this traceless approach, the low abundance of calcineurin in HeLa cells did not permit crosslinking of FKBP12 to endogenous calcineurin in live cell experiments.

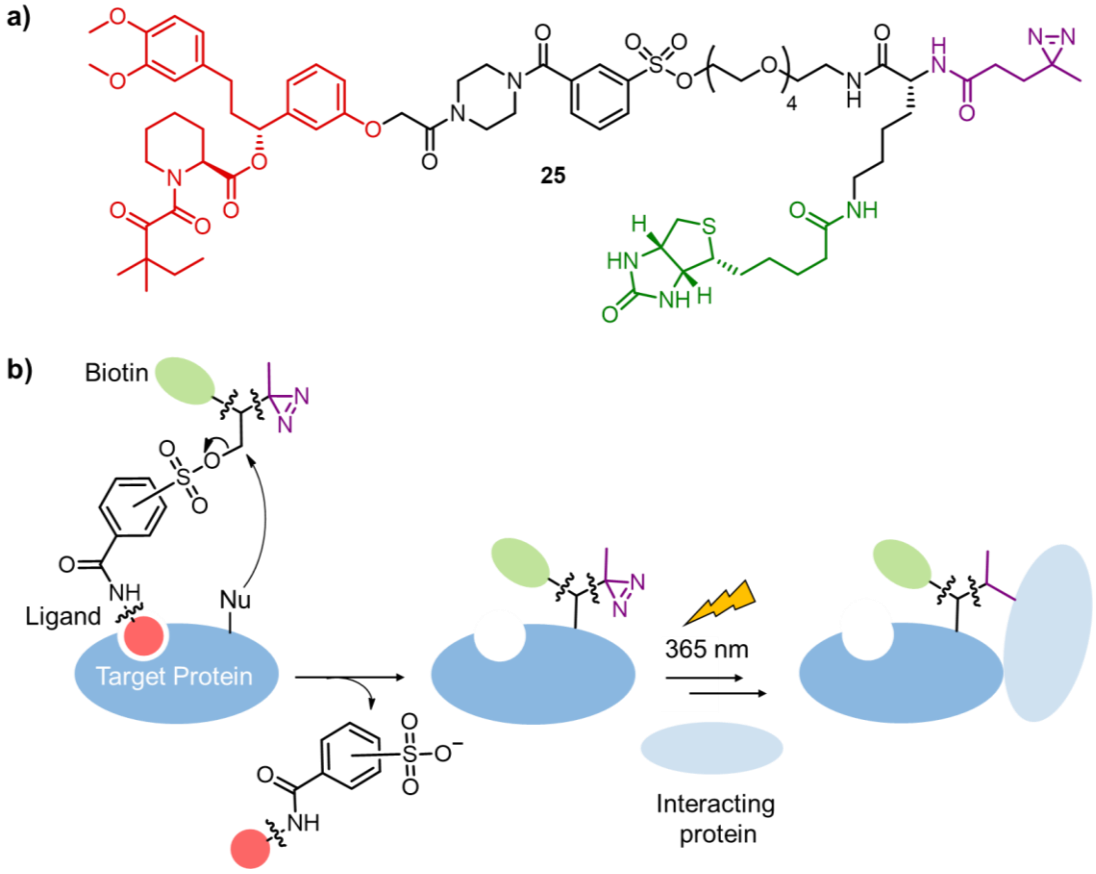


***Figure 5**: Identification of interacting partners of a target protein using bifunctional labelling reagent **25**. a) Chemical structure of bifunctional labelling reagent. b) Schematic representation of labelling approach using reagent **25**: LDT is the reactive group in step 1, diazirine is the reactive group in step 2, allowing covalent capture of interacting proteins.*

1.4.2 Affinity-guided DMAP catalysis

The use of a catalyst to aid the chemical labelling process has been adopted by the Hamachi group, leading to the development of a new class of ligand-directed chemistry: Affinity-Guided DMAP (AGD) catalysis.[53] Dimethylaminopyridine (DMAP) is a nucleophilic catalyst used in organic reactions, including acyl transfer.[59] Introduction of the DMAP group into ligand-directed reagents dramatically reduced labelling times relative to LDT-mediated labelling (<3 h versus >10 h).[57] Following binding of the ligand to the target protein, the DMAP moiety catalyses the transfer of an acyl group (and attached reporter) onto a nucleophilic residue proximal to the ligand-binding site (*Figure 6,* a). Once the label is attached to the protein surface, the ligand and attached DMAP group dissociate from the active site, preserving protein function. This chemistry allows the catalytic AGD reagent to be regenerated following cleavage from the target protein, reducing the quantity of ligand-DMAP conjugate administered to the cell and decreasing the risk of off-target labelling.

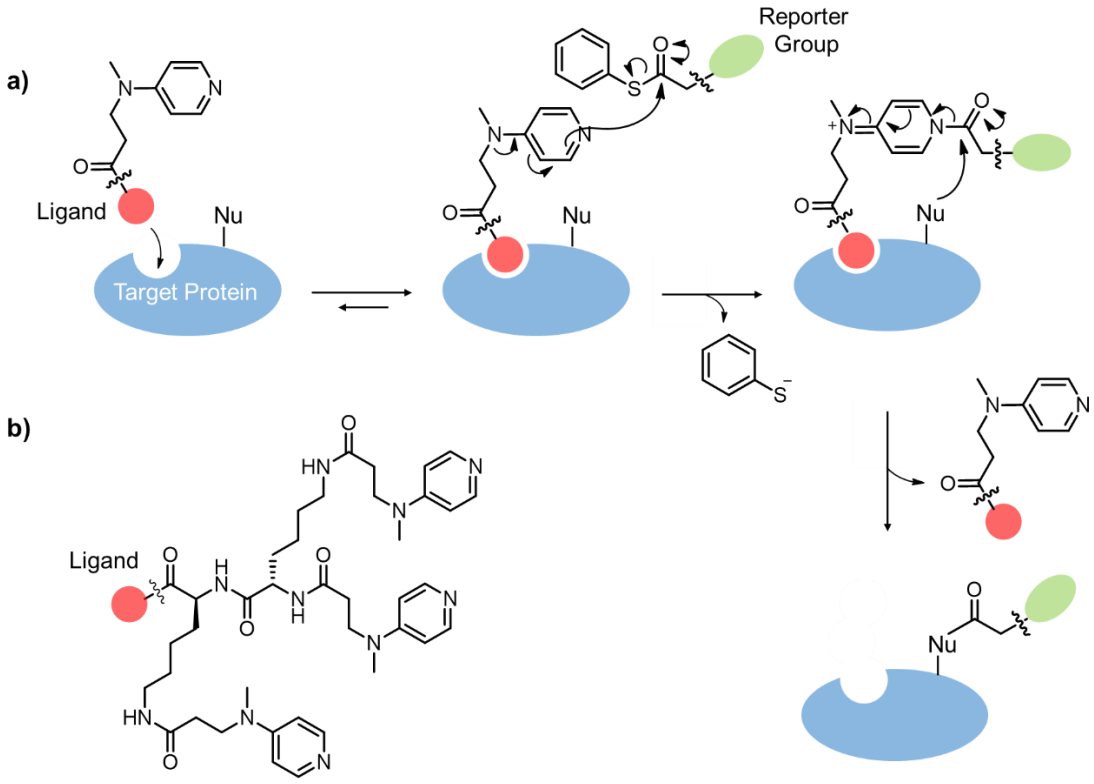


***Figure 6**: a) Schematic representing the mechanism for protein labelling using AGD chemistry. The reaction between a thioester label and AGD reagent results in a highly reactive acylpyridinium intermediate, which is attacked by a nucleophilic residue (Lys, Ser and Tyr have been reported) on the surface of the protein. b) General structure of a trivalent DMAP catalyst.*

This chemistry has been validated by the labelling of specific lectins, for instance congerin II (CongII) and concanavalin A (ConA).[53] Interestingly, the rate of the labelling reaction has been shown to increase with additional DMAP catalyst units tethered to the ligand, for example with the use of di- or tri-DMAP catalyst-ligand complexes (*Figure 6*, b).[55] This chemistry has also been used in the labelling of an SH2 domain (using a peptide ligand containing a phosphotyrosine),[60] with the tri-DMAP catalyst proving more effective for the labelling of this domain compared to proteins such as CongII and FKBP12.[55] Peptide mapping analysis revealed it was a lysine residue which had been acylated on SH2, an amino acid proximal to the ligand binding site. However, labelling was only attempted on purified portions of the SH2 domain *in vitro*. It is also not known whether the increase in labelling efficiency from using multivalent DMAP reagents is solely due to enhanced local concentration of the catalyst, or whether some DMAP moieties act as bases, accelerating acyl transfer *via* deprotonation of other DMAPs and/or of basic nucleophilic amino acid residues (both of which would be protonated in the physiological conditions of a cell).

Initially, the scope of this technique was reduced to labelling proteins on the surface of cells,[55,61] due to the high reactivity of the original acyl donor (an unsubstituted phenyl ring). However, methyl-substitution at the *ortho* and *para* positons of the phenyl ring, reducing the reactivity of the thioester by introducing a less effective leaving group, enabled the intracellular labelling of FKBP12.[54] Nevertheless, the requirement for basic pH conditions (pH >8) for efficient labelling and high concentrations of the thioester acyl donors due to their esterase-mediated decomposition, limits the use of AGD reagents in live cells and tissues.

1.4.3 Ligand-directed acyl-imidazole chemistry

Another recently developed type of traceless affinity labelling is Ligand-Directed Acyl Imidazole (LDAI) chemistry (*Figure 7*).[62] Analogous to LDT chemistry, the reaction is an exchange/cleavage type, however the linker connecting the ligand to the probe is an alkyl group, while the reactive group is an acyl imidazole moiety. Compared with the two previously described methods (LDT and AGD), this chemistry has the potential to label a broader range of targets, due to the tuneable electrophilicity of the acyl imidazole through the replacement of the X atom with either O or N. The increased reactivity of the reactive group (through variation of the X atom) means a faster rate of labelling reaction can be achieved, although alkylacylimidazole (X = C) variants have been shown to be too unstable to be isolated.[62]

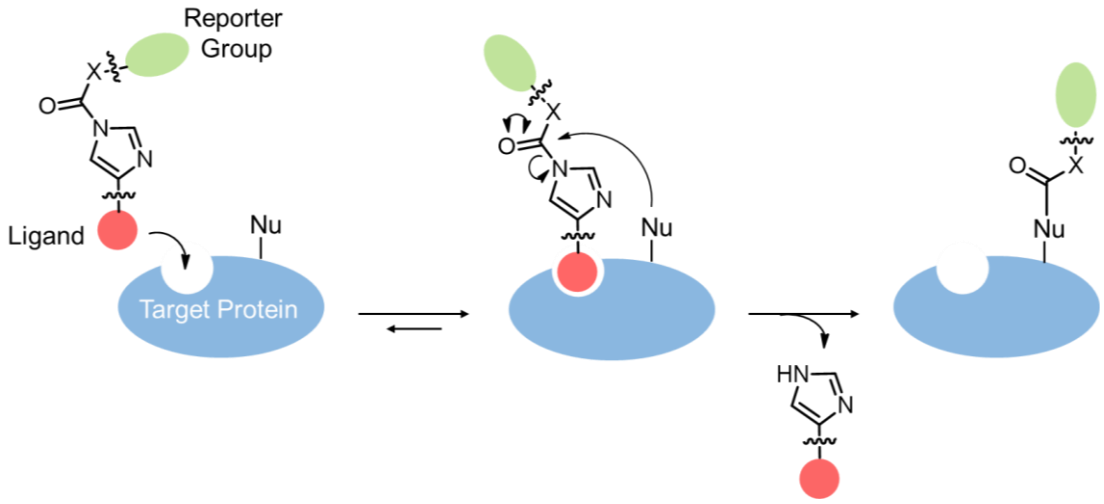


***Figure 7**: Schematic representing the mechanism of protein labelling using LDAI chemistry (X = O, N). Attack of Lys, Ser or Tyr residues onto the acyl imidazole group of the LDAI reagent facilitates acyl transfer onto the protein of interest.*

As of yet, LDAI chemistry has been limited to labelling proteins on the surface of live cells and tissues. However, this method has been shown to work where LDT chemistry was not effective, for example allowing kinetic studies of the endogenous folate receptor on a live cell surface to be carried out.[49] Optimisation of LDAI chemistry using carbonic anhydrase I (CAI) as a model enzyme allowed the reaction characteristics of the LDAI chemistry to be investigated, in addition to constructing a caged version of CAI (an enzyme whose activity can be controlled by light).[62] During these investigations, it was found that the acyl imidazole groups underwent non-productive hydrolysis under the aqueous conditions used for labelling. However, a LDAI reagent with an alkoxyacyl imidazole group, 5-carbon chain linker and coumarin fluorophore had a half-life of 16 hours (pH 7.2, 37 °C) according to HPLC analysis, and was not significantly affected by high concentrations (10 mM) of the nucleophile glutathione.

α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-type glutamate receptors (AMPARs) have been labelled in live cultured neurons and brain tissues, using LDAI reagents employing a small molecular antagonist as the ligand.[63,64] Selective modification of native AMPARs with fluorescein was achieved, permitting visualisation of AMPARs (using confocal imaging) in deeper regions of the brain compared with standard immunostaining methods. This was reasoned to be due to the enhanced tissue penetration of the small-molecule based LDAI reagent compared with antibodies.

1.4.4 Ligand-directed dibromophenyl benzoate chemistry

Ligand-Directed diBromophenyl Benzoate (LDBB) chemistry[65] was originally developed in an attempt to improve the labelling efficiency and aqueous stability of previously reported ligand-directed labelling reagents. However, it should be noted that the reactive entity is rather bulky and hydrophobic. On complexation of the ligand to the binding site of the target protein, a nucleophilic residue such as lysine can attack the ester, concomitantly transferring the acyl group (and attached reporter group) and cleaving the ligand component (*Figure 8*). A variety of substitutions on the phenyl ring in addition to different spacer lengths were investigated, with an *ortho*-dibromo phenyl ester proving to be the most efficient labelling reagent.

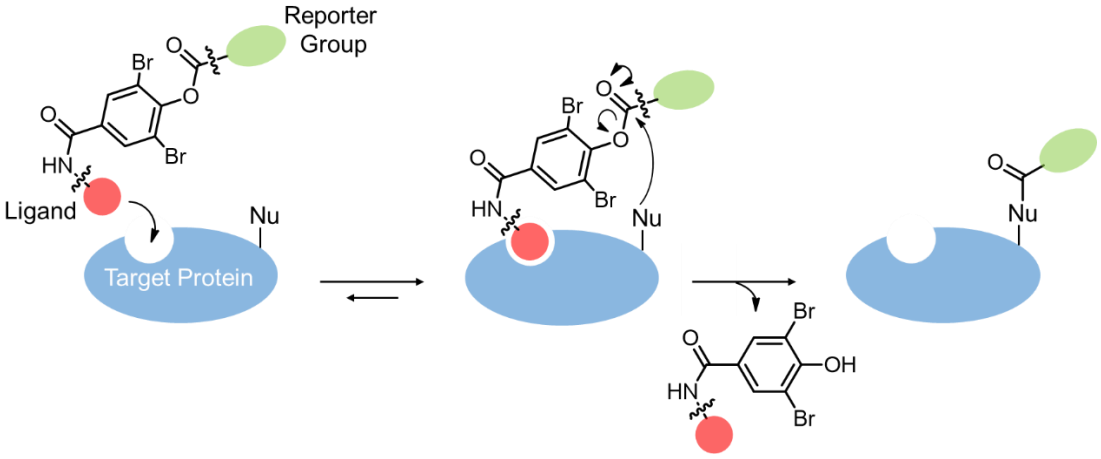


***Figure 8**: Schematic representing the mechanism of protein labelling using LDBB chemistry. Nucleophilic residues (e.g. Lys and His) attack the dibromophenyl ester, transferring the reporter group onto the protein of interest and releasing the ligand component.*

An experiment comparing the LDBB reagent with the analogous version based on acyl imidazole chemistry, found the former to have a labelling reaction 7-fold faster than the latter, *in vitro*.[65] Furthermore, the optimised reagent (with the fluorophore directly attached to the *ortho*-dibromophenyl ester) was found to be resistant towards autolysis in aqueous buffer, in addition to catalytic hydrolysis with esterase – possibly a result of the steric hindrance introduced by the bromine atoms on the phenyl ring. Overexpressed eDHFR (in mammalian cells) was shown to be selectively labelled (over 85% labelling yield), even within the complex environment of a cell. Finally, both cytosolic (human CAII) and membrane-bound (human CAXII) endogenous intracellular proteins were also labelled in live cells.[65] Peptide mapping analysis revealed that lysine and histidine were the residues modified upon labelling.

1.4.5 Affinity-guided oxime chemistry

Affinity-Guided OXime (AGOX) chemistry is the most recent catalyst-tethering type of traceless affinity labelling developed within the Hamachi group.[56] Analogous to AGD chemistry in that an acyl donor-oranocatalyst pair are utilised to selectively label proteins, a ligand-conjugated pyridinium oxime (PyOx) catalyst facilitates transfer of an acyl donor label onto the protein of interest (*Figure 9*, a). However, the labelling efficiency of AGOX chemistry is far superior to AGD-mediated labelling at physiological pH (pH 6.5−8.0), due to the ligand-conjugated pyridinium oxime (PyOx) catalyst possessing a lower pKa value than that of its DMAP counterpart (PyOx: 7.0−8.5, DMAP: 9.6). As a result, the corresponding oximate ion exhibits high nucleophilicity at neutral pH. Nevertheless, deprotonation of both oxime and amino acid side chains (such as Lys) facilitated by performing the reaction in basic conditions, contributed to increased rates of acyl transfer. A further advantage over AGD chemistry is the mild electrophilic nature of the *N*-acyl-*N*-alkylsulfonamide acyl donor, which reduces its susceptibility to enzymatic degradation. The reactivity of the acyl donor was optimised by substitution of the aryl sulfonyl (R^1) and *N*-alkyl positions (R^2) with electron withdrawing groups (*Figure 9*, b), enabling sufficient reduction in the electrophilicity of the acyl donor reducing non-productive labelling/decomposition.

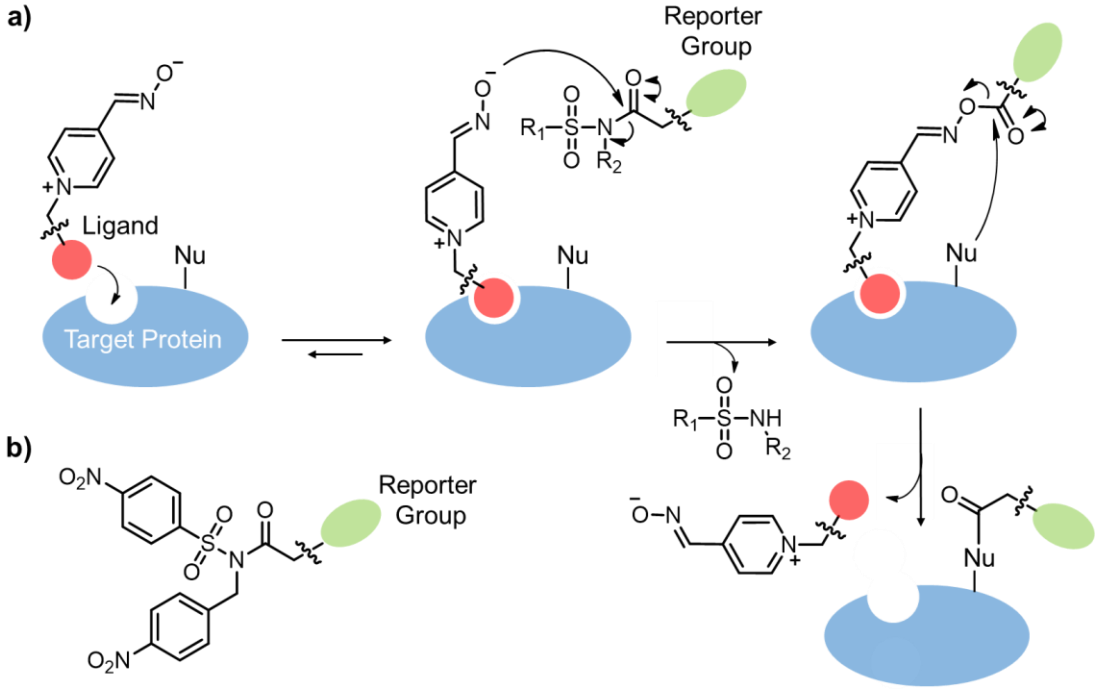


***Figure 9**: a) Schematic representing the mechanism of protein labelling using AGOX chemistry. Reaction of Ser and Lys amino acids with the activated acyl pyridinium group enables transfer of the reporter group onto the target protein. b) Structure of optimised N-acyl-N-alkylsulfonamide acyl donor.*

AGOX chemistry was used to selectively label endogenous membrane-bound proteins (such as CAXII and folate receptor) in live cells, allowing quantification of diffusion coefficients in their native environments. Akin to AGD chemistry, increasing the number of PyOx moieties also resulted in enhanced reaction rates. In addition, selective fluorescent modification of a native neurotransmitter receptor (AMPAR) in mouse brain slices was reported. Yet, this chemistry remains limited to extracellular proteins/binding sites on membrane proteins due in part to the low membrane permeability of the pyridinium cation.

1.4.6 Ligand-directed *N*-sulfonyl pyridone chemistry

Ligand-Directed *N*-Sulfonyl Pyridone (LDSP) chemistry is the most recent type of traceless affinity labelling developed within the Hamachi group.[52] Reaction of the *N*-sulfonyl pyridone reagent with an amino acid on the protein of interest results in sulfonylation of the protein surface (*Figure 10*, a). This chemistry was used to label CAII with a coumarin dye *in vitro*, with a labelling yield of ca. 100% after 12 h, analysed using MALDI-TOF MS (although a fluorescent band appeared after 1 hour). SDS-PAGE analysis revealed that the initial rate of the LDSP reaction (2.9 μM h^{-1}) was nearly 5-fold greater than that of LDT chemistry (0.53 μM h^{-1}) and 2-fold greater than LDAI chemistry (1.4 μM h^{-1}). Proteolytic digestion and tandem MS analysis of fluorescently labelled CAII identified Tyr and Lys residues proximal to the ligand binding site to be sulfonylated. Importantly, the enzymatic activity of the labelled CAII was almost identical to that of the unmodified enzyme, indicating a ‘traceless’ mode of labelling.

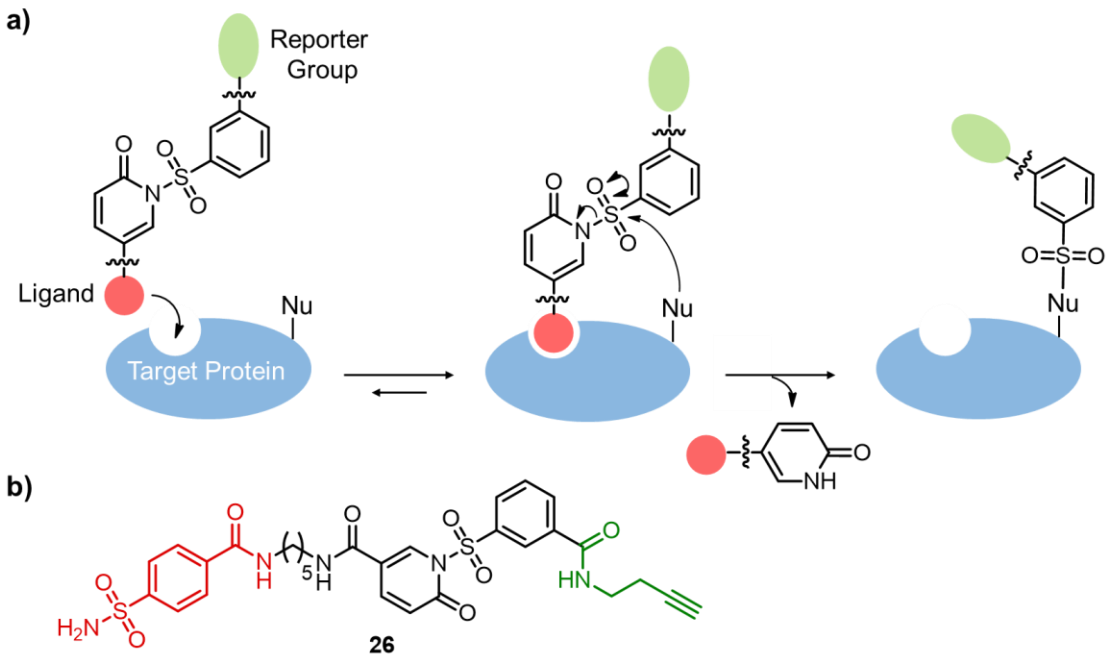


Figure 10: *a) Schematic representing the mechanism of protein labelling using LDSP chemistry. b) Chemical structure of alkyne-containing LDSP reagent designed to label CAII.*

The main advantage of this chemistry over previously reported techniques is its ability to label endogenous proteins in live mammalian cells. Both intracellular (CAII) and cell-surface proteins (CAXII) were labelled in MCF7 cells (a human breast cancer cell line) using a cell-permeable, alkyne-containing LDSP reagent. This allowed a two-step labelling approach whereby biotin-azide was attached *via* a CuACC reaction post-cell lysis, enabling detection of protein labelling using Western blotting. Notably, LSDP reagent **26** gave the most selective labelling of the intracellular protein CAII, compared with the corresponding LDT and LDAI reagents, in addition to the fastest and most conversion to modified protein. By changing the ligand-component of the LDSP reagent, intracellular heat shock protein 90 (HSP90) was also successfully labelled in MCF7 cells.

1.4.7 Comparison of traceless affinity labelling methods

A comparison of the various advantages and limitations of each ligand-directed labelling method developed by Hamachi and co-workers is given below (*Table 1*), along with the reported residues that were modified (identified *via* MS/MS analysis of labelled proteins). These may be referred to before the labelling of a certain protein is attempted using a particular traceless labelling chemistry.

Table 1. *Comparison of the different ligand-directed chemistries used in traceless affinity labelling*

Traceless labelling chemistry	Advantages	Limitations	Amino acid residues labelled
LDT	- Examples of use in live cells and live mice. - Sufficient stability in biological conditions.	- Slow rate of labelling reaction (>10 h). - Low labelling efficiencies.	Tyr, His, Glu, Asp, Cys
AGD	- Efficient rate of labelling reaction (and lower concentrations of reagents required). - Ease of synthesis of labelling reagents.	- Reagent is protonated under physiological pH, reduces labelling efficiency. - Thioester label not catalytic.	Lys, Ser, Tyr
LDAI	- Examples of use on live cell surfaces. - Tuneable reactivity (by altering chemical structure).	- Labile carbamate bond can undergo hydrolysis. - Restricted to labelling of proteins on cell surfaces.	Lys, Ser, Tyr
LDBB	- Labelled intracellular and membrane-bound proteins. - Examples of use in live cells.	- Bulky, hydrophobic reactive group.	Lys, His
AGOX	- Catalytic system. - Shortest labelling times (1 h). - Acyl donor label more stable in biological samples (than that used in AGD).	- Low membrane-permeability of pyridinium cation. - As a result, method has not yet been used to label intracellular proteins.	Lys, Ser
LDSP	- Labelled intracellular and membrane-bound proteins. - Use in live cells. - Cell-permeable reagents (alkyne-tagged). - Fastest rate of reaction compared with LDT and LDAI.	- In early stages of development, as a result only a limited range of proteins have been labelled using this chemistry.	Lys, Tyr

1.4.8 Affinity labelling using transition metal catalysis

1.4.8.1 Recognition-driven labelling

Specific modification of natural proteins has been achieved by combining transition metal catalysis with recognition-driven labelling. Ball and co-workers have reported the use of rhodium(II) metallopeptides to selectively modify side-chains on protein surfaces with functionalised diazo compounds (*Figure 11*).[66] A coiled peptide appended to a rhodium(II) tetracarboxylate catalyst was used to selectively label tryptophan, tyrosine and phenylalanine residues around the coiled-coil interface of interacting peptides. Notably, the relatively weak interaction between the E3/K3 coiled-coil (K_d = 9 µM) was sufficient to template the catalysis, with a >1000-fold increase in reaction rate seen when using the coiled metallopeptide versus the non-coiled variant. Proline-rich rhodium(II) metallopeptides have also been used to modify residues on natural SH3 domains, including Fyn, Src and Lck tyrosine kinases.[67,68] Although this method has been shown to selectively label proteins in cell lysates,[67,69] labelling in whole cells is yet to be reported.

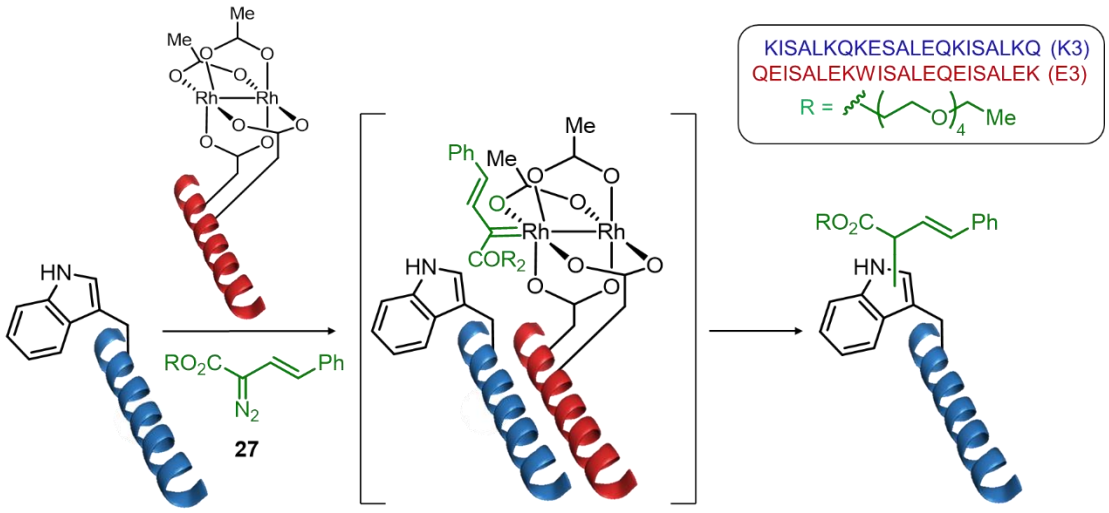


Figure 11: *Schematic illustration of proximity-driven labelling of peptides with diazo compounds using rhodium(II) metallopeptide catalysts, driven by peptide-peptide molecular recognition using the K3/E3 coiled-coil.*

1.4.8.2 Photo-induced transition metal catalysis

Light-activated metal complexes have also been used to label proteins, both *in vitro* and in live cell contexts. In 1999, Kodadek and co-workers reported the chemical crosslinking of native proteins mediated by palladium(II) porphyrins[70] **29** and tris-bipyridyl ruthenium(II) complexes **28**, *via* protein-centred radicals produced upon visible light irradiation.[71] This strategy was utilised by Finn and co-workers to modify tyrosine residues on the capsid proteins of cowpea mosaic virus, with biotin and alkyne reagents, *in vitro*.[72]

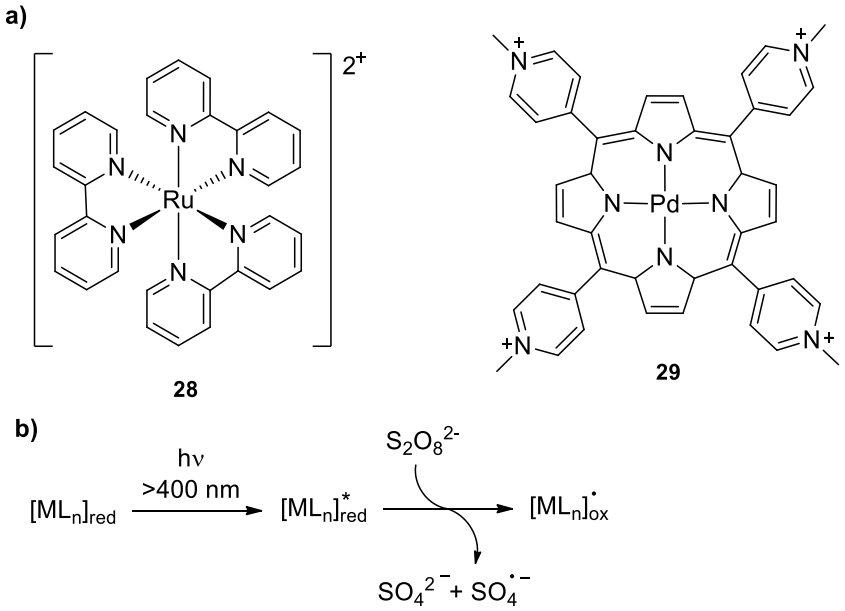
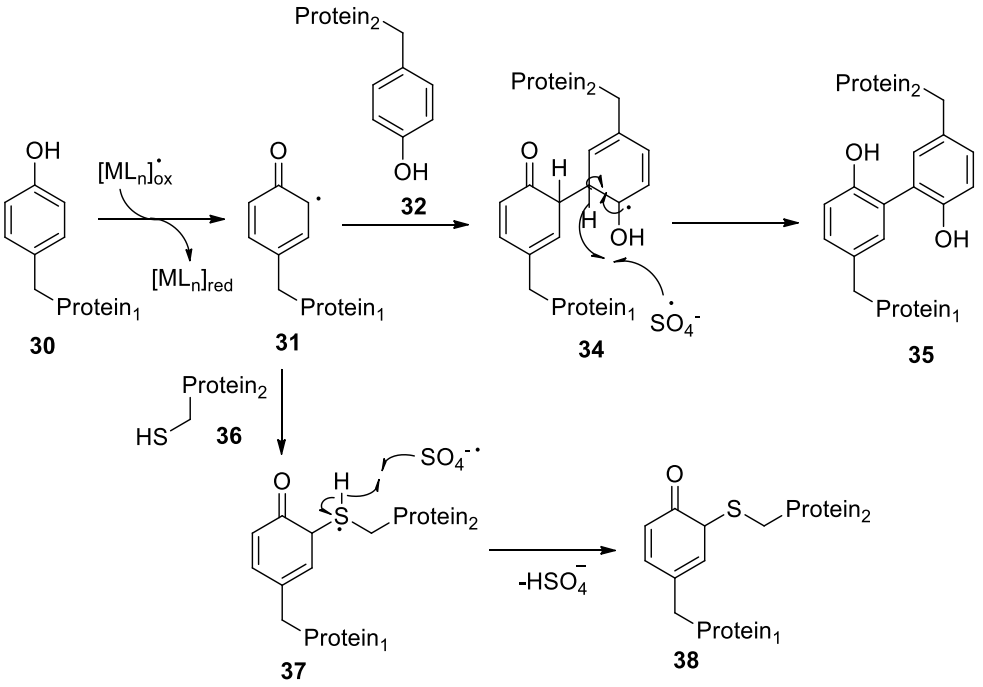


Figure 12: *a) Metal complexes used in catalytic crosslinking, b) mechanism for the photo-oxidation of a general metal complex, [MLₙ]. [MLₙ]red represents the reduced form of the metal complex, [MLₙ]red* denotes the excited state and [MLₙ]ox· indicates the oxidised form.*

A mechanism for this process has been proposed (*Scheme 4*).[71,73] Excitation of the metal complex using visible light irradiation (400-450 nm) is followed by transfer of an electron from the photo-excited state to an electron acceptor (*e.g.* persulfate). The resulting photo-oxidised metal complex abstracts an electron from tyrosine residues on the protein; the resultant tyrosyl radicals can covalently couple to side-chains on interacting proteins. Crosslinking was strongly inhibited by performing the reaction in the presence of excess tyrosine, tryptophan, cysteine and methionine, demonstrating the preferential reactivity towards these residues.[73]



Scheme 4: *Proposed mechanism for the formation of tyrosyl-amino acid crosslinks. Tyrosyl radical intermediate* ***31*** *can react with aromatic and nucleophilic amino acid residues such as tyrosine* ***32*** *and cysteine* ***36****, to form protein crosslinked adducts* ***35*** *and* ***38****, respectively.*

A proximity-driven, metal-mediated method that selectively labels proteins in live cells has been reported by the Nakamura group.[74] A ruthenium(II) tris(2,2'-bipyridyl) photoredox catalyst tethered to a ligand for the protein of interest was used to promote single-electron transfer (SET, *Figure 13*). Photo-oxidation of the metal complex upon irradiation with visible light (400-450 nm) results in an excited state $[Ru(II)(bpy)_3]^*$ complex, which loses an electron to a sacrificial oxidant such as ammonium persulfate or molecular oxygen. The activated metal complex then abstracts an electron from tyrosine residues proximal to the ligand binding site (SET takes place within a short distance *e.g.* a few nm)[75] leading to radical species that can be trapped by electron-rich dimethylaniline derivatives, such as fluorescent or biotinylated labels. However it should be noted that large excesses of both $Ru(II)(bpy)_3$ catalyst and label were used, meaning the catalytic nature of this chemistry has not been exploited.

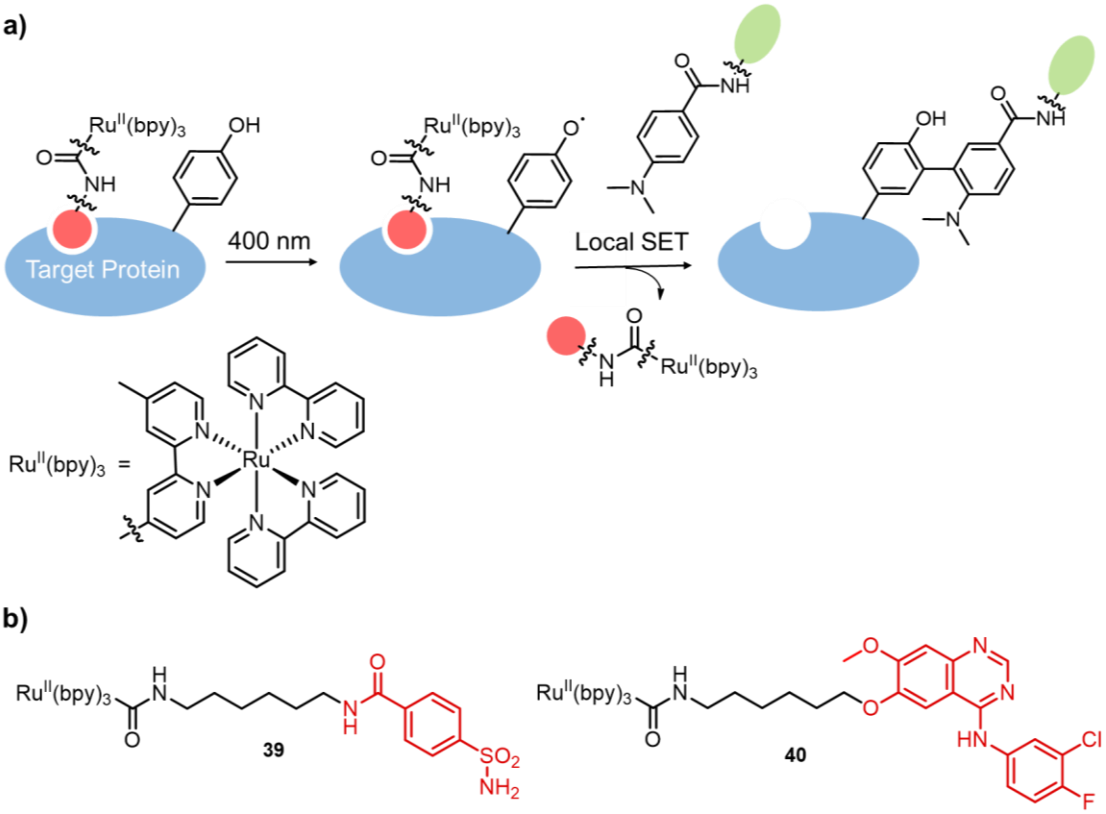


Figure 13: *Ligand-directed labelling mediated by Ru(II)(bpy)$_3$ catalysts. a) Schematic illustration of labelling reaction, which follows an analogous mechanism to that detailed in Scheme 4. b) Chemical structures of ligand-conjugated photocatalysts used for labelling of CAII (**39**) and EGFR (**40**).*

Akin to the labelling chemistries developed by the Hamachi group, this technique results in the 'traceless' labelling of target proteins. This approach has been used to label CAII in cell lysates and intact erythrocytes using benzenesulfonamide-derived complex **39**[74] in addition to modifying epidermal growth factor receptor (EGFR) using gefitinib-tethered [Ru(II)(bpy)$_3$]$^{2+}$ photo-catalyst **40**.[76] The group demonstrated that treatment of A431 cells with **40** and visible light irradiation, could result in either labelling or knockdown of EGFR. Additionally, both protein targeting group and photocatalyst have been immobilised onto affinity beads, allowing purification and labelling of CAII and dihydrofolate reductase (DHFR) from HeLa cell lysates.[77]

1.4.9 Comparison of photoaffinity labelling methods

The information summarised in *Table 2* details the advantages and limitations of the main photo-reactive groups used in affinity labelling described in *sections 1.3* and *1.4*.

Table 2: *Comparison of the most commonly used photo-reactive groups used in affinity labelling*

Photo-reactive group	Advantages	**Limitations**	**Amino acid residues labelled**
Aryl azide	- Easily synthesised. - Highly reactive and indiscriminate nitrene formed. - Photolysis of nitro-substituted derivatives occurs at ~340 nm (not damaging to proteins).	- Unsubstituted derivatives require activation at ~260 nm (damaging to proteins). - Singlet nitrene of non-perfluorinated derivatives prone to ring-expansion.	Insertion into CH, NH and OH bonds
Diazirine	- Highly reactive and “indiscriminate” carbene formed. - Activated at wavelengths not damaging to protein. - Small size (particularly alkyl diazirines).	- Generally low labelling efficiency due to short life-span of carbene/nitrene species in aqueous solution. - Synthesis can be challenging.	Insertion into CH, NH and OH bonds
Benzo-phenone	- Formation of diradical is reversible, leading to high crosslinking yields. - Chemically stable. - Reversible reaction with water.	- Large size. - Preferential reactivity for methionine residues.	Insertion into CH, NH and OH bonds (pref. Met)
Ru(II)$(bpy)_3$ complexes	- Catalytic process. - Activated at visible wavelengths: orthogonal to other photoaffinity groups, not damaging to proteins. - Stability in water (including the reactive intermediate).	- Regioselective reaction of tyrosyl radicals. - Large size.	Tyr, Trp, Cys, Met

1.5 Project aims

The work described in this chapter highlights the value of chemical tools for analysing the interactions of proteins, including determining molecular mechanisms of disease and of drug candidates. In particular, labelling reagents that modify native functionalities on endogenous proteins, in their natural environment (*i.e.* live cells) yield the most valuable information. By using a recognition element for the protein of interest, selective labelling of a target protein within a complex biological environment can be achieved. Additionally, probes activated upon irradiation with a particular wavelength of light allow temporal control of protein crosslinking and capture of non-covalent intermediates; these include protein-ligand interactions (*Figure 14*, a) and transient protein-protein interactions (*Figure 14*, b). Chapters 2 and 3 of this thesis detail the development of affinity-guided, photo-activatable chemical tools for:

a) the identification of the relevant molecular target of a small molecule (as part of a multidisciplinary target identification programme) and for elucidation of the binding site
b) the study of protein-protein interactions formed by BCL-2 family proteins, through labelling mediated by a ruthenium(II) bipyridyl-modified peptide.

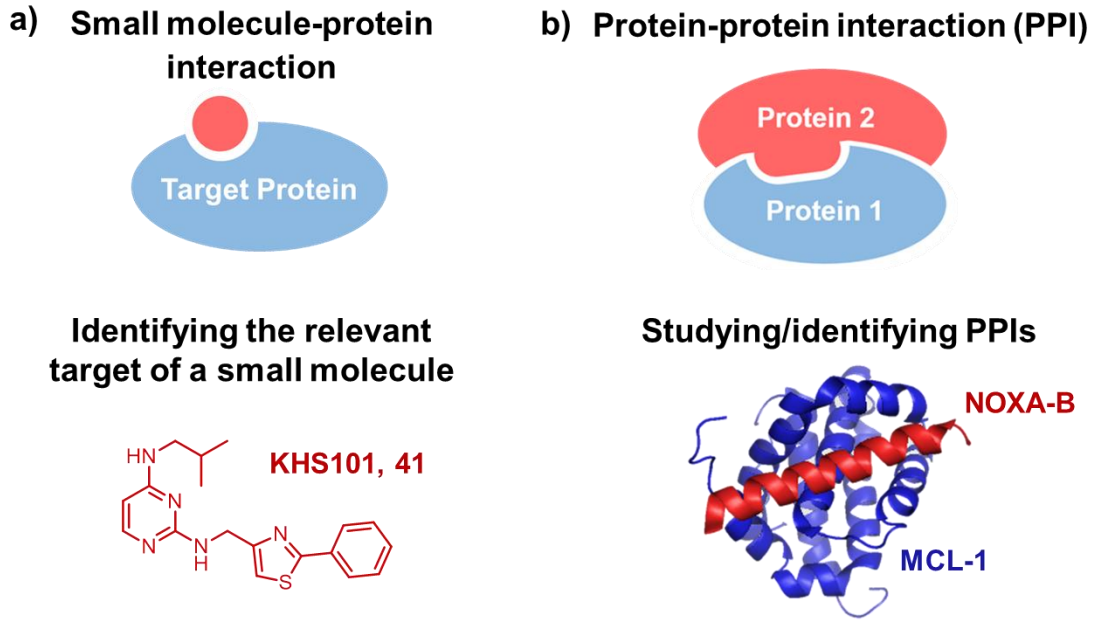


***Figure 14**: Systems studied using affinity-guided chemical tools developed in this work. a) The interaction between the small molecule KHS101 and its molecular target. b) Protein-protein interactions: NOXA-B/MCL-1 as a model system.*

Chapter 2 – Understanding the mechanism of action of the small molecule KHS101 in glioblastoma multiforme cells

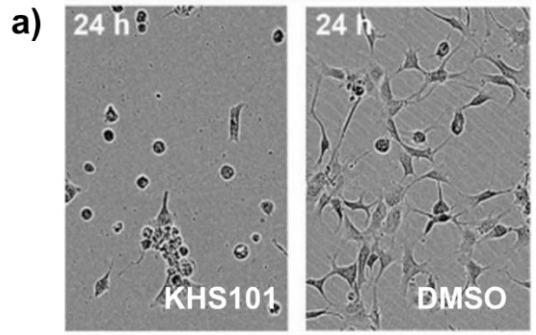
The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Portions of the work presented in this chapter were published as a preprint article in October 2017: 'The small molecule KHS101 induces bioenergetics dysfunction in glioblastoma cells through inhibition of mitochondrial HSPD1', bioRxiv 205203, 2017; https://doi.org/10.1101/205203; Polson, E.S., Kuchler, V.B., Abbosh, C., Ross, E.M., Mathew, R.K., Beard, H.A., Chuntharpursat-Bon, E., Williams, J., Da Silva, B., Shao, H., Patel, A., Davies, A.J., Droop, A., Griffiths, H.B.S., Chumas, P., Short, S.C., Lorger, M., Gestwicki, J., Roberts, L.D., Bon, R.S., Allison, S.J., Zhu, S., Markowetz, F., Wurdak, H. The contributions of the authors to the work described in this chapter *published in the preprint* were as follows: HAB (the candidate) performed all synthetic chemistry, ESP and VBK performed the phenotypic and chaperone activity assays, SZ performed the chemical proteomics and *in vitro* pull-down experiments. ESP, VBK, RKM, and HW prepared the manuscript and figures. RSB, JG, SJA, SZ, FM, and HW provided project leadership. The contributions of researchers that did not contribute to work detailed in the preprint are indicated throughout the chapter.

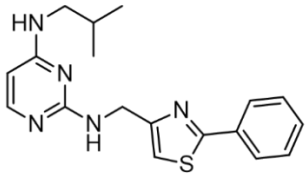
2.1 Introduction

Glioblastoma multiforme (GBM) is the most common and malignant form of brain cancer among adults.[78] Recurrence follows on average 7 months after primary treatment, with patients having a poor median survival of 12 - 15 months following diagnosis.[79] These statistics clearly highlight the lack of effective treatments for this devastating disease. Commonly, surgery to remove the major area of the tumour is followed by chemotherapy and radiotherapy, however problems with resistance and general cytotoxicity hinder their success.[80,81] Although the development of novel therapies for GBM remains challenging, a large body of evidence supports the notion that targeting stem-cell like features within tumours (*e.g.* the ability for the stem-like GBM cell population to self-renew) may reduce tumour progression.[78,82] In addition, it is recognised that stem-like GBM cells within tumours do not respond adequately to current chemotherapeutics.[83,82]

The brain-penetrable small molecule KHS101 **41** (*Figure 15*, b) targets neural stem cell differentiation *in vivo*, inducing the differentiation of rodent neural progenitor cells (NPCs) into mature neurons.[84] Transforming acidic coiled-coil containing protein 3 (TACC3),[85] a key regulator of cell division, was identified as the cellular target of KHS101 in NPCs (*via* photocrosslinking and RNAi knockdown experiments).[84,86] Due to the molecular similarities between NPC and stem-like GBM cells,[83] it was hypothesised that KHS101 may act as a potential modulator of GBMs – either by inducing differentiation into cells more vulnerable to therapy or by reducing their rate of growth. For instance, expression of TACC3 correlates positively with GBM disease severity and correlates negatively with glioma patient survival,[87] while KHS101 is known to induce degradation of TACC3 in NPCs.[86]

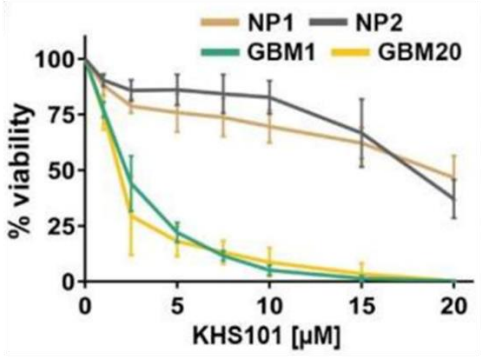


b)



41, KHS101

c)



GBM line	IC_{50} value
GBM1	2.23 (2.02;2.47)
GBM4	0.97 (0.90;1.04)
GBM11	5.05 (4.51;5.64)
GBM13	3.28 (2.77;3.89)
GBM14	1.98 (1.67;2.33)
GBM20	1.87 (1.58;2.21)
NP1	> 20
NP2	18.24 (14.83;22.43)
U87	1.20 (1.13;1.28)*

***Figure 15**: KHS101 selectively induces apoptosis of GBM cells. a) Still frames of live cell imaging of GBM1 cells treated with KHS101 (7.5 µM), or DMSO (0.1%) after 24 hours. b) Chemical structure of KHS101. c) Left, dose-response curve for cell viability (%) 5 days after KHS101 treatment (7.5 µM) for GBM and non-malignant NP cell lines. Right, IC_{50} values for KHS101-dependent cell viability for GBM and NP lines after 5 days of treatment (U87 values were obtained 2 days after treatment). Data are mean ± SD of 3 biological replicates, *, P<0.05; **, P<0.01 (95% confidence intervals in brackets).*

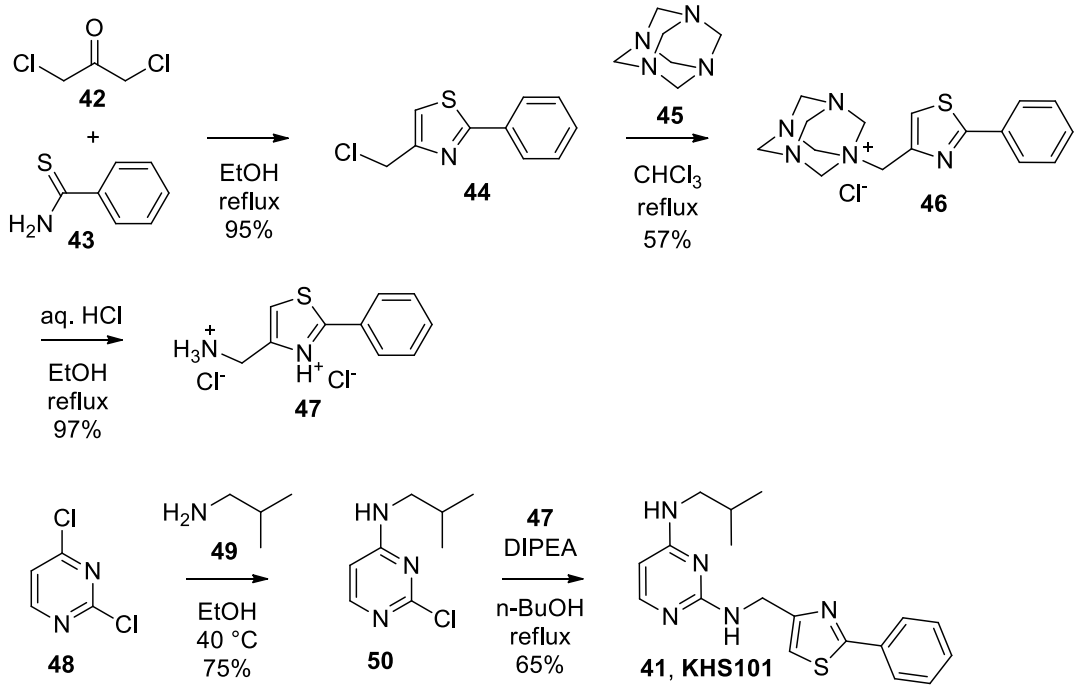
However, instead of the differentiation phenotype observed in NPCs, stem-like GBM1 cells treated with KHS101 (7.5 µM) showed signs of cell death through apoptosis, characterised by cell rounding (*Figure 15*, a, experiments performed by Dr Verena Kuchler at the Leeds Institute of Cancer and Pathology, LICAP). This effect was seen selectively for GBM cells over non-malignant brain cells (lines NP1 and NP2, derived from adult brain cells taken from epilepsy patients, *Figure 15*, c). Moreover, dose-response relationships were observed for a range of patient-derived GBM cell lines, including models derived from primary GBM (GBM1, 4 and 13), recurrent tumours (GBM20) and rare GBM subtypes (GBM11, GBM14). The broad-spectrum GBM cell activity of KHS101 is significant, due to the extensive cellular and molecular heterogeneity within and/or between GBM tumours (hence “multiforme”), which poses a large obstacle to the design of effective therapies.[82,88] The commonly used GBM cell line U87 was also tested in cell viability assays (*Figure 15*, c, table), however recent studies have revealed that this cell line is no longer a true representative of the tumour of origin,[89] highlighting the importance of establishing well-characterised cell lines used for biological research (such as those given in *Figure 15*, c, table).

Having a greater understanding of the molecular mechanisms underlying this specific cell death phenotype, including the biologically relevant target of KHS101 in GBM cells, could allow a novel therapeutic strategy for GBM to be uncovered. Therefore, scale-up of the synthetic route to KHS101 was necessary to obtain multiple grams of analytically pure KHS101 for testing in cells and *in vivo*. In addition, an inactive analogue of KHS101 was required. During the course of the project, the need for the synthesis of affinity probes for target identification and validation studies also became apparent. Together, these compounds proved to be valuable chemical tools for investigating the mechanism of action of KHS101 in GBM cells.

2.2 Synthesis of KHS101

2.2.1 Optimisation of synthetic route for scale-up

Optimisation of the synthetic route to KHS101 **41** was achieved through combination and modification of published procedures (*Scheme 5*).[84,90] Thiazole **44** was prepared *via* condensation of thiobenzamide and 1,3-dichloroacteone in a Hantzsch thiazole synthesis.[91] The procedure by Awoussong *et al.*[90] initially gave **47** as the partial hydrochloride salt (a mixture of solid and oil) and required a basic work-up to isolate **47** in a homogenous form. A Delépine reaction[92] was used to convert chloride **44** to the amine, *via* intermediate **46**. This method was chosen over synthesis *via* the *N*-alkylphthalimide, due to the solubility of starting materials **44** and **45** in chloroform, while urotropinium salt **46** remains insoluble (removing the need for purification). Hydrolysis of the quaternary ammonium salt under acidic conditions gave primary amine **47** as the dihydrochloride salt, after basic workup followed by treatment with HCl in ether. Pyrimidine **50** was prepared as described previously[84] *via* preferential S_NAr displacement of the 4-chloride in **48** using isobutylamine. Intermediates **50** and **47** were combined using a second S_NAr reaction, in the presence of DIPEA, affording final compound **41**.



***Scheme 5**: Optimised synthetic route to KHS101 via modification of procedures by Awoussong et al.[90] and Wurdak et al.[84] Synthetic route to intermediate **47** allows material to be obtained on gram-scale. Purification of final compound **41** using reverse-phase column chromatography gives product in high purity (>98%).*

A key obstacle faced when following the synthesis described by Wurdak *et al*.[84] was the unstable nature of amine **47** when used as the free base. By isolating **47** as the dihydrochloride salt, this intermediate could be obtained on gram-scale and stored for longer periods of time. In addition, purification of **41** using reverse-phase column chromatography was crucial for the removal of a side product (**51**) formed during the final step (*Figure 16,* a). This side product is absent from KHS101 purified by reverse phase chromatography, which allowed **41** to be obtained in adequate purity for biological testing (HPLC results suggest purity of >98%). Crystallisation of KHS101 from hot acetonitrile was also possible after reverse-phase purification.

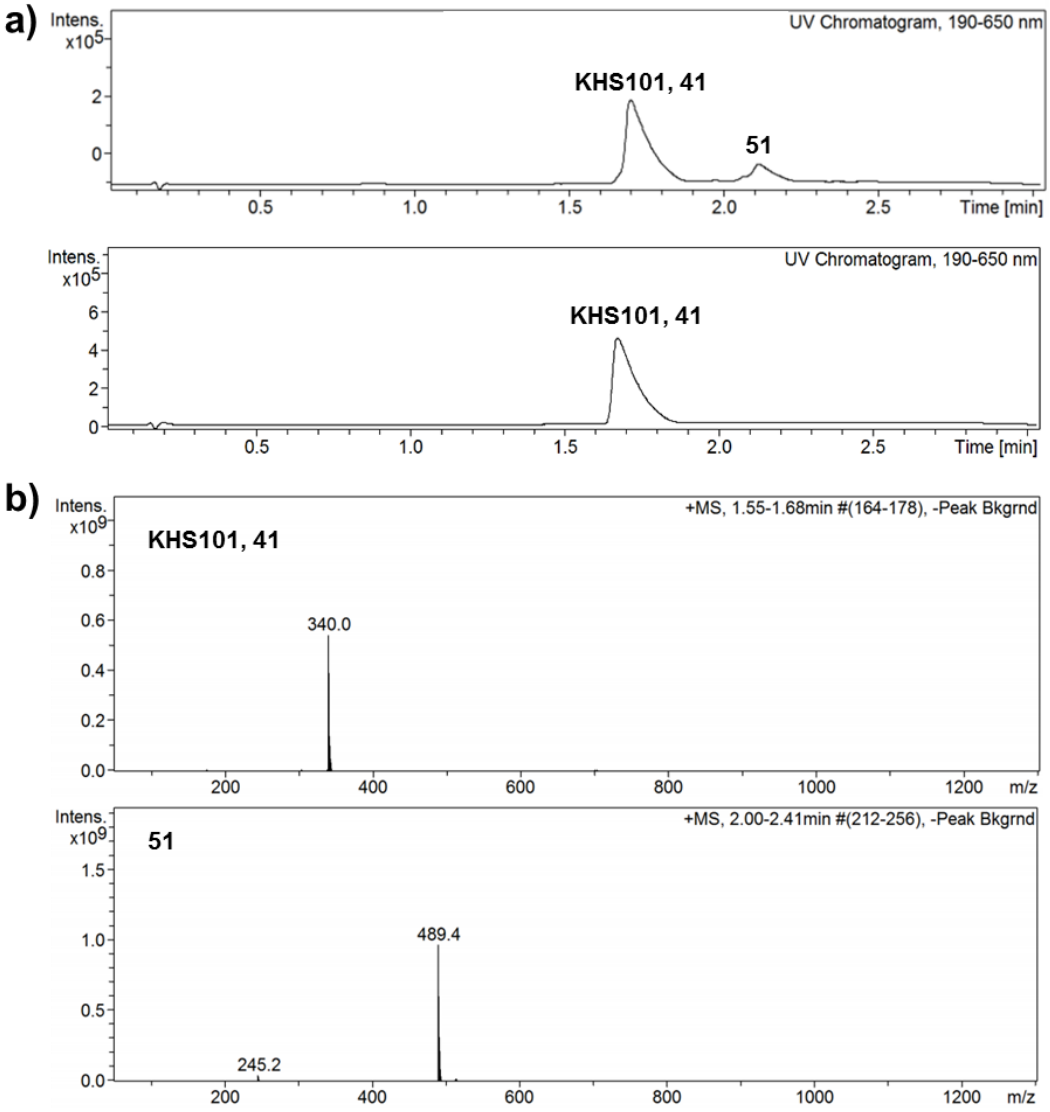


Figure 16: *a) UV traces from LC-MS of KHS101 purified by silica flash column chromatography (top) and reverse-phase column chromatography (bottom, optimised method). b) Mass spectra (LC-MS) of KHS101* ***41*** *(top) and side product* ***51*** *(bottom).*

2.2.2 Identification and structural determination of major side product

Investigation into the identity of the side product formed during the final step of the KHS101 synthesis was then undertaken. The LC-MS trace of the product purified as described by Wurdak *et al*.,[84] exhibited a mass corresponding to a product resulting from the double-addition of pyrimidine moiety **50** (*Figure 16,* b). At least two possible regioisomers of this side product may exist, the most likely of which are compounds in which the additional isobutyl-pyrimidine group is attached to one of the two secondary amines in KHS101 **41** (**51a** and **51b**, *Figure 17*). Deliberate synthesis of the side product was required in order to obtain a sample of **51** to determine its structure.

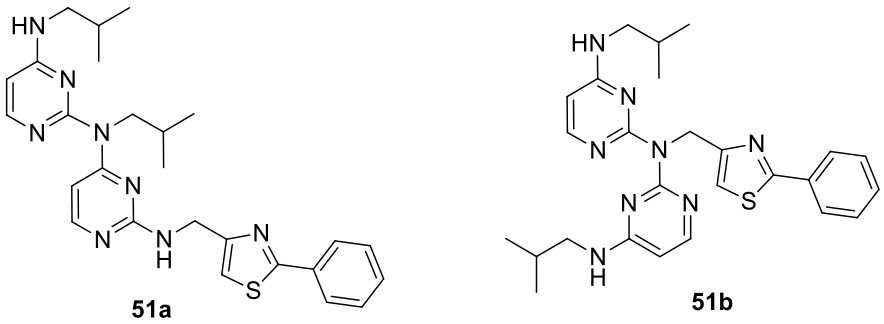
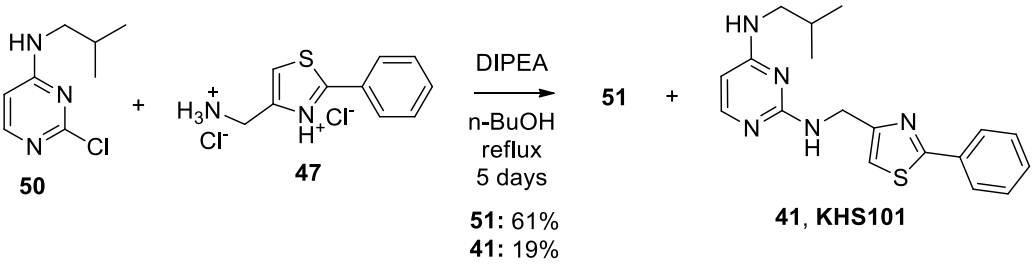


Figure 17: *Possible structures of side product **51** formed in the synthesis of KHS101 **41**, resulting from the double-addition of pyrimidine **50**.*

It was initially hypothesised that reaction of KHS101 **41** with two equivalents of pyrimidine **50**, would yield significant amounts of side product **51**. However, **41** was isolated as the major species from the reaction in 89% yield. The greatest conversion to di-substituted product **51** was observed when amine **47** was heated under reflux with an excess of pyrimidine **50** over 5 days (*Scheme 6*); **51** was obtained in 61% yield after reverse-phase purification.



Scheme 6: *Synthesis of side product **51**. 2 equivalents of pyrimidine **50** were allowed to react with 1 equivalent of amine **47**. Reflux over 5 days gave **51** in 61% yield.*

The structure of **51** was then determined using NMR. In the proton NMR spectrum of **51** (methanol-d_4), two separate environments for the isobutyl and pyrimidine protons were observed, suggesting the formation of asymmetric regioisomer **51a**. In addition, two separate N-H environments were seen in the proton NMR of **51** at 328 K in deuterated chloroform (in comparison to regioisomer **51b**, which would exhibit only one N-H environment). Consistent with this proposed structure, coupling between one N-H signal of **51** and the thiazole CH_2 (*Figure 18*, Ha) was observed in the COSY spectrum at 328 K, while coupling was also seen between a second N-H and the CH_2 of the external isobutyl group (*Figure 18*, Hb). Finally, separate peaks were observed for the carbon atoms on each pyrimidine ring, further confirming the asymmetric product **51a**. Side product **51a** was investigated as a potential PPI inhibitor, due to its structure resembling α-helix mimetics of HIF-1α (see *Figure 53*, Appendix I).

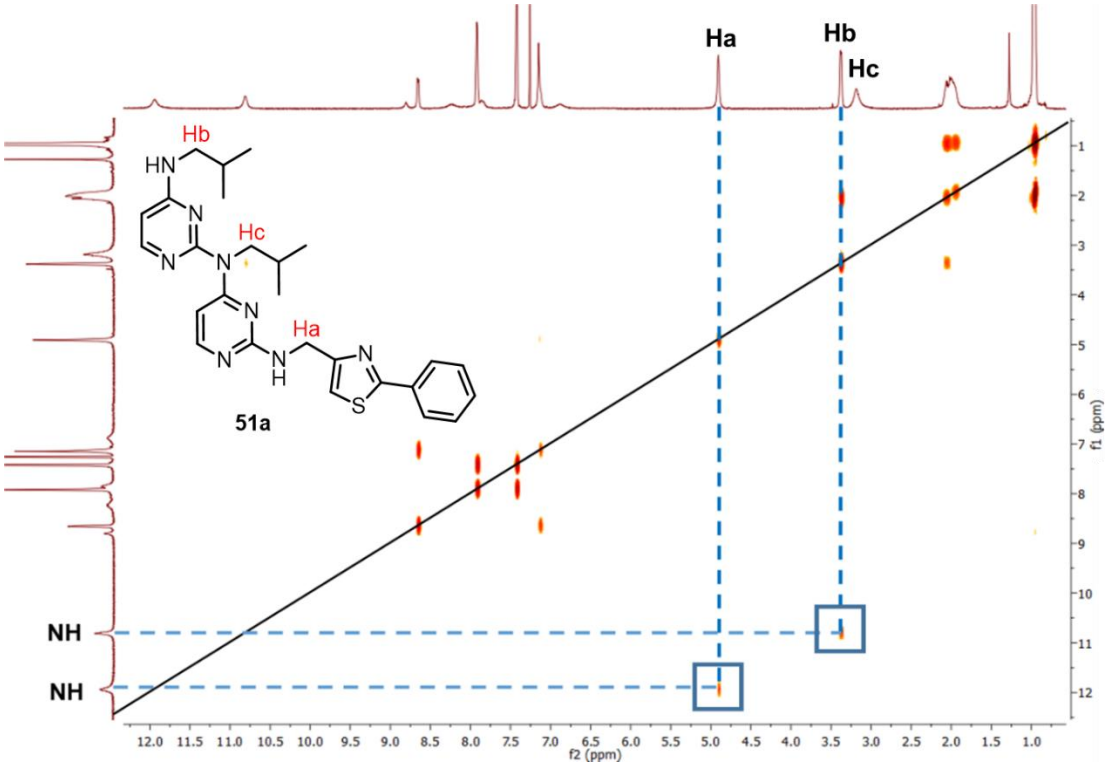
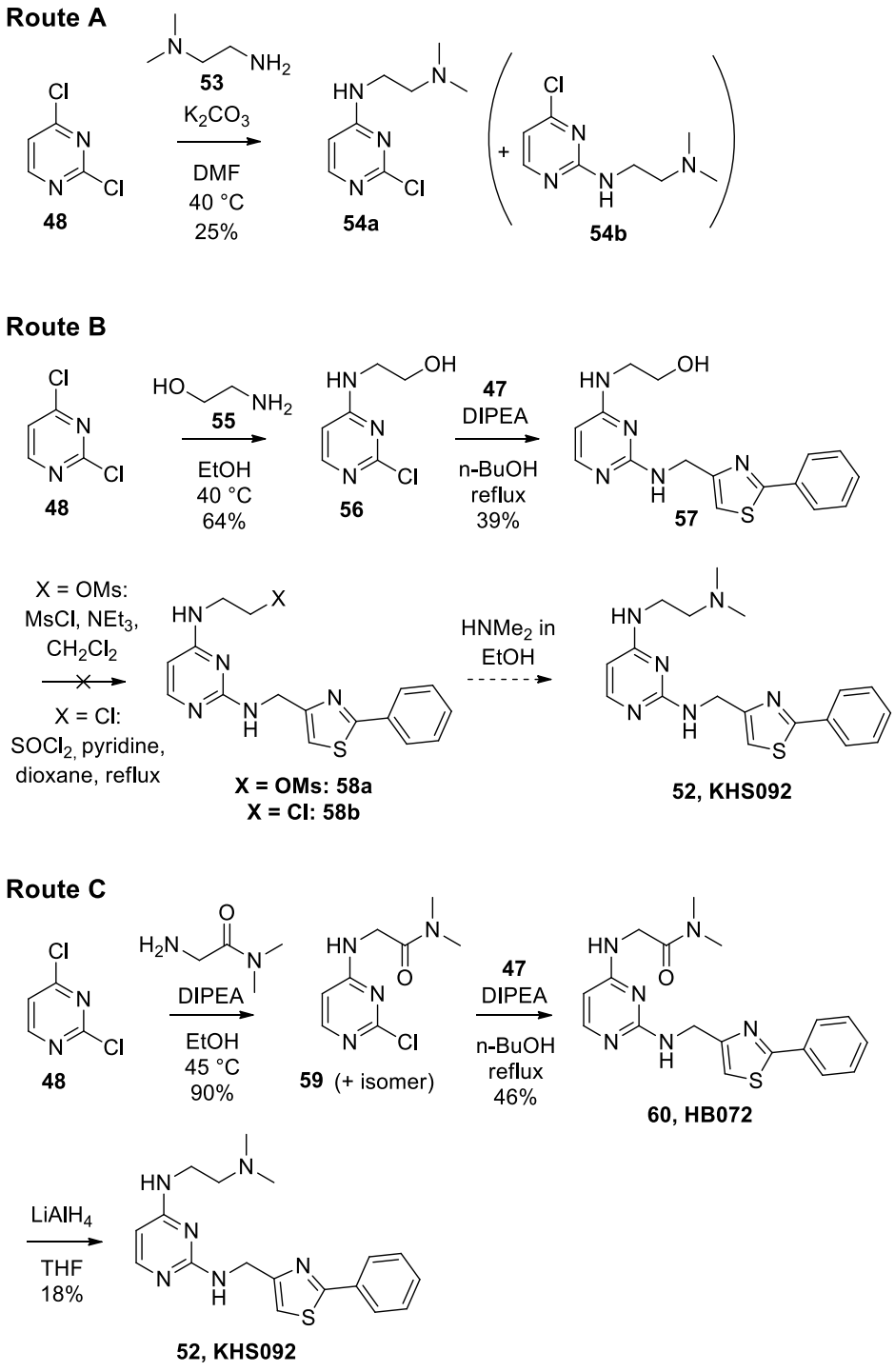


Figure 18: *COSY spectrum of* ***51*** *showing coupling of protons Ha and Hb to separate N-H environments (328 K, $CDCl_3$), confirming the asymmetric structure* ***51a****. Experiments were carried out at elevated temperatures due to the presence of rotamers (observed when NMR experiments were carried out at ambient temperature).*

2.3 Synthesis of analogues of KHS101

Comparison of the structure-activity relationship (SAR) of KHS101 in GBM cells with that seen in NPCs could give some insight into whether similarities exist between the mechanisms of action of KHS101 in both cell types. Therefore, an analogue of KHS101 that does not cause differentiation of NPCs[84] (KHS092 **52**) was synthesised, for testing in the new cell type (GBMs). As no synthesis of KHS092 was published, a new synthetic route to **52** had to be designed. The initial route to **52** (Route A, *Scheme 7*), involving direct S_NAr of *N,N*-dimethylaminoethylamine and pyrimidine **48** (work carried out by Justine Raymond), led to a mixture of regioisomers that could not be separated. The next strategy involved synthesis of alcohol **57**, functionalisation with a good leaving group (*via* intermediates **58a** or **58b**, Route B, *Scheme 7*) followed by displacement with dimethylamine. However, both mesylation and chlorination of alcohol **57** proved unsuccessful (possibly due to the instability of these intermediates). As with the initial approach, S_NAr of 2-amino-*N,N*-dimethylacetamide and pyrimidine **48** also gave inseparable isomers (Route C, *Scheme 7*), even after multiple re-crystallisations (work carried out by Dr Robin Bon). However, following reaction of **59** with amine **47**, isomers of **60** could be separated at this stage.

Reduction of amide **60** to the tertiary amine, using lithium aluminium hydride, successfully afforded KHS092 **52**.



Scheme 7: *Synthetic routes to KHS092* ***52*** *undertaken. Route A: Attempted synthesis of KHS092 via amine* ***54a*** *(carried out by Justine Raymond, ERASMUS student supervised by Hester Beard). Regioisomers of* ***54*** *proved inseparable. Route B: Attempted synthesis of KHS092 via mesylate* ***58a*** *and chloride* ***58b*** *derivatives (intermediate* ***56*** *was synthesised by Justine Raymond). Route C: Successful synthesis of KHS092 via reduction of amide* ***60*** *(intermediate* ***59*** *was synthesised by Dr Robin Bon).*

2.4 Biological testing of KHS101, KHS092 and HB072

Biological testing of KHS101, KHS092 and HB072 was carried out by Dr Verena Kuchler and Dr Euan Polson at LICAP. GBM1 cells treated with KHS101 (7.5 µM) showed a distinct phenotype of cell death through apoptosis, categorised by a >5-fold increase in caspase 3/7 activation (*Figure 19*, a). Electron microscopy (EM) images of cells treated with KHS101 revealed the formation of intracellular vacuoles, with magnification of these images highlighting their double-membranes (*Figure 19*, b). These double-membraned vacuoles are indicative of autophagy, the mechanism by which the cell degrades and recycles unwanted cellular components.[93] Importantly, these effects were not seen in non-malignant NP1 cells treated with KHS101, highlighting the selectivity of the small molecule for GBM cells. KHS092 **52**, which was inactive in the NPC differentiation assays,[84] did not exhibit complete inactivity in the GBM phenotypic assay (data not shown), suggesting engagement of different cellular targets in the NPC and GBM cell types. As precursor compound **60** (HB072) did not elicit the KHS101-induced phenotype in GBM cells, **60** was selected as the inactive control for future assays (see *section 2.6.3.2*).

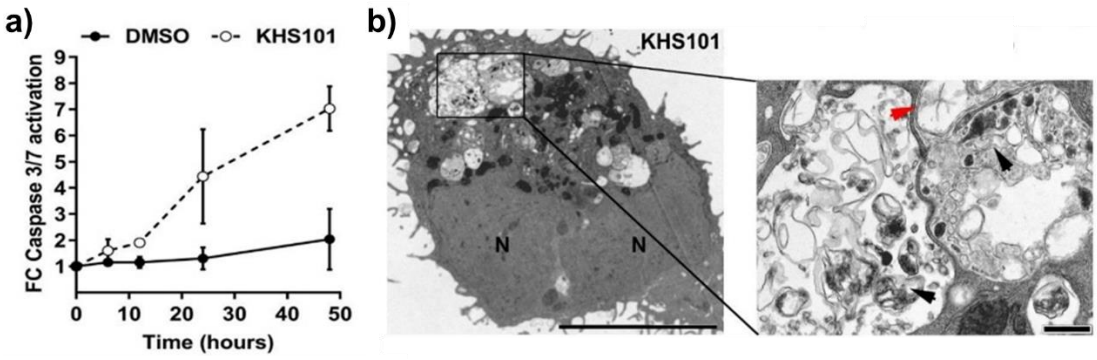


Figure 19: *KHS101 induces the self-destruction of GBM cells through hyper-vacuolisation. a) Caspase 3/7 activation in GBM1 cells treated with KHS101 (7.5 µM) is significantly increased compared to the DMSO control (data normalised to t_0). b) EM images of GBM1 cells treated with KHS101 (7.5 µM) after 12 hours show presence of double-membraned autophagosomes (red arrow). 'N' indicates nuclei. Scale bars 5 µm, 500 nm.*

Gene set enrichment analysis (a method which identifies groups of genes that are significantly enriched or depleted under certain conditions)[94] of DMSO- and KHS101-treated cells (7.5 µM), highlighted perturbation of genes encoding proteins involved in metabolic pathways such as the tricarboxylic acid (TCA) cycle and oxidative phosphorylation (OXPHOS). As the additional effects on metabolism were only seen in GBM1 cells treated with KHS101, it is unlikely that **41** shares the same target in both NPCs and GBMs. Furthermore, although shRNA mediated-knockdown of TACC3 reduced GBM cell proliferation, the autophagic vacuolisation phenotype

was not reproduced. Notably, the autophagic response was observed in KHS101-treated GBM cells before the levels of TACC3 were reduced by both RNAi and KHS101 (minutes vs. hours), suggesting a TACC3-independent mechanism of action. Various possible targets relating to mitochondrial metabolism were investigated, but none could be confirmed. As a result, a clear need arose for affinity-based probes to identify the protein target of KHS101 involved in the metabolic phenotype seen in GBM cells.

2.5 Design and synthesis of KHS101 target identification probes

2.5.1 Design of KHS101 target identification probes

An affinity-based target identification strategy using a photo-reactive KHS101 derivative was proposed, which would allow potential cellular target proteins of KHS101 to be identified (*Figure 20*, a). This approach required an analogue of KHS101 incorporating a reactive group and reporter group (see Chapter 1). Typically, extensive SAR studies are carried out to determine locations at the small molecule where substitution with such groups will be tolerated. Previous SAR studies indicated that modification at the *para*- position on the phenyl ring was tolerated (in NPCs).[84] Therefore, a previously described probe[84] (used to identify TACC3 as the cellular target of KHS101 in NPCs) was tested first – despite the activity of **61** in GBM cells not yet having been confirmed. KHS101-BP **61** (*Figure 20*, a) comprises a 'ligand' component (KHS101, red), which interacts with binding proteins, linked to a photo-reactive group (enabling covalent bond formation between KHS101 and the protein upon UV irradiation, purple). An alkyne handle (blue) enables the addition of a reporter group such as biotin *via* a click reaction, allowing protein identification.

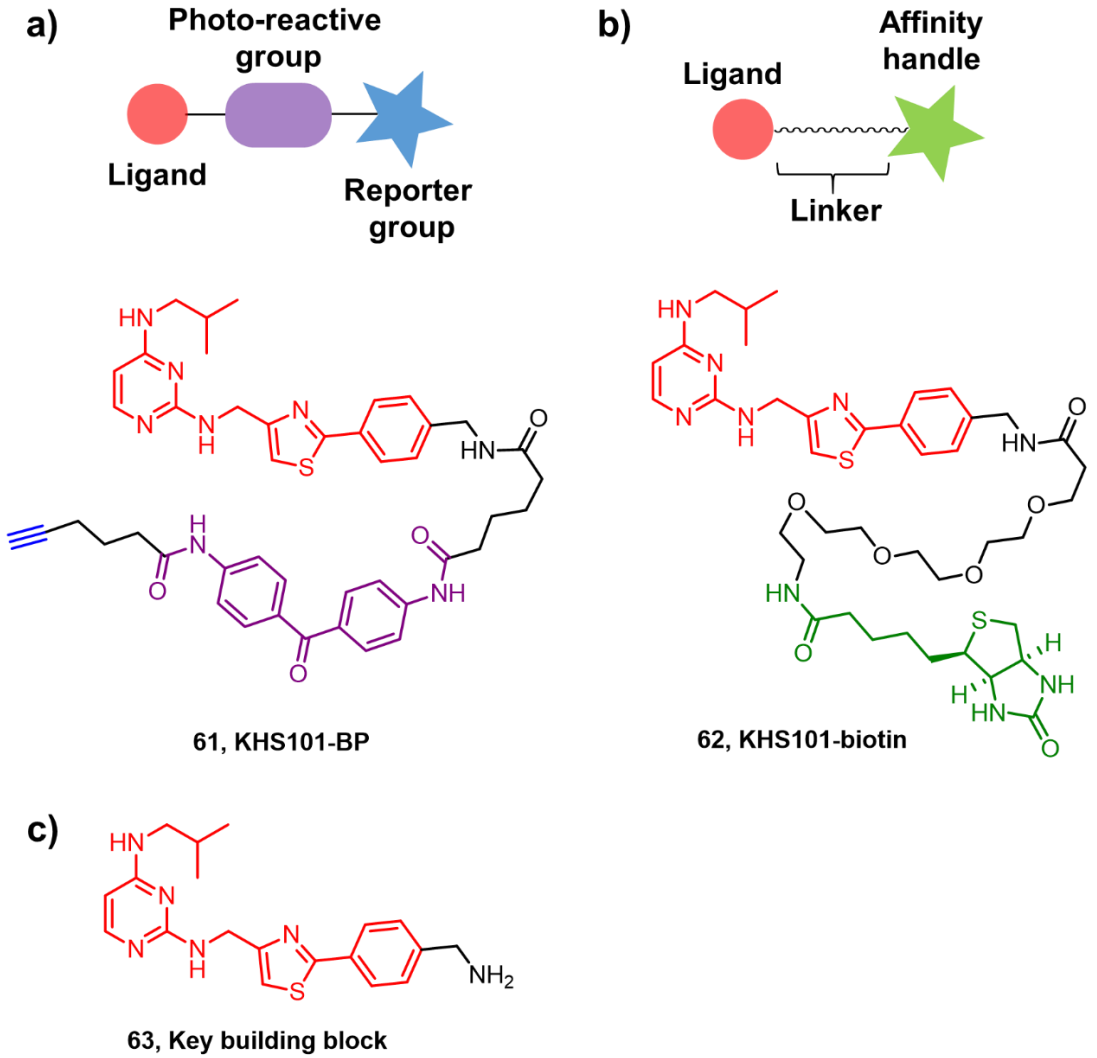
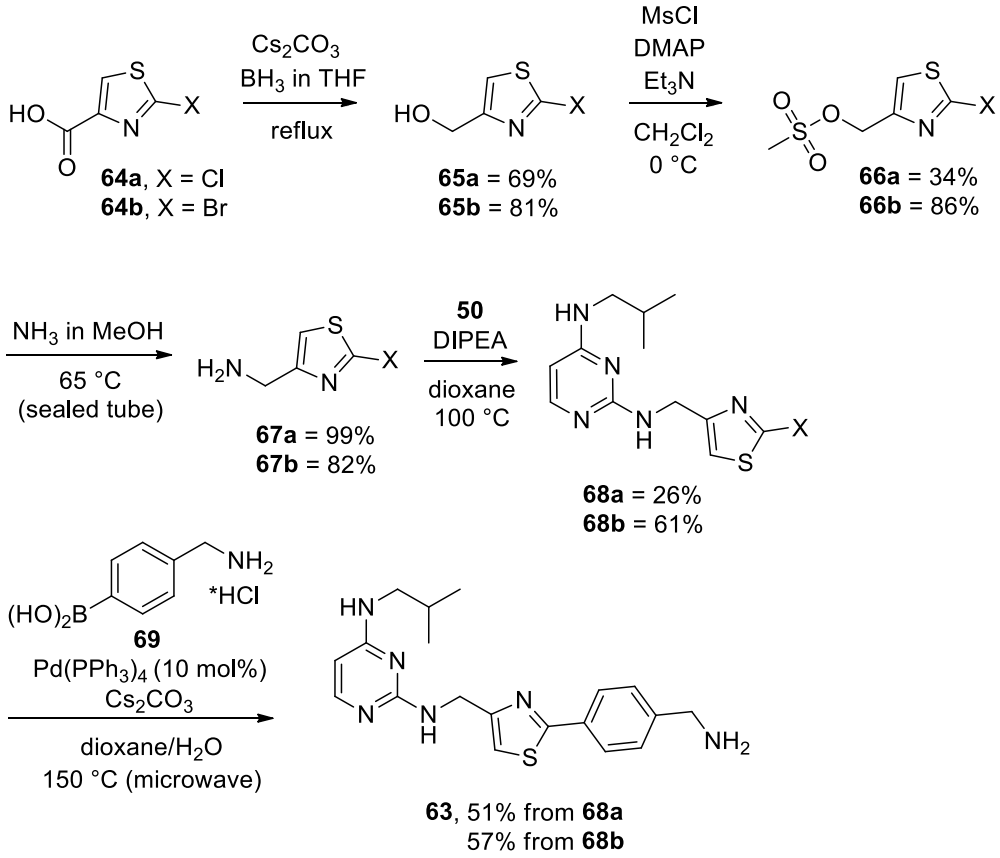


Figure 20: *KHS101-target identification probes. a) Top: general structure of photo-reactive probe used in photocrosslinking experiments, bottom: chemical structure of the benzophenone probe used in this work, KHS101-BP* ***61****. b) Top: general structure of affinity probe for pull-down experiments: chemical structure of affinity probe used in this work, KHS101-biotin* ***62****. c) Chemical structure of key building block* ***63*** *used to synthesise KHS101-BP* ***61*** *and KHS101-biotin* ***62****.*

In conjunction with photo-reactive probe KHS101-BP **61**, affinity probe KHS101-biotin **62** (*Figure 20*, b) can be used in pull-down experiments to validate possible KHS101-protein interactions identified through photocrosslinking. The design of KHS101-biotin was inspired by that described previously,[84] with a modification to the linker to allow synthesis directly from key building block **63** (*Figure 20*). So that KHS101-BP **61** and KHS101-biotin **62** could be synthesised rapidly, a divergent synthetic approach through key intermediate **63** was followed.

2.5.2 Synthesis of key building block for KHS101 target identification probes

In order to obtain a sufficient quantity of key building block **63** to synthesise functional probes **61** and **62**, a robust synthetic route was required. Initially, problems were met when using the synthesis published by Wurdak *et al*.,[84] which started from 2-chlorothiazole-4-carboxylic acid **64a** and resulted in low yields for the mesylation and final two steps (*Scheme 8*). It was proposed that replacing the chloride for a bromide, thus starting from 2-bromothiazole-4-carboxylic acid **64b** (*Scheme 8*), would not only decrease the susceptibility of the halide in species **66** and **67** to S_NAr substitution, but also increase the reactivity of pyrimidine **68** in the Suzuki coupling. As a result of this modification, significantly higher yields were achieved, resulting in a >6-fold increase in overall yield (from 3% to 20%).

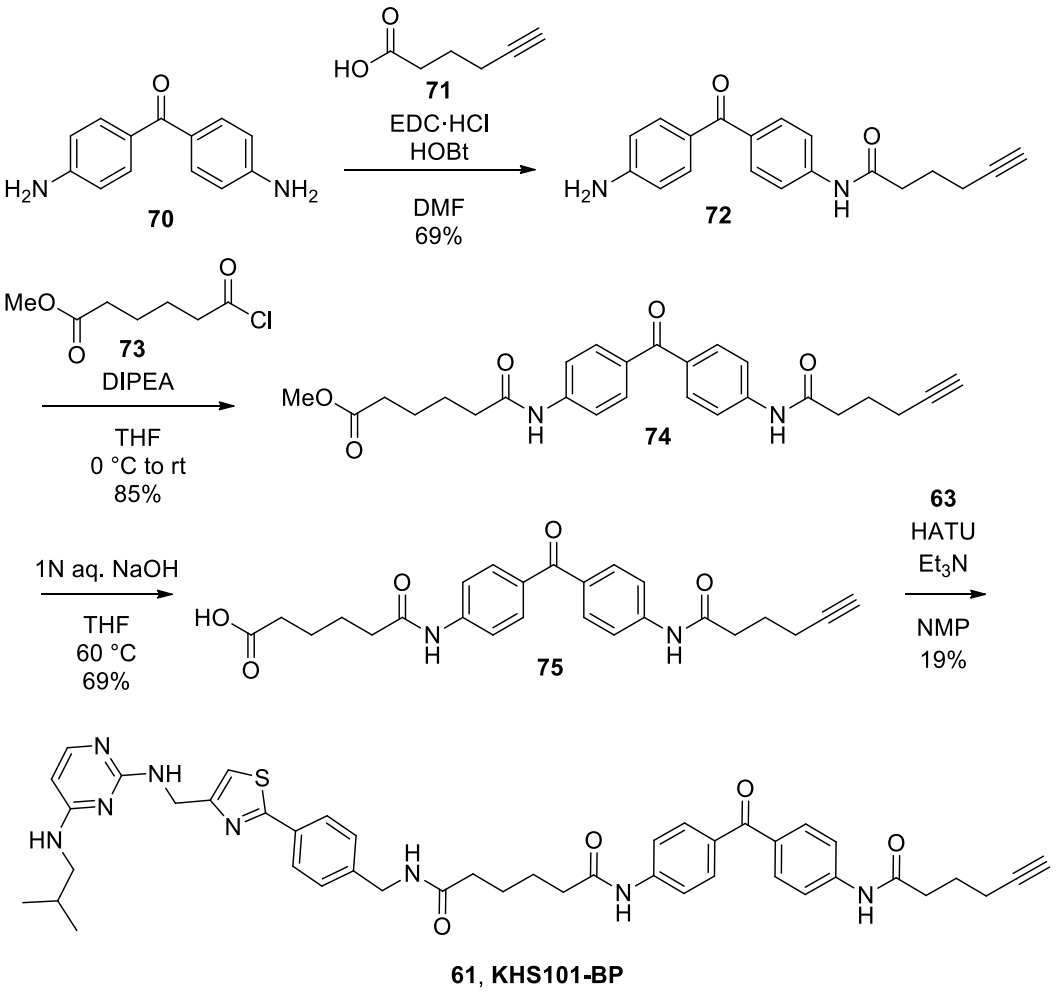


***Scheme 8**: Comparison of synthetic routes to key building block **63**. The route that starts from 2-bromothiazole-4-carboxylic acid **64b** gives significantly higher yields than that starting from 2-chlorothiazole-4-carboxylic acid **64a**.*[84]

Reduction of acid **64** to alcohol **65** using borane in THF, and subsequent mesylation gave **66**. The addition of the nucleophilic catalyst DMAP in the mesylation step greatly improved yields (from 34% for **66a** to 86% for **66b**). Mesylate **66** was then heated in ammonia-saturated methanol in a pressure vessel over 3 days, to afford amine **67**, which underwent S_NAr reaction with pyrimidine **50** to give **68**. Finally, boronic acid **69** was combined with pyrimidine **68** in a Suzuki coupling to afford key building block **63**.

2.5.3 Synthesis of KHS101-benzophenone

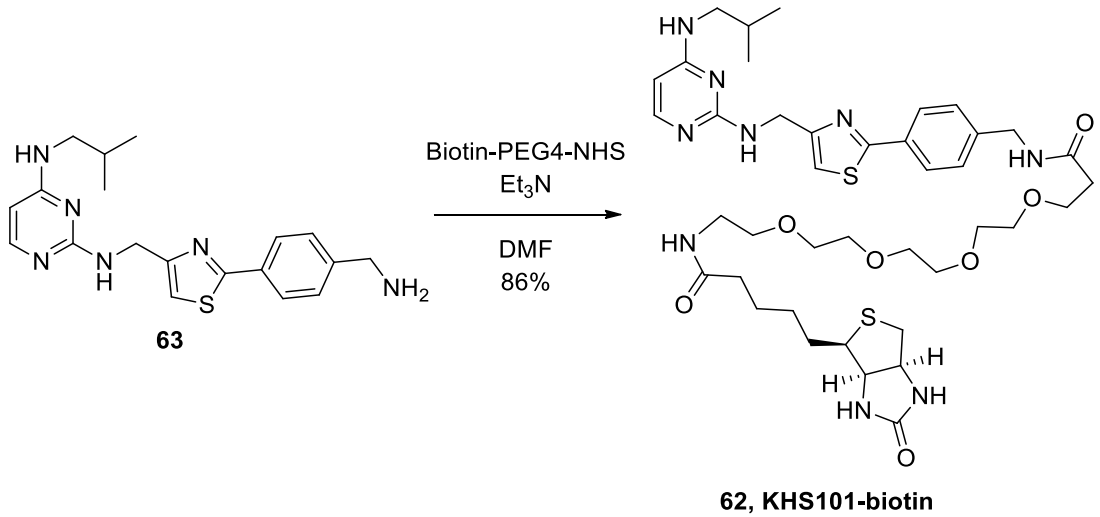
The benzophenone crosslinking moiety of KHS101-BP **61** was synthesised *via* adaptation of a procedure by Cravatt and Salisbury (*Scheme 9*).[95] Coupling of diaminobenzophenone **70** with 5-hexynoic acid **71** enabled incorporation of the alkyne handle in **72** for biotin-labeling *via* a click reaction. Coupling of the remaining amine in **72** with methyl adipoyl chloride under basic conditions, and subsequent hydrolysis of the methyl ester, functionalised **75** with a linker and carboxylic acid by which the KHS101 component **63** could be attached. Finally, key building block **63** and benzophenone **75** were coupled using HATU in the presence of triethylamine (*Scheme 9*) to afford KHS101-BP **61** (purity >98% by analytical HPLC).



Scheme 9: *Synthetic route to KHS101-benzophenone* ***61*** *via benzophenone crosslinker* ***75*** *(which was prepared through adapting a procedure by Salisbury and Cravatt).*[95]

2.5.4 Synthesis of KHS101-biotin

KHS101-biotin **62** was prepared through coupling of key building block **63** and commercially available biotin-PEG4-NHS under basic conditions, followed by reverse-phase HPLC purification (*Scheme 10*, purity >95% by analytical HPLC).



Scheme 10: *Synthesis of KHS101-biotin* ***62*** *from key building block* ***63*** *and commercially available biotin-PEG4-NHS.*

2.6 Identifying the molecular target of KHS101 in GBMs

2.6.1 Validation of KHS101-BP activity in GBM cells

Before any photocrosslinking experiments could be undertaken with KHS101-BP, the activity of **61** in GBM cells needed to be confirmed. This work was carried out by Dr Verena Kuchler at LICAP. Dose-dependent inhibition of GBM cell viability was observed in GBM1 cells treated with KHS101-BP **61** (*Figure 21*). Not only was the activity of the probe retained relative to unlabelled KHS101, but KHS101-BP showed a slightly increased cytotoxicity in GBM cells. Importantly, the same specific phenotype of autophagic vacuolisation was observed in GBM1 cells treated with KHS101-BP (7.5 µM) as with KHS101 (data not shown). Taken together, these data suggest that KHS101-BP is working through the same mechanism of action as KHS101.

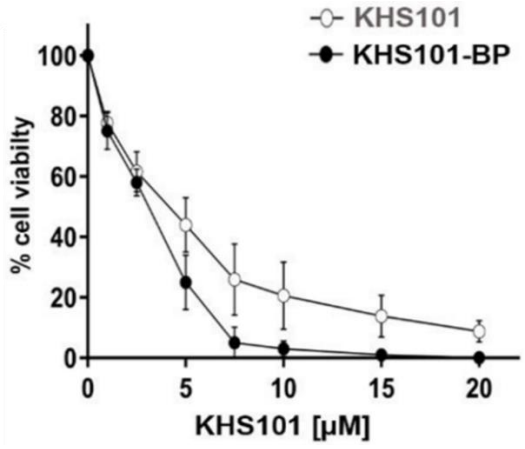


Figure 21: *Cell viability dose-response curves for GBM1 cells treated with KHS101 or KHS101-BP after 48 hours. Data are mean ± SD of 3 biological repeats.*

2.6.2 Affinity-based target identification using KHS101-BP

With the activity of KHS101-BP **61** in whole GBM cells established, an affinity-based target identification strategy could be undertaken. Experiments were performed by Dr Shoutian Zhu at the California Institute for Biotechnology (Calibr), where an established proteomics workflow is set up (*Figure 22*). GBM1 cells were incubated with KHS101-BP **61**, with or without an excess of KHS101 **41** (competition experiment). UV irradiation enabled covalent capture of KHS101-interacting proteins, which were detected through biotin-labelling *via* a copper-catalysed click reaction on GBM cell lysates. The cell lysate underwent fractionation, with each fraction separated on a two-dimensional gel (proteins are separated in the first dimension by isoelectric point and in the second dimension by molecular weight); biotin-modified proteins were identified using Western blotting. Spots that were competed away with unlabelled KHS101 **41** (added during a separate experiment in step 1, *Figure 22*) were excised, digested and analysed by LC-MS/MS to identify proteins contained within them.

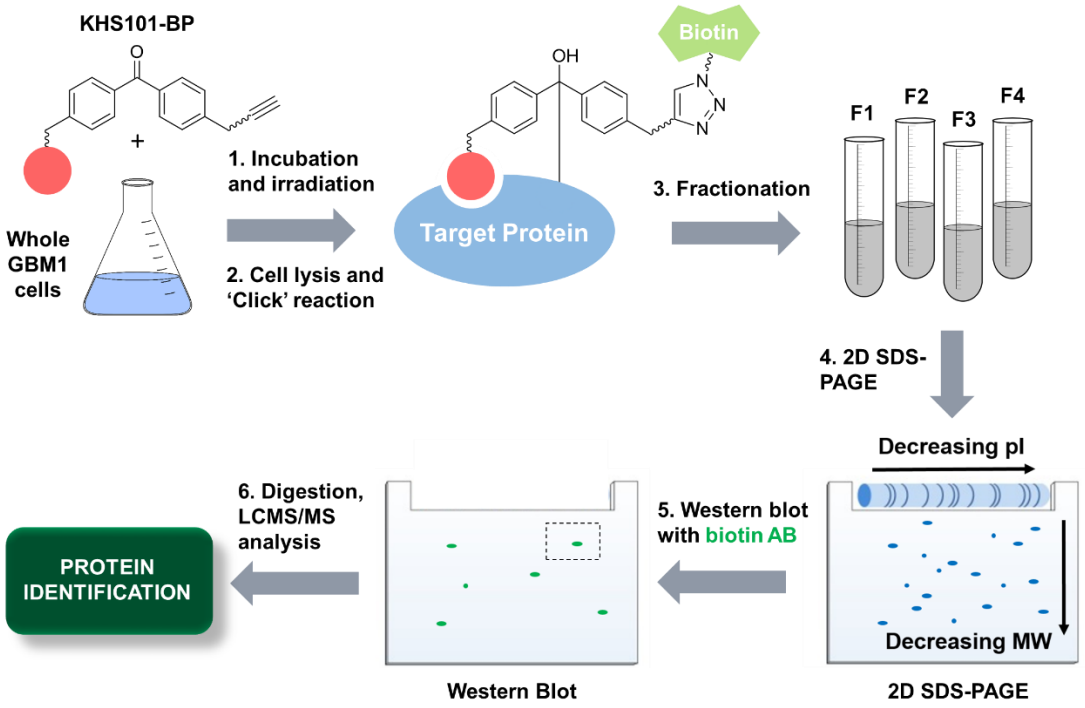


***Figure 22**: Affinity-based target identification workflow performed by Dr Shoutian Zhu. GBM1 cells were incubated with KHS101-BP **61** (with or without excess KHS101 **41**) and irradiated (365 nm, 30 minutes). 'Click' chemistry was then performed, followed by fractionation of the cell lysate and protein separation by 2D SDS-PAGE. Biotin-labelled proteins were identified through Western blotting, using an antibody against biotin. Competed spots were excised, digested and analysed by LC-MS/MS.*

Two-dimensional SDS-PAGE and Western blotting of GBM1 cell lysates (20-40% ammonium sulfate-precipitated fraction) highlighted a distinct spot at ~60 kDa (pI~5.7) that was reproducibly competed away with 50-fold excess of unlabelled KHS101 **41** (black arrow, *Figure 23*). LC-MS/MS analysis revealed that this spot corresponded to the mitochondrial 60 kDa heat shock protein 1 (HSPD1, also termed HSP60). It should be noted that there is not complete competition of the spot relating to HSPD1, possibly due to differences in affinity of KHS101-BP and the unlabelled small molecule. Perhaps more importantly, the covalent modification upon irradiation is not reversible, hindering competition of the covalently bound KHS101-BP by KHS101 (not an equilibrium situation). HSPD1 is a molecular chaperone responsible for the correct folding and import of proteins from the cytoplasm into the mitochondrial matrix, including refolding and assembly of unfolded polypeptides under mitochondrial stress conditions.[96,97] A strong positive correlation between the formation and development of gliomas in the brain and increased HSPD1 expression exists (which correlates negatively with patient survival). Notably, HSPD1 knockdown in primary GBM cell lines has been shown to supress cell proliferation.[98] As a result, successive experiments (*section 2.6.3*) focussed on validation of HSPD1 as the molecular target of KHS101 **41**, responsible for the GBM cell-death phenotype.

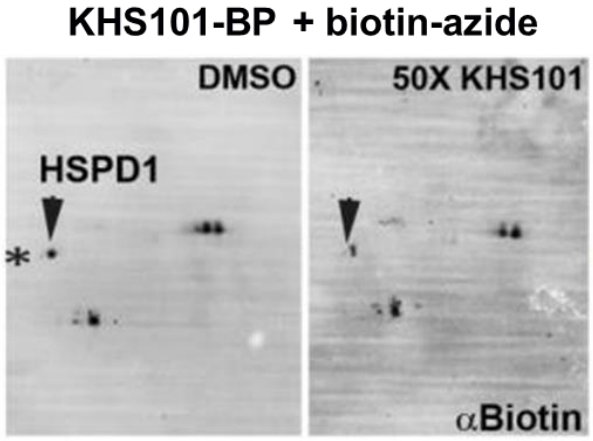


Figure 23: *Two-dimensional SDS-PAGE and Western blotting of 20-40% fraction of GBM1 cell lysates treated with KHS101-BP* ***61*** *alone or with 50-fold additional excess of unlabelled KHS101. Photocrosslinking was followed by biotin-labelling of protein-KHS101-BP complexes (1 mM $CuSO_4$, 100 µM TBTA, 1 mM TCEP and 25 µM biotin-azide). The spot at ~60 kDa (black arrow), which was reproducibly competed away with excess (50-fold) unlabelled KHS101 contained HSPD1 (using LC-MS/MS analysis).*

2.6.3 Validation of HSPD1 as the target of KHS101

2.6.3.1 *In vitro* pull-down of recombinant HSPD1 using KHS101-biotin

Direct binding of KHS101 to HSPD1 was confirmed by pull-down experiments using KHS101-biotin **62** (*Figure 24*), carried out by Shoutian Zhu at Calibr. Following incubation of recombinant human HSPD1 with biotinylated probe **62** in the presence or absence of unlabelled KHS101, HSPD1 was pulled-down using streptavidin-agarose beads, which were precipitated and washed with buffer. Bound protein was eluted and analysed by SDS-PAGE: silver staining and Western blotting allowed detection of HSPD1, the quantity of which was significantly reduced in the competition experiment (50-fold excess unlabelled KHS101, lane 6, *Figure 24*). It is important to note that HSPD1 could have been identified during the photocrosslinking experiments *via* an indirect interaction between KHS101 and the protein (*i.e.* mediated by interacting partners of HSPD1). Thus, these experiments using recombinant HSPD1 in buffer were crucial in verifying a direct HSPD1-KHS101 interaction, in an orthogonal manner to the photocrosslinking experiments.

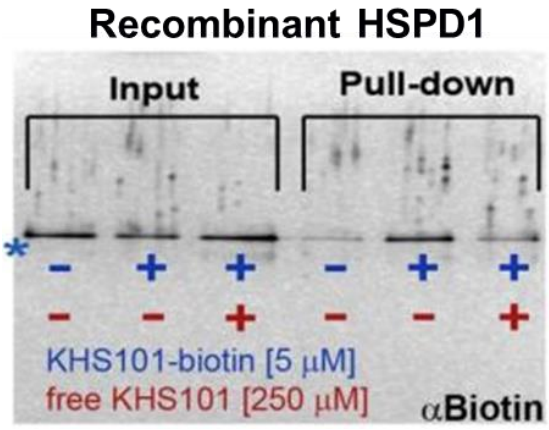


Figure 24: *Pull-down of recombinant human HSPD1 using KHS101-biotin* ***62*** *(in PBS containing 2 mM $MgCl_2$, 2 mM $MgCl_2$, and 0.1% Tween 20), precipitated with streptavidin-coated agarose beads. Silver staining of SDS-PAGE gel in presence and/or absence of unlabelled KHS101 enabled detection of direct binding of HSPD1 with KHS101-biotin.*

2.6.3.2 Effect of KHS101 and HB072 on HSPD1 chaperone activity

Assessment of the chaperone activity of HSPD1 upon KHS101 addition revealed a concentration-dependent inhibition of substrate re-folding *in vitro*: KHS101 inhibited the re-folding of an unfolded luciferase substrate by the HSPD1/HSPE1 complex (*Figure 25*, a, experiments performed by Dr Euan Polson, LICAP). Furthermore, this effect was recapitulated by the mitochondrial HSPD1-binding natural product myrtucommulone A[99] (MC, donated by Prof Johann Jauch, *Figure 25*, b). These data, in combination with the GBM cell death phenotype elicited by MC, further confirmed

HSPD1 as the target of KHS101. Interestingly, control compound HB072 **60** had no effect on the functional activity of HSPD1/HSPE1 (*Figure 25*, b). These data corroborate with the results from testing of HB072 in the phenotypic assay: the KHS101-induced autophagic phenotype was not recapitulated by HB072.

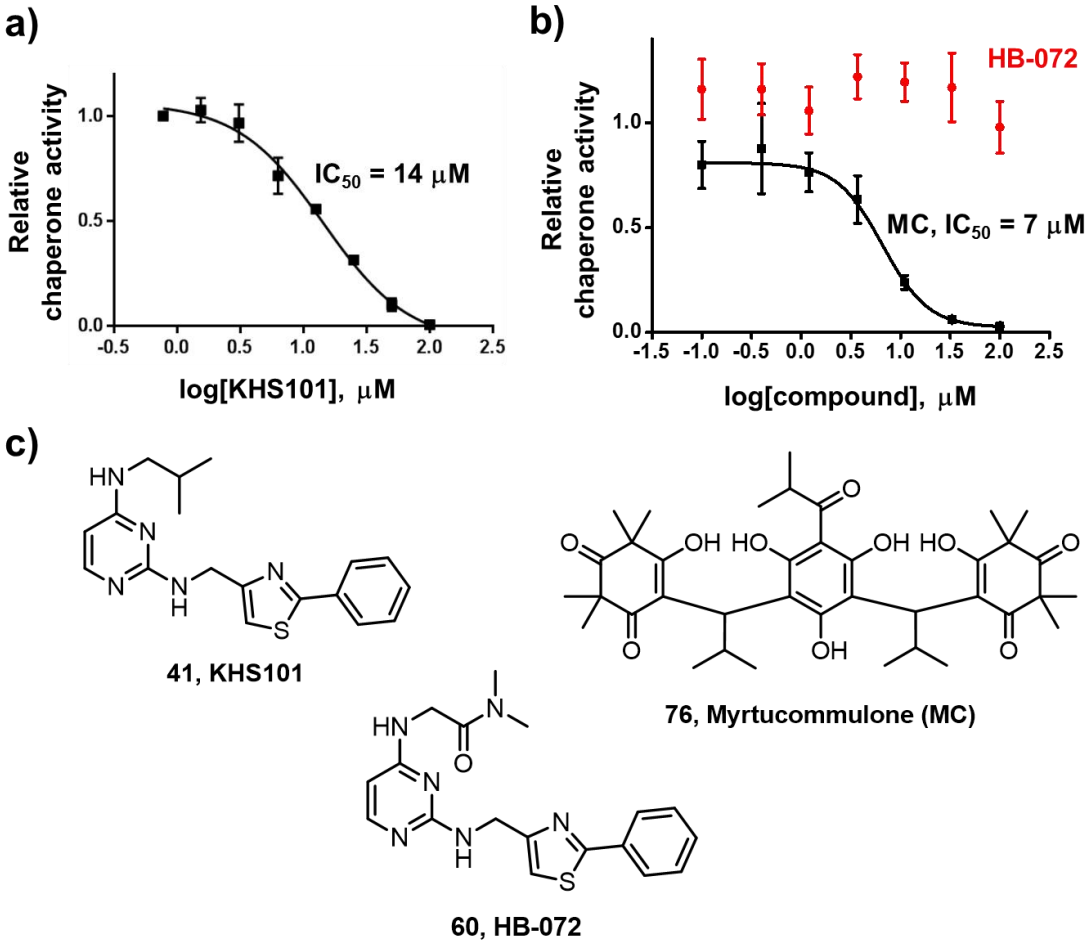


***Figure 25**: Effect of KHS101, HB072 and myrtucommulone A (MC) on HSPD1/HSPE1 chaperone activity. a) Inhibition of HSPD1/HSPE1 complex refolding activity by KHS101 **41** (IC_{50} = 14.4 μM). Error bars show standard deviation of three biological repeats. b) Inhibition of HSPD1/HSPE1 complex refolding activity by MC (IC_{50} = 6.7 μM). HB072 **60** had no effect on refolding activity of HSPD1/HSPE1. Error bars show standard deviation of three biological repeats. c) Chemical structures of KHS101 **41**, HB072 **60** and MC **76**.*

The effect of KHS101 on chaperone activity is significant, as HSPD1 is responsible for the folding of mitochondrial proteins including key enzymes involved in OXPHOS and glycolysis (required for ATP production). Depletion of cellular ATP levels was observed in GBM cells treated with KHS101 (performed by Dr Euan Polson), which was followed by autophagy. However, the addition of autophagy inhibitors did not rescue the cells, suggesting that autophagy is a stress response (as a means to preserve energy) rather than the primary cause of cell death. Ultimately, the selective cytotoxic activity of KHS101 in GBM cells highlights a dependency on mitochondrial HSPD1 chaperone activity for GBM cell energy metabolism.

2.7 Analysis of KHS101-HSPD1 interaction using surface plasmon resonance

Surface plasmon resonance (SPR) was carried out to analyse the KHS101-HSPD1 interaction *in vitro*, with the aim to measure a binding affinity. KHS101-biotin **62** was non-covalently immobilised onto a streptavidin-coated sensor chip at concentrations of 1 nM, 10 nM and 100 nM. Solutions of recombinant human HSPD1 (1 µM – 0.1 nM) were injected over the immobilised the small molecule, however no significant binding event between KHS101 and HSPD1 was observed (*Figure 26*, a, showing results from 1 µM HSPD1 injection). Despite the small change in refractive index detected upon protein injection (inset, *Figure 26*, a), the shape of the sensorgram is indicative of non-specific binding of the protein to the negatively-charged cyclodextran surface. A possible obstacle to the detection of a physical interaction between KHS101 and HSPD1 was that the linker between the small molecule and the biotin was not sufficient in length to allow KHS101 to extend above the bulky streptavidin-coated surface. Furthermore, a higher concentration of protein (>1 µM) may be required in order to detect a binding event, if the KHS101-HSPD1 interaction is relatively weak. In addition, the strength of the KHS101-HSPD1 interaction may be dependent on other factors, such as the presence of co-chaperone HSPE1, which HSPD1 works in conjunction with to perform its protein folding function.[100]

Although a smaller quantity of protein is consumed when employing the strategy described, published literature protocols for measuring HSPD1-small molecule/peptide interactions using SPR tend to immobilise the protein and flow over the ligand.[101,102] Therefore, the chip was derivitised with HSPD1 (donated by Dr Hao Shao and Prof Jason Gestwiki), by coupling the lysines of the protein onto the carboxy methyl groups on the surface of the CM5 chip (experiments were carried out in collaboration with David Klebl, Wellcome Trust student). However, injection of KHS101 (8 µM - 1 mM) over the protein-derivatised surface did not indicate binding (*Figure 26*, b). Although there is significant change in response units upon injection of KHS101, it is not sustained and rapidly decreases over the injection period. However, the reasonably harsh conditions used in the derivatisation process (10 mM acetate buffer, pH 5) may have affected the integrity of the HSPD1 complex.

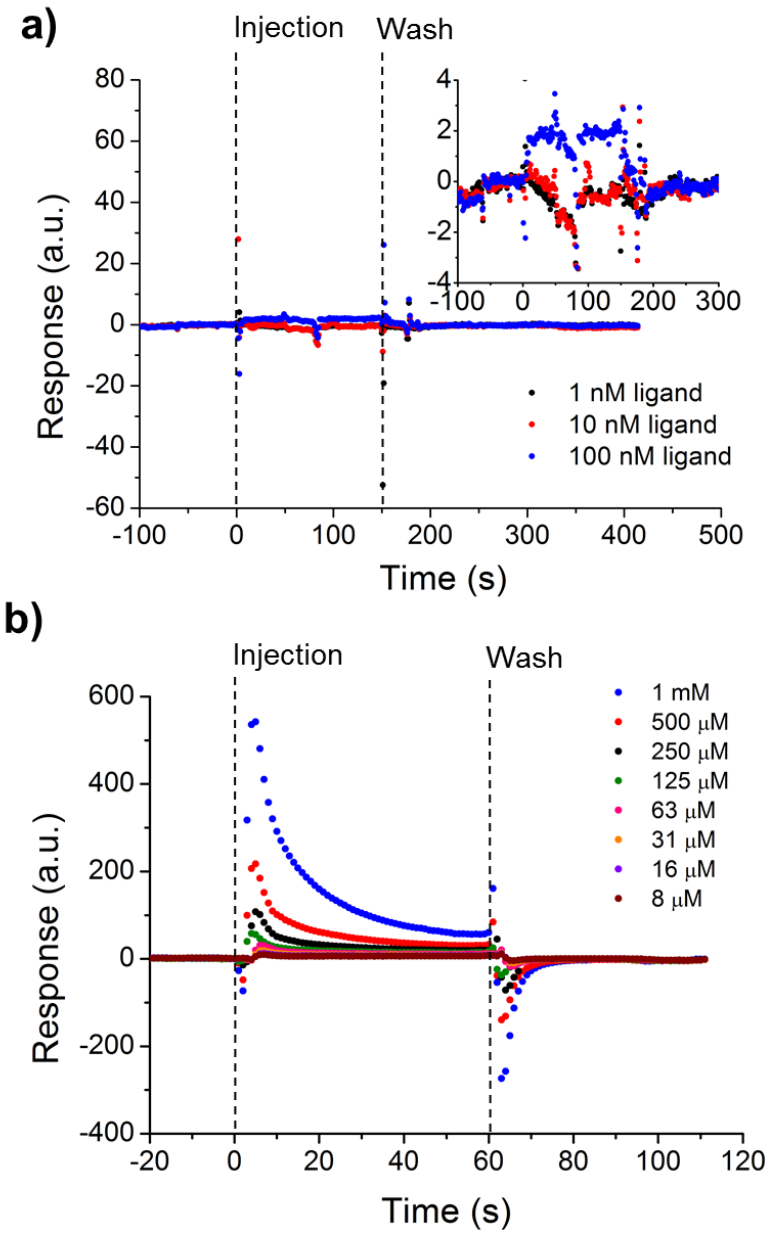
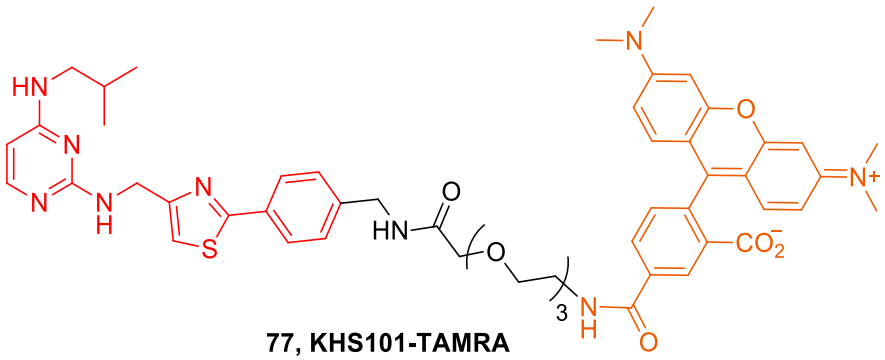


Figure 26: *SPR analysis of the interaction between HSPD1 and KHS101. a) 2 minute injection of 1 µM HSPD1 onto flow cells immobilised with KHS101 (via KHS101-biotin), followed by a 4 minute dissociation period. Sensorgram indicates no binding of the two. b) 1 minute injection of KHS101 (1 mM – 8 µM) onto flow cells derivitised with HSPD1, followed by a 1 minute dissociation period. Again, sensorgram indicates no binding of the two.*

2.8 Development of a fluorophore-labelled probe for biophysical analysis of the KHS101-HSPD1 interaction

2.8.1 Design of fluorophore-labelled probe

In addition to SPR, fluorescence anisotropy is a complementary biophysical technique that can be used to measure protein-small molecule interactions. Furthermore, providing the compound is cell-permeable, cell imaging studies can be carried out to determine the cellular localisation of the small molecule, for example using FRET or FRET/FLIM (when protein target is also labelled). The rhodamine dye 5-(and-6)-carboxytetramethylrhodamine (TAMRA) was chosen due to its stability towards photo-bleaching and high quantum yield. The addition of an OEG linker between the small molecule (KHS101, red, *Figure 27*) and the fluorophore (TAMRA, orange) enhanced probe solubility in aqueous environments. Furthermore, the design of KHS101-TAMRA **77** allowed synthesis directly from key building block **63**.



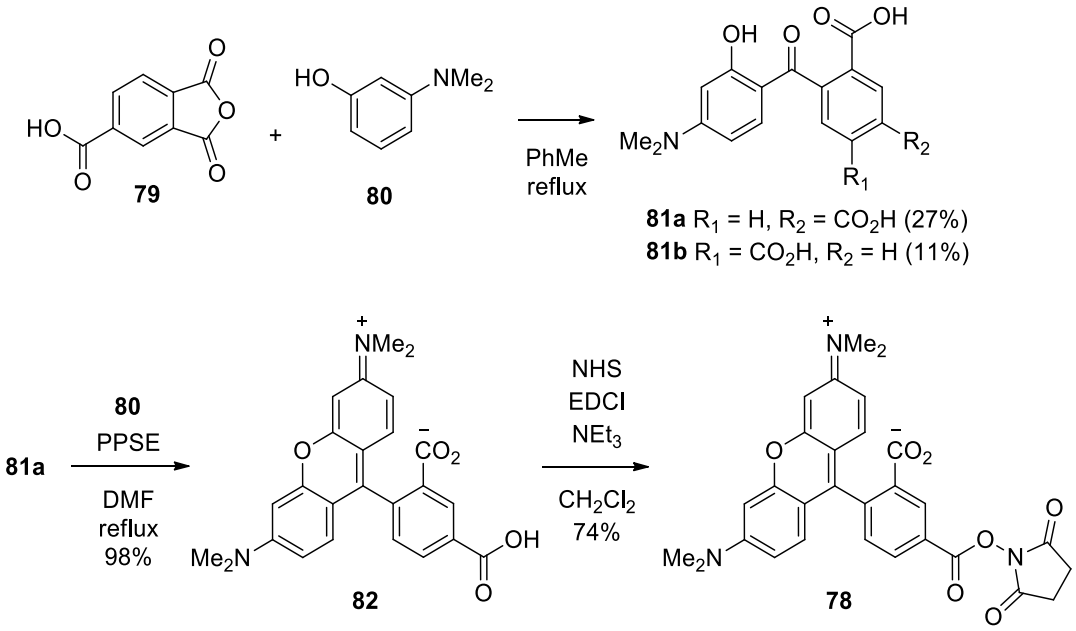
77, KHS101-TAMRA

Figure 27: *Chemical structure of KHS101-TAMRA **77**. The KHS101 component (red) is attached via an OEG linker to the fluorophore (TAMRA).*

2.8.2 Synthesis of KHS101-TAMRA

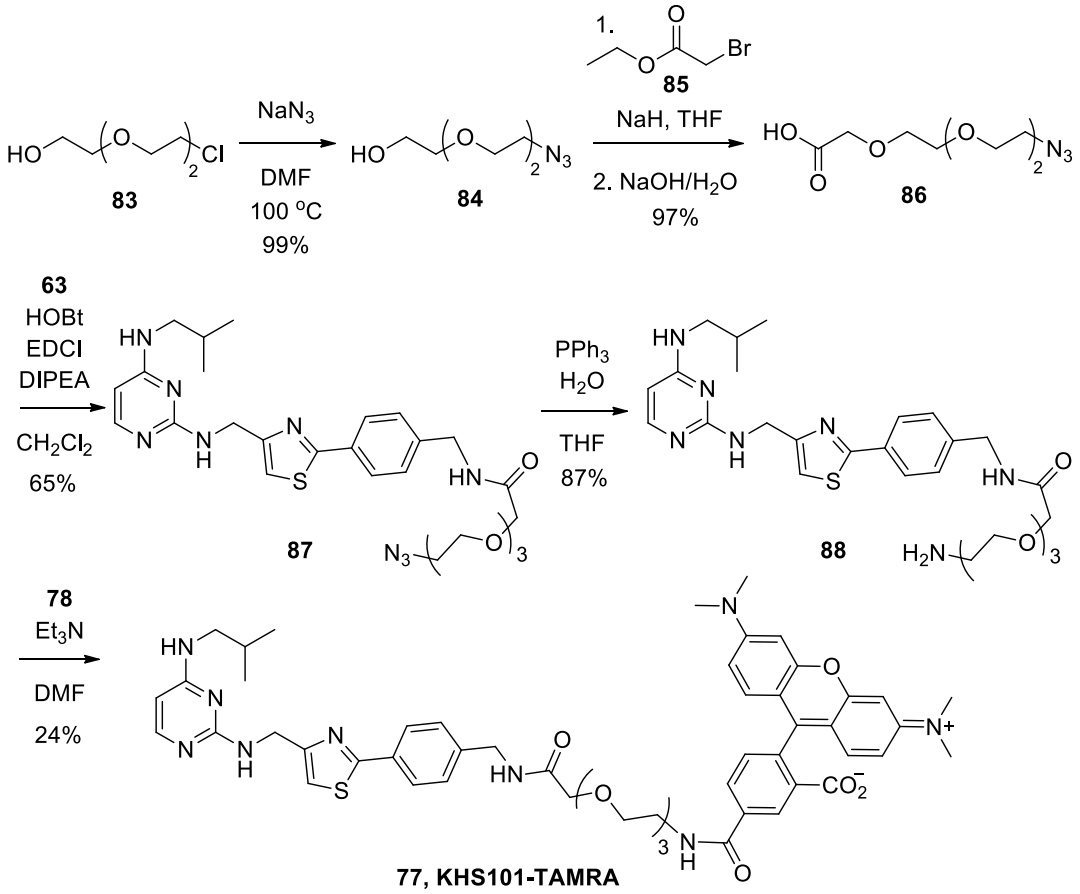
The fluorescent component of the KHS101-TAMRA probe **77** was prepared following a method by Kvach and co-workers.[103] TAMRA is commonly purchased from suppliers as a mixture of 5- and 6- isomers, with isomerically pure material fetching a high price. Individual isomers are preferred for analytical applications such as FRET-FLIM.[104] Therefore, TAMRA-NHS **78** was synthesised. On refluxing 3-dimethylaminophenol **80** and 1,2,4-benzetricarboxylic anhydride **79** in toluene, **81a** and **81b** were obtained; these were separated by sequential recrystallization from methanol and acetic acid, respectively (*Scheme 11*). To assemble the tetramethylrhodamine core, regioisomer **81a** was allowed to react with a further equivalent of **80** using trimethylsilyl polyphosphate as a dehydrating reagent and weak acid catalyst;[105] **82** was obtained in nearly quantitative yield.

To allow attachment of the KHS101 component, *N*-hydroxysuccinimide ester **78** was prepared according to a method by Brunet *et al*.[106] Initially, synthesis of **78** was attempted using *N*-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDCI), however the reaction appeared to give a mixture of starting material and di-NHS-ester product. To achieve regioselectivity for the 5-carboxylic acid over the 3-carboxylic acid, the rhodamine must be present in the closed lactone form during the reaction, which is observed under basic conditions. Therefore, by using an excess of base (5 equivalents of triethylamine) the desired regioselectivity was afforded, giving ester **78** in good yield.



Scheme 11: *Synthetic route to TAMRA-NHS* ***78****. For the cyclisation step, trimethylsilyl polyphosphate (PPSE) solution in chloroform was prepared according to the method of Yokoyama and co-workers.*[105] *TAMRA is drawn in the open-chain form: TAMRA can exist as two tautomeric open-chain and closed chain (spirolactone) forms. In polar protic solvents, the fluorescent open-chain structure predominates. In aprotic solvents, the non-fluorescent lactone form prevails.*

The OEG linker component of KHS101-TAMRA **77** was synthesised following a procedure by Liu *et al.* (*Scheme 12*).[107] Installation of the azide functionality in **84** was achieved by reaction of chloride **83** with sodium azide. Acid **86** was prepared by S_N2 reaction of alcohol **84** with ethyl bromoacetoactate, followed by hydrolysis of the ethyl ester under basic conditions. OEG linker **86** was then attached to the KHS101 component **63** *via* amide bond formation using 1-hydroxybenzotriazole (HOBt) and EDCI. Reduction of azide **87** to amine **88** under Staudinger conditions liberated the amine handle, enabling attachment to the TAMRA component through reaction with TAMRA-NHS **78**.



Scheme 12: *Synthetic route to KHS101-TAMRA* ***77****. OEG linker* ***86*** *was prepared as described by Liu and co-workers.*[107]

Binding between KHS101-TAMRA **77** and HSDP1 was measured using a fluorescence anisotropy assay: HSPD1 (10 µM – 5 nM) was titrated into a constant concentration of KHS101-TAMRA tracer (100 nM). However, no significant change in anisotropy was observed with increasing concentrations of protein, suggesting no binding between the two (*Figure 28*). The minimal difference in anisotropy was reasoned to be due to the flexible OEG linker between KHS101 and the rhodamine dye, providing **77** did in fact bind to HSPD1. Fluorescent KHS101-derivative **77** was also used in microscale thermophoresis (MST) experiments (performed by David Klebl), however only a small shift in thermophoretic mobility was observed upon binding of **77** to HSPD1 and a reliable K_d could not be determined. It should be noted that the stability and oligomerisation state of HSPD1 was found to be dependent on a range of factors (including temperature, presence of cofactors such as ATP/ATP and its co-chaperone HSPE1), complicating biophysical measurements with purified protein (work carried out by David Klebl).

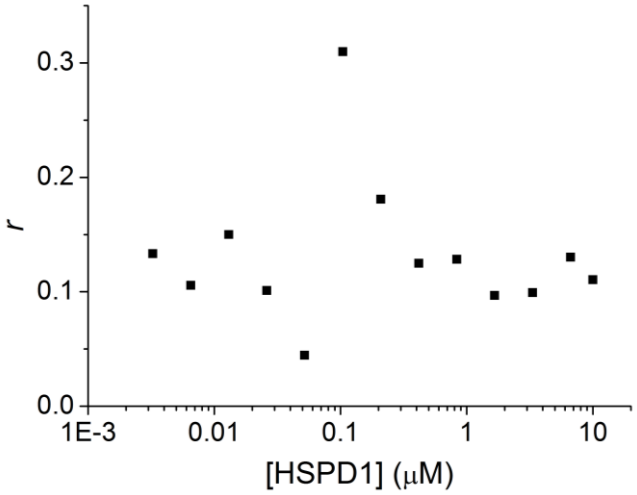


Figure 28: *Fluorescence anisotropy direct binding assay of TAMRA-KHS101* ***77*** *to HSPD1. Assay run in PBS (2 mM $MgCl_2$, 2 mM DTT, 0.1% Tween, pH 7). HSPD1 (starting at 10 µM) was serially diluted over a ½ regime into a solution of KHS101-TAMRA (fixed at 100 nM). Plates were read after 1 h and 20 h incubation, data shown from 20 h incubation.*

2.9 Studies towards identifying the KHS101-binding site on HSPD1

Now that the molecular target of KHS101 had been identified, elucidating the binding site on the protein could be investigated. An alternative photocrosslinker to KHS101-BP **61** with a much smaller modification relative to the unlabelled small molecule was designed. Photocrosslinker probe **89** incorporates a cleavable ester linker between the ligand (KHS101, red, *Figure 29, a*) and photo-reactive (diazirine) elements (full structure given in *Figure 29, b*). Upon irradiation of the sample with UV light, a reactive carbene is revealed, which inserts into X-H bonds (where X is carbon or a heteroatom) in the vicinity of the ligand-binding site (*Figure 29, a*). Aminolysis or hydrolysis of the ester leaves a minimal modification on the protein, which can be digested into peptides and analysed by tandem MS.

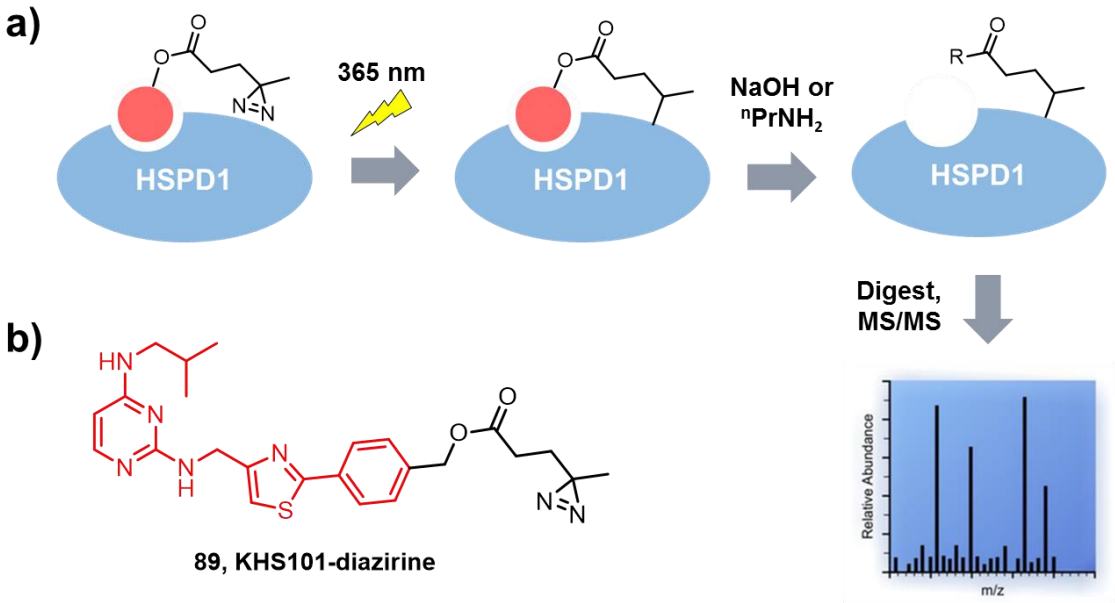
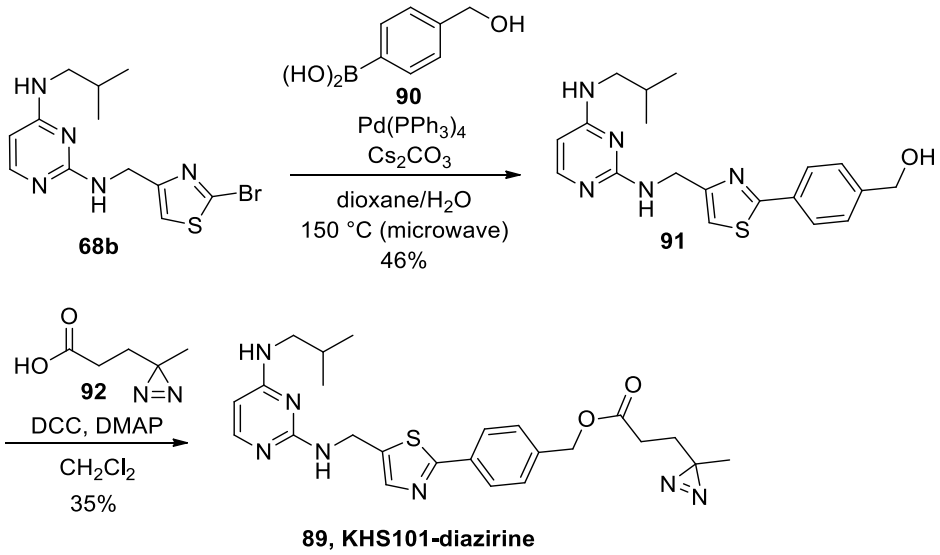


Figure 29: *Identifying the binding site of KHS101 on HSPD1. a) Proposed workflow for binding site identification using diazirine analogue of KHS101 (red). R = OH when NaOH is used, R = NH^{n}Pr when ^{n}PrNH$_2$ is used. b) Chemical structure of photo-reactive analogue of KHS101, KHS101-diazririne* ***89****.*

KHS101-diazirine **89** was prepared in two steps from previously synthesised pyrimidine intermediate **68b** (*Scheme 13*). Boronic acid **90** was combined with pyrimidine **68b** in a Suzuki coupling to afford alcohol-functionalised KHS101 derivative **91**. Acid-functionalised diazirine **92** was activated using DCC and DMAP, and reacted with alcohol **91** in a Steglich esterification to give ester **89**.[108]



Scheme 13: *Synthetic route to KHS101-diazirine* ***89*** *from previously synthesised pyrimidine building block* ***68b****.*

Preliminary photocrosslinking experiments were then undertaken. KHS101-diazirine (30 µM) was added to a buffered solution of recombinant HSPD1 (10 µM, in 50 mM NH_4HCO_3, 10 mM $MgCl_2$, pH 7). Following incubation for 10 minutes at room temperature, the mixture was irradiated at 365 nm for 30 minutes on ice. As an initial experiment to determine whether crosslinking between HSPD1 and **89** could be detected, the photocrosslinking mixture was not subjected to hydrolysis or aminolysis (as described in *Figure 29*). Excess KHS101-diazirine was removed using a centrifugal filter prior to proteolytic digestion. Q-TOF MS/MS analysis of the tryptic peptides revealed a distinct peptide which was modified upon photocrosslinking with **89** (199-223, *Figure 30,* performed by Dr Rachel George). Interestingly, two modifications were found on each of the three peptides detected (*Table 3*, 2 × +451.21 Da). However, despite good sequence coverage of both unmodified (75%) and modified HSPD1 (80%) the sequence coverage of amino acids 199-223 in the unmodified protein was poor (*Figure 56*, Appendix I). The fact that the sequence coverage of the modified protein in this region (199-223) was enhanced, most likely due to the presence of the modifications proximal to lysine residues (causing miscleavages), was further evidence that crosslinking had occurred.

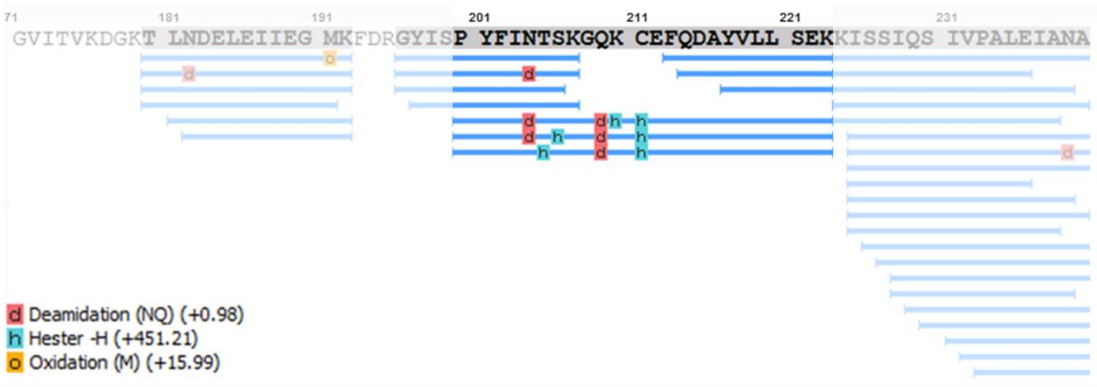


Figure 30: *MS/MS data for crosslinked HSPD1 shows a distinct modified peptide: 199-223. Note: only a section of the whole sequence in shown in figure (residues 171-240), see Figure 57, Appendix I for peptide mapping of whole protein. Modified residues are represented by a blue 'h', however the exact sites of modification cannot be confirmed.*

Table 3: *Tabulated MS data of peptides modified with crosslinker **89**.*

	Modified HSPD1 peptide	Mass (Da)
1)	PYFIN(+0.98)TSKGQ(+0.98)**K**(+451.21)**C**(+451.21)EFQDAYVLLSEK	3711.7759
2)	PYFIN(+0.98)T**S**(+451.21)KGQ(+0.98)K**C**(+451.21)EFQDAYVLLSEK	3711.7759
3)	PYFIN**T**(+451.21)SKGQ(+0.98)K**C**(+451.21)EFQDAYVLLSEK	3710.7920

It should be noted that the MS/MS spectra for the modified peptides (amino acid sequences given in *Table 3*) were not sufficiently strong to confirm the exact amino acids that the modifications resided on, potentially due to low crosslinking yields. However, the data obtained indicates a binding site in a certain region on HSPD1 which appears to be particularly flexible, following mapping of the modified residues identified through MS/MS onto cryo-EM and X-ray crystal structures of apo (*Figure 31*) and ADP/HSPE1-bound HSPD1 (*Figure 32*). Interestingly, the potential binding site is not proximal to the ATP/ADP binding site, corroborating with data demonstrating that KHS101 does not affect the ATPase activity of the chaperone (data not shown). Furthermore, the potential binding site appears to be at the interface between HSPD1 and its co-chaperone HSPE1 in the ADP/HSPE1-bound HSPD1 structure (*Figure 32*).

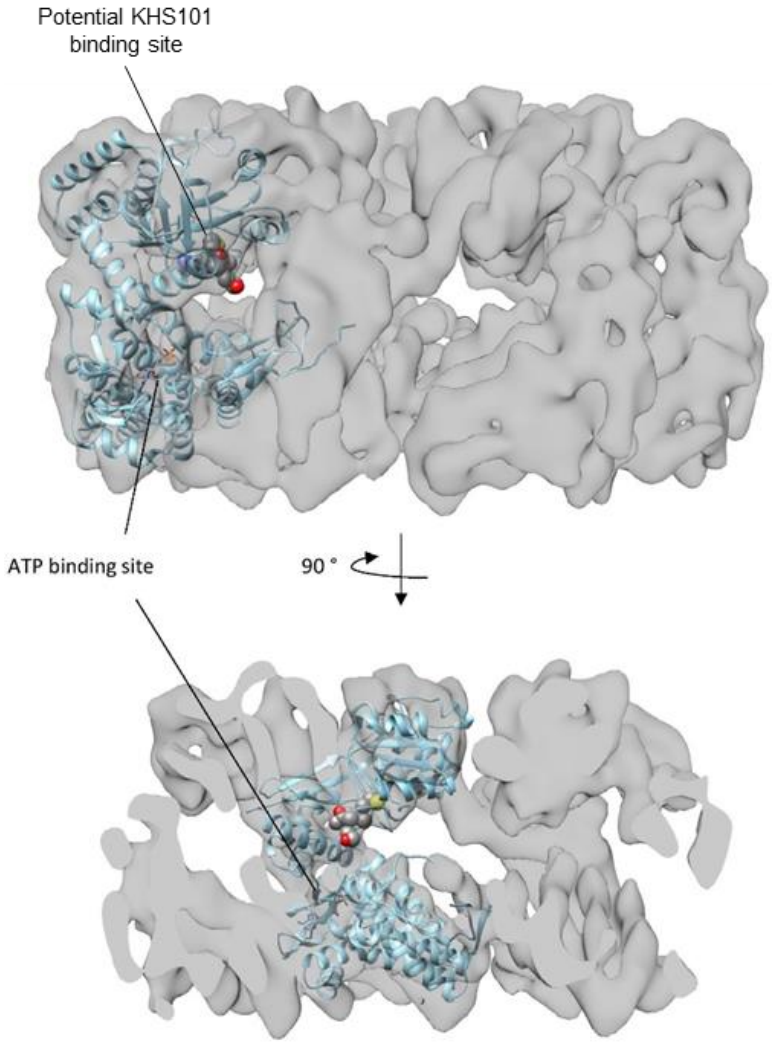


Figure 31: *Residues modified by KHS101-diazirine identified through MS/MS experiments mapped onto HSPD1 structure. Grey: EM density map of HSPD1 (7 Å) obtained by David Klebl, Emma Hesketh and Prof Neil Ranson. Blue: model for apo HSPD1. Spheres: amino acid residues modified through photocrosslinking with KHS101-diazirine. Sticks: ATP binding site.*

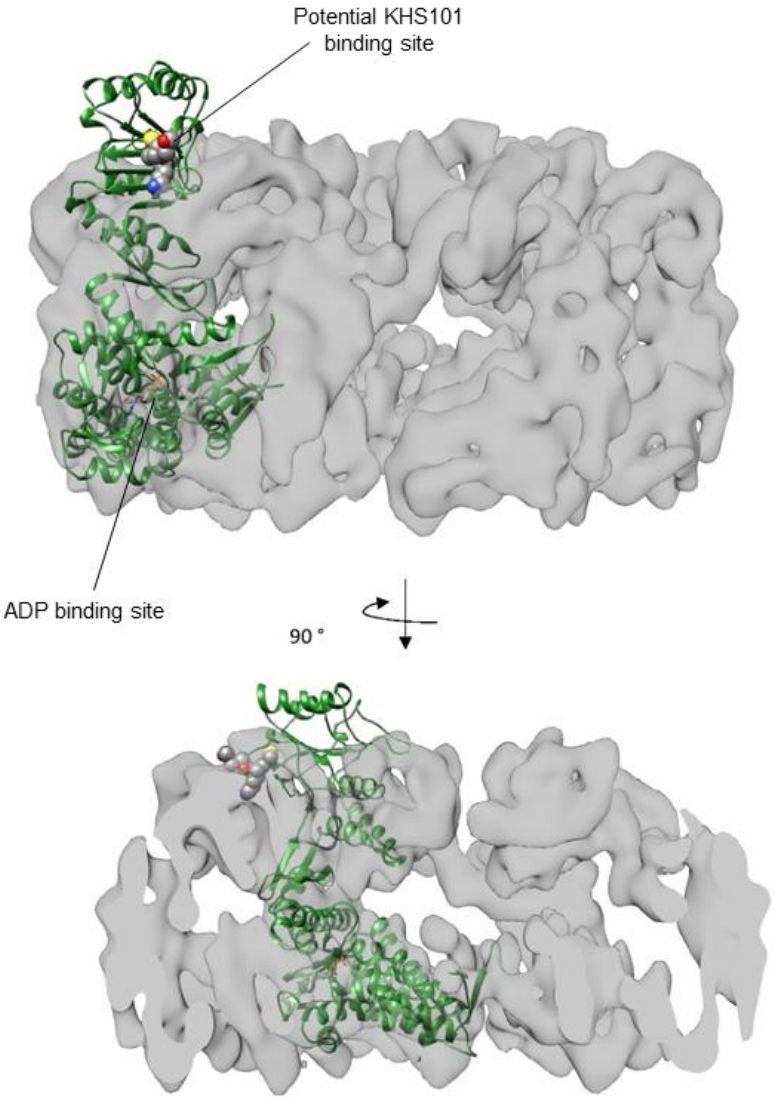


Figure 32: *Residues modified by KHS101-diazirine identified through MS/MS experiments mapped onto HSPD1 structure. Grey: EM density map of HSPD1 (7 Å) obtained by David Klebl, Emma Hesketh and Prof Neil Ranson. Green: crystal structure (4PJ1) of ADP/HSPE1 bound HSPD1. Spheres: amino acid residues modified through photocrosslinking with KHS101-diazirine. Sticks: ADP binding site.*

2.10 Conclusions

In summary, the range of chemical tools synthesised herein have proved essential in elucidating details of the mechanism of action of KHS101, including the direct molecular target: HSPD1. Importantly, a robust synthetic route to KHS101 has been developed, which has enabled grams of analytically pure material to be obtained, used for cellular and *in vivo* testing (in addition to allowing rapid synthesis for future studies). Furthermore, the design of a synthetic route to KHS101 analogue HB072 has provided a useful inactive control compound. Results comparing the phenotype in KHS101 and KHS092-treated GBM cells suggested engagement of different cellular targets in both NPCs and GBMs, while additional biological data disproved TACC3 as the molecular target responsible for the specific cell death phenotype observed in KHS101-treated GBM cells.

The development of ligand-directed affinity probes KHS101-BP **61** and KHS101-biotin **62** for photocrosslinking in whole GBM cells and *in vitro* pull-down, respectively, was crucial in identifying and validating HSPD1 as the biologically relevant target of KHS101. Not only was the synthetic route to a key building block for their synthesis optimised (by a >6-fold increase in yield relative to the original procedure), but this also enabled a route to a fluorescent-KHS101 derivative KHS101-TAMRA **77**, for future use in cellular imaging studies. Although a binding affinity for the KHS101-HSPD1 interaction could not be obtained using SPR and fluorescence anisotropy assays, binding may require the presence of co-chaperone HSPE1 or other co-factors such as ATP/ADP. However, having tool compounds KHS101-biotin **62** and KHS101-TAMRA **77** at hand, may allow the rapid screening of new compounds and lead optimisation studies through development of such biophysical assays. Validation of the KHS101-binding site on HSPD1 should reveal further details of the mechanism of action of the small molecule and enable structure-based design of small molecule inhibitors.

A significant outcome from this study is the discovery of KHS101 as the first known drug-like inhibitor of HSPD1. Finally, in gaining a greater understanding of the mechanism of action of KHS101, this has validated specific metabolic vulnerabilities of cancer cells as a potential therapeutic strategy, with the discovery of HSPD1 as a possible druggable target for glioblastoma treatment.

Chapter 3 – Traceless labelling of proteins for the study of protein-protein interactions

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

In this chapter, reference is made to work from a research article that was published in September 2016: ('Economical and scalable synthesis of 6-amino-2-cyanobenzothiazole'; Hauser, J.R., Beard, H.A., Bayana, M.E, Jolley, K.E., Warriner, S.L., Bon, R.S, *Beilstein J Org Chem*., 2016, **12**, 2019-2025). HAB (the candidate) and JRH performed initial synthesis studies, JRH synthesised the final compounds, JRH performed calorimetry experiments with assistance from MEB and KEJ, SLW and RSB provided supervision, JRH prepared a draft which was edited into its final form by RSB.

3.1 Introduction and aims

3.1.1 Protein-protein interactions

Protein-protein interactions (PPIs) mediate a large number of important cellular processes; these interactions can be classified as stable or transient.[109] While stable interactions are usually associated with multi-subunit protein complexes, transient interactions are temporary in nature. Transient PPIs control the majority of cellular processes, including protein modification, intracellular signalling and programmed cell death and are therefore implicated in a variety of disease states.[110] Thus, it is essential that new methods are developed to probe these pathways, to enhance our understanding of the molecular mechanisms of disease and for the discovery of new therapeutic targets.

3.1.2 Methods for the identification of protein-protein interactions

3.1.2.1 Affinity-based methods

Traditionally, PPIs have been investigated using affinity purification methods. Co-immunoprecipitation (co-IP) is regarded as the gold-standard method for analysing PPIs;[111] this works by using an antibody for a known member of a protein complex to pull unknown members of the complex out of solution (*e.g.* cell lysate). Approaches such as mass spectrometry and Western blotting are then used to identify retrieved binding partners. However, this technique is limited to PPIs that are sufficiently stable to resist washes, and also requires the epitope on the protein to be exposed, meaning that weak interactions as well as proteins in complexes which mask the epitope are not identified. Furthermore, generating an antibody that specifically binds to a protein of interest is challenging. To overcome this, a variety of tags (such as FLAG, GST and GFP tags) have been developed,[112] which can be attached to the *C*- or *N*-termini of proteins; tagged-proteins can then be isolated from the mixture using a solid support. However, incorporation of these tags can affect the localisation of proteins, in addition to their PPIs (if these depend on their *C*- and *N*- termini for the interaction).

3.1.2.2 Genetic methods

Introduced by Fields and Song in 1989,[113] the yeast two-hybrid (Y2H) system has been used to screen libraries of proteins against a known ‘bait’ protein, in addition to validating PPIs *in vivo*. Y2H is based on the reconstitution of a transcription factor upon the interaction of a ‘bait’ protein (fused to a DNA-binding domain) with ‘prey’ proteins (fused to an activation domain), leading to the activation of a downstream

reporter gene.[114] Y2H is considered as a high-throughput technique and can be used to detect weaker interactions as the reporter gene strategy results in signal amplification. However, this method is notorious for generating large numbers of false positives (estimated as 45 – 80%)[115] and requires proteins to be relocated to the nucleus meaning membrane proteins are difficult to investigate.[116] Variations on the system have been developed, as a means to overcome some of these drawbacks. For example the split-ubiquitin system, which is based on the reconstitution of an active ubiquitin and cleavage of a reporter protein.[117]

3.1.2.3 Chemical crosslinking

Transient interactions that occur briefly as part of a signalling cascade for example, can be trapped by covalently coupling spatially proximate amino acid side chains of interacting proteins. Stabilisation of PPIs through crosslinking prior to purification by affinity techniques can aid detection of transient or low affinity interactions, while specific interaction sites on proteins can be revealed using MS/MS analysis.[118] Homobifunctional, amine reactive crosslinkers (*Figure 33*) such as disuccinimidyl suberate (DSS, **93a**) and its cleavable analogue dithiobis(succinimidylpropionate) (DSP, **94a**) are membrane permeable, and can be used to stabilize intracellular PPIs.[118] Non-membrane permeable analogues bissuccinimidyl suberate (BS3, **93b**) and dithiobis(sulfosuccinimidyl propionate) (DTSSP, **94b**) can be used to crosslink cell-surface proteins. Cleavable crosslinkers (*e.g.* disuccinimidyl sulfoxide, DSSO, **95**) which have specific fragmentation patterns in the mass spectrometer have the advantage of facilitating identification of crosslinked products based on characteristic fragment ions.[119] Yet, all of these highly reactive bifunctional reagents described are susceptible to hydrolysis in aqueous solutions, limiting crosslinking yields and hindering detection of low abundance proteins. In addition, as their reactivity is biased towards lysine residues, crosslinking between interacting partners requires two lysine residues to be in proximity on the interacting surfaces.

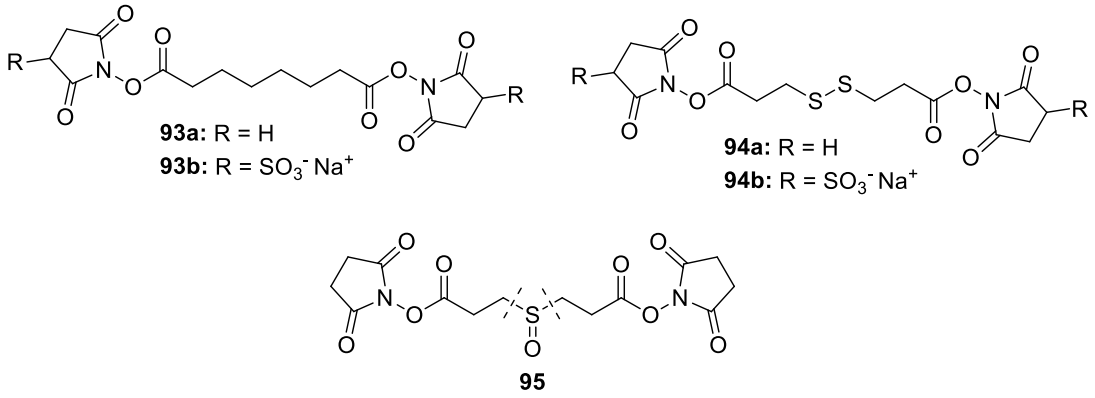


Figure 33: *Examples of homobifunctional chemical crosslinkers, including non-cleavable (**93a** and **93b**), thiol-cleavable (**94a** and **94b**) and MS-cleavable reagents (**95**).*

Photo-reactive reagents which are inert under biological conditions, yet activated upon irradiation with UV light, hold promise for conducting crosslinking under *in vivo* conditions. Heterobifunctional crosslinkers (*Figure 34*) such as SDA (**96a**) or Sulfo-SDA (**96b**) can be used to fix protein complexes upon UV irradiation, allowing temporal control over crosslinking. Photo-reactive amino acid analogues such as photo-leucine (**97**) and photo-methionine (**98**) can be metabolically incorporated into proteins, enabling PPIs to be trapped *in vivo* upon exposure to UV light.[15] The use of palladium(II) porphyrins[70] and $[Ru(II)(bpy)_3]^{2+}$ complexes[120,121,71] for protein-protein crosslinking has been described in Chapter 1 (*section 1.4.8.2*).

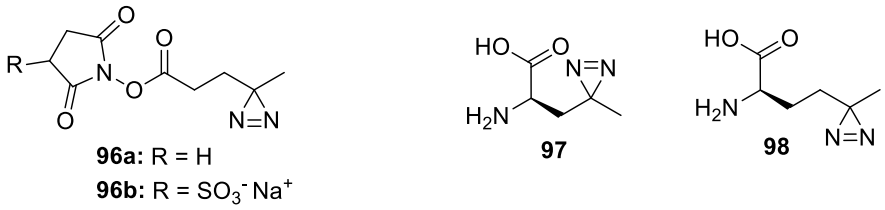


Figure 34: *Examples of photoreactive chemical crosslinkers, such as exogenous heterobifunctional crosslinkers SDA (**96a**) and Sulfo-SDA (**96b**) and photoreactive amino acids analogues **97** and **98**.*

Crosslinking reagents bearing an electrophilic group, photoreactive group, affinity tag and cleavable linker can be used in a 'label transfer' method to identify PPIs (*Figure 35*). A typical workflow starts with reaction of a protein 1 *via* the *N*-terminus or lysine residues with the NHS ester moiety of the labelling reagent (*e.g.* Sulfo-SBED, **99**, *Figure 35,* a). Protein 1 is then allowed to interact with prey proteins *in vitro*, to form stable or transient complexes, which can be covalently crosslinked upon irradiation with UV light using a photo-reactive group on the crosslinker (*e.g.* aryl azide, *Figure 35,* b). Cleavage of the spacer arm between the two proteins using reducing conditions releases the attached bait protein. Western blot and mass spectrometry analysis can be used to identify protein 2.

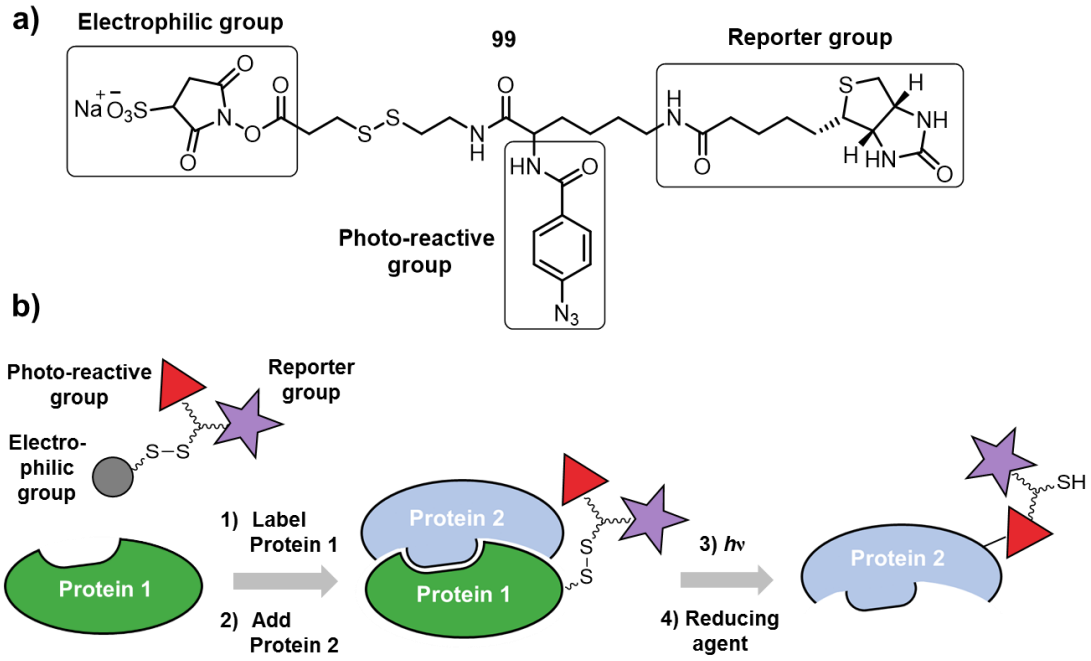


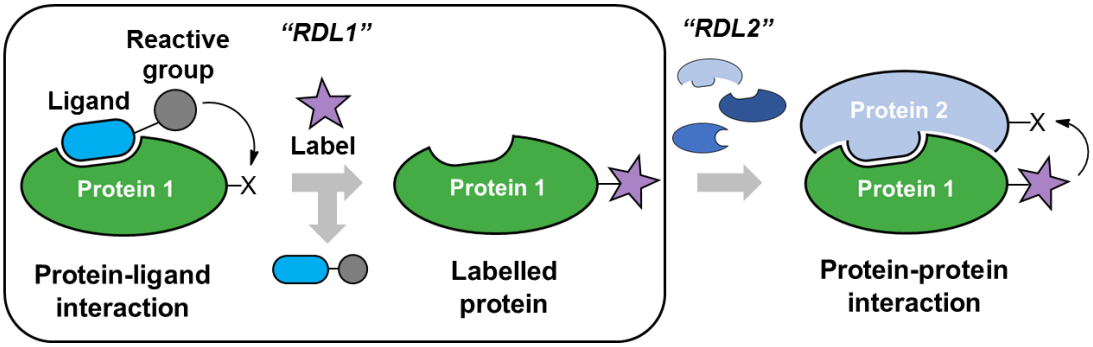
Figure 35: *Label transfer reagents for PPI identification. a) Chemical structure of Sulfo-SBED label transfer reagent. b) Schematic illustration of workflow for PPI identification using a label transfer reagent, such as* ***99****.*

Compared to high-throughput techniques such as Y2H or phage display, chemical crosslinking can afford more detailed structural information such as the interacting surfaces on proteins, even down to the amino acid. However, the non-specific reactivity of many reagents means that mass spectrometry analysis of whole proteomes can be challenging, with low abundance proteins often difficult to detect. Therefore, using a recognition element to direct labelling onto a specific protein would be advantageous for analysing the interactome of a protein of interest (see *section 1.4.1*, Chapter 1, for example).

3.1.3 Project aims

There exists an unmet need for methods by which interacting partners of a protein of interest can be identified in an unbiased manner (*section 3.1.2.*), particularly those that are transient and/or weak. As a result, the proposed concept would use traceless labelling of proteins at the endogenous level, to ultimately identify PPIs *in cells* (*Scheme 14*). Labelling of ‘protein 1’ in a manner whereby the ligand component of the probe is removed post-labelling, permits subsequent interactions with protein partners (*i.e.* ‘protein 2’); these proteins can be isolated following a second labelling event. For example, incorporation of a second reactive group into the label transferred onto ‘protein 1’ (*e.g.* a diazirine, a functional group with orthogonal

reactivity to that used in the first labelling step) would enable trapping of transient interactions.



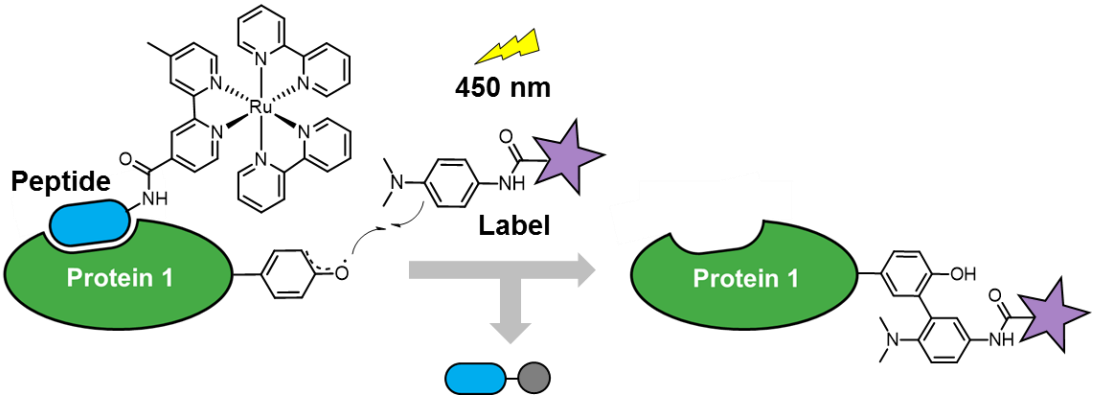
Scheme 14: *Schematic illustrating traceless labelling of proteins to identify PPIs. Two different recognition-driven labelling steps are involved: "RDL1" and "RDL2". Traceless labelling of protein 1 (via RDL1) enables interaction with protein 2. Complex formed by proteins 1 and 2 is linked by a covalent bond in RDL2, allowing transient/weak interactions to be trapped. Box indicates RDL step developed in this project: RDL1.*

3.1.4 Choice of labelling chemistry to be employed in RDL1

The proposed method outlined in *Scheme 14* requires optimisation of two separate recognition-driven labelling reactions, *RDL1* and *RDL2*. This project focuses on the development of the initial, 'traceless' labelling step (*RDL1*). The traceless affinity labelling methods developed by Hamachi and co-workers were explored in Chapter 1 (*section 1.4.*). Requirements for the chemistry employed in *RDL1* include a rapid and biocompatible reaction, synthetically tractable 'tag transfer' reagents and that labelling of protein 1 results in a protein with the capacity to bind interacting partners.

Initially, LDAI chemistry was explored (*Figure 61*, Appendix II) due to the tuneable reactivity of the acyl-imidazole moiety and faster reaction time (3 – 6 h) compared with LDT chemistry (>10 h). While synthesis of intermediates for the LDAI reagent were successful (resulting in optimisation of the route to a key building block),[122] difficulties were encountered in the final synthetic step. Furthermore, the disadvantages of this technique also became apparent: LDAI chemistry is limited to labelling proteins on the surface of live cells,[57] and acyl imidazole reagents are susceptible to non-productive hydrolysis of the carbamate bond and reactive with other nucleophiles in the cell (*i.e.* glutathione).[50] A catalytic labelling method using AGD reagents was also investigated, in collaboration with Jacob Hauser (University of Leeds), due to the lower reactant concentrations required (less toxic to cell) and the faster reaction kinetics (labelling times of <3 h).[57] Yet, the requirement for basic conditions (pH > 8) to label intracellular proteins due to protonation of the DMAP groups under physiological pH limits the use of AGD chemistry for labelling in live cells.

An alternative labelling chemistry based on metal-complex mediated photocatalysis was investigated, based on recent literature (see Chapter 1, *section 1.4.8.2*).[74,76] Ru(II)(bpy)$_3$-conjugated small molecules have been used to label intracellular proteins such as CA II and EGFR in live cells, with short reaction times (5-15 minutes irradiation, at visible wavelengths). Using light to induce protein labelling presents an advantage over the acyl transfer chemistries developed by Hamachi and co-workers, as equilibration can be reached prior to initiation of label transfer, resulting in reduction of both background labelling (of proteins other than the protein of interest) and consumption of the reagent due to non-productive reactions with nucleophiles in the cell. However, Ru(II)(bpy)$_3$-mediated photolabelling of proteins using ligand-directed reagents has not been performed at sub-stoichiometric concentrations of labelling reagents.[74,76] The chemistry employed in *RDL1* utilises a Ru(II)(bpy)$_3$-conjugated peptide to mediate transfer of a label containing a dimethylaniline moiety onto an amino acid proximal to the ligand binding site (*Scheme 15*).



Scheme 15: *Schematic representation of Ru(II)(bpy)$_3$-mediated labelling of protein 1. The peptide-catalyst component is regenerated post-labelling, leaving the peptide-binding site on protein 1 free. Label (purple) contains a fluorophore or biotin.*

3.1.5 BCL-2 family

The B-cell lymphoma 2 (BCL-2) family of proteins regulate apoptotic cell death through a variety of interactions between pro-apoptotic and anti-apoptotic members (*Figure 36*); this well studied interactome presents a suitable model system to assess the labelling chemistry detailed in *section 3.1.4*. The BCL-2 members can be sub-divided based on which BH (BCL-2 homology) regions they possess, for example the multiple-domain anti-apoptotic and pro-apoptotic proteins contain *up to four* BH regions (BH1-4), while BH3-only pro-apoptotic proteins contain a single BH3 α-helical domain. Anti-apoptotic proteins such as BCL-2 (proper), B-cell lymphoma-extra-large (BCL-X$_L$) and myeloid cell leukaemia 1 (MCL-1) sequester the apoptosis-inducing

BCL-2 homology domain 3 (BH3) of pro-apoptotic BCL-2 proteins upon binding, promoting cell survival.[123] Overexpression of anti-apoptotic proteins enables cancer cells to evade programmed cell death resulting in tumour initiation, progression and resistance to therapies.[124]

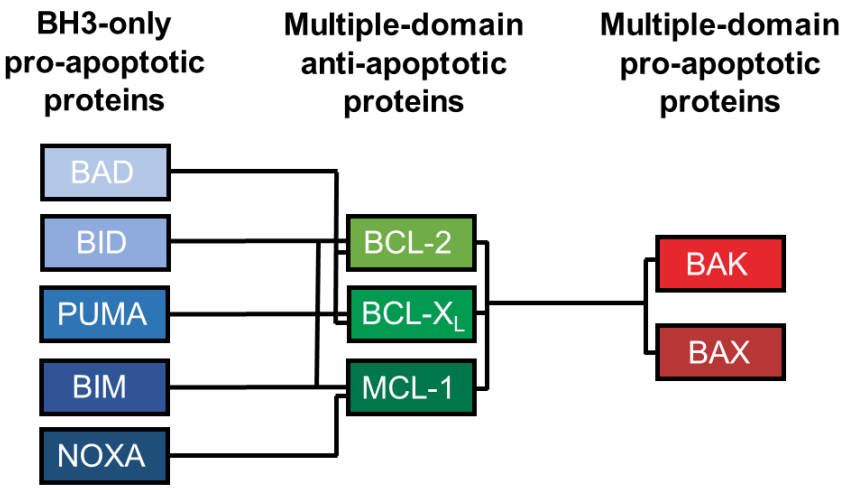


Figure 36: *Schematic illustrating the interactions between the BH3-only and multiple-domain pro-apoptotic proteins with the multiple-domain anti-apoptotic family members.*

MCL-1 is neutralised by BH3-only proteins BID, PUMA, BIM and NOXA, which bind to a hydrophobic groove on the surface of MCL-1 (common to all anti-apoptotic members) *via* their BH3 α-helices.[125,126] While BID, BIM and PUMA engage the BH3-binding groove of the multiple-domain anti-apoptotic BCL-2 proteins with comparable affinities, the NOXA BH3 domain selectively binds MCL-1, inducing MCL-1 degradation by the proteasome.[127] A NOXA-$B_{75\text{-}93(C75A)}$ peptide containing 19 amino acids from the BH3 binding region of NOXA has been shown to bind to MCL-1 with high affinity and selectivity (*Figure 37*).[128]

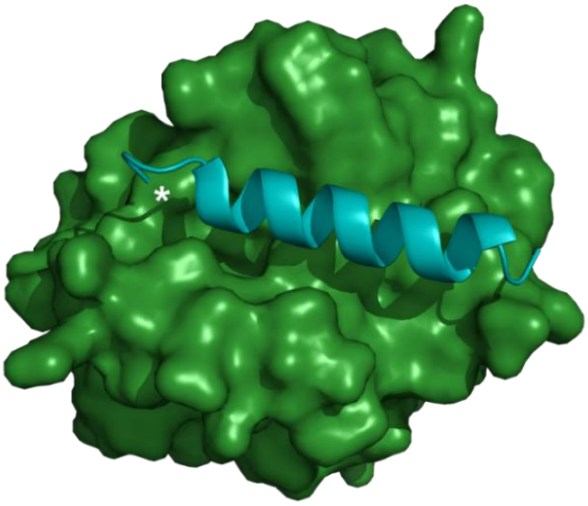


Figure 37: *Model system used in this work (MCL-1/NOXA-B, PDB:2NLA, MCL-1 in green, NOXA-B in cyan). NOXA-B$_{75\text{-}93(C75A)}$ peptide will be used as the 'ligand' component of the tag transfer reagent, to direct labelling onto MCL-1. N-terminus of NOXA-B peptide denoted by a white star is where Ru(II)(bpy)$_3$ will be attached.*

The BCL-2 family of proteins presents a more challenging range of targets for selective protein labelling using Ru(II)(bpy)$_3$ photocatalysis compared with those labelled previously (CA II, EGFR and DHFR),[74,76] as BH3-only members bind to the same hydrophobic cleft on multidomain BCL-2 proteins. Although many of these interactions have been validated (*Figure 36*), additional interactions between family members and/or unrelated proteins may exist, presenting an opportunity to gain biological insights into a protein family with complex PPIs. Therefore, the ability of ligand-conjugated Ru(II)(bpy)$_3$ reagents to selectively label proteins other than those described previously[74,76] was investigated, in addition to whether labelling can be performed catalytically (and whether this results in traceless labelling of the BCL-2 protein). As many of the interactions between BCL-2 proteins are regarded as transient, the use of a photo-initiated reagent allows trapping of these intermediate complexes. Labelling of proteins such as MCL-1 in a traceless manner (using the reagent depicted in *Scheme 15*) should allow subsequent interactions with binding partners, which may be isolated directly through a pull-down. Alternatively, the peptide-conjugated Ru(II)(bpy)$_3$ reagents may be used to mimic a binding partner within a PPI, enabling identification of interacting partners of the peptide.

3.1.6 Design of Tag Transfer Reagent for labelling of MCL-1

The Tag Transfer Reagent designed to label MCL-1 comprises a ruthenium bipyridine complex as the 'photo-reactive group', linked *via* a 6-carbon linker to the 'ligand' component, the NOXA-B peptide: NOXA-$B_{75\text{-}93(C75A)}$ (*Figure 38*). Substitution at the *N*-terminus of this NOXA-B sequence has been shown to be tolerated, demonstrated by investigating the affinity of FITC-labelled NOXA-B (FITC-Ahx-AAQLRRIGDKVNLRQKLLN-$CONH_2$) in fluorescence anisotropy assays established by Miles *et al.* at the University of Leeds.[129]

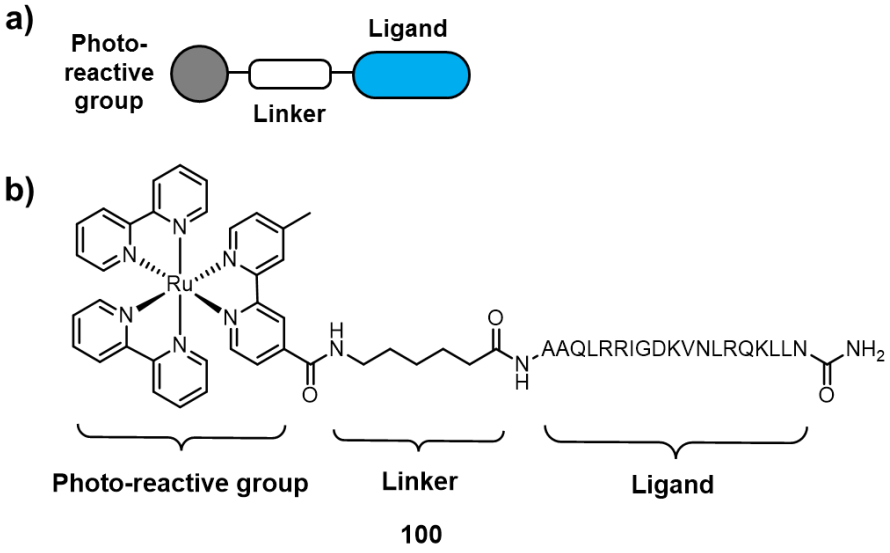
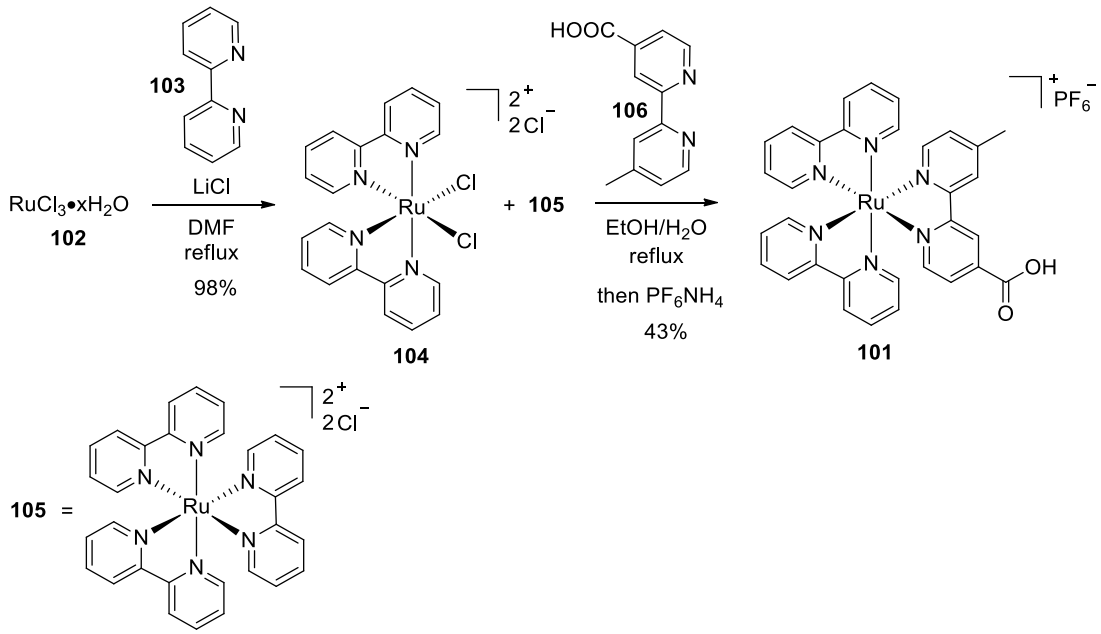


Figure 38: *a) General structure of Tag Transfer Reagent, composed of a photo-reactive element attached though a linker to a ligand for the protein of interest. b) Chemical structure of Tag Transfer Reagent* ***100*** *used in this work: Ru(II)(bpy)$_3$ complex as the photo-reactive group, aminohexanoic acid linker and NOXA-$B_{75\text{-}93(C75A)}$ as the ligand.*

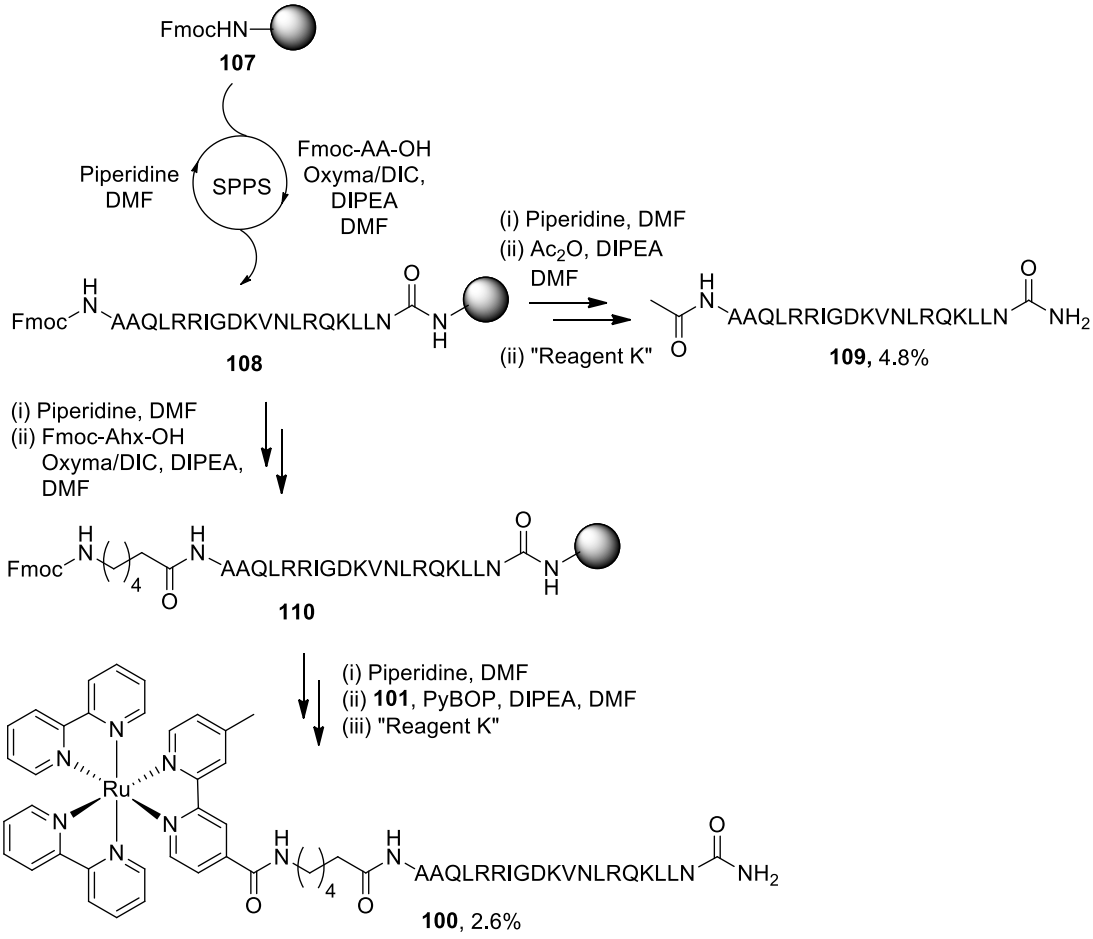
3.2 Synthesis of proof-of-principle Tag Transfer Reagent

Ruthenium(II) bipyridine complex **101** was prepared in two steps from ruthenium(III) trichloride hydrate (*Scheme 16*). Following a procedure by Meyer and co-workers,[130] dichloride complex **104** was synthesised by heating ruthenium(III) trichloride hydrate with 2,2'-bipyridine in the presence of lithium chloride (isolated as a crude mixture with the homoleptic Ru(II)$(bpy)_3$ complex **105**). Addition of acid-functionalised bipyridine ligand **106** to the mixture gave heteroleptic complex **101**. Purified **101** was isolated using silica flash column chromatography, followed by precipitation of the red hexafluorophosphate salt. Ruthenium(II) complex **101** was used without separation of enantiomers (Λ- and Δ-forms).



Scheme 16: *Synthetic route to heteroleptic ruthenium(II) bipyridine complex* ***101*** *via dichloride complex* ***104****.*

A divergent synthesis in which both acetylated ("wild-type", **109**) and ruthenium(II)-modified (**100**) NOXA-B peptides could be synthesised was devised (*Scheme 17*). Peptidyl-resin intermediate **108** was prepared using Fmoc automated solid phase peptide synthesis. Capping of **108** *via N*-terminal acetylation followed by cleavage from the resin using "Reagent K" afforded NOXA-B peptide **109**. Elongation of **108** through the addition of an aminohexanoic acid linker gave **110**, which was manually capped with Ru(II)(bpy)$_3$ complex **101**. The Ru(II)(bpy)$_3$-modifed peptide was simultaneously side-chain deprotected and cleaved from the resin using "Reagent K", precipitated and then purified by preparative mass-directed HPLC. Analysis of purified **100** by HRMS and analytical HPLC confirmed the identity and purity (>95%) of the final Tag Transfer Reagent, respectively (*Figure 39*).



Scheme 17: *Synthetic route to acetylated and Ru(II)(bpy)$_3$ derivatives of NOXA-B peptide. Cleavage cocktail "Reagent K" is composed of TFA:EDT:thioanisole:phenol:H_2O, 82:3:5:5:5.*

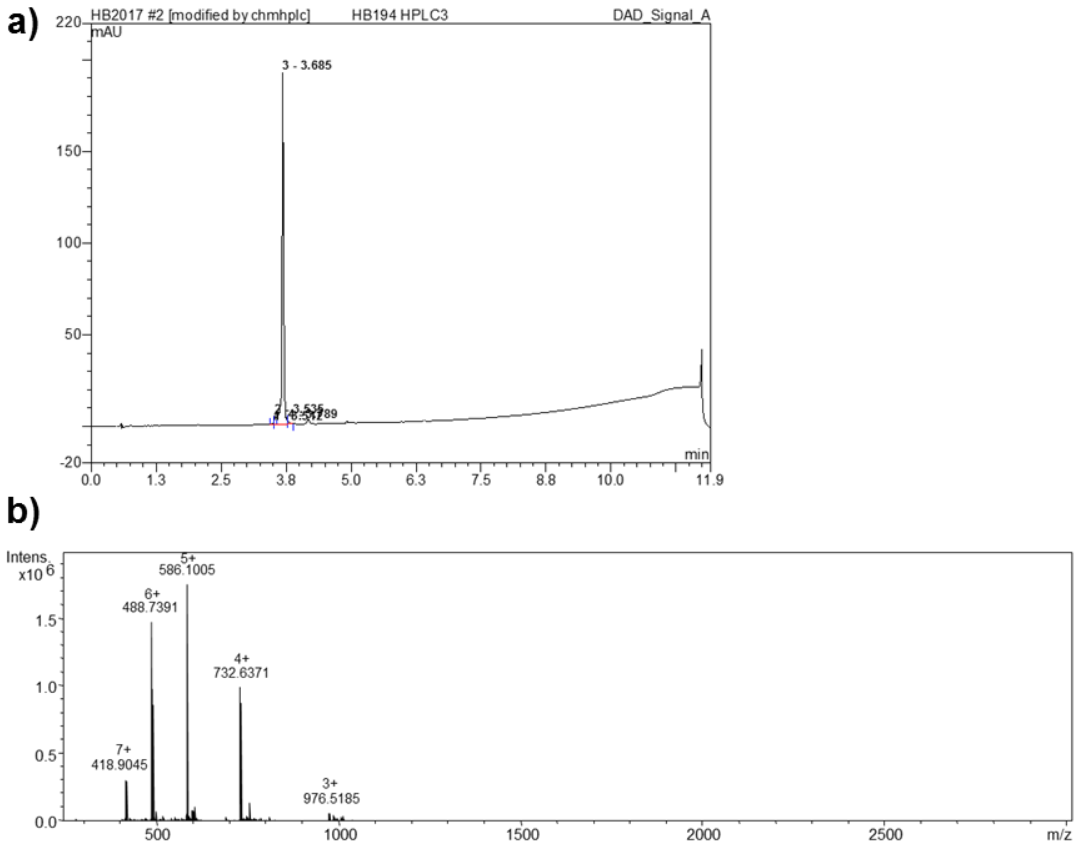


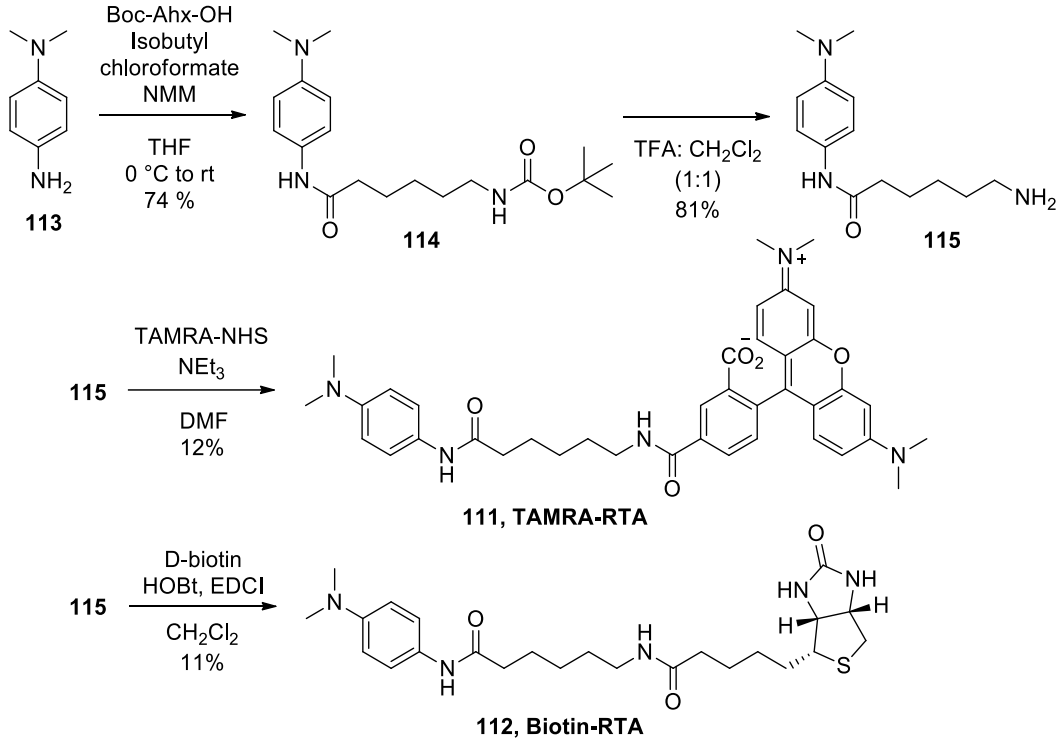
Figure 39: *Analytical data for Ru(II)(bpy)$_3$-NOXA-B peptide* ***100****. a) Analytical HPLC data for* ***100****. b) HRMS (ESI) spectrum for* ***100****.*

Table 4: *Tabulated HRMS data for* ***100*** *showing the expected (Exp^d) and observed (Obs^d) masses for the multiple charge states.*

Peptide	$[M+3H]^{3+}$ Obs^d	$[M+3H]^{3+}$ Exp^d	$[M+4H]^{4+}$ Obs^d	$[M+4H]^{4+}$ Exp^d
Ru(II)bpy$_3$-NOXA-B	976.5187	976.5162	732.6371	732.6391

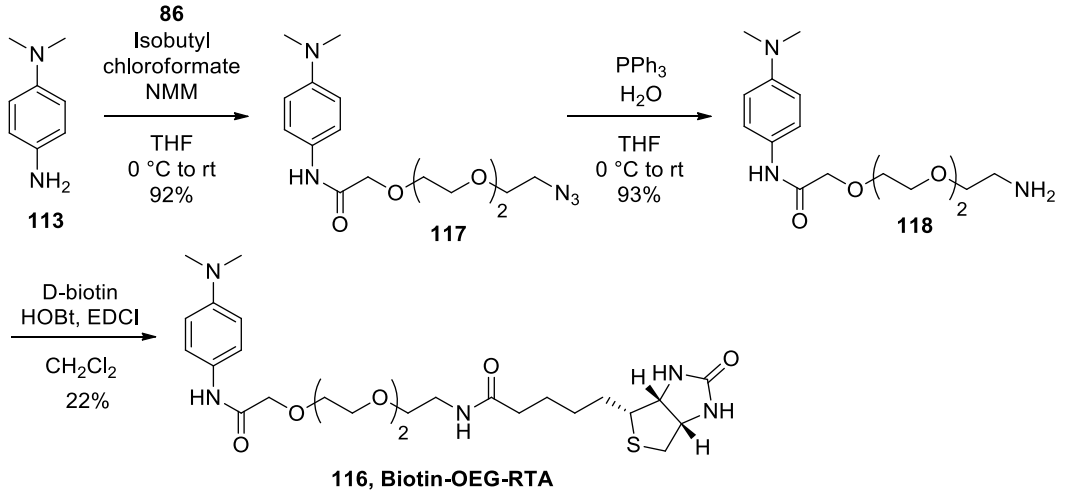
3.3 Synthesis of radical trapping agents

Novel fluorescent and biotinylated radical trapping agents (RTA's) **111** and **112**, were synthesised from common dimethylaniline intermediate **115** (*Scheme 18*). Coupling of aniline **113** with Boc-protected aminohexanoic acid *via* the mixed carbonic anhydride using isobutyl chloroformate, gave **114**, which was deprotected to give **115**. Reaction of building block **115** with TAMRA-NHS **78** (synthesis detailed in, *Scheme 11*, *section 2.8.2*, Chapter 2) or D-biotin, gave access to TAMRA-RTA **111** and Biotin-RTA **112**, respectively.



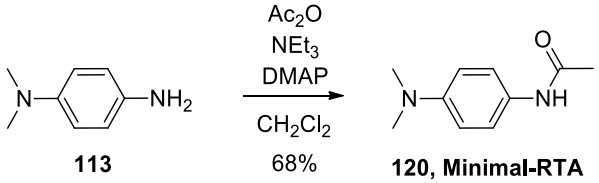
Scheme 18: *Synthetic routes to radical trapping agents **111** and **112**. Both fluorescent (**111**, TAMRA-RTA) and biotinylated (**112**, Biotin-RTA) labels were synthesised from common building block **115**, in a divergent manner.*

In an aim to increase the aqueous solubility of the biotinylated RTA, an alternative label was designed (*Scheme 19*). Incorporating an oligoethyleneglycol (OEG) linker between the dimethylaniline and biotin moieties facilitates a greater number of hydrogen-bonding interactions of the label with water, enhancing aqueous solubility. Biotin-OEG radical trapping agent **116** (Biotin-OEG-RTA) was synthesised over 3 steps. Coupling of aniline **113** with OEG linker **86** (synthesis described in *Scheme 12*, *section 2.8.2*, Chapter 2) afforded **117**. Reduction of azide **117** using Staudinger conditions gave access to amine **118**, which was coupled with D-biotin as described for **112** to give **116**.



Scheme 19: *Synthetic route to biotin OEG radical trapping agent* ***116*** *(Biotin-OEG-RTA).*

A minimal RTA was also synthesised to enable facile MS/MS analysis of modified MCL-1, to determine the site of labelling (*Scheme 20*). Acetylation of *N,N*-dimethyl-*p*-phenylenediamine **113** using acetic anhydride and DMAP gave RTA **120**, following a procedure by Sato and Nakamura.[74]



Scheme 20: *Synthesis of minimal RTA* ***120****, via acetylation of N,N-dimethyl-p-phenylenediamine* ***113****.*

3.4 Expression and purification of MCL-1 (172-327)

MCL-1 (172-327) was expressed and purified as described previously,[131] performed in collaboration with Dr Fruzsina Hobor. Overnight starter cultures produced from a glycerol stock (His-SUMO-MCL-1, prepared by Kirsten Spence) were used to produce larger scale cultures (3 L), which were induced by the addition of 0.8 mM IPTG (18 °C, overnight). Following harvesting of cells by centrifugation and lysis by sonication, the His-SUMO-MCL-1 fusion protein was purified by affinity column chromatography (HisTrap, *Figure 40, a*). The His-SUMO-MCL-1 fusion protein was cleaved using Smt3 protease, Ulp1, overnight at 4 °C. A second HisTrap column was used to remove any uncleaved MCL-1, His-SUMO and Ulp1 (*Figure 40, b*).

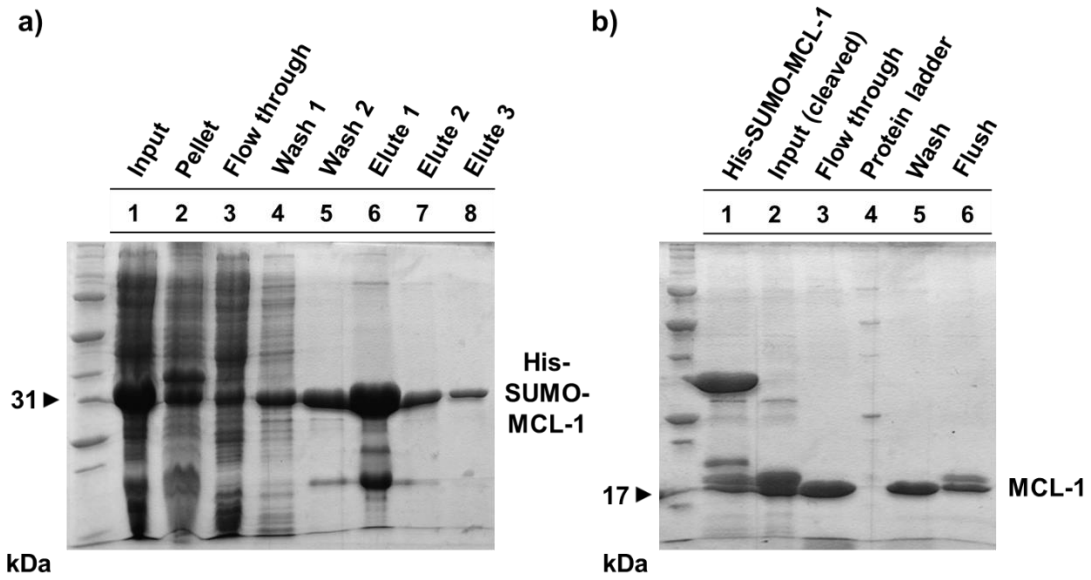


Figure 40: *SDS-PAGE gels from MCL-1 purification. a) Purification of His-SUMO-MCL-1 using a nickel column. Column was washed with 15 mM imidazole (lane 4) and 50 mM imidazole (lane 5). His-SUMO-MCL-1 was eluted in three fractions using 300 mM imidazole (lanes 6-8). b) Purification of MCL-1 using a second nickel column. Column was washed with dialysis buffer (lane 5) and flushed with 300 mM imidazole (lane 6). Fractions containing purified MCL-1 (lanes 3 and 5) were combined, concentrated and applied to a gel filtration column.*

Fractions containing MCL-1 were combined and concentrated, then purified further using a Superdex 75 gel filtration column to give 8 mL, 5.62 mg/mL (318 µM) of purified MCL-1. The purity and identity of MCL-1 was confirmed by SDS-PAGE and mass spectrometry analysis (*Figure 41, a, b*). Circular dichroism was used to assess whether the protein was correctly folded: a spectrum indicative of α-helical secondary structure was obtained (*Figure 41, c*).

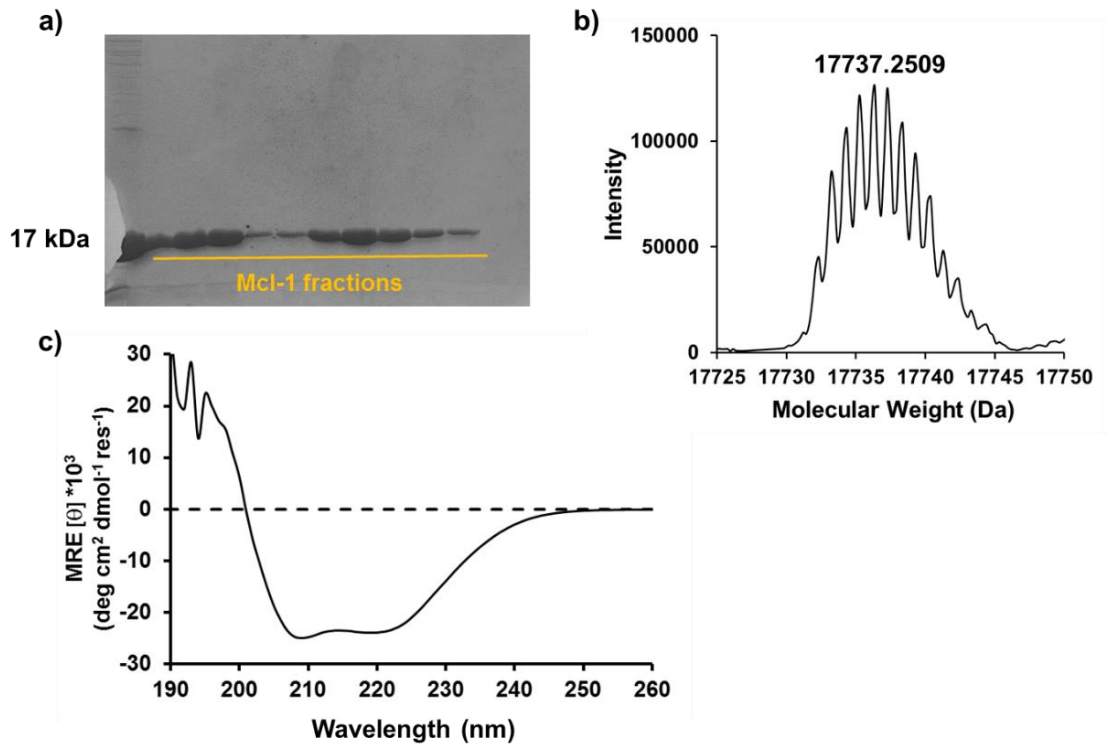


Figure 41: *Analytical data for MCL-1: a) Coomassie stained SDS-PAGE gel of fractions from gel filtration column purification b) deconvoluted HRMS of MCL-1 giving a mass of 17.73725 kDa (expected: 17.73720 kDa) c) CD spectra at 20 ºC, 50 mM sodium phosphate buffer, pH 7.5.*

3.5 Biophysical analysis of Ru(II)(bpy)$_3$-NOXA-B binding

To analyse the effect of the addition of Ru(II)(bpy)$_3$ onto NOXA-B on MCL-1 binding, fluorescence anisotropy (FA) experiments were carried out. FA works on the premise that, upon binding of a fluorescent tracer to a larger macromolecule, an increase in anisotropy is observed.[132] Titration of MCL-1 into a constant concentration of fluorescent tracer should result in a binding curve (an example of which is given in *Figure 42*, a), from which a value for K_d can be determined. Initial investigations utilised the inherent fluorescence of the Ru(II)(bpy)$_3$ complex, using Ru(II)(bpy)$_3$-NOXA-B as the fluorescent tracer in a direct binding experiment (*Figure 42, b*). However, despite the expected general trend of increased anisotropy at higher concentrations of MCL-1 (indicating a higher percentage of bound tracer), the data could not be fitted using a logistic model (*Figure 40, b*). The low quantum yield of the Ru(II)(bpy)$_3$ fluorophore (Φ_F: 0.04 in water)[133] compared to fluorescein (Φ_F: 0.79 in ethanol)[134] for example, may have accounted for the difficulties in detecting binding between MCL-1 and **100**.

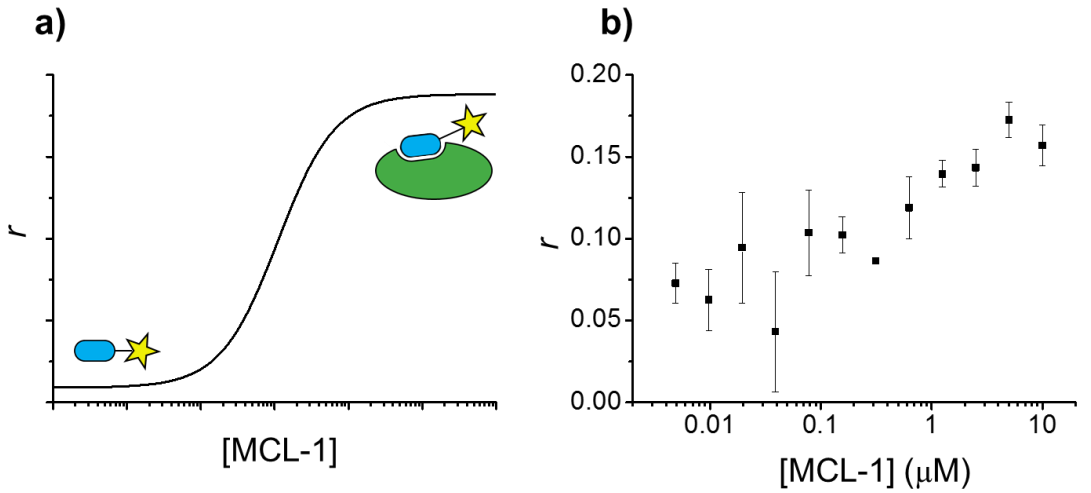


Figure 42: *Fluorescence anisotropy direct binding assays. a) Expected binding curve from fluorescence anisotropy direct titration experiment: low values of anisotropy observed when unbound tracer is in solution, high anisotropy when fluorescent tracer is bound by protein. b) Direct titration using Ru(II)(bpy)$_3$-NOXA-B as fluorescent tracer (25 nM) with [MCL-1] starting at 10 μM, diluted over 12 points in a ½ regime, showing data from 20 h incubation. Error bars represent standard deviation of three technical replicates.*

It was reasoned that a competition experiment in which peptide **100** would be used to displace a FITC-labelled version of NOXA-B (synthesised by Dr Katherine Horner) could be an alternative method to analyse binding to MCL-1.[131] Due to the sufficient difference in absorption and emission wavelengths of the two fluorophores (Ru(II)(bpy)$_3$-NOXA-B: Abs$_{max}$ 450 nm, Em$_{max}$ 620 nm; FITC-NOXA-B: Abs$_{max}$ 495 nm, Em$_{max}$ 516 nm), the FITC tracer could be excited and its emission detected without interference from the Ru(II)(bpy)$_3$ complex. Comparing the IC$_{50}$ value obtained for **100** (225 ± 21 nM, *Figure 43 and Table 5*) with that of the "wild-type" sequence (Ac-NOXA-B **109**, 1800 ± 30 nM, *Table 5*, Dr Katherine Horner, unpublished results) a nearly 10-fold increase in inhibitory potency for MCL-1 was observed, possibly the result of additional interactions between the ruthenium complex and MCL-1. BCL-X$_L$ is a related BCL-2 family protein to which NOXA-B does not bind.[131] Indeed, no effect on inhibitory potency was observed for the BCL-X$_L$/FITC-BID interaction (*Figure 41*, b), indicating selectivity of **100** for MCL-1 over BCL-X$_L$.

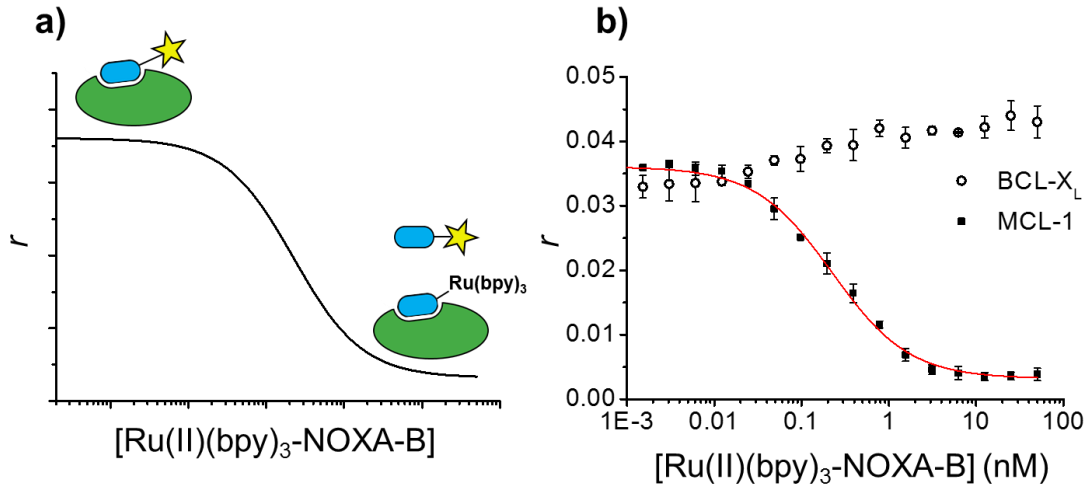


Figure 43: *Fluorescence anisotropy competition experiments. a) Expected binding curve for fluorescence anisotropy competition experiment: low levels of anisotropy are observed when competitor peptide displaces fluorescent tracer. b) Competition experiments for the inhibition of MCL-1/FITC-NOXA-B interaction (black) and BCL-X_L/FITC-BID interaction (white) by Ru(II)(bpy)$_3$-NOXA-B. Conditions for MCL-1/FITC-NOXA-B competition assay: Tris Buffer (50 mM, Tris, 150 mM NaCl, pH 7.4) + 0.01% Triton-X-100 with [MCL-1] fixed at 75 nM, data shown from 20 h incubation.* ***100*** *was serially diluted over 16 or 24 points in a 1/2 regime with [tracer] fixed at 25 nM. Conditions for BCL-X_L/FITC-BID interaction: Tris buffer (50 mM, Tris, 150 mM NaCl, pH 7.4) + 0.02 mg/mL BSA with [BCL-X_L] fixed at 50 nM, data shown from 20 h incubation. Error bars represent the standard deviation of three technical replicates.*

Table 5: *IC_{50} values for NOXA-B-derived peptides obtained from fluorescence anisotropy competition assays against MCL-1/FITC-NOXA-B and BCL-X_L/FITC-BID.*

Peptide	MCL-1/FITC-NOXA-B IC_{50} (nM)	BCL-X_L/FITC-BID IC_{50} (nM)
Ru(II)(bpy)$_3$-NOXA-B	225 ± 21	Cannot fit curve
Ac-NOXA-B	1800 ± 30	N/A

These data suggest that the addition of Ru(II)(bpy)$_3$ does not compromise binding of the modified NOXA-B peptide to MCL-1, but in fact increases the binding potency compared with the wild-type sequence by nearly 10-fold. Furthermore, Ru(II)(bpy)$_3$-NOXA-B retained selectivity for MCL-1 over a BCL-2 family member it should not bind (BCL-X_L). With the binding potency and selectivity of Tag Transfer Reagent **100** established, investigations into the ability of **100** to label MCL-1 with radical trapping agents **111**, **112** and **120** could then be undertaken.

3.6 *In vitro* photolabelling of MCL-1

3.6.1 Fluorescent modification of MCL-1

Initial labelling experiments using purified recombinant MCL-1 (172-327) were carried out *in vitro*, to optimise photolabelling conditions. Ru(II)$(bpy)_3$-NOXA-B **100**, TAMRA-RTA **111** and ammonium persulfate (APS) were added to a buffered solution of MCL-1, and irradiated for 1 minute at 450 nm. An aqueous solution of dithiothreitol (DTT, 10 μM) was then added to quench the reaction, acting as a radical scavenger.[74] Optimisation of the labelling conditions resulted in fluorescent modification of MCL-1 using 20 mol% of peptide-catalyst **100** and an equimolar concentration (relative to [MCL-1]) of fluorescent label **111** (lane 1, *Figure 44, b*). Modification of MCL-1 was only observed when both peptide and label components were used, along with irradiation of the reaction mixture. Yet, a small amount of background labelling of MCL-1 occurred in the absence of **100**, when larger amounts of protein were loaded onto the SDS-PAGE gel (*e.g. Figure 67*, Appendix III), due to the ability of rhodamine dye **111** to act as a photoredox catalyst.[135,136] Furthermore, the addition of APS to the reaction mixture was necessary for efficient labelling of MCL-1 (*Figure 66*, Appendix III). It should be noted that the concentrations of Ru(II)$(bpy)_3$ reagents used were dramatically reduced compared to previous literature, which did not exploit the catalytic nature of this chemistry ([Ru(II)$(bpy)_3$ reagent] reduced from 200 mol% to <20 mol%).[74,76]

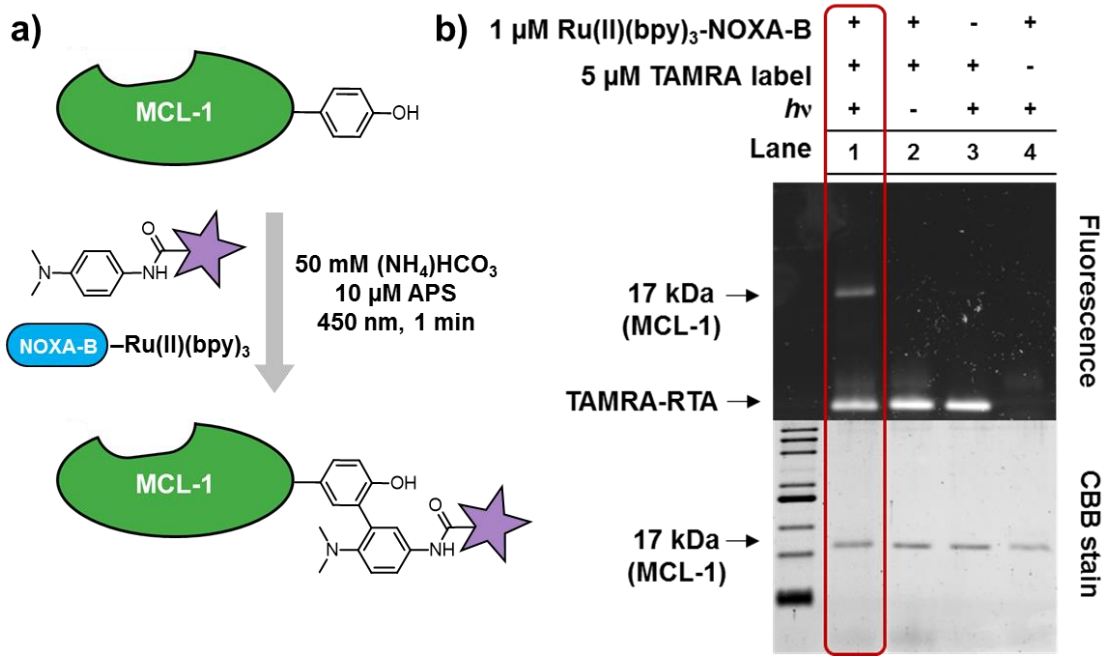


Figure 44: *Labelling of MCL-1 with TAMRA-RTA* ***111****. a) Schematic representation of labelling reaction. b) Fluorescence image and Coomassie Brilliant Blue (CBB) stained SDS-PAGE gels; fluorescence image shows MCL-1 only labelled when 1 min hv, 5 µM TAMRA-RTA and 1 µM Ru(II)(bpy)$_3$-NOXA-B are present (5 µM MCL-1 and 10 µM APS used), CBB stain shows MCL-1 for all conditions. Repeats are shown in Appendix III.*

As an orthogonal method to confirm the identity of the species indicated by the fluorescent band in the SDS-PAGE gel (*Figure 44*, b), intact protein electrospray ionisation mass spectrometry (ESI-MS) was carried out. Prior to irradiation, the deconvoluted mass spectrum of the reaction mixture exhibited unmodified MCL-1 (17737 Da, *Figure 45*, top). However, when the sample was irradiated with visible light for 1 minute, the mass spectrum displayed masses corresponding to the unmodified (17737 Da, MCL-1) and fluorescently modified (18396 Da, MCL-1 + TAMRA-RTA) protein *(Figure 45,* bottom). These data confirm the main labelled species produced results from the addition of one TAMRA-RTA label, indicating labelling of a single tyrosine residue on MCL-1. Several additional peaks were evident in the MS trace of the irradiated sample (denoted by a red star, *Figure 45,* bottom), presumably resulting from the oxidation of amino acid residues on MCL-1 proximal to the Ru(II)(bpy)$_3$ complex (indicated by masses of +16 Da, see *Figure 74****,*** Appendix III). Oxidation of proteins using ligand-directed Ru(II)(bpy)$_3$ reagents in the absence of a ‘radical trapper’ has been reported previously: chromophore-assisted light inactivation (CALI) has been used to knockdown target proteins *in vitro* and within cells.[76,137]

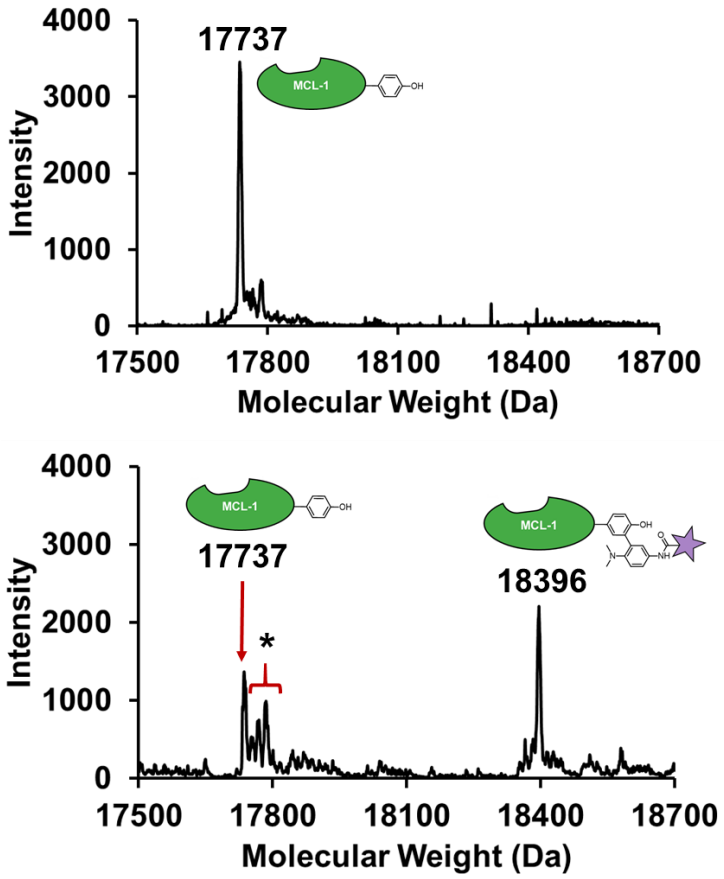


Figure 45: *Deconvoluted ESI-MS spectrum of photolabelling reaction mixture pre-irradiation (top) and post-irradiation (bottom). Unmodified MCL-1 (17737 Da) is observed in the spectrum pre-irradiation, whereas both unmodified (17737 Da) and fluorescently modified MCL-1 (18396 Da) are observed post-irradiation. Red star denotes possible oxidised protein species. Conditions: 5 µM protein, 1 µM* ***100****, 5 µM TAMRA-RTA, 10 µM APS, 1 min irradiation, 50 mM* $(NH_4)HCO_3$ *(pH 7.4).*

The data presented within *section 3.6.1.* demonstrated that peptide **100** could be used to mediate labelling of purified MCL-1 using fluorescent label **111** *in vitro* (50 mM ammonium bicarbonate buffer, pH 7.5). Analysis of the labelled mixture using in-gel fluorescence indicated that ruthenium-modified peptide **100**, label **111**, APS and visible light irradiation were necessary for efficient labelling of MCL-1, while intact protein MS analysis confirmed the identity of the labelled species visualised on the SDS-PAGE gel. With the conditions for photolabelling of MCL-1 with TAMRA-RTA **111** using peptide **100** optimised, investigations into the labelling of MCL-1 using different functional labels could be carried out.

3.6.2 Labelling of MCL-1 with biotinylated RTA

To allow labelled MCL-1 to be isolated from a mixture of proteins (*i.e.* a cell/cell lysate), MCL-1 was modified with Biotin-RTA **112**. Analysis of the crude reaction mixture using ESI-MS revealed the appearance of a distinct labelled species (18209 Da, *Figure 46,* b) relating to the mass of MCL-1 with the addition of **112** (again, modified by a single label as with TAMRA-RTA **111**, see *Figure 45*). As with the fluorescent modification, the mass spectrum of the reaction mixture prior to visible light irradiation depicted only the unmodified protein (17737 Da, *Figure 46,* b, top). Post-irradiation, a second peak became visible, relating to MCL-1 with a single Biotin-RTA label attached (18209 Da, *Figure 46*, b, bottom). It should be noted that irradiation times greater than 1 minute resulted in high degrees of conversion to oxidised protein species – of both unmodified and modified MCL-1 (see *Figure 75*, Appendix III). This effect is significant if labelling of a target protein is to be considered 'traceless' *i.e.* protein function and/or ability to bind to interacting partners is preserved. Therefore, irradiation times should be kept to a maximum of 1 minute to prevent oxidative damage of the labelled protein.

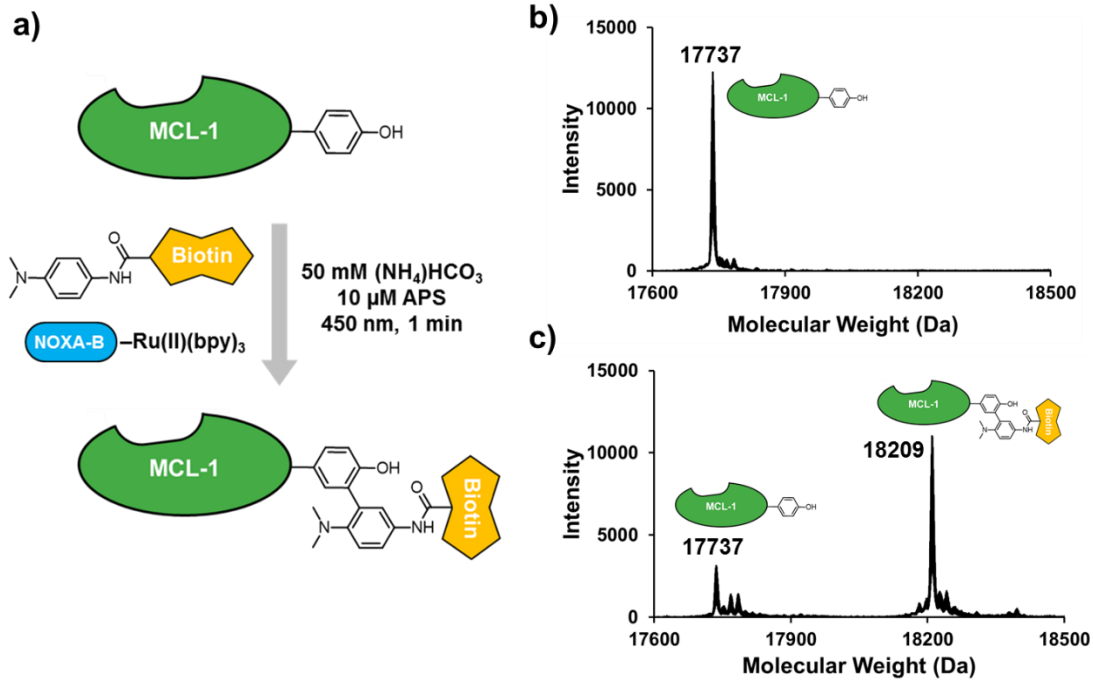


Figure 46: *Labelling of MCL-1 with Biotin-RTA **112**. a) Schematic depicting photolabelling reaction. b) ESI-MS spectrum of reaction mixture prior to irradiation (top) and post-irradiation (bottom), showing the appearance of the labelled species modified with one biotin label. Conditions: 5 µM MCL-1, 10 µM Biotin-RTA **122**, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min irradiation, 50 mM (NH$_4$)HCO$_3$ (pH 7.4).*

Biotinylated MCL-1 was subsequently isolated from the crude reaction mixture through a pull-down using avidin-agarose beads (*Figure 47*). It was reasoned that excess biotin label **112** may prevent efficient recovery of biotinylated MCL-1 from the reaction mixture. Indeed, pull-down of the modified protein following photolabelling with 100 µM of biotin-RTA **112** did not result in any detectable protein by gel staining with Coomassie Brilliant Blue (*Figure 47*, lane 1). However, removal of excess **112** prior to incubation with avidin beads (protein concentrator, 10 kDa MWCO) allowed recovery of labelled protein *via* pull-down (*Figure 47*, lane 5). In addition, photolabelling with lower concentrations of Biotin-RTA **112** allowed biotinylated MCL-1 to be subsequently detected following pull-down (*Figure 47*, lanes 2 and 3). Optimisation of the pull-down procedure though these experiments may be used to inform conditions and amounts of reagents (*e.g.* avidin-agarose beads) for isolation of labelled MCL-1 in more complex biological environments such as cell lysates.

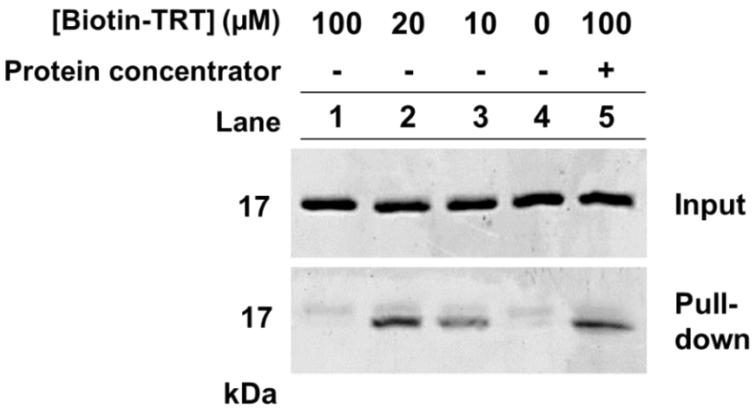


Figure 47: *Pull-down of biotinylated MCL-1 from crude labelling mixture (CBB stained gel). Conditions: 5 µM MCL-1, 0-100 µM Biotin-RTA* ***112****, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min irradiation, 50 mM (NH$_4$)HCO$_3$ (pH 7.4). Biotinylated MCL-1 is pulled-down when 20 µM and 10 µM Biotin-RTA* ***112*** *is used (lanes 2 and 3, respectively), in addition to when 100 µM Biotin-RTA* ***112*** *is used and the excess Biotin-RTA* ***112*** *is removed using a protein concentrator (lane 5). 50 µL Avidin-agarose beads used in pull-down.*

3.6.3 MS/MS analysis to determine site of labelling on MCL-1

Tandem mass spectrometry was used to identify the amino acid residues on MCL-1 that were modified with the RTA, mediated by Ru(II)(bpy)$_3$-NOXA-B. RTA **120** (*Figure 46*, a) was chosen over TAMRA-RTA **111** and Biotin-RTA **112** due to the relative simplicity of the chemical structure, avoiding complication of MS/MS spectra due to fragmentation of the label. For the photolabelling experiment, 10 µM RTA **120** was used to maximise conversion to the labelled species. Intact protein mass

spectrometry (ESI-MS) was used to analyse the labelled mixture prior to proteolytic digestion (*Figure 48*, b), both revealing the degree of labelling and confirming that a single label was added to MCL-1.

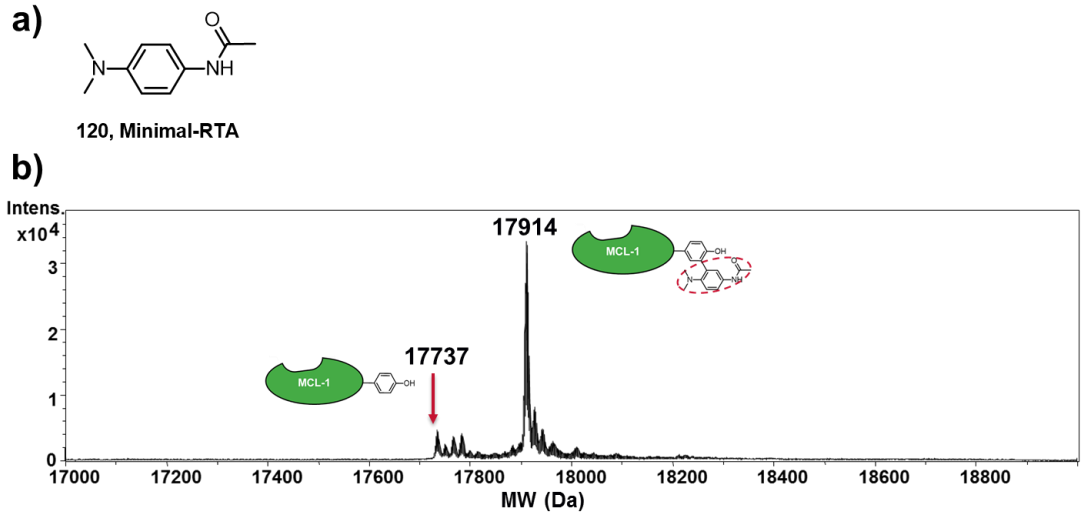


Figure 48: *Photolabelling of MCL-1 with RTA* ***120****. a) Chemical structure of RTA* ***120****. b) HRMS (ESI) spectrum of irradiated mixture: 5 µM MCL-1, 10 µM RTA 37, 1 µM Ru(bpy)$_3$-NOXA-B, 10 µM APS, 1 min irradiation, 50 mM (NH$_4$)HCO$_3$ (pH 7.4), showing unmodified MCL-1 (17737 Da) and modified MCL-1 (17914 Da).*

Proteolytic digestion of RTA-labelled MCL-1 using a mixture of Glu-C and trypsin proteases resulted in peptide fragments which were analysed using reverse-phase HPLC and Q-TOF MS/MS (performed by Dr Rachel George). Peptide mapping analysis revealed Cys286 to be the site on MCL-1 modified with **120** (*Figure 49*), despite the presence of two tyrosine residues in the MCL-1 protein sequence (Tyr175 and Tyr185, the expected sites of modification). It should be noted that Kodadek and co-workers previously reported that protein-protein crosslinking using Ru(II)(bpy)$_3$ reagents was inhibited by the addition of excess amino acids Tyr, Trp, Cys and Met,[73] suggesting these amino acids were involved in crosslink formation. However, this is the first example of a cysteine residue being modified by dimethylaniline-derived labels, mediated by a Ru(II)(bpy)$_3$ reagent. Furthermore, the single modification identified through tandem MS analysis corroborates with the intact protein MS data, which indicated the addition of a single RTA molecule onto MCL-1 (*Figure 48*).

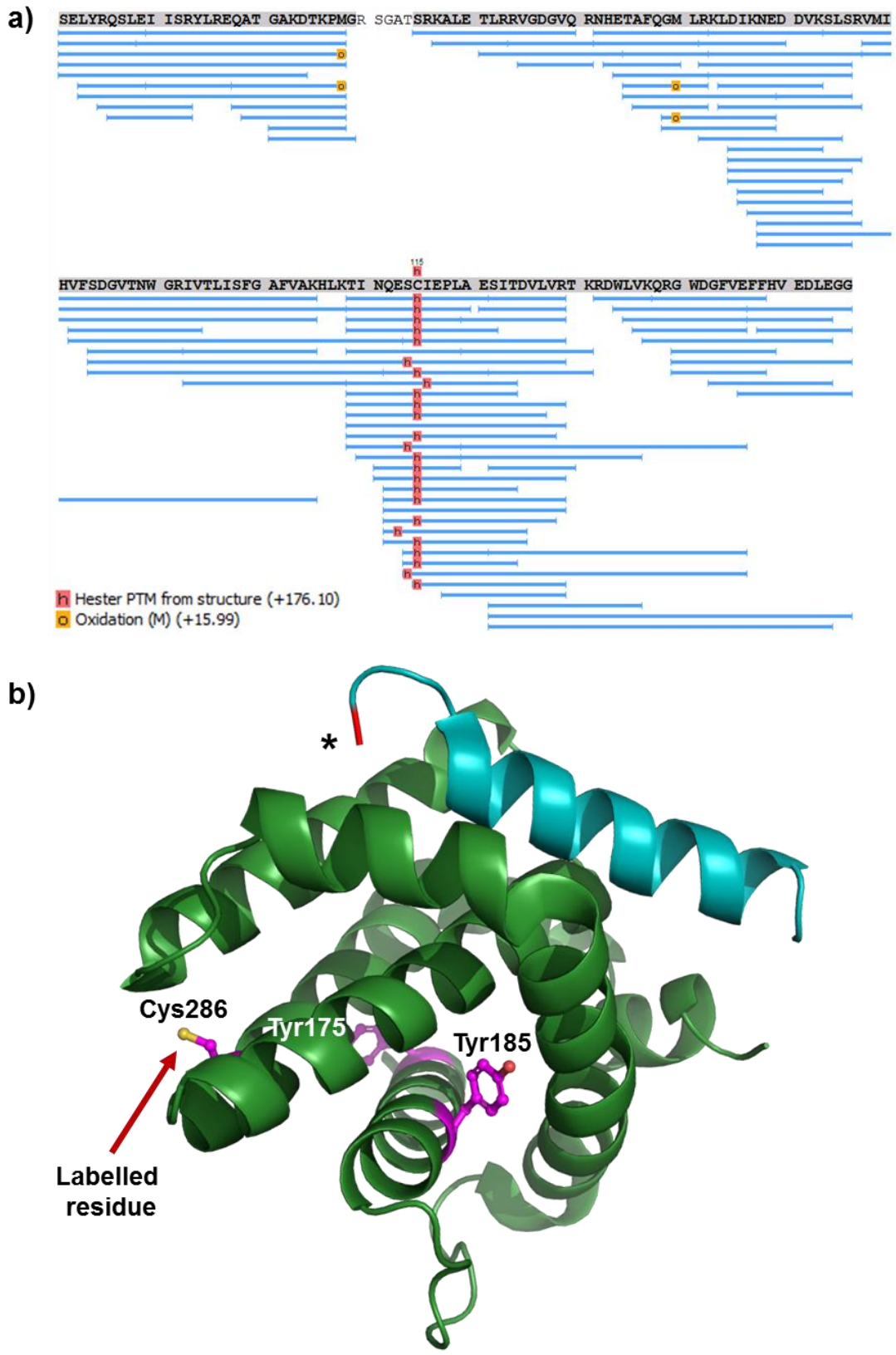


Figure 49: *MS/MS analysis of MCL-1 modified with RTA* ***120****. a) Peptide mapping of modified MCL-1 showing sequence coverage (peptides in blue) and modified residue in red (denoted by red 'h'). Sequence coverage of modified MCL-1 using a mixture of Glu-C and trypsin proteases: 94%. b) Mapping of labelled cysteine residue Cys286 (and unmodified tyrosine residues Tyr175 and Tyr185) onto MCL-1/NOXA-B structure (PDB:2NLA). MCL-1 shown in green, NOXA-B peptide shown in cyan. Star at N-terminus of NOXA-B peptide denotes position of attachment of linker and* $Ru(II)(bpy)_3$ *complex.*

3.6.4 Selective labelling of MCL-1 in protein mixtures

Previous experiments demonstrated the selectivity of peptide **100** for MCL-1 over related BCL-2 family protein BCL-X_L: no binding was observed in the fluorescence anisotropy competition experiment (*Table 5*, *section 3.5*). The ability of peptide **100** to selectively label MCL-1 *in vitro* was then explored. In a stoichiometric mixture of MCL-1 and BCL-X_L, MCL-1 was labelled to a significantly higher degree than BCL-X_L (lane 1, *Figure 50*). Yet, no fluorescent modification of BCL-X_L was observed in a solution of solely BCL-X_L (lane 5, *Figure 50*). However, it should be noted that labelling of BCL-X_L was observed when higher concentrations of TAMRA label **111** were used (see *Figure 71*, Appendix III). Again, a small amount of background labelling was seen when Ru(II)$(bpy)_3$-NOXA-B **100** was omitted. MCL-1 was also selectively labelled in a stoichiometric mixture of MCL-1 and *h*DM2, a regulator of the p53 tumour suppressor (*Figure 70*, Appendix III).[138] These results indicate a 'ligand-directed' mode of labelling, *i.e.* binding of Ru(II)$(bpy)_3$-NOXA-B to MCL-1 facilitates selective labelling of MCL-1 over proteins it does not bind to. However, competition experiments were carried out to confirm this hypothesis.

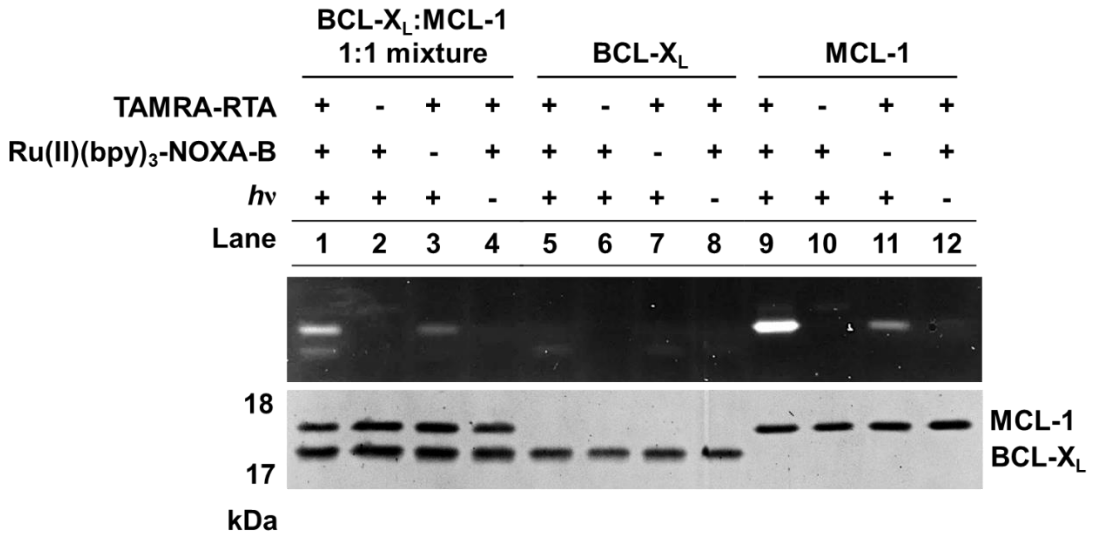


Figure 50: *MCL-1 is selectively labelled over BCL-X_L in a 1:1 mixture of proteins (lane 1). A small amount of background labelling of BCL-X_L is observed (lane 1). BCL-X_L is not labelled in solution of the protein on its own. Replicate experiments are shown in Appendix II. Conditions: 5 µM protein, 1 µM **100**, 5 µM TAMRA-RTA, 10 µM APS, 1 min irradiation, 50 mM $(NH_4)HCO_3$ (pH 7.4).*

3.6.5 Competition experiments to confirm ligand-directed nature of labelling

To confirm that labelling of MCL-1 was indeed mediated by peptide binding, bringing the ruthenium complex in proximity to the tyrosine residue to be labelled, competition experiments were undertaken. Initially, Ac-NOXA-B **109** was used as the competitor peptide. However, due to the reduced binding potency of **109** compared with ruthenium(II)-modified peptide **100** (*Table 5*, *section 3.5*), excess Ac-NOXA-B failed to compete off the more strongly binding **100**. As a result, a peptide that is known to bind in the same hydrophobic groove on MCL-1 as NOXA-B, but that possesses a more comparable IC_{50} value to that of **100** (Ac-BID, IC_{50} 390 ± 80 nM,[131] synthesised by Dr Claire Grison) was chosen. Increasing concentrations (0-1000 µM) of Ac-BID were added to the reaction mixture prior to visible light irradiation, resulting in decreasing amounts of fluorescently modified protein - with labelling completely abolished using 1000X Ac-BID (*Figure 51*, a). The intensity of the fluorescent and Coomassie-stained bands was quantified using ImageJ software, and the data from four independent repeats were plotted against the concentration of competitor peptide, revealing a dose-dependent inhibition of labelling (*Figure 51*, b).

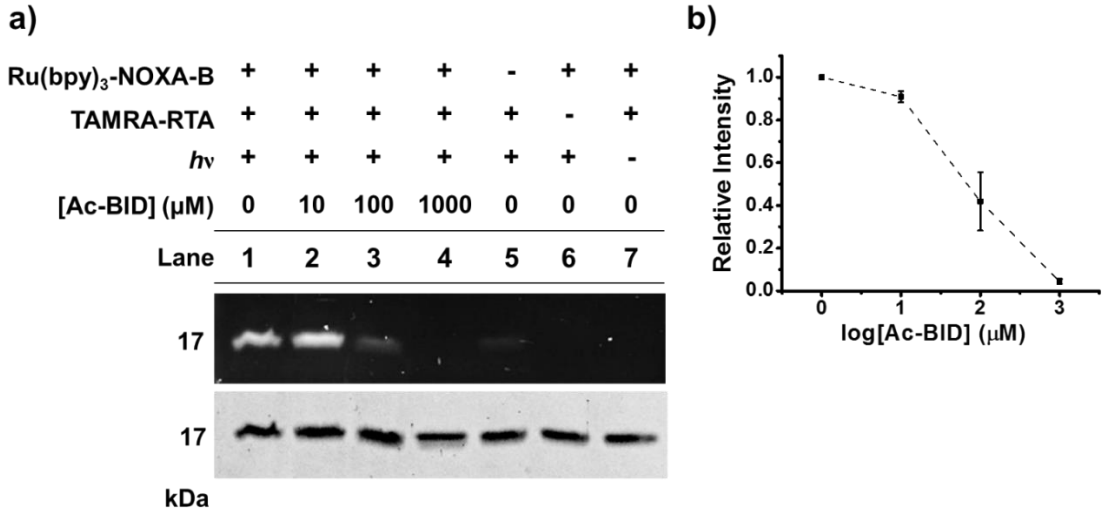


Figure 51: *Competition experiments using Ac-BID. Conditions: 5 µM MCL-1, 1 µM* ***100****, 5 µM TAMRA-RTA, 10 µM APS, 0-1000 µM Ac-BID, 1 min irradiation, 50 mM* $(NH_4)HCO_3$ *(pH 7.4). a) SDS-PAGE gel showing results of competition experiment. b) Plot of relative fluorescence intensity for the competition experiment with different concentrations of Ac-BID. Error bars represent the standard deviations from 4 independent repeats.*

3.7 Conclusions

In summary, a novel reagent for traceless labelling of NOXA-B-interacting proteins has been developed; Ru(II)(bpy)$_3$-NOXA-B was designed to label anti-apoptotic BCL-2 protein, MCL-1. Following synthesis of Ru(II)(bpy)$_3$-modified peptide **100** using solid-phase peptide synthesis, fluorescence anisotropy competition experiments confirmed that the binding potency and selectivity for MCL-1 were retained relative to wild-type peptide **109**. Rapid and selective photolabelling of recombinant MCL-1 with novel fluorescent (TAMRA-RTA **111**) and biotinylated (Biotin-RTA **112**) dimethylaniline derivatives was achieved, with irradiation times <1 min. In-gel fluorescence and ESI-MS analysis of fluorescently modified MCL-1 enabled quantification of labelling *in vitro*, and confirmed selective labelling over BCL-X_L. Photolabelling of MCL-1 was inhibited in a dose-dependent manner upon addition of competitor peptide Ac-BID, confirming the ligand-directed nature of the labelling reaction. Finally, tandem MS experiments revealed a single cysteine residue (Cys185) was modified with minimal dimethylaniline label **120**.

The work presented in this chapter builds on previous literature describing selective photolabelling of proteins using Ru(II)(bpy)$_3$-based ligand-directed reagents.[74,76,77] Notably, prior investigations did not exploit the catalytic nature of this photoredox chemistry, employing large excesses of both Ru(II)(bpy)$_3$ reagents (200 mol%, relative to [protein]) and labels (100-200X [protein]) for photolabelling of recombinant proteins *in vitro*. Conversely, 20 mol% of Ru(II)(bpy)$_3$-NOXA-B was sufficient for *in vitro* labelling of MCL-1 with stoichiometric concentrations of labels (TAMRA-RTA **111**, Biotin-RTA **112** and minimal-RTA **120**) relative to MCL-1 concentration. Furthermore, due to the conserved sequence homology between BCL-2 family members, this network of PPIs presented a significant challenge for selective protein labelling compared with structurally distinct proteins CA II, EGFR and DHFR labelled previously in the literature. In addition, photolabelling of a target protein has been achieved using a peptide to mimic an interacting partner within a PPI.

In comparison to traditional chemical crosslinking methods to analyse PPIs, photolabelling of proteins mediated by Ru(II)(bpy)$_3$-modified peptides such as **100** presents a number of advantages. Compared with non-specific reagents such as DSSO **95**, SDA **96a** and Sulfo-SBED **99** (described in *section 3.1.2.3*), ligand-directed labelling of a protein of interest within a complex mixture may facilitate

analysis of its individual interactome. Additionally, reagent **100** is unreactive in biological media, activated only upon irradiation at a specific time point, while electrophilic groups such as NHS esters or maleimides are susceptible to hydrolysis/reaction with bulk nucleophiles in the cellular environment. Finally, the sub-stoichiometric quantities of $Ru(II)(bpy)_3$ peptide required for labelling are less likely to perturb the system under study.

A number of points still require consideration; these will form the basis of future studies. For example, the cell-permeability and stability (*e.g.* in the presence of thiols such as glutathione) of reagent **100** in a cellular context, despite previous examples of $Ru(II)(bpy)_3$ photolabelling in cells.[76,74,77] Experiments described in this chapter found that APS was necessary for efficient protein labelling *in vitro*, exposing a possible limitation of this method for labelling in whole cells. However, previous literature describing intracellular protein labelling in the absence of APS comment on an alternative pathway whereby molecular oxygen acts as the electron acceptor.[73,74] Furthermore, the distance limits of the SET reaction have not been fully explored due to the limited number of protein targets that have been labelled using $Ru(II)(bpy)_3$ ligand-directed reagents. Development of a FRET assay to measure the ability of TAMRA-labelled MCL-1 to bind FITC-NOXA-B is currently underway, to establish whether labelling is truly ‘traceless’. Further developments in photolabelling using $Ru(II)(bpy)_3$-modified peptides may ultimately enable identification of novel transient/weak PPIs, *in cellulo* at the endogenous level.

Thesis summary and future directions

This thesis has explored two projects in which affinity-guided labelling of proteins was employed for different purposes. In both instances, photo-reactive probes were used to label target proteins using a recognition unit for the protein of interest; a small-molecule was employed for this purpose in Chapter 2, and a peptide mimicking a protein within a PPI was used in Chapter 3.

Chapter 2 described how chemical tools were used as part of a multidisciplinary approach to elucidate the mechanism of action of the small molecule KHS101 in glioblastoma cells, including the molecular target responsible for the cell-death phenotype. The development of robust synthetic routes to KHS101, inactive analogue HB072, functional probes KHS101-BP, KHS101-biotin, KHS101-TAMRA and KHS101-diazirine enabled investigations into the interaction between KHS101 and its target protein. Importantly, photocrosslinking in whole GBM cells using KHS101-BP identified HSPD1 as a candidate target, which was validated *in vitro* using pull-downs with KHS101-biotin. Preliminary experiments involving photocrosslinking of recombinant HSPD1 with KHS101-diazirine and analysis by tandem MS revealed a potential KHS101-binding site on HSPD1, although additional experiments are necessary to validate these findings.

Future efforts will focus on understanding the KHS101-HSPD1 interaction in greater detail. Although a binding affinity for the KHS101-HSPD1 interaction could not be determined through the SPR and FA methods described in Chapter 2, the use of an alternative fluorescent tracer containing a more rigid linker between the fluorophore and ligand may enable determination of a K_d (work currently underway). Additionally, development of such a biophysical assay would offer a more high-throughput alternative to the phenotypic assays currently used to test analogues of KHS101, allowing for the rapid screening of new compounds against HSPD1. Validation of the binding site may help to further understand how KHS101 functionally inhibits the activity of HSPD1 and could aid structure-based design of small molecule inhibitors. Ultimately, this project resulted in the discovery of HSPD1 as a potential therapeutic target for the treatment of glioblastoma multiforme.

Chapter 3 detailed the development of a photocatalytic reagent for the traceless labelling of anti-apoptotic protein MCL-1, to study the PPIs formed by BCL-2 family members. Sub-stoichiometric quantities of Ru(II)$(bpy)_3$-NOXA-B were used to rapidly label anti-apoptotic protein MCL-1 with a variety of novel functional tags, *in vitro*. Fluorescent labelling of MCL-1 using TAMRA-RTA was analysed by in-gel

fluorescence, while intact protein MS was used to confirm the identity of the labelled species. The fact that MCL-1 was fluorescently modified over BCL-X_L (and *h*DM2) and that labelling was inhibited by the addition of competitor peptide Ac-BID points towards a 'ligand-directed' mode of labelling (*i.e.* labelling of the target protein is mediated by ligand binding, bringing the reactive group in proximity to the residue to be labelled). Cys286 was established as the amino acid modified upon labelling, consistent with the expected sites of modification presented in the literature.[73] This work presents the first example of Ru(II)$(bpy)_3$-mediated photolabelling of a protein using a peptide to mimic an interacting partner within a PPI.

Certainly, a number of features must be addressed before this technique can be used as a general method for identifying PPIs *in cellulo*. Future studies will therefore involve testing the ability of Ru(II)$(bpy)_3$-NOXA-B to label MCL-1 in complex biological mixtures such cell lysates and then in whole cells. Potential limitations of using peptide **100** to label proteins in cells include the unknown cell-permeability of Ru(II)$(bpy)_3$-NOXA-B, the necessity for APS as a co-oxidant for efficient labelling of MCL-1 (*in vitro*) and whether the presence of thiols in the cellular environment hampers labelling. Furthermore, it remains to be established whether the labelled MCL-1 retains the ability to interact with binding partners post-labelling. This point is currently being investigated through the optimisation of a FRET assay in which the ability of TAMRA-labelled MCL-1 (FRET acceptor) to bind FITC-NOXA-B (donor dye) is measured. Finally, it would be interesting to apply this method to labelling other proteins (and PPIs) to establish the distance limits of the SET reaction and how applicable this method is to other systems.

Experimental section

5.1 General materials and methods for organic synthesis

All chemical reagents were purchased from commercial suppliers and used without further purification. When used as reaction solvents, THF and CH_2Cl_2 were dried and deoxygenated using an Innovative Technology Inc. PureSolv® solvent purification system.

Flash column chromatography was carried out using silica (Merck Geduran silica gel, 35–70 μm particles) according to the method described by Still, Kahn and Mitra.[139] Thin layer chromatography was carried out on commercially available pre-coated aluminium plates (Merck silica 2 8 8 0 Kieselgel 60 F_{254}). Lyophilisation of compounds was performed using a Virtis Benchtop K freeze dryer. Automated RP (C_{18}) chromatography was performed using a Biotage Isolera™ Prime Advanced Flash Purification system with RediSep R_f C_{18} columns (26 g or 130 g; Teledyne Isco). Compounds were eluted with the indicated solvent mixtures/gradients; elution was monitored by UV detection and fractions were analysed by LC-ESI-MS before combination and concentration of fractions containing pure product.

Preparative high performance liquid chromatography (HPLC) was performed on an Agilent 1260 Infinity Series equipped with a UV detector and Ascentis Express C_{18} reverse phase column using MeCN/water (5→95%) containing 0.1% formic acid. A flow rate of 0.5 mL min^{-1} was used over a period of 8 minutes. Analytical HPLC was performed on an Agilent 1290 Infinity Series equipped with a UV detector and a Hyperclone C_{18} reverse phase column using MeCN/water (5→95%) containing 0.1% formic acid, at either 0.5 mL min^{-1} over a period of five minutes or 1.0 mL min^{-1} over a period of 30 minutes.

High resolution electrospray (ESI+) mass spectrometry was performed on a Bruker MaXis Impact QTOF mass spectrometer, and *m/z* values are reported in Daltons to four decimal places. LC-ESI-MS data were recorded on an Agilent Technologies 1200 series HPLC combined with a Bruker HCT Ultra ion trap using 50 × 20 mm C_{18} reverse phase columns using MeCN/water (5→95%) containing 0.1% formic acid. A flow rate of 1.5 mL min^{-1} was used and *m/z* values are given in Daltons to one decimal place.

Microwave reactions were performed using a Discovery Explorer 48 Autosampler (CEM Corporation) single-mode microwave instrument producing controlled irradiation at 2450 MHz, using standard sealed microwave glass vessels.

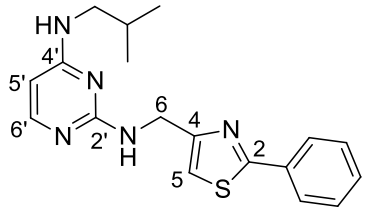
The instrument was operated in temperature control mode and reaction temperatures were monitored with an IR sensor on the outside wall of the reaction vessel. Reaction times refer to hold times at the indicated temperatures, not to total irradiation times.

^{1}H and ^{13}C NMR spectra were recorded in deuterated solvents on a Bruker Avance 500 or Bruker Avance DPX 300. Chemical shifts are quoted in parts per million downfield of tetramethylsilane and referenced to residual solvent peaks ($CDCl_3$: ^{1}H = 7.26 ppm, ^{13}C = 77.16 ppm, DMSO-d_6: ^{1}H = 2.50 ppm, ^{13}C = 39.52 ppm) and coupling constants (*J*) are reported to the nearest 0.1 Hz. Assignment of spectra was based on expected chemical shifts and coupling constants, aided by COSY, HSQC and HMBC measurements where appropriate.

5.2 Experimental section for Chapter 2

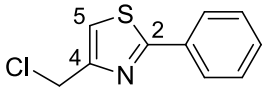
5.2.1 Organic synthesis

*N*4-(2-methylpropyl)-*N*2-[(2-phenyl-1,3-thiazol-4-yl)methyl]pyrimidine-2,4- diamine (KHS101) 41[84]



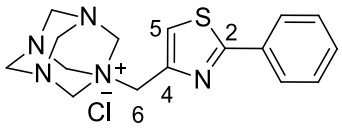
To a solution of 1-(2-phenyl-1,3-thiazole-4-yl)methanamine dihydrochloride (784 mg, 2.90 mmol) in 1-butanol (8 mL) was added 2-chloro-*N*-(2-methylpropyl)pyrimidin-4-amine (460 mg, 2.48 mmol) and DIPEA (1.47 mL, 8.43 mmol). The reaction mixture was heated at reflux for 21 hours. After cooling to ambient temperature, the reaction mixture was concentrated *in vacuo*. The crude product was purified by reverse phase column chromatography (C_{18}, eluting with H_2O (+0.1% HCOOH)-MeCN (+0.1% HCOOH) 5:95→50:50 to give the *title compound* as a colourless solid (546 mg, 65%). δ_H (500 MHz, DMSO-d_6) 8.23 (1H, app. s, 6'-C*H*), 8.00–7.86 (2H, m, Ar-C*H*), 7.65 (1H, s, N*H*), 7.55–7.43 (3H, m, Ar-C*H*), 7.35 (1H, s, 5-C*H*), 5.84 (1H, d, *J* 6.1, 5'-C*H*), 4.61 (2H, d, *J* 5.6, 6-CH_2), 3.06 (2H, app. s, isobutyl-CH_2), 1.75 (1H, app. s, isobutyl-C*H*), 0.83 (6H, app. s, isobutyl-CH_3); δ_C (125 MHz, DMSO-d_6) 206.6 (*C*-2), 164.2 (*C*-4), 162.9 (*C*H-6'), 157.0 (*C*-Ar), 133.4 (*C*H-Ar), 130.2 (*C*H-5), 129.3 (*C*H-Ar), 126.1 (*C*H-Ar), 114.9 (*C*H-5'), 41.41 (*C*H_2-isobutyl), 30.8 (*C*H_2-6), 27.89 (*C*H-isobutyl), 20.31 (*C*H_3-isobutyl), *C*-2' and *C*-4' not observed; HRMS (ESI): calcd. for $C_{18}H_{22}N_5S$ $[M+H]^+$ 340.1590, found 340.1594.

4-(chloromethyl)-2-phenyl-1,3-thiazole 44[90]



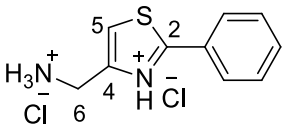
Thiobenzamide **43** (5.4 g, 39.4 mmol) was added to a solution of 1,3-dichloroacetone **42** (5.0 g, 39.4 mmol) in absolute EtOH (170 mL). The reaction mixture was heated at reflux for 2 hours. After removal of the solvents *in vacuo*, the yellow residue was stirred in Et_2O (150 mL) and saturated aqueous $NaHCO_3$ (150 mL) for 5 minutes. The layers were separated, and the aqueous layer was extracted with Et_2O (3 × 60 mL). The combined organic layers were washed with brine (180 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo* to give the *title compound* as an orange oil (7.8 g, 95%). δ_H (300 MHz, $CDCl_3$) 7.98–7.91 (2H, m, Ar-C*H*), 7.47–7.39 (3H, m, Ar-C*H*), 7.29 (1H, t, *J* 0.8, 5-C*H*), 4.75 (2H, d, *J* 0.8, CH_2); δ_C (125 MHz, $CDCl_3$) 169.2 (*C*-2), 154.0 (*C*-4), 131.1 (*C*-Ar), 129.5 (*C*H-Ar), 127.3 (*C*H-Ar), 127.0 (*C*H-Ar), 116.4 (*C*H-5), 31.2 (*C*H_2).

1-[(2-phenyl-1,3-thiazol-4-yl)methyl]-1,3,5,7-tetraazatricyclo[3.3.1.1^{3,7}] decan-1-ium chloride 46



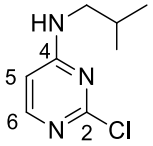
To a solution of **44** (7.5 g, 35.9 mmol) in $CHCl_3$ (55 mL) was added hexamethylene tetramine **45** (5.0 g, 35.9 mmol) and the resulting reaction mixture was stirred for 2 hours at reflux. After cooling to 0 °C, the resulting white solid was isolated by filtration, washed with cold $CHCl_3$ (1 × 35 mL) and Et_2O (1 × 35 mL). The filtrate was concentrated to ~5 mL and cold Et_2O (20 mL) was added, giving a white precipitate, which was isolated by filtration as a white powder. This process was repeated a total of four times and the solids from all filtrations were combined and dried *in vacuo* to give the *title compound* as a white powder (7.2 g, 57%). δ_H (300 MHz, $CDCl_3$) 8.21 (1H, s, 5-C*H*), 8.02–7.86 (2H, m, Ar-C*H*), 7.55–7.40 (3H, m, Ar-C*H*), 5.57 (6H, s, methenamine-CH_2), 4.90 (2H, s, 6-CH_2), 4.70– 4.43 (6H, m, methenamine-CH_2); δ_C (125 MHz, $CDCl_3$) 170.3 (*C*-2), 143.3 (*C*-4), 133.0 (*C*-Ar), 131.2 (*C*H-Ar), 129.5 (*C*H-Ar), 129.9 (*C*H-Ar), 125.6 (*C*H-5), 79.8 (*C*H_2-methenamine), 72.0 (*C*H_2-methenamine), 55.0 (*C*H_2-6); HRMS (ESI): calcd. for $C_{16}H_{20}N_5S$ $[M]^+$ 314.1434, found 314.1435.

1-(2-phenyl-1,3-thiazole-4-yl)methanamine dihydrochloride 47



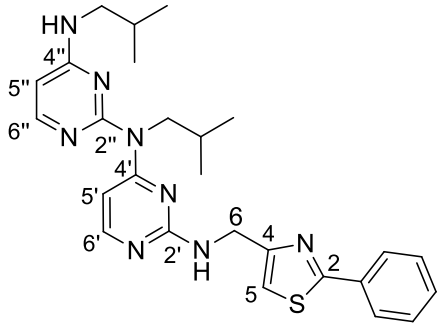
To a solution of **46** (7.1 g, 20.4 mmol) in absolute EtOH (75 mL) was added concentrated hydrochloric acid (10 mL) and the reaction mixture was stirred at reflux for 1 hour. After cooling to ambient temperature, the solid (ammonium chloride) was filtered off, washed with EtOH (3 × 20 mL) and the filtrate was concentrated *in vacuo*. The crude solid was re-dissolved in Et_2O (50 mL), saturated $NaHCO_3$ (100 mL) was added and the mixture stirred for 5 minutes. The layers were separated, and the aqueous layer was extracted with Et_2O (1 × 30 mL). The combined organics were dried over $MgSO_4$, filtered and the solution was treated with 2M HCl in Et_2O (5 mL). The white precipitate was filtered off and dried *in vacuo* to give the *title compound* as an off-white solid (5.2 g, 97%). δ_H (500 MHz, DMSO-d_6) 8.78 (3H, s, N*H*$_3$), 8.00–7.93 (2H, m, Ar-C*H*), 7.87 (1H, s, 5-C*H*), 7.58–7.44 (3H, m, Ar-C*H*), 4.17 (2H, m, 6-C*H*$_2$), 3.81 (1H, s, N*H*); δ_C (125 MHz, DMSO-d_6) 167.5 (*C*-2), 150.0 (*C*-4), 132.6 (*C*-Ar), 130.5 (*C*H-Ar), 129.3 (*C*H-Ar), 126.2 (*C*H-Ar), 118.8 (*C*H-5), 38.2 (*C*H$_2$-6); HRMS (ESI): calcd. for $C_{10}H_{11}N_2S$ $[M+H]^+$ 191.0637, found 191.0634.

2-chloro-*N*-(2-methylpropyl)pyrimidin-4-amine 50[84]



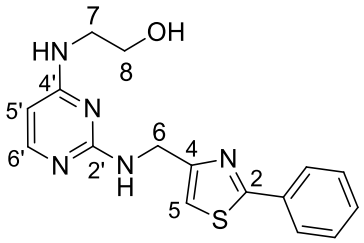
To a solution of 2,4-dichloropyrimidine **48** (4.0 g, 26.8 mmol) in EtOH (400 mL) was added 2-methylpropan-1-amine **49** (5.34 mL, 53.6 mmol). The reaction mixture was stirred at 40 °C for 3.5 hours. After this time, the reaction mixture was concentrated *in vacuo*, and the residue dissolved in EtOAc (250 mL). After cooling to 0 °C, the resulting solid (2-methylpropan-1-ammonium chloride) was filtered off and the filtrate was concentrated *in vacuo* to yield a colourless oil. The crude product was purified by flash column chromatography (SiO_2; eluting with Hexane-EtOAc 4:1→2:1) to give the *title compound* as a colourless oil, which slowly crystallised upon storage in the fridge to give colourless needles (3.7 g, 75%). δ_H (500 MHz, $CDCl_3$) 8.06 (1H, app. s, 6-C*H*), 6.26 (1H, d, *J* 6.0, 5-C*H*), 5.51 (1H, N*H*), 3.14 (2H, app. s, isobutyl-C*H*$_2$), 1.96–1.80 (1H, m, isobutyl-C*H*), 1.00 (6H, d, *J* 6.5, isobutyl-C*H*$_3$); δ_C (125 MHz, $CDCl_3$) 160.3 (*C*H-6), 109.6 (*C*H-5), 48.9 (*C*H$_2$-isobutyl), 28.3 (*C*H-isobutyl), 20.1 (*C*H$_3$-isobutyl), *C*-2 and *C*-4 not observed; HRMS (ESI): calcd. for $C_8H_{13}{}^{35}ClN_3$ $[M + H]^+$ 186.0793, found 186.0798.

_N_2,_N_4-bis(2-methylpropyl)-_N_2-(2-{[(2-phenyl-1,3-thiazol-4yl)methyl] amino}pyrimidin-4-yl)pyrimidine-2,4-diamine 51a



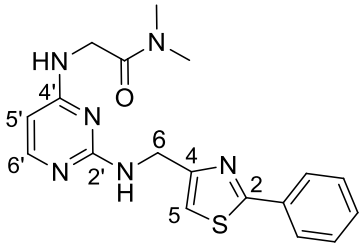
To a solution of **47** (784 mg, 2.90 mmol) in 1-butanol (8 mL) was added 2-chloro-*N*-isobutylpyrimidin-4-amine **50** (460 mg, 2.48 mmol) and DIPEA (1.47 mL, 8.43 mmol). The reaction mixture was heated at reflux for 5 days. After cooling to ambient temperature, the reaction mixture was concentrated to give a brown residue. A portion of the crude product was purified by flash column chromatography (SiO_2; eluting with 95:5 CH_2Cl_2-MeOH), then by reverse-phase column chromatography (C_{18}; eluting with H_2O (+0.1 HCOOH)-MeCN (+0.1 HCOOH) 5:95→20:80 to give the *title compound* as an off-white powder (47 mg, 61%). δ_H (500 MHz, methanol-d_4, 332 K) 9.27 (1H, d, *J* 8.0, pyrimidine-C*H*), 8.46 (1H, app. s, pyrimidine-C*H*), 8.37–8.28 (2H, m, Ar-C*H*), 7.90–7.81 (3H, m, Ar-C*H*), 7.81 (1H, s, 5-C*H*), 6.92 (1H, s, app. pyrimidine-C*H*), 6.72 (1H, d, *J* 8.0, pyrimidine-C*H*), 5.36 (2H, s, 6-CH_2), 3.74–3.71 (2H, m, isobutyl-CH_2), 3.67–3.52 (2H, m, isobutyl-CH_2), 2.36–2.24 (1H, m, isobutyl-C*H*), 1.46–1.24 (13H, m, isobutyl-C*H*, isobutyl-CH_3); δ_C (125 MHz, methanol-d_4, 332 K) 199.3 (*C*-4'), 170.5 (*C*-2), 164.7 (*C*-2''), 162.3 (*C*-4''), 157.7 (*C*-2'), 154.7 (*C*-4), 140.7 (*C*-Ar), 134.7 (*C*H-Ar), 131.5 (*C*H-6', *C*H-6''), 130.2 (*C*H-Ar), 127.5 (*C*H-Ar), 116.5 (*C*H-5), 100.9 (*C*H-5', *C*H-5''), 50.1 (*C*H_2-isobutyl), 43.3 (*C*H_2-6), 29.3 (*C*H-isobutyl), 20.5 (*C*H_3-isobutyl); IR (neat, ν_{max}/cm^{-1}) 1582, 1398, 1218; HRMS (ESI): calcd. for $C_{26}H_{33}N_8S$ $[M+H]^+$ 489.2543, found 489.2541.

2-[(2-{[(2-phenyl-1,3-thiazol-4-yl)methyl]amino}pyrimidin-4-yl)amino] ethan-1-ol 57



To a solution of amine **47** (490 mg, 1.87 mmol) in *n*-butanol was added DIPEA (0.97 mL, 5.60 mmol) and pyrimidine **56** (270 mg, 1.55 mmol) and the reaction mixture was heated at reflux for 18 hours. After cooling to ambient temperature, the solvent was removed *in vacuo* to give a brown oil. The crude product was purified by flash column chromatography (SiO_2, eluting with EtOAc-MeOH-NEt_3 99:0:1→89:10:1 to afford the *title compound* as an off-white solid (198 mg, 39%). δ_H (500 MHz, methanol-d_4) 7.99-7.87 (2H, m, Ar-C*H*), 7.60 (1H, s, 5-C*H*), 7.51-7.41 (3H, m, Ar-C*H*), 7.36 (1H, app. s, 6'-C*H*), 5.99 (1H, d, *J* 6.7, 5'-C*H*), 4.72 (2H, s, 6-CH_2), 3.65 (2H, t, *J* 5.7, 8-CH_2), 3.51 (2H, app. s, 7-CH_2); δ_C (500 MHz, methanol-d_4), 170.1 (*C*-2), 164.7 (*C*H-6'), 156.3 (*C*-4), 134.7 (*C*-Ar), 131.4 (*C*H-Ar), 130.2 (*C*H-Ar), 127.4 (*C*H-Ar), 116.3 (*C*H-5), 98.4 (*C*H-5'), 61.3 (*C*H_2-8), 44.3 (*C*H_2-7), 42.3 (*C*H_2-6), *C*-2' and *C*-4' not observed; IR (neat, ν_{max}/cm^{-1}) 3306, 3093, 2921, 2332, 1633, 1521, 1057; HRMS (ESI): calcd. for $C_{16}H_{17}N_5NaOS$ $[M+Na]^+$ 350.1046 found 350.1047.

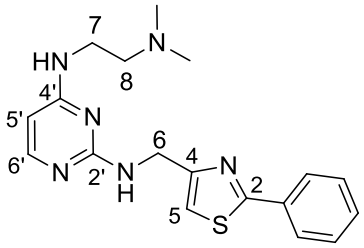
***N,N*-dimethyl-2-[(2-{[(2-phenyl-1,3-thiazol-4-yl)methyl]amino}pyrimidin-4-yl)amino]acetamide (HB072) 60**



To a solution of pyrimidine **59** (140 mg, 0.65 mmol) and amine **47** (197 mg, 0.75 mmol) in *n*-butanol was added DIPEA (0.34 mL, 1.96 mmol) and the reaction mixture was stirred at reflux overnight. After cooling to ambient temperature, the yellow precipitate in the reaction mixture was filtered off (DIPEA.HCl), and the filtrate was concentrated *in vacuo*. The solid was dissolved in a mixture of MeOH and CH_2Cl_2 (100 mL), washed with saturated aqueous $NaHCO_3$ (2 × 20 mL) and water (1 × 20 mL), dried over Na_2SO_4 and concentrated *in vacuo* to afford the *title compound* as an

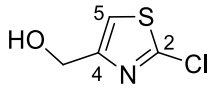
off-white solid (110 mg, 46%). δ_H (500 MHz, methanol-d_4:$CDCl_3$ 10:1) 7.96 – 7.84 (2H, m, Ar-C*H*), 7.68 (1H, d, *J* 6.0, 6'-C*H*), 7.49 – 7.37 (3H, m, Ar-C*H*), 7.17 (1H, s, 5-C*H*), 5.90 (1H, d, *J* 6.0, 5'-C*H*), 4.68 (2H, d, *J* 6.1, 6-CH_2), 4.09 (2H, s, amide-CH_2), 2.94 (3H, s, N-CH_3), 2.89 (3H, s, N-CH_3); δ_C (500 MHz, methanol-d_4-$CDCl_3$ 10:1) 171.0 (*C*-2), 169.9 (*C*=O), 164.0 (*C*-4'), 162.8 (*C*-2'), 157.6 (*C*-4), 155.6 (*C*H-6'), 134.6 (*C*-Ar), 131.4 (*C*H-Ar), 130.2 (*C*H-Ar), 127.6 (*C*H-Ar), 115.5 (*C*H-5), 96.4 (*C*H-5'), 43.0 (*C*H_2-amide), 42.8 (*C*H_2-6), 37.0 (*C*H_3-N), 36.5 (*C*H_3-N); IR (neat, ν_{max}/cm^{-1}) 3099, 2927, 1641, 1577, 1422, 1001; HRMS (ESI): calcd. for $C_{18}H_{20}N_6NaOS$ $[M+Na]^+$ 391.1312 found 391.1312.

*N*4-[2-(dimethylamino)ethyl]-*N*2-[(2-phenyl-1,3-thiazol-4-yl)methyl] pyrimidine-2,4-diamine (KHS092) 52[84]



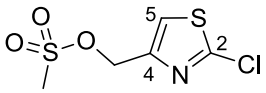
To a solution of $LiAlH_4$ (31 mg, 0.81 mmol) in THF, was added a solution of amide **60** (100 mg, 0.27 mmol) in THF (6 mL) at 0 °C. The reaction mixture was allowed to warm to ambient temperature and stirred for 45 minutes. The reaction mixture was quenched by pouring into a solution of saturated aqueous Na_2SO_4 (100 mL) at 0 °C, and stirred at ambient temperature for 1 hour. The solution was filtered and the organic solvent was removed *in vacuo*. The solution was extracted with EtOAc (2 × 30 mL), washed with an aqueous solution of Rochelle's salt (50% wt/wt), dried over Na_2SO_4 and concentrated *in vacuo* to give an off-white solid. The crude product was purified by flash column chromatography (SiO_2, eluting with EtOAc-EtOH-NH_3 91:8:1) to afford the *title compound* as a colourless solid (16 mg, 18%). R_f 0.28 (EtOAc-EtOH-NH_3 91:8:1); δ_H (500 MHz, $CDCl_3$) 8.00 – 7.86 (2H, m, Ar-C*H*), 7.81 (1H, d, *J* 5.8, 5'-C*H*), 7.50 – 7.35 (3H, m, Ar-C*H*), 7.12 (1H, s, 5-C*H*), 5.72 (1H, d, *J* 5.8, 6'-C*H*), 5.34 (1H, s, N*H*), 4.77 (2H, d, *J* 5.9, 6-CH_2), 3.35 (2H, app. s, 7-CH_2), 2.46 (2H, t, *J* 6.0, 8-CH_2), 2.22 (6H, s, N-CH_3); δ_C (125 MHz, $CDCl_3$) 168.6 (*C*-2), 163.3 (*C*H-6'), 162.3 (*C*-4), 156.6 (*C*H-5'), 134.1 (*C*H-Ar), 130.3 (*C*-Ar), 129.3 (*C*H-Ar), 126.9 (*C*H-Ar), 114.5 (*C*H-5), 58.2 (*C*H_2-8), 45.5 (*C*H_3-N), 42.4 (*C*H_2-6) 38.4 (*C*H_2-7), *C*-2' and *C*-4' not observed; IR (neat, ν_{max}/cm^{-1}) 3295, 1555, 1416, 1248; HRMS (ESI): calcd. for $C_{18}H_{24}N_6S$ $[M+2H]^{2+}$ 178.0886 found 178.0889.

(2-chlorothiazol-4-yl)methanol 65a[84]



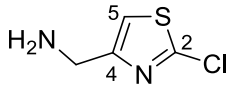
To a solution of 2-chlorothiazole-4-carboxylic acid **64a** (3 g, 18.3 mmol) in THF (22 mL) was added borane in THF (1M, 55 mL, 55.0 mmol), causing formation of bubbles during the addition. The solution was heated at reflux overnight. Then, the solution was cooled to 0 °C and MeOH (30 mL) was added using a dropping funnel. The reaction mixture was concentrated *in vacuo*. The resulting orange residue was taken up in CH_2Cl_2 (150 mL) and washed with saturated aqueous $NaHCO_3$ (3 x 25 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo* to afford an orange oil. The crude product was purified by flash column chromatography (SiO_2; eluting with EtOAc-Hexane 1:1) to give the *title compound* as a pale yellow oil (1.88 g, 69%). R_f 0.23 (EtOAc-Hexane 1:1); δ_H (500 MHz, $CDCl_3$) 7.11 (1H, s, 5-C*H*), 4.68 (2H, s, CH_2); δ_C (125 MHz, $CDCl_3$) 155.4 (*C*-4), 152.4 (*C*-2), 116.8 (*C*-5), 60.9 (*C*H_2); HRMS (ESI): calcd. for C_4H_4ClNOS $[M+H]^+$ 149.9775 found 149.9771.

(2-chloro-1,3-thiazol-4-yl)methyl methanesulfonate 66a[84]



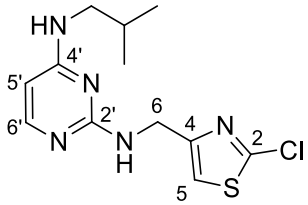
To a solution of (2-chlorothiazol-4-yl)methanol **65a** (1.8 g, 12.03 mmol) and NEt_3 (1.09 mL, 37.7 mmol) in CH_2Cl_2 (85 mL) at –5 °C, was added methanesulfonyl chloride (1.46 mL, 18.9 mmol) dropwise, over a 10-minute period. The solution was then warmed to ambient temperature and stirred for 3 hours. The resulting solution was washed with saturated aqueous $NaHCO_3$ (3 × 20 mL) and the organic layer dried over anhydrous Na_2SO_4, filtered and the filtrate concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, eluting with 40:60 EtOAc-Hexane) to afford the *title compound* as a pale orange oil (927 mg, 34%). R_f 0.28 (40:60 EtOAc-Hexane); δ_H (500 MHz, $CDCl_3$) 7.35 (1H, d, *J* 0.9, 5-C*H*), 5.19 (2H, d, *J* 0.9, CH_2), 3.02 (3H, s, CH_3); δ_C (125 MHz, $CDCl_3$) 152.7 (*C*-4), 148.2 (*C*-2), 121.5 (*C*-5), 65.9 (*C*H_2), 38.2 (*C*H_3); IR (neat, ν_{max}/cm^{-1}) 1335, 1168, 1049, 929, 524; HRMS (ESI): calcd. for $C_5H_7{}^{35}ClNO_3S_2$ $[M+H]^+$ 227.9550, found 227.9546.

(2-chloro-1,3-thiazol-4-yl)methanamine 67a[84]



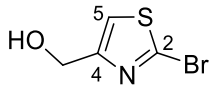
Mesylate **66a** (490 mg, 2.15 mmol) was dissolved in ammonia-saturated methanol (20 mL) and transferred to a sealed pressure vessel. The reaction mixture was warmed to 65 °C (behind a blast shield) and stirred for 72 hours. The resulting mixture was slowly cooled to –78 °C before opening the vessel. The solution was then warmed to ambient temperature, filtered and the filtrate concentrated to afford the *title compound* as an orange residue (318 mg, 99%). δ_H (500 MHz, methanol-d_4) 7.57 (1H, t, *J* 0.8, 5-C*H*), 4.14 (2H, d, *J* 0.8, CH_2); δ_C (125 MHz, methanol-d_4) 153.7 (*C*-4), 148.4 (*C*-2), 122.1 (*C*H-5), 39.2 (*C*H_2); IR (neat, ν_{max}/cm^{-1}) 3419, 1636, 1142, 745; HRMS (ESI): calcd. for $C_4H_6{}^{35}ClN_2S$ $[M+H]^+$ 148.9935, found 148.9929.

***N*2-[(2-chloro-1,3-thiazol-4-yl)methyl]-*N*4-(2-methylpropyl)pyrimidine-2,4-diamine 68a**[84]



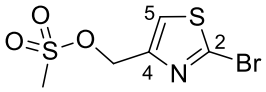
To a solution of 2-chloro-*N*-isobutylpyrimidin-4-amine **50** (400 mg, 2.16 mmol) in 1,4-dioxane (2.7 mL) was added DIPEA (1.3 mL, 7.24 mmol) and **67a** (361 mg, 2.44 mmol). The reaction mixture was stirred for 72 hours at 100 °C. After cooling to ambient temperature, the reaction mixture was concentrated to give a brown residue, which was dissolved in EtOAc (15 mL). The organic layer was washed with brine (3 × 5 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2; eluting with CH_2Cl_2-EtOH-Et_3N 90:10:1→80:20:1) to give the *title compound* as a brown solid (161 mg, 26%). R_f 0.24 (CH_2Cl_2-EtOH-Et_3N 80:20:1); δ_H (500 MHz, $CDCl_3$) 7.79 (1H, app. s, 6'-*H*), 7.05 (1H, s, 5-C*H*), 5.76 (1H, d, *J* 5.9, 5'-C*H*), 4.65 (2H, dd, *J* 6.1, 1.1, isobutyl-CH_2), 3.08 (2H, s, 6-CH_2), 1.83 (1H, m, isobutyl-C*H*), 0.95 (6H, m, isobutyl-CH_3); δ_C (125 MHz, MeOD-d_4) 155.7 (*C*-2', *C*-4'), 152.8 (*C*-4), 152.6 (*C*H-6'), 118.4 (*C*-2), 117.8 (*C*H-5), 106.5 (*C*H-5'), 47.7 (*C*H_2-isobutyl), 42.5 (*C*H_2-6), 29.5 (*C*H-isobutyl), 20.5 (*C*H_3-isobutyl); IR (neat, ν_{max}/ cm^{-1}) 3258, 1640, 748; HRMS (ESI): calcd. for $C_{12}H_{17}{}^{35}ClN_5S$ $[M+H]^+$ 298.0888, found 298.0896.

(2-bromothiazol-4-yl)methanol 65b



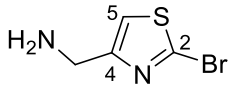
To a solution of 2-bromothiazole-4-carboxylic acid **64b** (2.5 g, 15.0 mmol) in THF (20 mL) was added borane in THF (1M, 50 mL, 50.0 mmol), causing formation of bubbles during the addition. The solution was heated at reflux overnight. Then, the solution was cooled to 0 °C, and MeOH (25 mL) was added using a dropping funnel. The reaction mixture was concentrated *in vacuo*. The resulting orange residue was taken up in CH_2Cl_2 (125 mL) and washed with saturated aqueous $NaHCO_3$ (3 x 20 mL). The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo* to afford an orange oil. The crude product was purified by flash column chromatography (SiO_2; eluting with EtOAc-Hexane 1:1) to give the *title compound* as an orange oil (1.89 g, 81%). R_f 0.24 (EtOAc-Hexane 1:1); δ_H (500 MHz, $CDCl_3$) 7.35 (1H, s, 5-C*H*), 4.55 (2H, s, CH_2); δ_C (125 MHz, $CDCl_3$) 155.9 (*C*-4), 136.9 (*C*-2), 119.2 (*C*-5), 59.7 (*C*H_2); HRMS (ESI): calcd. for C_4H_5BrNOS $[M+H]^+$ 193.9270 found 193.9268.

(2-bromo-1,3-thiazol-4-yl)methyl methanesulfonate 66b[140]



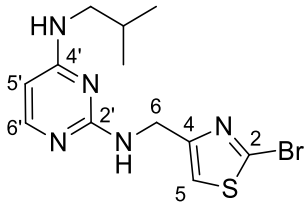
To a solution of (2-bromothiazol-4-yl)methanol **65b** (2.08 g, 10.7 mmol), NEt_3 (2 mL, 13.4 mmol) and DMAP (128 mg, 1.04 mmol) in CH_2Cl_2 (20 mL) at 0 °C, was added methanesulfonyl chloride (2.92 mL, 37.8 mmol) dropwise, over a 10-minute period. The resulting mixture was stirred at 0 ºC for 1 hour. The reaction was quenched by the addition water (60 mL) dropwise over 15 minutes. The mixture was separated, the aqueous layer extracted with CH_2Cl_2 (2 x 30 mL) and the combined organic layers dried over Na_2SO_4, filtered and concentrated *in vacuo* to afford an orange oil. The crude product was purified by flash column chromatography (Biotage, 40g SiO_2; eluting with 8%→66% EtOAc in hexanes) to afford the *title compound* as a white solid (2.52 g, 86%). R_f 0.27 (EtOAc-Hexane 1:1); δ_H (500 MHz, $CDCl_3$) 7.40 (1H, s, 5-C*H*), 5.26 (2H, s, CH_2), 3.05 (3H, s, CH_3); δ_C (125 MHz, $CDCl_3$), 149.3 (*C*-4), 136.9 (*C*-2), 123.2 (*C*H_2), 65.5 (*C*H-5), 38.1 (*C*H_3); HRMS (ESI): calcd. for $C_5H_7BrNO_3S$ $[M+H]^+$ 273.9025 found 273.9029.

(2-bromo-1,3-thiazol-4-yl)methanamine 67b



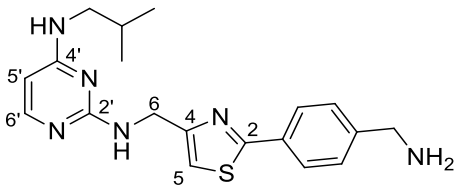
Mesylate **66b** (485 mg, 1.78 mmol) was dissolved in ammonia-saturated methanol (20 mL) and transferred to a sealed pressure vessel. The reaction mixture was warmed to 65 °C (behind a blast shield) and stirred for 48 hours. The resulting mixture was slowly cooled to –78 °C before opening the vessel. The solution was then warmed to ambient temperature, and concentrated *in vacuo* to give an orange solid. The crude product was purified by flash column chromatography (SiO_2, eluting with 90:10:1 EtOAc-MeOH-NEt_3) to afford the *title compound* as a yellow oil, which crystallised into a yellow solid (277 mg, 81 %). R_f 0.17 (90:10:1 EtOAc-MeOH-NEt_3); δ_H (500 MHz, methanol-d_4) 7.36 (1H, t, *J* 0.8, 5-C*H*), 3.88 (2H, d, *J* 0.8, CH_2); δ_C (125 MHz, methanol-d_4) 158.8 (*C*-4), 137.3 (*C*-2), 119.7 (*C*-5), 42.3 (CH_2); HRMS (ESI): calcd. for $C_4H_6BrN_2S$ $[M+H]^+$ 194.9410 found 194.9410.

*N*2-[(2-bromo-1,3-thiazol-4-yl)methyl]-*N*4-(2-methylpropyl)pyrimidine-2,4-diamine 68b



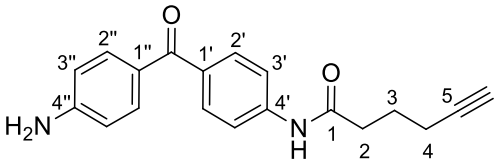
To a solution of 2-chloro-*N*-isobutylpyrimidin-4-amine **50** (171 mg, 0.92 mmol) in 1,4-dioxane (1.5 mL) was added DIPEA (0.54 mL, 3.08 mmol) and **67b** (200 mg, 1.04 mmol). The reaction mixture was stirred for 72 hours at 100 °C. After cooling to ambient temperature, the reaction mixture was concentrated to give a brown residue, which was dissolved in EtOAc (15 mL). The organic layer was washed with brine (3 × 5 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2, eluting with CH_2Cl_2-EtOH-Et_3N 90:10:1) to give the *title compound* as a light brown solid (190 g, 61%). R_f 0.16 (CH_2Cl_2-EtOH-Et_3N 90:10:1); δ_H (500 MHz, $CDCl_3$) 7.83 (1H, d, *J* 5.8, 6'-C*H*), 7.08 (1H, t, *J* 1.1, 5-C*H*), 5.71 (1H, d, *J* 5.8, 5'-C*H*), 5.42 (1H, s, N*H*), 4.77 (1H, s, N*H*), 4.68 (2H, dd, *J* 6.2, 1.1, 6-CH_2), 3.07 (2H, app. s, isobutyl-CH_2), 1.83 (1H, m, isobutyl-C*H*), 0.93 (6H, d, *J* 6.7, isobutyl-CH_3); δ_C (125 MHz, methanol-d_4) 163.4 (*C*H-6'), 162.0 (*C*-2), 156.1 (*C*H-5'), 135.5 (*C*-4), 118.2 (*C*H_2-5), 49.0 (*C*H_2-isobutyl), 41.7 (*C*H_2-6), 28.5 (*C*H-isobutyl), 20.3 (*C*H_3-isobutyl), *C*-2' and *C*-4' not observed; HRMS (ESI): calcd. for $C_{12}H_{17}BrN_5S$ $[M+H]^+$ 344.0363 found 344.0372.

*N*2-({2-[4-(aminomethyl)phenyl]-1,3-thiazol-4-yl}methyl)-*N*4-(2-methylpropyl)pyrimidine-2,4-diamine 63[84]



A mixture of **68b** (90 mg, 0.30 mmol), cesium carbonate (266 mg, 0.82 mmol), 4-(aminomethyl)phenylboronic acid hydrochloride **69** (112 mg, 0.59 mmol), and tetrakis(triphenylphosphine)palladium(0) (30 mg, 0.03 mmol) in 1,4-dioxane (5.4 mL) and water (0.9 mL) was placed in a microwave reactor vessel and purged with nitrogen for 10 minutes. The reaction mixture was heated at 150 °C for 25 minutes (hold time) in a microwave reactor. The resulting black mixture was filtered through a plug of celite and washed through with EtOAc (50 mL). The filtrate was washed with brine (3 × 20 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2; eluting with CH_2Cl_2-MeOH-NH_3 94:6:1) to give the *title compound* as a brown amorphous solid (64 mg, 57%). R_f 0.20 (CH_2Cl_2-MeOH-NH_3 94:6:1); δ_H (300 MHz, $CDCl_3$) 7.97-7.81 (2H, m, Ar-C*H*), 7.43–7.29 (3H, m, Ar-C*H*, 6'-C*H*), 7.10 (1H, t, *J* 1.0, 5-C*H*), 5.72 (1H, d, *J* 5.9, 5'-C*H*), 4.75 (2H, dd, *J* 5.9, 1.1, 6-CH_2), 3.92 (2H, m, NH_2-CH_2), 3.08 (2H, d, *J* 6.4, isobutyl-CH_2), 1.91–1.75 (1H, m, isobutyl-C*H*), 0.94 (6H, d, *J* 6.6, isobutyl-CH_3); δ_C (125 MHz, $CDCl_3$) 168.8 (*C*-2), 163.4 (*C*H-6'), 162.1 (*C*-4), 156.2 (*C*H-5'), 145.2 (*C*-Ar), 132.3 (*C*-Ar), 127.5 (*C*H-Ar), 126.9 (*C*H-Ar), 133.9 (*C*H-5), 49.2 (*C*H_2-isobutyl), 46.9 (*C*H_2-6), 41.9 (*C*H_2-NH_2), 28.3 (*C*H-isobutyl), 20.2 (*C*H_3-isobutyl), *C*-4' and *C*-2' not observed; HRMS (ESI): calcd. for $C_{19}H_{25}N_6S$ $[M+H]^+$ 369.1856, found 369.1860.

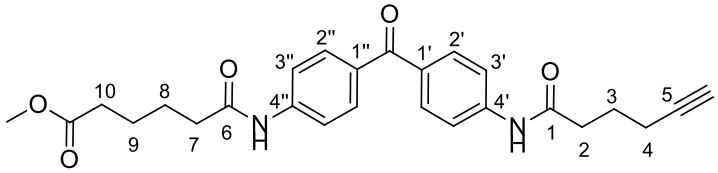
N-[4-(4-aminobenzoyl)phenyl]hex-5-ynamide 72[95]



To a solution of 5-hexynoic acid **71** (0.52 mL, 4.71 mmol) in DMF (8 mL) was added EDC·HCl (903 mg, 4.71 mmol) and HOBt (721 mg, 4.71 mmol). The mixture was stirred for 5 minutes at ambient temperature. Then, 4,4'-diaminobenzophenone **70** (1 g, 4.71 mmol) was added, and the brown reaction mixture was stirred for 16 hours at ambient temperature. The reaction mixture was poured into EtOAc (50 mL), then washed saturated aqueous $NaHCO_3$ (3 × 15 mL) and brine (1 × 15 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to give

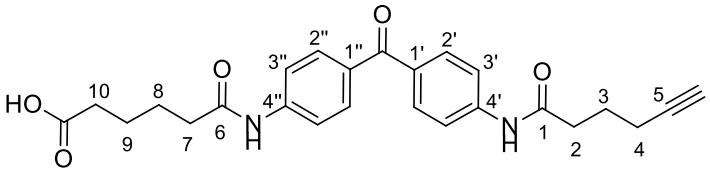
the *title compound* as a brown amorphous solid (998 mg, 69%). δ_H (500 MHz, $CDCl_3$) 8.42 (1H, broad s, N*H*), 7.81–7.51 (6H, m, 2'-C*H*, 3'-C*H*, 2''-C*H*), 6.64 (2H, dd, *J* 8.6, 4.7, 3''-C*H*), 4.26 (2H, broad s, NH_2), 2.52 (2H, t, *J* 7.4, 2-CH_2), 2.27 (2H, td, *J* 6.8, 2.6, 4-CH_2), 1.96 (1H, t, *J* 2.7, alkyne-C*H*), 1.93 (2H, m, 3-CH_2); δ_C (125 MHz, $CDCl_3$) 195.0 (*C*-ketone), 171.6 (*C*-1), 151.5 (*C*-4''), 141.8 (*C*-4'), 133.2 (*C*H-2''), 132.8 (*C*H-2'), 131.3 (*C*H-3'), 127.6 (*C*-1'), 119.2 (*C*-1''), 114.0 (*C*H-3''), 83.8 (*C*-5), 69.8 (*C*H-alkyne), 36.3 (CH_2-2), 24.3 (CH_2-3), 18.2 (CH_2-4); HRMS (ESI): calcd. for $C_{19}H_{19}N_2O_2$ $[M+H]^+$ 307.1441, found 307.1441.

Methyl 5-(4-[4-(hex-5-ynamido)benzoyl]phenyl)carbamoyl pentanoate 74[95]



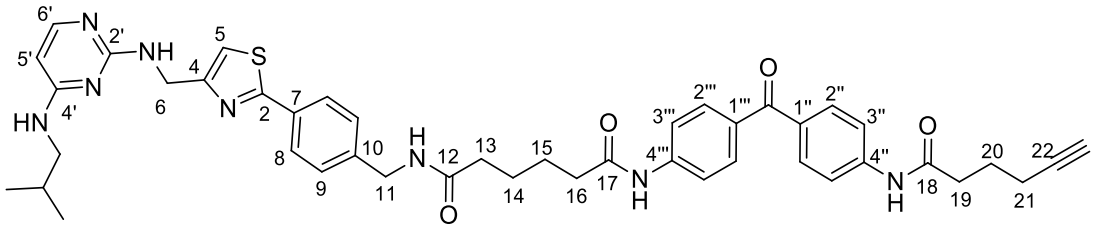
Methyl adipoyl chloride **73** (594 µL, 3.82 mmol) was added dropwise over 10 minutes to a cooled solution (0 °C) of **72** (900 mg, 2.94 mmol) and DIPEA (1.02 mL, 5.87 mmol) in THF (30 mL). The solution was allowed to warm to ambient temperature and stirred overnight. The reaction mixture was diluted with CH_2Cl_2 (100 mL), then washed with saturated aqueous $NaHCO_3$ (3 × 20 mL), saturated aqueous citric acid solution (2 × 20 mL) and brine (1 × 20 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* to give a yellow solid. The crude product was purified by flash column chromatography (SiO_2; eluting with EtOAc-Hexane 7:3), to give the *title compound* as a white solid (1.12 g, 85%). R_f 0.31 (EtOAc-Hexane 7:3); δ_H (500 MHz, $CDCl_3$) 10.29 (1H, s, N*H*), 10.24 (1H, s, N*H*), 7.88–7.62 (8H, m, 2'-C*H*, 3'-C*H*, 2''-C*H*, 3''-C*H*), 3.59 (3H, s, O-CH_3), 2.82 (1H, t, *J* 2.6, alkyne-C*H*), 2.48 (2H, t, *J* 6.6, 10-CH_2), 2.36 (4H, dt, *J* 13.6, *J* 6.6, 7-CH_2, 2-CH_2), 2.24 (2H, td, *J* 7.0, *J* 2.6, 4-CH_2), 1.78 (2H, m, 3-CH_2), 1.59 (4H, m, 8-CH_2, 9-CH_2); δ_C (125 MHz, $CDCl_3$) 198.1 (*C*-ketone), 178.4 (*C*-ester), 176.8 (*C*-6), 176.4 (*C*-1), 148.3 (*C*-4', *C*-4''), 136.9 (*C*-1', *C*-1''), 136.1 (*C*H-2', *C*H-2''), 123.4 (*C*H-3', *C*H-3''), 89.2 (*C*-5), 76.9 (*C*H-alkyne), 56.4 (CH_3-ester), 41.3 (CH_2-7), 40.4 (CH_2-2), 38.3 (CH_2-10), 29.7 (CH_2-3), 29.3 (CH_2-8), 30.0 (CH_2-9), 22.6 (CH_2-4); HRMS (ESI): calcd. for $C_{26}H_{29}N_2O_5$ $[M+H]^+$ 449.2071, found 449.2079.

5-({4-[4-(hex-5-ynamido)benzoyl]phenyl}carbamoyl)pentanoic acid 75[95]



To a solution of **74** (480 mg, 1.01 mmol) in THF (45 mL) was added an aqueous solution of NaOH (1M, 330 mL) and the reaction mixture was stirred at 60 °C for 3 hours. The organic solvent was removed *in vacuo*, and the remaining aqueous solution was acidified to pH 2 using 1M HCl. A white precipitate formed, which was filtered off and dried *in vacuo*. The aqueous layer was extracted with EtOAc (3 × 30 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to give the *title compound* (combined with the precipitate) as a yellow solid (323 mg, 69%). δ_H (500 MHz, $CDCl_3$) 7.75 (8H, app. s, 2'-C*H*, 3'-C*H*, 2''-C*H*, 3''-C*H*), 2.57 (2H, t, *J* 7.4, 2-CH_2), 2.45 (2H, t, *J* 7.2, 10-CH_2), 2.37 (2H, t, *J* 7.1, 7-CH_2), 2.34–2.27 (3H, m, 4-CH_2, alkyne-C*H*), 1.96–1.87 (2H, m, 3-CH_2), 1.81–1.66 (4H, m, 8-CH_2, 9-CH_2); δ_C (125 MHz, $CDCl_3$) 196.6 (*C*-ketone), 177.2 (*C*-(O)OH), 174.5 (*C*-6), 174.0 (*C*-1), 144.2 (*C*-4', *C*-4''), 134.0 (*C*-1', *C*-1''), 132.2 (*C*H-2', *C*H-2''), 120.1 (*C*H-3', *C*H-3''), 84.1 (*C*-5), 70.3 (*C*H-alkyne), 37.7 (*C*H_2-7), 36.7 (*C*H_2-2), 34.6 (*C*H_2-10), 26.2 (*C*H_2-8), 25.6 (*C*H_2-9), 25.5 (*C*H_2-3), 18.6 (*C*H_2-4); HRMS (ESI): calcd. for $C_{25}H_{27}N_2O_5$ $[M+H]^+$ 435.1914, found 435.1927.

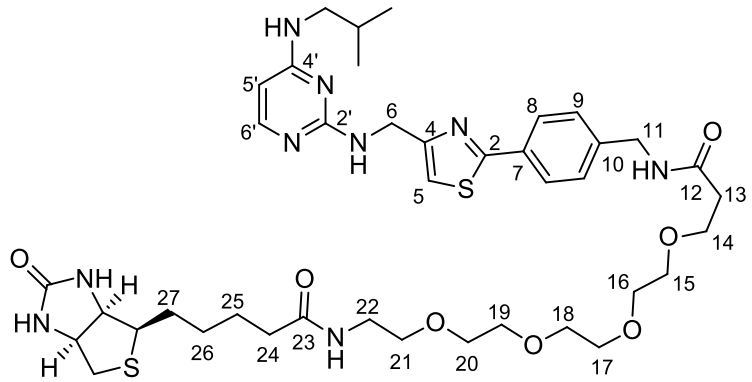
***N*-{4-[4-(hex-5-ynamido)benzoyl]phenyl}-*N'*-[(4-{4-[({4-[(2-methylpropyl) amino] pyrimidin-2-yl}amino)methyl]-1,3-thiazol-2-yl}phenyl)methyl] hexanediamide (KHS101-BP) 61**[84]



To a solution of benzophenone derivative **75** (60 mg, 0.14 mmol) in *N*-methyl-2-pyrrolidone (NMP) (3 mL) was added HATU (76 mg, 0.20 mmol) and NEt_3 (75 µL, 0.54 mmol) and the reaction mixture was stirred at ambient temperature for 5 minutes. To this mixture a solution of amine **63** (51 mg, 0.14 mmol) in NMP (1.3 mL) was added, and the reaction mixture was stirred at ambient temperature for 1.5 hours. Then, the reaction mixture was diluted with EtOAc (15 mL), washed with brine (3 × 5 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (SiO_2; eluting with

CH_2Cl_2-MeOH-NH_3 89:10:1, then a second column eluting with CH_2Cl_2-MeOH-NH_3 94:5:1) to give a clear yellow solid, which was recrystallised from methanol to give the *title compound* as colourless crystals (21 mg, 19%). R_f 0.23 (CH_2Cl_2-MeOH-NH_3 94:5:1); δ_H (500 MHz, DMSO-d_6, 348 K) 10.29 (1H, s, amide-N*H*), 10.27 (1H, s, amide-N*H*), 8.42 (1H, t, *J* 6.0, amide-N*H*), 7.81 (2H, dd, *J* 8.2, 1.7, 8-C*H*), 7.72 (4H, dd, *J* 8.8, 6.0, benzophenone-C*H*), 7.66 (4H, dd, *J* 8.8, 2.2, benzophenone-C*H*), 7.59 (1H, d, *J* 6.0, 6'-C*H*), 7.35 (2H, m, 9-C*H*), 7.23 (1H, s, 5-C*H*), 5.84 (1H, d, *J* 6.0, 5'-C*H*), 4.53 (2H, d, *J* 5.8, 11-CH_2), 4.28 (2H, d, *J* 6.0, 6-CH_2), 3.06–2.90 (1H, m, isobutyl-C*H*), 2.73 (1H, t, *J* 2.6, alkyne-C*H*), 2.45 (2H, t, *J* 7.4, 13-CH_2), 2.40–2.30 (2H, m, 16-CH_2), 2.26–2.13 (4H, m, 19-CH_2, 21-CH_2), 1.81–1.65 (4H, m, 20-CH_2, isobutyl-CH_2), 1.61–1.53 (4H, m, 15-CH_2, 14-CH_2), 0.81 (6H, app. s, isobutyl-CH_3); HRMS (ESI): calcd. for $C_{44}H_{48}N_8O_4SNa$ $[M+Na]^+$ 807.3417, found 807.3411.

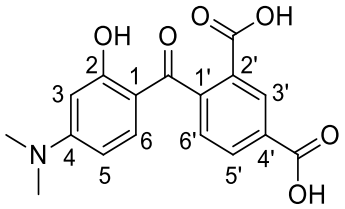
1-{5-[(3aR,4S,6aS)-2-oxo-hexahydro-1H-thieno[3,4-d]imidazolidin-4-yl] pentanamido}-*N*-[(4-{4-[({4-[(2-methylpropyl)amino]pyrimidin-2-yl} amino)methyl]-1,3-thiazol-2-yl}phenyl)methyl]-3,6,9,12-tetraoxapenta decan-15-amide (KHS101-biotin) 62



To a solution of biotin-PEG_4-NHS (22 mg, 0.04 mmol) in DMF (37 µL), was added a solution of amine **63** (17 mg, 0.04 mmol) in DMF (112 µL). The reaction mixture was stirred at ambient temperature for 20 hours, then concentrated *in vacuo* to afford a colourless solid. The crude product was purified by preparative HPLC (C_{18}, eluting with H_2O (+0.1% HCOOH)-MeCN (+0.1% HCOOH) 95:5→5:95) to give the *title compound* as a white fluffy solid (27 mg, 86%). δ_H (500 MHz, methanol-d_4) δ 8.57 (1H, s, N*H*), 7.98 – 7.89 (2H, m, d, *J* 8.1, Ar-C*H*), 7.61 (1H, app. s, 6'-C*H*), 7.45 (2H, d, *J* 8.1, Ar-C*H*), 7.37 (1H, s, 5-C*H*), 6.05 (1H, d, *J* 6.9, 5'-C*H*), 4.78 (2H, s, 6-CH_2), 4.55 – 4.44 (3H, m, 11-CH_2, biotin-C*H*), 4.32 (1H, dd, *J* 7.9, 4.5, biotin-C*H*), 3.82 (2H, t, *J* 6.0, 21-CH_2), 3.71 – 3.58 (12H, m, 20,19,18,17,16,15-CH_2), 3.55 (2H, t, *J* 5.5, 14-CH_2), 3.37 (2H, t, *J* 5.5, 22-CH_2), 3.31 – 3.17 (3H, m, isobutyl-CH_2, biotin-C*H*), 2.94 (1H, dd, *J* 12.7, 5.0, biotin-C*H*), 2.73 (1H, d, *J* 12.8, biotin-C*H*), 2.57 (2H, t, *J*

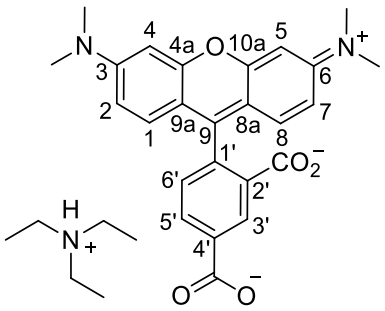
6.0, 13-CH_2), 2.23 (2H, t, *J* 7.4, 24-CH_2), 1.92 – 1.80 (1H, m, isobutyl-C*H*), 1.80 – 1.54 (4H, m, 25-CH_2, 27-CH_2), 1.52 – 1.40 (2H, m, 26-CH_2), 0.91 (6H, d, *J* 6.2, isobutyl-CH_3); δ_C (125 MHz, methanol-d_4) 176.0 (*C*-23), 174.0 (*C*-12), 169.8 (*C*-2), 166.0 (*C*(O)-biotin), 156.0 (*C*H-6'), 142.7 (*C*-10), 133.5 (*C*-7), 129.0 (*C*H-8), 127.5 (*C*H-9), 115.9 (*C*H-5), 107.2 (*C*H-5'), 71.5 – 68.2 (*C*H_2-14, 15, 16, 17, 18, 19, 20, 21), 63.3 (*C*H-biotin), 61.6 (*C*H-biotin), 56.9 (*C*H_2-isobutyl), 43.6 (*C*H_2-11), 41.0 (*C*H_2-6), 40.9 (*C*H_2-biotin), 40.3 (*C*H_2-22), 37.7 (*C*H_2-13), 36.7 (*C*H_2-24), 29.7 – 29.3 (*C*H_2-25, 26, 27), 26.8 (*C*H-isobutyl), 20.4 (2*C*H_3-isobutyl), *C*-2', *C*-4', *C*-4 and *C*H-S not observed; HRMS (ESI): calcd. for $C_{40}H_{60}N_9O_7S_2$ $[M+H]^+$ 842.4052, found 842.4059.

4-dimethylamino-2-hydroxy-2',4'-dicarboxy-benzophenone[103] **81a**



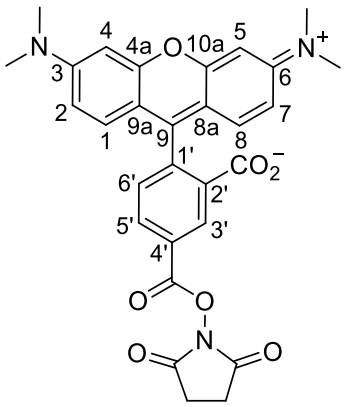
To a solution of 3-dimethylaminophenol **80** (500 mg, 3.64 mmol) in toluene (11 mL) heated to 60 °C, was added pounded benzene-1,2,4-tricarboxylic anhydride **79** (840 mg, 4.37 mmol) and the reaction mixture was heated at reflux for 24 hours. The reaction mixture was allowed to cool to ambient temperature, then the residue was filtered off, washed with toluene (3 × 5 mL), dissolved in MeOH (20 mL) and refluxed for 10 minutes. Then, acetic acid (6 mL) was added and the mixture was evaporated to dryness. The solid was recrystallized from MeOH (10 mL) to afford the *title compound* as a purple crystalline solid (323 mg, 27%). m.p. 251.5 °C (dec.), MeOH; δ_H (500 MHz, DMSO-d_6) 12.36 (1H, s, O*H*), 8.49 (1H, d, *J* 1.7, 3'-C*H*), 8.20 (1H, dd, *J* 7.8, 1.7, 5'-C*H*), 7.52 (1H, d, *J* 7.8, 6'-C*H*), 6.81 (1H, d, *J* 9.1, 6-C*H*), 6.21 (1H, dd, *J* 9.2, 2.5, 5-C*H*), 6.11 (1H, d, *J* 2.4, 3-C*H*), 3.01 (6H, s, N-CH_3); δ_C (125 MHz, DMSO-d_6) 197.6 (*C*(O)-ketone), 166.1 (*C*(O)OH-2', *C*(O)OH-4'), 164.2 (*C*-2), 155.8 (*C*-4), 143.7 (*C*-1'), 133.8 (*C*H-6), 132.6 (*C*H-5'), 131.6 (*C*-2' or *C*-4'), 130.7 (*C*H-3'), 130.0 (*C*-2' or *C*-4'), 128.2 (*C*H-6'), 109.4 (*C*-1), 104.5 (*C*H-5), 97.0 (*C*H-3); IR (neat, ν_{max}/cm^{-1}) 2937, 1729, 1633, 1495, 1350, 1212; HRMS (ESI): calcd. for $C_{17}H_{14}NO_6$ $[M-H]^-$ 328.0827 found 328.0827.

3,6-bis(dimethylamino)-9-[2,4-dicarboxylatephenyl]xanthylium, triethyl ammonium salt 82[103]



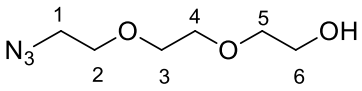
To a solution of **81a** (2.5 g, 7.60 mmol) in DMF (60 mL) was added **80** (1.35 g, 9.88 mmol) and trimethylsilylpolyphosphate solution in chloroform (15 mL, prepared according to Yokoyama *et al.*[105]). The reaction mixture was stirred at reflux for 3 hours. The solvents were removed *in vacuo*, and the residue was dissolved in an aqueous solution of NaOH (5%, 53 mL) and stirred at ambient temperature for 20 hours. The solution was diluted with water (100 mL) and the rhodamine was precipitated with concentrated HCl (~5 mL). The solid was filtered off, washed with water (40 mL) and dried *in vacuo* to give a black solid. The crude product was purified by flash column chromatography (SiO_2, eluting with Et_3N-MeOH-$CHCl_3$ 5:15:80) to afford the *title compound* as a dark purple solid (1.8 g, 98%). R_f 0.18 (Et_3N-MeOH-$CHCl_3$ 5:15:80); δ_H (500 MHz, methanol-d_4) 8.73 (1H, d, *J* 1.7, 3'-C*H*), 8.18 (1H, dd, *J* 7.8, 1.7, 5'-C*H*), 7.27 (3H, m, 1-C*H*, 8-C*H*, 6'-C*H*), 7.00 (2H, dd, *J* 9.5, 2.5, 2-C*H,* 7-C*H*), 6.88 (2H, d, *J* 2.5, 4-C*H,* 5-C*H*), 3.26 (12H, s, N-CH_3), 3.12 (9H, q, *J* 7.3, N-CH_2), 1.26 (13H, t, *J* 7.3, CH_2-CH_3); δ_C (125 MHz, methanol-d_4) 173.7 (*C*(O)OH-4'), 173.2 (*C*(O)OH-2'), 160.0 (*C*-9), 158.8 (*C*-4a, *C*-10a), 158.4 (*C*-3, *C*-6), 140.8 (*C* 1'), 136.4 (*C*-2', *C*-4'), 132.7 (*C*H-1, *C*H-8), 131.7 (*C*H-3'), 131.3 (*C*H-5'), 129.8 (*C*H-6'), 114.69, 114.65, (*C*H-2, *C*H-7, *C*-8a, *C*-9a), 97.3 (*C*H-4, *C*H-5), 47.5 (*C*H_3-N), 40.8 (*C*H_2-N), 9.2 (*C*H_3-CH_2); IR (neat, ν_{max}/cm^{-1}) 2978, 2604, 1745, 1633, 1344, 1246, 1118; HRMS (ESI): calcd. for $C_{25}H_{22}N_2NaO_5$ $[M+Na]^+$ 453.1421 found 453.1426.

2,5-dioxopyrrolidin-1-yl 3',6'-bis(dimethylamino)-3-oxo-3H-spiro[iso benzofuran-1,9'-xanthene]-5-carboxylate (named as closed form) 78[106]



To a solution of **82** (100 mg, 0.19 mmol) in CH_2Cl_2 (10 mL) was added DMAP (115 mg, 0.94 mmol) and NEt_3 (131 µL, 0.94 mmol). *N*-disuccinimidyl carbonate (96 mg, 0.38 mmol) was added and the reaction mixture was stirred at ambient temperature for 1 hour. The reaction was quenched with acetic acid (188 µL), reduced in volume *in vacuo* (to ~2 mL), loaded onto an equilibrated flash chromatography column and eluted with AcOH-Acetone 1:100 to afford the *title compound* as a dark purple solid (91 mg, 92%). R_f 0.10 (AcOH-Acetone 1:100); δ_H (500 MHz, $CDCl_3$) 8.83 (1H, d, *J* 1.5, 3'-C*H*), 8.30 (1H, dd, *J* 8.0, 1.6, 5'-C*H*), 7.29 (1H, d, *J* 8.0, 6'-C*H*), 6.75 (2H, d, *J* 8.8, 1-C*H*, 8-C*H*), 6.57 – 6.46 (4H, m, 2-C*H,* 4-C*H,* 5-C*H,* 7-C*H*), 3.04 (12H, s, N-CH_3), 2.91 (4H, t, *J* 3.1, NHS-CH_2); δ_C (125 MHz, $CDCl_3$) 176.6 (*C*(O)OH-2'), 169.1 (*C*(O)-NHS), 168.2 (*C*(O)O-4'), 161.0 (*C*-9), 153.5 (*C*-4a, *C*-10a), 153.0 (*C*-3, *C*-6), 135.7 (*C*-1'), 129.1 (*C*H-1, *C*H-8), 128.3 (*C*H-3'), 126.6 (*C*H-5'), 125.7 (*C*H-6'), 109.6 (*C*H-2, *C*H-7), 106.5 (*C*-8a, *C*-9a), 98.4 (*C*H-4, *C*H-5), 40.4 (*C*H_3-N), 25.8 (*C*H_2-NHS), *C*-2' and *C*-4' not observed; IR (neat, ν_{max}/cm^{-1}) 3392, 2924, 1703, 1588, 1340, 1182; HRMS (ESI): calcd. for $C_{29}H_{26}N_3O_7$ $[M+H]^+$ 528.1765 found 528.1764.

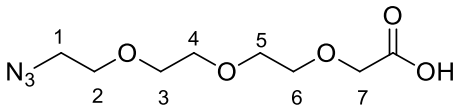
2-(2-(2-azidoethoxy)ethoxy)ethanol 84[107]



To a solution of [2-[2-[2-chloroethoxy]ethoxy]ethanol **83** (3.0 g, 17.8 mmol) in DMF (50 mL) was added NaN_3 (2.31 g, 35.5 mmol) and the reaction mixture was heated to 100 °C with stirring for 16 hours. The solution was allowed to cool to ambient temperature, then the solvent was co-evaporated with toluene (3 × 30 mL) *in vacuo*. The residue was re-suspended in EtOAc (30 mL), filtered through celite and the celite plug was washed with EtOAc (2 × 20 mL). The solvent was removed *in vacuo* to afford the *title compound* as a pale yellow oil (3.1 g, 99%). R_f 0.18 (Hexane-EtOAc

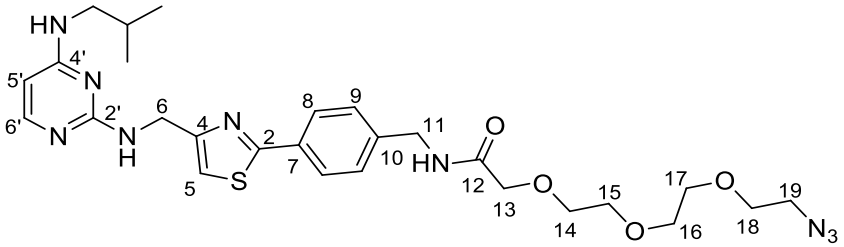
1:1); δ_H (500 MHz, $CDCl_3$) 3.73 – 3.75 (2H, m, 6-C*H*$_2$), 3.61 – 3.63 (6H, m, 3,4,5-C*H*$_2$), 3.62 (2H, t, *J* 5.0, 2-C*H*$_2$), 3.40 (2H, t, *J* 5.0, 1-C*H*$_2$), 2.28 (s, 1H, O*H*); δ_C (125 MHz, $CDCl_3$) 72.6 (*C*H_2-4), 70.8 – 70.2 (*C*H_2-2, *C*H_2-3, *C*H_2-5), 61.9 (*C*H_2-6), 50.8 (*C*H_2-1); IR (neat, ν_{max}/cm^{-1}) 3418, 2870, 2096, 1064; HRMS (ESI): calcd. for $C_6H_{13}N_3O_3$ $[M+Na]^+$ 198.0855 found 198.0873.

3-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)propanoic acid 86[107]



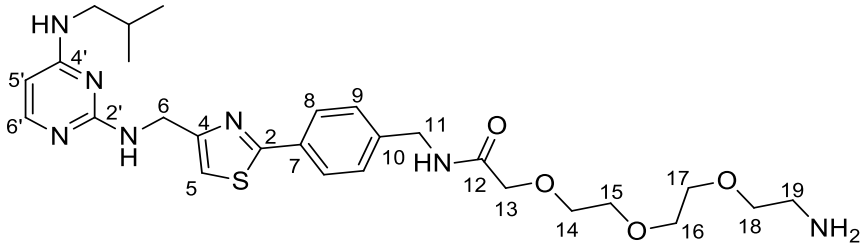
To a solution of 2-(2-(2-azidoethoxy)ethoxy)ethanol **84** (1 g, 5.71 mmol) in THF (30 mL) was added NaH (285 mg, 7.13 mmol) at 0 °C (ice-bath). After stirring for 1 hour at ambient temperature, ethyl bromoacetate (1.27 mL, 11.4 mmol) was added dropwise over 10 minutes and the reaction mixture was stirred at ambient temperature for 48 hours. The solvents were removed *in vacuo*, then water (20 mL) was added and the aqueous layer was extracted with CH_2Cl_2 (3 × 10 mL). The combined organic layers were dried over anhydrous Na_2SO_4 and concentrated *in vacuo* to give an orange oil (1.89 g). The ethyl ester was hydrolysed with an aqueous solution of NaOH (3M, 5 mL) at ambient temperature for 18 hours. After extracting the aqueous layer with CH_2Cl_2 (10 × 5 mL) to remove mineral oil and other impurities, the aqueous layer was acidified to pH 1, and extracted with EtOAc (3 × 50 mL). The combined organic layers were dried (Na_2SO_4), filtered and concentrated to afford the *title compound* as a pale orange oil (1.29 g, 97%). δ_H (500 MHz, $CDCl_3$) δ 3.40 (2H, t, *J* 5.0, 1-C*H*$_2$), 3.39 – 3.41 (8H, m, 3-C*H*$_2$, 4-C*H*$_2$, 5-C*H*$_2$, 6-C*H*$_2$), 3.68 – 3.70 (2H, d, *J* 5.0 2-C*H*$_2$), 4.20 (2H, s, 7-C*H*$_2$); δ_C (125 MHz, $CDCl_3$) 50.6 (*C*H_2-7), 68.4 (*C*H_2-1), 69.9 – 71.1 (*C*H_2-2, *C*H_2-3, *C*H_2-4, *C*H_2-5, *C*H_2-6), 174.0 (*C*-(O)OH); IR (neat, ν_{max}/cm^{-1}) 2874, 2101, 1732, 1224, 1092; HRMS (ESI): calcd. for $C_8H_{14}N_3O_5$ $[M-H]^-$ 232.0939 found 232.0943.

1-azido-N-[(4-{4-[({4-[(2-methylpropyl)amino]pyrimidin-2-yl}amino) methyl]-1,3-thiazol-2-yl}phenyl)methyl]-3,6,9,12-tetraoxapentadecan-15-amide 87



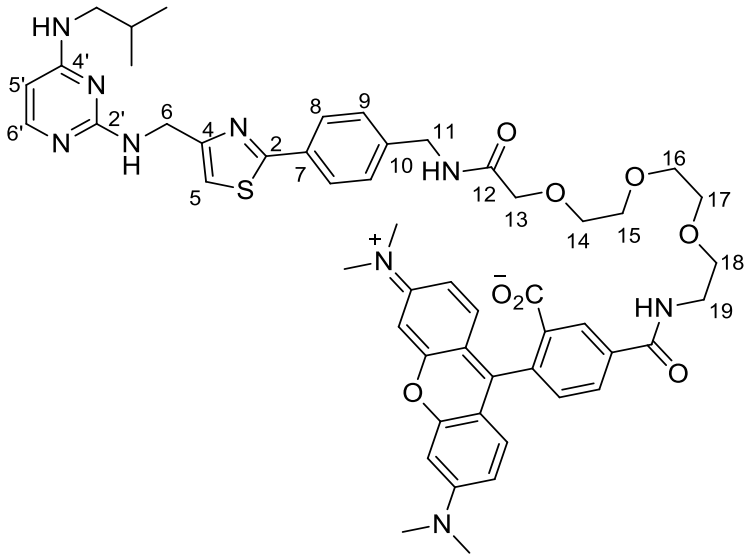
To a solution of amine **63** (20 mg, 0.05 mmol) in CH_2Cl_2 (2 mL), was added acid **86** (13 mg, 0.05 mmol), HOBt (7 mg, 0.06 mmol), DIPEA (20 µL, 0.12 mmol) and EDC·HCl (11 mg, 0.06 mmol). The reaction mixture was stirred at ambient temperature for 20 hours, then washed with water (3 × 0.5 mL), dried over anhydrous Na_2SO_4 and concentrated to give a pale orange oil. The crude product was purified by flash column chromatography (SiO_2; eluting with CH_2Cl_2-MeOH 9:1) to afford the *title compound* as a colourless oil (21 mg, 65%). R_f 0.23 (CH_2Cl_2-MeOH 9:1); δ_H (500 MHz, $CDCl_3$) 7.89 (2H, d, *J* 8.3, 8-C*H*), 7.80 (1H, s, N*H*), 7.44 (1H, t, *J* 6.0, 6'-C*H*), 7.35 (2H, d, *J* 8.3, 9-C*H*), 7.11 (1H, s, 5-C*H*), 5.73 (1H, d, *J* 6.0, 5'-C*H*), 4.99 (1H, s, N*H*), 4.74 (2H, d, *J* 4.8, 6-CH_2), 4.52 (2H, d, *J* 6.0, 11-CH_2), 4.07 (2H, s, 13-CH_2), 372 – 3.62 (4H, m, 17-CH_2,18-CH_2), 3.61 – 3.48 (6H, m, 14-CH_2,15-CH_2,16-CH_2), 3.31 – 3.25 (2H, m, 19-CH_2), 3.08 (2H, app. s, isobutyl-CH_2), 1.83 (1H, m, isobutyl-CH), 0.92 (6H, d, *J* 6.7, isobutyl-CH_3), δ_C (125 MHz, methanol-d_4) 170.2 (*C*-12), 167.9 (*C*-2), 163.4 (*C*H-6'), 156.1 (*C*-4), 140.4 (*C*-10), 133.0 (*C*-7), 128.3 (*C*H-9), 127.0 (*C*H-8), 114.3 (*C*H-5), 71.3 – 70.0 (13-CH_2,14-CH_2,15-CH_2,16-CH_2,17-CH_2, 18-CH_2,19-CH_2), 50.7 (*C*H_2-isobutyl), 42.6 (*C*H_2-11), 42.0 (*C*H_2-6), 28.5 (*C*H-isobutyl), 20.4 (*C*H_3-isobutyl), *C*-2', *C*-4' and C-5' not observed; HRMS (ESI): calcd. for $C_{27}H_{37}N_9NaO_4S$ $[M+Na]^+$ 606.2585 found 606.2581.

2-{2-[2-(2-aminoethoxy)ethoxy]ethoxy}-N-[(4-{4-[({4-[(2-methylpropyl) amino]pyrimidin-2-yl}amino)methyl]-1,3-thiazol-2-yl}phenyl)methyl] acetamide 88



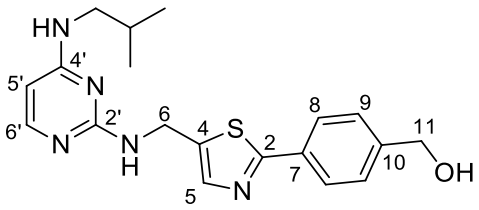
To a solution of azide **87** (50 mg, 0.086 mmol) in THF (0.5 mL), cooled to 0 °C (ice-bath), was added triphenylphosphine (25 mg, 0.094 mmol), and the mixture was stirred at ambient temperature for 20 hours. The reaction mixture was diluted into water (1 mL), then washed with toluene (3 × 0.25 mL). The aqueous layers were concentrated *in vacuo* to give the *title compound* as a pale orange oil (36 mg, 75%). δ_H (500 MHz, $CDCl_3$) 7.87 (2H, d, *J* 8.1, 8-C*H*), 7.37 – 7.22 (3H, m, 6'-C*H,* 9-C*H*), 7.10 (1H, s, 5-C*H*), 5.72 (1H, d, *J* 5.9, 5'-C*H*), 5.04 (1H, s, N*H*), 4.74 (2H, s, 6-CH_2), 4.52 (2H, *J* 6.0, 11-CH_2), 4.07 (2H, s, 13-CH_2), 3.76 – 3.34 (8H, m, 14-CH_2,15-CH_2,16-CH_2,17-CH_2), 3.18 – 2.96 (2H, m, isobutyl-CH_2), 2.95 – 2.63 (4H, m, 18-CH_2, 19-CH_2), 1.82 (1H, m, isobutyl-C*H*), 0.92 (6H, d, *J* 6.7, isobutyl-CH_3); δ_C (125 MHz, $CDCl_3$) 170.2 (*C*-12), 167.9 (*C*-2), 163.5 (*C*H-6'), 156.2 (*C*-4), 140.5 (*C*-10), 132.9 (*C*-7), 128.7 (*C*H-9), 128.3 (*C*H-8), 114.3 (*C*H-5), 72.8 (*C*H_2-13), 71.2 – 70.2 (*C*H_2-14, *C*H_2-15, *C*H_2-16, *C*H_2-17, *C*H_2-18, *C*H_2-19), 42.6 (*C*H_2-isobutyl), 42.0 (*C*H_2-11), 41.5 (*C*H_2-6), 28.5 (*C*H-isobutyl), 20.4 (*C*H_3-isobutyl), *C*-2', *C*-4' and C-5' not observed; IR (neat, v_{max}/cm^{-1}) 3385, 3245, 1653, 1260, 1055; HRMS (ESI): calcd. for $C_{27}H_{40}N_7O_4S$ $[M+H]^+$ 558.2857 found 558.2854.

2-(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)-5-((1-(4-(4-(((4-(isobutylamino)pyrimidin-2-yl)amino)methyl)thiazole-2-yl)phenyl)-3-oxo -5,8,11-trioxa-2-azatridecan-13-yl)carbamoyl)benzoate (KHS101-TAMRA) 77



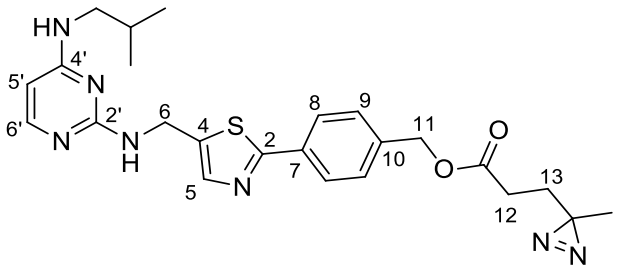
To a solution of amine **88** (35 mg, 0.06 mmol) in DMF (0.5 mL) was added TAMRA-NHS **78** (43 mg, 0.08 mmol) and NEt_3 (18 µL, 0.13 mmol) and the reaction mixture was stirred at ambient temperature overnight. Solvents were removed *in vacuo* and the residue was dissolved in DCM (3 mL), washed with saturated aqueous $NaHCO_3$ (aq., 3 × 1 mL), brine (1 × 2 mL), dried over anhydrous Na_2SO_4, filtered and concentrated to give a purple solid. The crude product was purified using preparative HPLC (C_{18}, eluting with H_2O (+0.1% HCOOH)-MeCN (+0.1% HCOOH) 95:5→50:50) to give the *title compound* as a purple solid (17 mg, 24%). δ_H (500 MHz, 99:1 $CDCl_3$-methanol-d_4) 8.68 (1H, s, rhodamine-C*H*), 8.17 (1H, dd, *J* 7.9, 1.6, rhodamine-C*H*), 7.75 (2H, d, *J* 8.0, 8-C*H*), 7.31 – 7.22 (5H, m, 5-C*H*, 6'-C*H*, 9-C*H*, rhodamine-C*H*), 7.00 (2H, d, *J* 9.4, rhodamine-C*H*), 6.77 (2H, d, *J* 9.4, rhodamine-C*H*), 6.71 (2H, app. s, rhodamine-C*H*), 5.97 (1H, d, *J* 5.9, 5'-C*H*), 4.64 (2H, s, 6-CH_2), 4.38 (2H, s, 11-CH_2), 4.00 (2H, s, 13-CH_2), 3.81 – 3.34 (8H, m, 14-CH_2,15-CH_2,16-CH_2,17-CH_2), 3.20 (14H, app. s, N-CH_3, isobutyl-CH_2), 3.11 (4H, m, 18-CH_2, 19-CH_2), 1.71 (1H, app. s, isobutyl-C*H*), 0.75 (6H, d, *J* 6.7, isobutyl-CH_3); HRMS (ESI): calcd. for $C_{52}H_{60}N_9O_8S$ $[M+H]^+$ 970.4280, found 970.4336.

(4-(5-(((4-(isobutylamino)pyrimidin-2-yl)amino)methyl)thiazole-2-yl) phenyl)methanol 91



A mixture of pyrimidine **68b** (150 mg, 0.44 mmol), cesium carbonate (573 mg, 1.76 mmol), (4-(hydroxymethyl)phenyl)boronic acid **90** (134 mg, 0.88 mmol), and tetrakis(triphenylphosphine)palladium(0) (45 mg, 0.04 mmol) in 1,4-dioxane (13.5 mL) and water (2.4 mL) was placed in a microwave reactor vessel and purged with nitrogen for 10 minutes. The reaction mixture was heated at 150 °C for 25 minutes (hold time) in a microwave reactor. The resulting dark brown solution was filtered through a plug of celite and washed through with EtOAc (150 mL). The filtrate was washed with brine (3 × 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to give a brown oil. The crude product was purified by FCC (SiO_2, eluting with EtOAc-MeOH 49:1→6:1) to afford the title compound as a colourless oil (74 mg, 46%). R_f 0.21 (EtOAc-MeOH 49:1); δ_H (400 MHz, $CDCl_3$) 7.81 (2H, d, *J* 8.1, 8-C*H*), 7.72 (1H, app. s, 6'-C*H*), 7.34 (2H, d, *J* 8.1, 9-C*H*), 6.99 (1H, s, 5-C*H*), 5.68 (2H, d, *J* 5.9, 5'-C*H*), 5.09 (1H, s, N*H*), 4.67 – 4.57 (4H, m, 6-CH_2, 11-CH_2), 3.03 (2H, app. s, isobutyl-CH_2), 1.87-1.72 (1H, m, isobutyl-C*H*), 0.89 (6H, d, *J* 6.7, isobutyl-CH_3); δ_C (100 MHz, $CDCl_3$) 168.0 (*C*-2), 163.3 (*C*H-5'), 161.7 (*C*-4), 155.9 (*C*H-6'), 143.5 (*C*-Ar), 132.6 (*C*-Ar), 127.2 (*C*H-Ar), 126.5 (*C*H-Ar), 113.9 (*C*H-5), 64.3 (*C*H_2-11), 48.8 (*C*H_2-isobutyl), 41.8 (*C*H_2-6), 28.4 (*C*H-isobutyl), 20.2 (*C*H_3-isobutyl), *C*-2' and *C*-4' not observed; IR (neat, ν_{max}/cm^{-1}) 3270, 2956, 1642, 1597, 1409, 1004; HRMS (ESI): calcd. for $C_{19}H_{24}N_5OS$ $[M+H]^+$ 370.1696, found 370.1701.

4-(5-(((4-(isobutylamino)pyrimidin-2-yl)amino)methyl)thiazole-2-yl) benzyl 3-(3-methyl-3*H*-diazirin-3-yl)propanoate (KHS101-diazirine) 89



To a solution of alcohol **91** (51 mg, 0.14 mmol) in CH_2Cl_2 (2 mL), was added *N,N'*-dicyclohexylcarbodiimide (DCC) (33 mg, 0.16 mmol), DMAP (20 mg, 0.16 mmol) and a solution of 3-methyl-diazirine-3-propanoic acid **92** (18 mg, 0.14 mmol) in CH_2Cl_2 (1 mL). The reaction mixture was stirred at ambient temperature overnight. The reaction mixture was washed with water (2 × 1 mL), dried (Na_2SO_4) and concentrated *in vacuo* to give a yellow oil. The crude product was purified by FCC (SiO_2, eluting with EtOH-MeOH 19:1) to afford the title compound as a pale yellow oil (24 mg, 35%). R_f 0.25 (EtOH-MeOH 19:1); δ_H (400 MHz, 100:1 $CDCl_3$-methanol-d_4) 8.00 – 7.86 (2H, m, 8-C*H*), 7.76 (1H, d, *J* 6.0, 6'-C*H*), 7.52 – 7.36 (2H, m, 9-C*H*), 7.17 (1H, s, 5-C*H*), 5.72 (1H, d, *J* 6.0, 5'-C*H*), 5.16 (2H, s, 11-CH_2), 4.73 (2H, s, 6-CH_2), 3.07 (2H, d, *J* 6.7, isobutyl-CH_2), 2.27 (2H, t, *J* 7.7, 12-CH_2), 1.93 – 1.68 (3H, m, 13-CH_2, isobutyl-C*H*), 1.03 (3H, s, CH_3), 0.93 (6H, d, *J* 6.7, isobutyl-CH_3); δ_C (100 MHz, $CDCl_3$) 172.2 (C=O), 167.7 (*C*-2), 163.5 (*C*H-6'), 162.0 (*C*-4), 156.4 (*C*H-5'), 137.6 (*C*-Ar), 133.8 (*C*-Ar), 128.8 (*C*H-Ar), 126.8 (*C*H-Ar), 114.4 (*C*H-5), 66.1 (*C*H_2-11), 49.2 (*C*H_2-isobutyl), 42.1 (*C*H_2-6), 29.8 (*C*H_2-13), 28.9 (*C*H_2-12), 28.5 (*C*H-isobutyl), 20.4 (*C*H_3-isobutyl), 19.8 (*C*H_3), *C*-2' and *C*-4' and *C*-diazirine not observed; IR (neat, ν_{max}/cm^{-1}) 3315, 2955, 1735, 1588, 1230, 1165; HRMS (ESI): calcd. for $C_{24}H_{30}N_7O_2S$ $[M+H]^+$ 480.2176, found 480.2179.

5.2.2 Affinity-based target identification in GBM1 cells

Experiments performed by Shoutian Zhu. GBM1 cells were incubated with KHS101-BP (5 µM) in the presence or absence of unlabelled KHS101 (250 µM) for 30 minutes, and irradiated with UV light (365 nm) for 30 minutes. Cells were lysed using 0.5% Triton X-100 and protease inhibitor cocktail (Sigma, cat # P8340). Cell lysates were incubated with 25 µM biotin azide (ThermoFisher, cat # B10184), 1 mM TCEP, 100 mM ligand (TBTA), and 1 mM aqueous copper sulfate at 4 °C overnight. Subsequently, proteins were fractionated using ammonium sulfate and the 20-40% fractions were subject to 2D SDS/PAGE. Biotin-labeled proteins were detected through Western blotting using an antibody against biotin (Abcam ab1227). Protein spots corresponding to the specific biotin-labeled proteins were visualized with silver staining on parallel gels. A distinct spot was excised and protein identified using liquid chromatography tandem mass spectrometry (LC-MS/MS).

5.2.3 In vitro pull-down of recombinant HSPD1

Experiments performed by Shoutian Zhu. For human HSPD1 and KHS101 *in vitro* binding assays, a total of 1 µg recombinant human HSPD1 (Abcam, ab113192) was diluted in 1 mL PBS (with 2 mM $MgCl_2$, 2 mM DDT and 0.1% Tween 20) and incubated with 5 µM biotinylated KHS101 at 4°C overnight in the presence or the absence of non-labelled KHS101. Streptavidin agarose was added to the incubation mixture and rotated at 4°C for 2 hours. The beads were then precipitated and washed three times in PBS. Bound proteins were eluted with 2X SDS sample buffer and analysed with SDS/PAGE followed by silver staining and Western blotting.

5.2.4 Surface Plasmon Resonance

5.2.4.1 Immobilisation of KHS101-biotin

Experiments performed by Hester Beard and Dr Iain Manfield. Binding affinities were analysed by SPR (Biacore 3000). Prior to immobilisation, the chip was conditioned with 3 consecutive 1 minute injections of 1M NaCl in 50 mM NaOH. KHS101-biotin was immobilised on a SA (streptavidin) chip non-covalently in PBS with 2 mM $MgCl_2$, 2 mM DTT and 0.1% Tween 20 (running buffer) at concentrations of 1 nM, 10 nM and 100 nM. For the binding analyses, recombinant human HSPD1 (Abcam ab113192) was diluted to 100 µM, 10 µM, 1 µM, 100 nM, 10 nM, 1 nM and 0.1 nM in PBS (2 mM $MgCl_2$, 2 mM DTT, 0.1% Tween 20, 2 mM ATP). HSPD1 solutions were injected over the KHS101-derivitised and reference surfaces for 180 seconds,

followed by a 10 minute dissociation period. Data analysis was performed using BIA-evaluation software.

5.2.4.2 Immobilisation of HSPD1

Experiments performed by Hester Beard, David Klebl and Dr Iain Manfield. Binding affinities were analysed by SPR (Biacore 3000). Recombinant human HSPD1 was immobilised onto a CM5 Sensor Chip (GE Healthcare) using standard amine coupling: flow cells were activated with *N*-hydroxysuccinimide/1-ethyl-3-(3-dimethylaminopropyl)carbodiimide treatment. Each flow cell was then subjected to a solution of HSPD1 (5 µg/mL) in acetate buffer (10 mM acetate, 2 mM $MgCl_2$, pH 5.0) until the signal had increased by 6000 units. The blank flow cell was subjected to acetate buffer and no signal change was observed. Flow cells were washed with acetate buffer, followed by capping with 1 M ethanolamine·HCl. Binding experiments with KHS101 (8 µM - 1 mM) were performed in 10 mM acetate, 2 mM $MgCl_2$, 0.1 % Tween 20, pH 5.0 at a flow rate of 5 µL/min. 1 minute injections followed by 2 minute dissociation period. Data analysis was performed using BIA-evaluation software.

5.2.5 Fluorescence anisotropy

5.2.5.1 General remarks

Fluorescence anisotropy assays were run in 384 well Optiplates and scanned using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader. KHS101-TAMRA used an excitation and emission wavelength of 555 nm and 610 nm, respectively (5 nm bandwidths). Assays were run in PBS (2 mM $MgCl_2$, 2 mM DTT, 0.1% Tween 20, pH 7.0).

5.2.5.2 Processing of fluorescence anisotropy data

The data obtained for both the *P* (perpendicular intensity) and *S* (parallel intensity) channels were corrected by subtracting the corresponding control wells, and the resulting values were used to calculate intensity (Eq. 1) and anisotropy (Eq. 2) for each well (using Microsoft Excel). These data were transferred into OriginPro 8.5, where a plot of anisotropy against protein concentration was plotted. However, the data could not be fitted using a logistic model (Eq. 3) to obtain the minimum (r_{min}) and maximum (r_{max}) values of anisotropy to allow determination of a K_d.

$$I = 2PG + S \qquad \textit{Equation 1}$$

Where *I* is the total intensity, *G* is an instrument factor which was set to 1, and *r* is the anisotropy.

$$r = \frac{S - PG}{I} \qquad \textit{Equation 2}$$

$$y = r_{max} + \frac{r_{min} - r_{max}}{1 + \left({}^{x}/_{x_0}\right)^{p}} \qquad \textit{Equation 3}$$

Where x_0 is the midpoint and *p* is the power.

5.2.5.3 Direct binding fluorescence anisotropy assay

Protein was serially diluted into a solution of KHS101-TAMRA, with the concentration of HSPD1 starting at 10 µM, diluted over a ½ regime with [KHS101-TAMRA] fixed at 100 nM. Plates were read after 1 h and 20 h of incubation. The assays consisted of three test rows (containing protein and tracer), and three control rows (where the tracer was replaced with buffer). As in section 5.2.5.2., the intensity, *I*, and anisotropy, *r*, were calculated using equations 1 and 2, respectively. A plot of anisotropy against HSPD1 concentration was plotted in OriginPro 8.5, but could not be fitted using Eq. 3.

5.2.6 Photocrosslinking with recombinant HSPD1

5.2.6.1 Photocrosslinking

To a solution of HSPD1 (10 µM) in ammonium acetate buffer (50 mM NH_4HCO_3, 10 mM $MgCl_2$, pH 7.0) was added KHS101-diazirine (final concentration 0 - 100 µM), and the mixture was incubated for 10 minutes at ambient temperature. The mixture was irradiated for 30 min on ice, 10 cm from light source (365 nm, CL-1000 UV crosslinker, UVP). Small molecules were removed using a centrifugal filter (MWCO: 10,000, Amicon Ultra, 0.5 mL).

5.2.6.2 In-solution digestion

In-solution digestion was performed by Dr Rachel George. DTT was added to the sample to a final concentration of 10 mM and heated to 56 ºC for 30 min with shaking. The Eppendorf was left to cool before adding iodoacetic acid to a final concentration of 55 mM and incubated at room temperature in the dark for 30 min with shaking. Trypsin solution (20 ng μL^{-1} in 25 mM ammonium bicarbonate) was added in a 1:50 ratio (protease:total protein content). Samples were incubated at 37 ºC with shaking for 18 h. The digest reaction was stopped by adding 5 µL of 1% HCOOH.

5.2.6.3 Sep-pak C18 purification

Sep-pak purification was performed by Dr Rachel George. The Sep-pak column was equilibrated with 1 mL 0.1% TFA. 500 µL of 0.1% TFA was added to the peptide digest, the mixture was passed through the column and the column was washed with 1 mL 0.1% TFA. Peptides were eluted from the column with 500 µL MeCN-H_2O 1:1 + 0.1% HCOOH. The eluant was dried by vacuum centrifugation and the peptides were reconstituted in 20 µL 0.1% TFA.

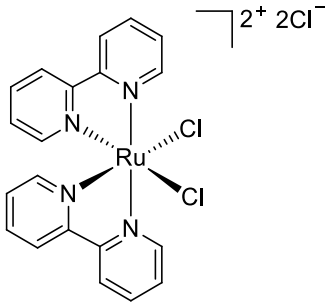
5.2.6.4 LC-MS analysis

LC-MS analysis was performed by Dr Rachel George, on an ACQUITY M-Class UPLC (Waters UK, Manchester). 1 µL sample was loaded onto a Symmetry C18 trap column and washed with 1% MeCN/0.1% HCOOH for 5 min at 5 µL min^{-1}, then the peptides were separated on a HSS T3 C18 analytical column (Waters UK, Manchester) by gradient elution of 1-60% solvent B (0.1% HCOOH in MeCN) in A (0.1% HCOOH in H_2O) over 30 min at 0.3 µL min^{-1}. The peptides were analysed using a Xevo G2-XS Q-TOF mass spectrometer. Data processing was performed using the MassLynX v4.1 suite of software. Peptide MS/MS data were processed with PEAKS Studio (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada).

5.3 Experimental section for Chapter 3

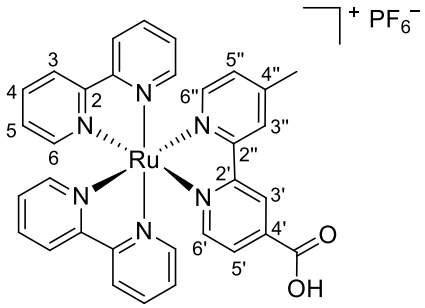
5.3.1 Organic synthesis

Tris(2,2'-bipyridine)ruthenium(II) dichloride 104[130]



To a solution of ruthenium(III) chloride hydrate (500 mg, 2.41 mmol) in DMF (7.5 mL) was added lithium chloride (717 mg, 16.90 mmol), followed by 2,2'-bipyridine (753 mg, 4.82 mmol) in three equal portions, and the reaction mixture was heated at reflux overnight. The reaction mixture was cooled to room temperature, acetone was added (40 mL) and the mixture was left at -4°C overnight. The brown microcrystals were filtered, washed with Et_2O (30 mL) and water (10 mL) to afford a brown solid (1.15 g, 98%). The crude product was used without further purification. HRMS (ESI): calcd. for $C_{20}H_{16}Cl_2N_4NaRu$ $[M+Na]^+$ 506.9686, found 506.9685. Data consistent with literature.

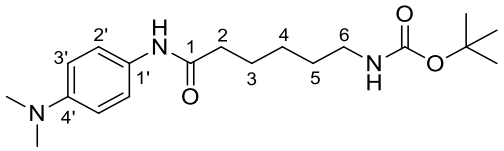
Bis-(2,2'-bipyridine)-(4'-methyl-[2,2'-bipyridine]-4-carboxylic acid)ruthenium (II) bis-(hexafluorophosphate) 101[141]



To a solution of crude **104** (905 mg, 1.87 mmol) in a mixture of EtOH (42 mL) and water (18 mL), was added 4-methyl-2,2'-bipyridine-4'-carboxylic acid (250 mg, 1.17 mmol), and the reaction mixture was heated at reflux for 24 hours. The reaction mixture was cooled to room temperature, the solvents were removed *in vacuo* and the residue was re-suspended in water (30 mL) and filtered. The crude product was purified by FCC (SiO_2, eluting with 100:10:1→100:20:1, v/v/v, MeCN:H_2O:KNO_3(sat.)), solvents were removed *in vacuo*, and then re-dissolved in water (30 mL). Ammonium hexafluorophosphate was added (500 mg) in portions,

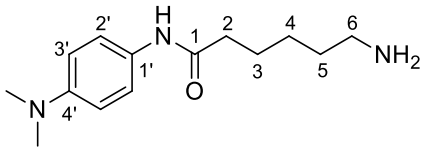
resulting in the formation of a red-orange precipitate, which was collected *via* vacuum filtration and washed with H_2O (20 mL) to give the product as dark red crystals (740 mg, 43%). R_f 0.15 (100:20:1, v/v/v, MeCN:H_2O:KNO_3(sat.)); δ_H (500 MHz, acetonitrile-d_3) 8.93 (1H, s, 3'-C*H*), 8.53 – 8.45 (5H, m, 3-C*H,* 3''-C*H*), 8.05 (4H, m, 6-C*H*), 7.82 – 7.69 (6H, m, 4-C*H,* 6'-C*H,* 6''-C*H*), 7.54 (1H, d, *J* 5.8, 5'-C*H*), 7.44 – 7.34 (4H, m, 5-C*H*), 7.24 (1H, d, *J* 5.5, 5''-C*H*), 2.53 (3H, s, CH_3); δ_C (125 MHz, acetonitrile-d_3) 167.6 (*C*(O)OH), 158.1 – 157.9 (*C*-2, *C*-2', *C*-2''), 152.8 (*C*H-6'), 152.7 (*C*H-4), 152.6 (*C*-4'), 151.65 (*C*H-5''), 146.8 (*C*-4''), 138.7 (*C*H-6), 129.4 (*C*H-5''), 128.6 (*C*H-5), 127.7 (*C*H-6''), 126.3 (*C*H-3''), 125.2 (*C*H-3), 124.3 (*C*H-3'), 21.15 (*C*H_3); IR (neat, ν_{max}/cm^{-1}) 3320, 1726, 1618, 1424, 1233, 825; HRMS (ESI): calcd. for $C_{32}H_{26}N_6O_2{}^{102}Ru$ $[M]^{2+}$ 315.0581, found 315.0583. Data consistent with literature.

Tert-butyl(6-((4-(dimethylamino)phenyl)amino)-6-oxohexyl)carbamate 114



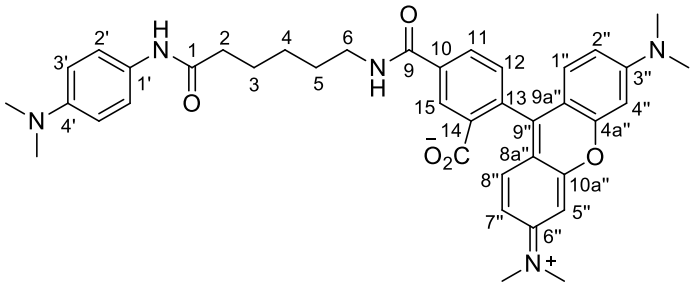
To a solution of Boc-Ahx-OH (500 mg, 2.16 mmol) in THF (8 mL) at 0 °C (ice-bath) was added isobutyl chloroformate (305 µL, 2.36 mmol) slowly, then *N*-methylmorpholine (402 µL, 3.65 mmol) slowly. The resulting solution was stirred for 30 min at 0 °C, then allowed to warm to ambient temperature. After 1 h, a solution of *N,N*-dimethyl-*p*-phenylenediamine (478 mg, 3.51 mmol) in THF (0.5 mL) was added (solution turned dark green/brown) and the reaction mixture was stirred at ambient temperature overnight. The reaction mixture was quenched with 1 N HCl (ca. 5 mL) and washed with EtOAc (3 × 3 mL). A solution of 6 N NaOH was added to the mixture to become basic (pH 8). The product was extracted with EtOAc (5 × 10 mL), the organic layers were dried (Na_2SO_4), filtered and concentrated *in vacuo*. The crude product was purified by FCC (SiO_2, Biotage, 40 g, eluting with 12-70% EtOAc-Hex) to give the *title compound* as a purple solid (558 mg, 74%). R_f 0.31 (1:1 EtOAc-Hex); δ_H (400 MHz, $CDCl_3$) 7.43 (1H, s, N*H*), 7.38 – 7.32 (2H, app. d, *J* 9.0, 2'-C*H*), 6.74 – 6.62 (2H, app. d, *J* 9.0, 3'-C*H*), 4.70 – 4.55 (1H, s, N*H*), 3.09 (2H, m, 6-CH_2), 2.89 (6H, s, N-CH_3), 2.30 (2H, t, *J* 7.5, 2-CH_2), 1.71 (2H, p, *J* 7.5, 3-CH_2), 1.54 – 1.30 (13H, m, O^{t}Bu-CH_3, 4-CH_2, 5-CH_2); δ_C (100 MHz, $CDCl_3$) 171.1 (*C*-1), 156.2 (*C*(O)-^{t}Bu), 148.0 (*C*-4'), 128.2 (*C*-1'), 121.9 (*C*H-2'), 113.3 (*C*H-3'), 79.2 (*C*-^{t}Bu), 41.1 (*C*H_3-N), 40.5 (*C*H_2-6), 37.4 (*C*H_2-2), 29.9 (*C*H_2-3), 28.6 (*C*H_3-^{t}Bu), 26.5 (*C*H_2-5), 25.4 (*C*H_2-4); IR (neat, ν_{max}/cm^{-1}) 3373, 2932, 1649, 1513, 1159; HRMS (ESI): calcd. for $C_{19}H_{31}N_3NaO_3$ $[M+Na]^+$ 372.2258, found 372.2257.

6-amino-*N*-(4-(dimethylamino)phenyl)hexanamide 115



To a solution of **114** (1 g, 2.87 mmol) in CH_2Cl_2 (10 mL) was added trifluoroacetic acid (10 mL) and the reaction mixture was stirred at ambient temperature for 1.5 hours. The solvents were removed *in vacuo* to give a purple oil. The crude product was purified by FCC (SiO_2, eluting with CH_2Cl_2-MeOH-NEt_3 95:4:1→90:9:1) to afford the *title compound* as a pale purple waxy solid (681 mg, 95%). R_f 0.08 (CH_2Cl_2-MeOH-NEt_3 90:9:1); δ_H (400 MHz, $CDCl_3$) 7.70 (1H, s, N*H*), 7.41 – 7.30 (2H, m, 2'-C*H*), 6.74 – 6.62 (2H, app. d, *J* 9.0, 3'-C*H*), 2.88 (6H, m, N-CH_3), 2.68 (2H, t, *J* 6.9, 6-CH_2), 2.30 (4H, m, 2-CH_2, NH_2), 1.70 (2H, p, *J* 7.5, 3-CH_2), 1.55 – 1.30 (4H, m, 4-CH_2, 5-CH_2); δ_C (100 MHz, $CDCl_3$) 171.4 (*C*-1), 148.0 (*C*-4'), 128.2 (*C*-1'), 121.9 (*C*H-2'), 113.2 (*C*H-3'), 41.8, (*C*H_2-6), 41.1 (*C*H_3-N), 37.4 (*C*H_2-2), 32.9 (*C*H_2-5), 26.5 (*C*H_2-4), 25.6 (*C*H_2-3); IR (neat, ν_{max}/cm^{-1}) 3282, 2930, 1644, 1517, 1315; HRMS (ESI): calcd. for $C_{14}H_{24}N_3O$ $[M+H]^+$ 250.1914, found 250.1912.

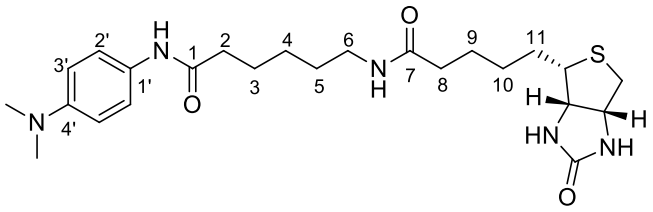
2-(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)-5-((6-((4-dimethylamino) phenyl)amino)-6-oxohexyl)carbamoyl)benzoate (TAMRA-RTA) 111



To a solution of amine **115** (50 mg, 0.20 mmol) in DMF (1 mL) was added TAMRA-NHS **78** (138 mg, 0.26 mmol) and NEt_3 (56 µL, 0.40 mmol) and the mixture was stirred at ambient temperature for 20 hours. The solvent was removed *in vacuo* and the residue was re-dissolved in $CHCl_3$ (5 mL). The organic layer was washed with water (1 × 2 mL), saturated aqueous $NaHCO_3$ (aq., 3 × 2 mL), brine (1 × 2 mL), dried over anhydrous Na_2SO_4, filtered and concentrated to give a purple solid. The crude product was purified by FCC (SiO_2, eluting with CH_2Cl_2-MeOH 19:1→9:1) to afford the title compound as a purple solid (11 mg, 12%). R_f 0.12 (CH_2Cl_2-MeOH 9:1); δ_H (500 MHz, $CDCl_3$) 8.43 (1H, s, 15-C*H*), 8.08 (1H, d, *J* 8.2, 11-C*H*), 7.37 – 7.30 (2H, m, 2'-C*H*), 7.14 (1H, d, *J* 7.9, 12-C*H*), 6.90 (2H, d, *J* 8.9, 1''-C*H*, 8''-C*H*),

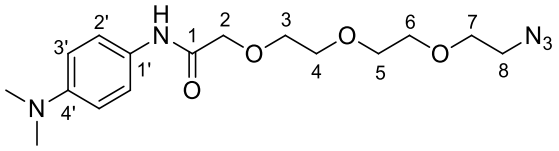
6.66 – 6.60 (2H, m, 3'-C*H*), 6.61 – 6.46 (5H, m, 2''-C*H*, 4''-C*H*, 5''-C*H*, 7''-CH, N*H*), 3.42 (2H, t, *J* 6.8, 6-CH_2), 3.02 (6H, s, rhodamine-N-CH_3), 2.89 (6H, s, rhodamine-N CH_3), 2.82 (6H, s, N-CH_3), 2.28 (2H, t, *J* 7.5, 2-CH_2), 1.68 (4H, m, 3-CH_2, 5-CH_2), 1.44 (2H, p, *J* 7.7, 4-CH_2); IR (neat, ν_{max}/cm^{-1}) 3496, 2998, 2853, 1651, 1493, 1320, 1058; HRMS (ESI): calcd. for $C_{39}H_{45}N_5O_5$ $[M+2H]^{2+}$ 331.6705, found 331.6705.

N-(4-(dimethylamino)phenyl)-6-(5-((4*S*)-2-oxohexahydro-1*H*-thienol[3,4-*d*]imidazol-4-yl)pentanamido)hexanamide (Biotin-RTA) 112



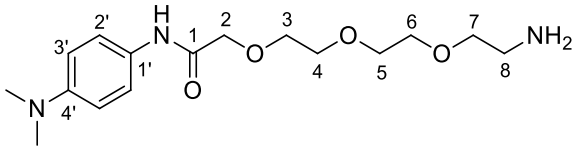
To a solution of D-biotin (421 mg, 1.72 mmol) in $CHCl_3$ (20 mL) was added EDCI (331 mg, 1.72 mmol), HOBt (233 mg, 1.72 mmol), amine **115** (286 mg, 1.15 mmol) and DIPEA (300 µL, 1.72 mmol), and the reaction mixture was stirred at ambient temperature overnight. The solvents were removed *in vacuo* to give a pale purple solid. The crude product was purified by FCC (SiO_2, eluting with CH_2Cl_2-MeOH 19:1→9:1) to afford the *title compound* as a colourless solid (182 mg, 33%). R_f 0.32 (CH_2Cl_2-MeOH 9:1); δ_H (400 MHz, $CDCl_3$) 7.56 – 7.39 (2H, app. d, *J* 9.0, 2'-C*H*), 6.85 (2H, app. d, *J* 8.9, 3'-C*H*), 4.59 (1H, dd, *J* 7.9, 4.8, biotin-C*H*), 4.39 (1H, dd, *J* 7.9, 4.4, biotin-C*H*), 3.48 (1H, s, biotin-C*H*), 3.28 (3H, m, 6-CH_2, biotin-C*H*), 3.16 – 2.86 (6H, m, N-CH_3), 2.82 (1H, d, *J* 12.8, biotin-C*H*), 2.45 (2H, t, *J* 7.4, 2-CH_2), 2.29 (2H, t, *J* 7.1, 8-CH_2), 1.90 – 1.59 (8H, m, 3-CH_2, 9-CH_2, 10-CH_2, 11-CH_2), 1.48 – 1.34 (4H, m, 4-CH_2, 5-CH_2); δ_C (100 MHz, $CDCl_3$) 175.3 (*C*-1), 173.4 (*C*-7), 165.2 (*C*(O)-biotin), 148.7 (*C*-4'), 129.4 (*C*-1'), 122.5 (*C*H-2'), 114.2 (*C*H-3'), 62.7(*C*H-biotin), 60.9 (*C*H-biotin), 56.4 (*C*H-biotin), 41.1 (*C*H_3-N), 40.8 (*C*H_2-biotin), 39.7 (*C*H_2-6), 37.2 (*C*H_2-2), 36.4 (*C*H_2-11), 29.5 – 28.8 (*C*H_2-3, *C*H_2-9, *C*H_2-10), 27.0 (*C*H_2-11), 26.3 (*C*H_2-4 or *C*H_2-5), 26.0 (*C*H_2-4 or *C*H_2-5); IR (neat, ν_{max}/cm^{-1}) 3283, 2927, 1697, 1645, 1518, 1458, 1260; HRMS (ESI): calcd. for $C_{24}H_{38}N_5O_3S$ $[M+H]^+$ 476.2690, found 476.2688.

2-(2-(2-(2-azidoethoxy-ethoxy)ethoxy)-*N*-(4-(dimethylamino)phenyl) acetamide 117



To a solution of azide **86** (300 mg, 1.29 mmol) in THF (5 mL) at 0 °C (ice-bath) was added isobutyl chloroformate (184 µL, 1.42 mmol) slowly and then NMM (241 µL, 2.19 mmol) slowly. The resulting solution was stirred for 30 min at 0 °C, then allowed to warm to ambient temperature. After 1 h, a solution of *N,N*-dimethyl-*p*-phenylenediamine (284 mg, 2.08 mmol) in THF (0.5 mL) was added (solution turned dark green) and the reaction mixture was stirred at ambient temperature overnight. The reaction mixture was quenched with 1 N HCl (ca. 5 mL) and washed with EtOAc (3 × 3 mL). A solution of 6 N NaOH was added to the mixture to become basic (pH 8). The product was extracted with EtOAc (5 × 10 mL), the organic layers were dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. The crude product was purified by FCC (SiO_2, eluting with 60-80% EtOAc-Hex) to give the *title compound* as a brown oil (324 mg, 92%). R_f 0.18 (60:40 EtOAc-Hex); δ_H (400 MHz, $CDCl_3$) 8.55 (1H, s, N*H*), 7.48 – 7.36 (2H, m, 2'-C*H*), 6.79 – 6.64 (2H, m, 3'-C*H*), 4.10 (2H, s, 2-CH_2), 3.79 – 3.70 (6H, m, 4-CH_2, 5-CH_2, 6-CH_2), 3.69 – 3.64 (2H, m, 3-CH_2), 3.61 (t, *J* 6.0, 7-CH_2), 3.32 (t, *J* 6.0, 8-CH_2), 2.91 (6H, s, N-CH_3); δ_C (100 MHz, $CDCl_3$) 167.9 (*C*-1), 148.3 (*C*-4'), 127.4 (*C*-1'), 122.0 (*C*H-2'), 113.2 (*C*H-3'), 71.3, (*C*H_2-2), 70.9 – 70.2 (*C*H_2-3, *C*H_2-4, *C*H_2-5, *C*H_2-6, *C*H_2-7), 50.8 (*C*H_2-8), 41.1 (*C*H_3-N); IR (neat, ν_{max}/cm^{-1}) 3017, 2911, 2126, 1681, 1454, 1086; HRMS (ESI): calcd. for $C_{16}H_{25}N_5NaO_4$ $[M+Na]^+$ 374.1799, found 374.1802.

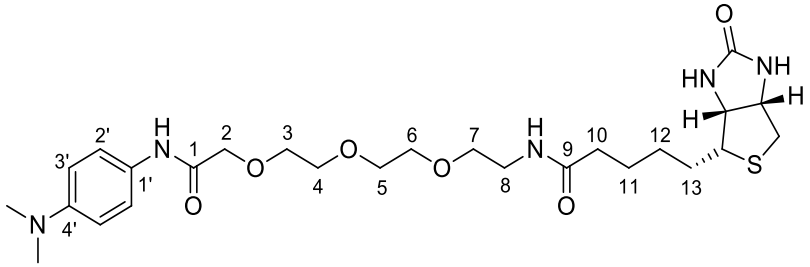
2-(2-(2-(2-aminoethoxy-ethoxy)ethoxy)-*N*-(4-(dimethylamino)phenyl) acetamide 118



To a solution of azide **117** (300 mg, 0.85 mmol) in THF (10 mL) at 0 °C (ice-bath) was added triphenylphosphine (246 mg, 0.94 mmol) and the reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted into water (20 mL) and washed with toluene (3 × 5 mL). The aqueous layer was lyophilised to give a purple solid (256 mg, 93%). δ_H (400 MHz, $CDCl_3$) 7.48 – 7.33 (2H, m, 2'-C*H*), 6.34 – 6.23 (2H, m, 3'-C*H*), 4.06 (2H, s, 2-CH_2), 4.12 – 3.99 (12H, m, 3-CH_2, 4-CH_2,

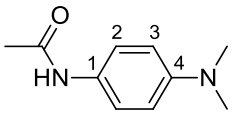
5-CH_2, 6-CH_2, 7-CH_2, 8-CH_2), 2.88 (6H, s, N-CH_3); δ_C (400 MHz, $CDCl_3$) 167.7 (*C*-1), 148.1 (*C*-4'), 127.4 (*C*-1'), 121.7 (*C*H-2'), 113.0 (*C*H-3'), 77.4 (CH_2-7), 73.4 (CH_2-2), 71.1 – 70.1 (CH_2-3, CH_2-4, CH_2-5, CH_2-6), 41.6 (CH_2-8), 40.9 (CH_3-N); IR (neat, ν_{max}/cm^{-1}) 3308, 2913, 1673, 1519, 1100; HRMS (ESI): calcd. for $C_{16}H_{28}N_3O_4$ $[M+H]^+$ 326.2074, found 326.2070.

***N*-(2-(2-(2-(2-((4-(dimethylamino)phenyl)amino)-2-oxoethoxy)ethoxy) ethoxy)ethyl)-5-((4R)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl) pentanamido (Biotin-OEG-RTA) 116**



To a solution of amine **118** (140 mg, 0.43 mmol) in DMF (5 mL), was added D-biotin (158 mg, 0.65 mmol), HOBt (87 mg, 0.65 mmol) and EDCI (124 mg, 0.65 mmol), and the reaction mixture was stirred at ambient temperature overnight. The solvents were removed *in vacuo*, and the residue was re-dissolved in CH_2Cl_2 (10 mL). The organic layer was washed with an aqueous solution of LiCl (1N, 5 × 3 mL), saturated aqueous $NaHCO_3$ (aq., 2 × 3 mL), brine (1 × 3 mL), dried over anhydrous Na_2SO_4, filtered and concentrated to give a red solid. The crude product was purified by FCC (SiO_2, eluting with CH_2Cl_2-MeOH-NEt_3 90:9:1) to give the *title compound* as an off-white solid (24 mg, 11%). R_f 0.22 (CH_2Cl_2-MeOH-NEt_3 90:9:1); δ_H (400 MHz, $CDCl_3$) 8.62 (1H, s, N*H*-amide), 7.46 – 7.41 (2H, app. d, *J* 9.0, 2'-C*H*), 6.74 – 6.66 (2H, app. d, *J* 9.0, 3'-C*H*), 6.35 (1H, s, N*H*-amide), 5.68 (1H, s, N*H*-urea), 4.85 (1H, s, N*H*-urea), 4.47 (1H, m, biotin-C*H*), 4.31 – 4.24 (1H, m, biotin-C*H*), 4.16 (1H, s, 2-CH_2), 3.81 – 3.60 (8H, m, 3-CH_2, 4-CH_2, 5-CH_2, 6-CH_2), 3.50 (2H, dd, *J* 9.6, 4.7, 7-CH_2), 3.39 (2H, td, *J* 5.8, 4.1, 8-CH_2), 3.11 (1H, td, *J* 7.5, 4.6, biotin-C*H*), 2.92 (6H, s, N-CH_3), 2.81 (1H, dd, *J* 12.9, 4.6, biotin-C*H*), 2.70 (1H, d, *J* 12.9, biotin-C*H*), 2.12 (1H, td, *J* 7.3, 2.3, 10-CH_2), 1.72 – 1.58 (4H, m, 11-CH_2, 13-CH_2), 1.48 – 1.32 (2H, m, 12-CH_2); δ_C (400 MHz, $CDCl_3$) 173.0 (C(O)-amide), 167.8 (C(O)-amide), 147.7 (*C*-4'), 127.3 (*C*-1'), 121.7 (*C*H-2'), 112.9 (*C*H-3'), 70.87 – 69.7 (CH_2-2, CH_2-3, CH_2-4, CH_2-5, CH_2-6, CH_2-7), 61.6 (*C*H-biotin), 60.0 (*C*H-biotin), 55.2 (*C*H-biotin), 45.7 (CH_3-N), 40.8 (CH_2-biotin), 38.9 (CH_2-8), 35.5 (CH_2-10), 29.5 (CH_2-12), 27.9 (CH_2-11), 25.2 (CH_2-13); IR (neat, ν_{max}/cm^{-1}) 3289, 2923, 2603, 1664, 1519, 1444, 1094; HRMS (ESI): calcd. for $C_{26}H_{42}N_5O_6S$ $[M+H]^+$ 552.2850, found 552.2853.

***N*-[4-(Dimethylamino)phenyl]acetamide (minimal-RTA) 120**[74]



To a solution of *N,N*-dimethyl-*p*-phenylenediamine (100 mg, 0.73 mmol) in CH_2Cl_2 (0.5 mL) was added triethylamine (50 µL, 0.36 mmol) and *N,N*-dimethyl-4-aminopyridine (2 mg, 0.02 mmol), and the reaction mixture was stirred for 5 min at ambient temperature. Acetic anhydride (140 µL, 1.50 mmol) was added dropwise to the reaction mixture, and the resulting mixture was stirred for 2 hours at ambient temperature. The reaction was quenched with water (20 mL) and the product was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic layers were washed with water (1 × 5 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo* to give a brown solid. The crude product was purified by FCC (SiO_2, eluting with 50:50→100:0 EtOAc-Hex) to afford the *title compound* as a grey powder (89 mg, 68%). R_f 0.15 (EtOAc-Hexane 1:1); δ_H (400 MHz, DMSO-d_6) 9.58 (1H, s, N*H*), 7.42 – 7.31 (2H, app. d, *J* 9.1, 2-CH_2), 6.73 – 6.60 (2H, app. d, *J* 9.0, 3-CH_2), 2.83 (6H, s, N-CH_3), 1.97 (3H, s, CH_3); δ_C (100 MHz, DMSO-d_6) 167.3 (*C*=O), 146.9 (*C*-4), 129.3 (*C*-1), 120.5 (*C*H_2-2), 112.7 (*C*H_2-3), 40.5 (*C*H_3-N), 23.7 (*C*H_3); IR (neat, ν_{max}/cm^{-1}) 3183, 2803, 1641, 1519, 1320; HRMS (ESI): calcd. for $C_{20}H_{28}N_4NaO_2$ $[2M+Na]^+$ 379.2106, found 379.2104. Data consistent with literature.

5.3.2 Peptide synthesis

5.3.2.1 General remarks

All amino acids and resins were purchased from either Novabiochem (Merck) or Sigma-Aldrich. All amino acids were *N*-Fmoc protected and side chains protected with Boc (Lys); ^{t}Bu (Asp, Glu, Ser, Thr, Tyr); Trt (Asn, Gln); Pbf (Arg). Synthesis of all peptides was performed using a microwave assisted automated peptide synthesiser (CEM, Liberty Blue). Peptide acetylation was performed manually. DMF used in peptide synthesis was of HPLC grade and from Sigma Aldrich. Peptides were synthesised on a 0.125 mmol scale. Lyophilisation was performed using a BenchTop Pro with Omnitronics™ (VirTis SP Scientific). Preparative HPLC was performed on an Agilent Technologies 1260 infinity controller in conjunction with a diode array detector. Analytical HPLC was performed on an Agilent Technologies 1260 infinity controller in conjunction with a diode array detector. Mass spectrometry data were obtained on a Bruker Daltonics microTOF using electrospray ionisation (ESI) MS instruments as appropriate. FITC-NOXAB (FITC-Ahx-AAQLRRIGDKVNLRQKLLN-

NH_2, prepared by Dr Katherine Horner), FITC-BID (FITC-Ahx-E^{80}DIIRNIARHLAQVGDSN$_L$DRSIW-NH_2, prepared by Dr Katherine Horner) and Ac-BID (Ac-E^{80}DIIRNIARHLAQVGDSN$_L$DRSIW-NH_2, prepared by Dr Claire Grison) were used from stocks prepared previously.

5.3.2.2 Procedure for automated SPPS

Resin Loading: Clean reaction vessel; wash with DMF; transfer resin to reaction vessel; wash with DMF:CH_2Cl_2 (1:1); vessel draining.

Deprotection and Coupling: Add 20% piperidine in DMF (4 mL); microwave method (30 sec); wash with DMF (4 × 4 mL); drain; add amino acid (2.5 mL); add coupling reagent (1 mL); add base (0.5 mL); microwave method (1 min); wash through manifold to waste (2 mL); drain. For the deprotection and coupling of methods that did not use microwave assistance, these were replaced by agitation of the resin at r.t. for 10 min and 90 min, respectively. After the final residue, the resin was ejected from the reaction vessel and cleavage and deprotection was performed manually using methods A to D (*section 5.3.2.3*).

5.3.2.3 Methods for manual solid phase *N*-terminal chain elongation and capping

Method A: Deprotection of N-Fmoc protecting groups

N-terminal Fmoc protecting groups were removed by the addition of 20% piperidine: DMF (5 × 2 mL × 2 min), followed by rinsing the resin with DMF (5 × 2 mL × 2 min). Successful deprotection was determined by a positive colour test (Method B).

Method B: Kaiser Test

The Kaiser Test was used to determine successful coupling or deprotection of manually coupled residues. A few beads of resin were placed in a vial, two drops of each solution (1-3, see below) was added to the beads and the solution was heated to ca. 100 °C for 1 minute. Successful coupling was indicated by no change in colour of the beads, whereas successful deprotection was indicated by bright blue beads.

1) Ninhydrin (5% w/v) in ethanol
2) Phenol (80% w/v) in ethanol
3) 1 mM KCN $_{(aq.)}$ in pyridine (2% v/v)

Method C: N-terminal acetylation

Acetic anhydride (10 equiv.) and DIPEA (10 equiv.) were dissolved in DMF (1 mL) and the solution was transferred to the resin. After 2 h, the resin was drained, washed with DMF (3 × 2 mL × 2 min) and successful capping determined by a negative colour test (Method B).

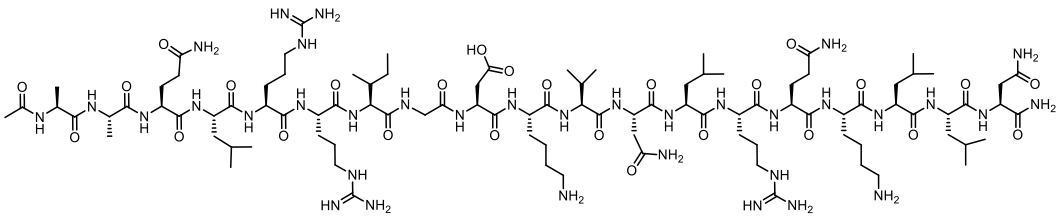
Method D: Cleavage and deprotection of Rink amide MBHA resin

After elongation and *N*-terminal capping was complete, the resin was washed with DMF (5 × 2 mL × 2 min), CH_2Cl_2 (5 × 2 mL × 2 min), Et_2O (5 × 2 mL × 2 min) and dried under vacuum for 1 h. Peptides were simultaneously cleaved and side-chain deprotected using 'Reagent K' TFA:EDT:thioanisole:phenol:H_2O, 82:3:5:5:5 (1 × 2 mL × 2 h). The solution was precipitated in ice-cold Et_2O (50 mL), placed in a centrifuge (3000 rpm × 1 min) and the supernatant was removed. The precipitate was washed with ice-cold Et_2O (3 × 30 mL) and the washed precipitate was dried under a stream of nitrogen (1 h), before being dissolved in H_2O and lyophilised.

5.3.2.4 Peptide purification

Peptides were purified by preparative mass-directed HPLC using a Jupiter Proteo C_{18} preparative column on an increasing gradient of acetonitrile in water + 0.1% HCOOH (v/v) at a flow rate of 10 mL min^{-1}. Crude peptides were suspended in acetonitrile at an approximate concentration of 20 mg mL^{-1}. Purification runs injected a maximum of 0.4 mL of crude peptide solution and were allowed to run for 30 min, with acetonitrile increasing at a stated gradient. The mass directed chromatography software Masshunter by ChemStation (Agilent) was used to allow the collection of the desired peptide by mass, with the eluent split into an Agilent 6120 Quadropole LC-MS which triggers collection of eluent at a programmed m/z. Fractions containing purified peptide were combined, concentrated *in vacuo* and lyophilised.

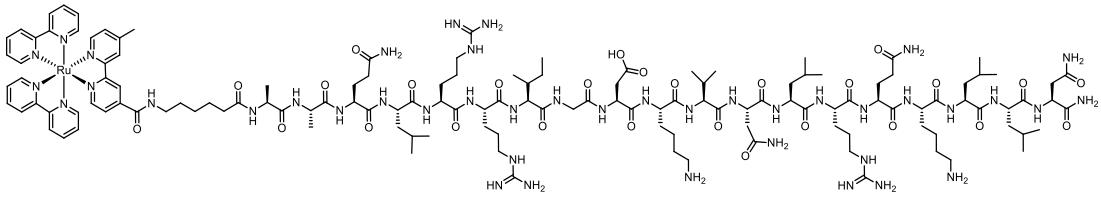
5.3.2.5 Synthesis of *wt* NOXA-B peptide



Ac-AAQLRRIGDKVNLRQKLLN-NH_2

wt NOXA-B was synthesised on a 0.125 mmol scale on rink amide MBHA LL resin (0.35 mmol g^{-1}) using procedures described in *section 5.3.2.2*, all residues following isoleucine were double coupled. The *N*-terminus was acetylated as described in *section 5.3.2.3* (Method C) and the peptide was simultaneously cleaved and side-chain deprotected using Method D. *wt* NOXA-B was purified using mass-directed HPLC over a gradient of 5-50% acetonitrile in H_2O + 0.1% HCOOH, to give (4.1 mg, 4.8%) of the final peptide.

5.3.2.6 Synthesis of Ru(II)(bpy)$_3$-NOXA-B peptide



Ru(II)(bpy)$_3$-Ahx-AAQLRRIGDKVNLRQKLLN-NH_2

Ru(II)(bpy)$_3$-NOXA-B was synthesised on a 0.125 mmol scale on rink amide MBHA LL resin (0.35 mmol g^{-1}) using procedures described in *section 5.3.2.2***,** all residues following isoleucine were double coupled. The *N*-terminus was elongated by dissolving Fmoc-6-aminohexanoic acid (5 equiv.), HCTU (5 equiv.) and DIPEA (5 equiv.) in DMF (2 mL) and the mixture was added to the resin, followed by agitation for 2 h. After removal of the reagents by filtration, the resin was washed with DMF (3 × 2 mL × 2 min) and the success of coupling determined by a negative colour test (Method B). Deprotection of the Fmoc-protected *N*-terminus then followed (Method A). The *N*-terminus was elongated by dissolving Ru(II)(bpy)$_3$$PF_6$ (1.5 equiv.), PyBOP (3 equiv.) and DIPEA (5 equiv.) in DMF (2 mL) and the mixture was added to the resin, followed by agitation for 2 h (double coupling). After removal of the reagents by filtration, the resin was washed with DMF (3 × 2 mL × 2 min) and the success of coupling determined by a negative colour test (Method B). The peptide was simultaneously cleaved and side-chain deprotected using Method D. Ru(II)(bpy)$_3$-NOXA-B was purified using mass-directed HPLC over a gradient of 15-25% acetonitrile in H_2O + 0.1% HCOOH, to give (3.9 mg, 2.6%) of the final peptide.

5.3.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

A resolving gel of appropriate percentage was prepared according to *Table 6*: 12% gels were used in analysis of MCL-1 expression and purification, 15% gels were used

in the analysis of all photolabelling reactions. Following addition of tetramethylethylenediamine (TEMED), the solution was thoroughly mixed and immediately added to the BioRad tetragel apparatus. A layer of propan-2-ol (1 mL) was applied to the top of the gel. When the resolving gel had set, the propan-2-ol was removed and a stacking gel was prepared according to *Table 6.* Once the TEMED was added the solution was thoroughly mixed, immediately added to the top of the resolving gel and comb with a suitable number of lanes was inserted into the stacking gel.

Table 6: *SDS-PAGE gel recipes*

Components	**Resolving gel**		**Stacking gel**
	12%	**15%**	
1.5 M Tris (pH 8.8)	2.53 mL	2.53 mL	-
0.5 M Tris (pH 6.8)	-	-	0.95 mL
40 % (w/v) acrylamide	3.00 mL	3.75 mL	0.62 mL
10% SDS	100 µL	100 µL	50 µL
H_2O	4.25 mL	3.50 mL	3.30 mL
10% APS	100 µL	100 µL	50 µL
TEMED	10 µL	10 µL	10 µL

Once set, the gel was placed inside an electrophoresis tank and SDS running buffer was added. Samples were mixed with an equal volume of loading buffer, heated to 95 °C for 10 minutes and loaded onto the gel. Electrophoresis was performed at a constant voltage of 180V for approximately 50 minutes, or until the loading buffer had reached the bottom of the gel. Proteins were visualised using Coomassie Blue stain (45% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R-250) and destained (30% (v/v) methanol, 10% (v/v) acetic acid), or alternatively using Instant Blue (TripleRed).

5.3.4 Expression and purification of MCL-1

MCL-1 protein expression and purification was performed in collaboration with Dr Fruszina Hobor. The pet28a His-SUMO MCL-1 (172-327) construct was over-expressed in the E.coli strain Rosetta 2. 10 mL of overnight starter culture was used to inoculate 3 L 2 x YT containing 50 µg/mL Kanamycin and 50ug/mL Chloramphenicol. Cultures were grown at 37 °C plus shaking until OD600 ~ 0.6 –

0.8, the temperature was then switched to 18 °C and protein expression induced by the addition of 0.8 mM IPTG. Induced cultures were grown at 18 °C plus shaking overnight before harvesting by centrifugation. Cells were resuspended in 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole and lysed by sonication in the presence of 10 µL of 1 U.ml-1 DNase I per litre of over-expression culture and 5 mM $MgCl_2$. The cell lysate was centrifuged (Beckman JA25.50 rotor, 17,000 rpm, 45 min, 4 °C) and the supernatant filtered (0.22 µM syringe filter) before application onto a 5 mL HisTrap that had previously been equilibrated with 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole. The cleared cell lysate was then allowed to flow through the HisTrap with the aid of a peristaltic pump. The HisTrap was then washed with 10 CV of 50 mM TRIS pH 8.0, 500 mM NaCl, 15 mM imidazole followed by 10 CV 50 mM TRIS pH 8.0, 500 mM NaCl, 50 mM imidazole. The His-SUMO-MCL-1 fusion protein was then eluted from the HisTrap with 50 mM TRIS pH 8.0, 500 mM NaCl, 300 mM imidazole. The His-SUMO-MCL-1 fusion protein was cleaved overnight in dialysis into 50 mM TRIS pH 8.0, 250 mM NaCl in the presence of Smt3 protease, Ulp1 (50 µL/10 mg protein) at 4 °C. To remove any uncleaved MCL-1, His-SUMO and Ulp1, the sample was reapplied to a HisTrap in 50 mM TRIS pH 8.0, 250 mM NaCl and the flow through containing MCL-1 collected. The flow through containing cleaved MCL-1 was concentrated (Amicon Ultra centrifugal filter, MWCO 10,000 Da) to approximately 5 mL. The sample was then filtered before further purification on a Superdex 75 column (GE healthcare) equilibrated in 50 mM TRIS pH 8.0, 250 mM NaCl, 0.5 mM DTT, 2.5% Glycerol. The purified protein (8 mL, 5.62 mg/mL, 318 µM) was concentrated and stored at -80°C.

5.3.5 Protein characterisation

5.3.5.1 Mass spectrometry

High resolution electrospray (ESI+) mass spectrometry was performed on a Bruker MaXis Impact QqTOF mass spectrometer, and *m/z* values are reported in Daltons to four decimal places. LC-ESI-MS data was obtained using an Acquity BEH C_4 column using MeCN/water (1→95%) containing 0.1% HCOOH, over 4 mins in one step at a flow rate of 0.7 ml/min.

5.3.5.2 Circular Dichroism (CD) Spectroscopy

Spectra were recorded on a Chirascan Plus circular dichroism spectropolarimeter (Applied Photophysics) using the following parameters: 190-260 nm range; temperature 20 °C; 5 nm bandwidth; 1 mm path length; scan speed of 5 nm/min. Samples were dissolved in 50 mM sodium phosphate buffer, pH 7.5 at a

concentration of 0.2 mg/mL. The spectra were averaged over 3 repeats with a buffer baseline subtraction.

5.3.6 Fluorescence anisotropy assays: general remarks

Fluorescence anisotropy assays were run in 384 well Optiplates and scanned using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader. Fluorescein labelled peptides used an excitation and emission wavelength of 480 nm (30 nm bandwidth) and 535 nm (40 nm bandwidth) respectively. All assays were run in Tris buffer (50 mM Tris, 150 mM NaCl, 0.01% Triton X-100, pH 7.4) with additives where described.

5.3.7 Direct binding assays

5.3.7.1 Processing of fluorescence anisotropy data

The data obtained for both the *P* (perpendicular intensity) and *S* (parallel intensity) channels were corrected by subtracting the corresponding control wells, and the resulting values were used to calculate intensity (Eq. 1) and anisotropy (Eq. 2) for each well (using Microsoft Excel). These data were transferred into OriginPro 8.5, where a plot of anisotropy against protein concentration was fitted using a logistic model (Eq. 3) to obtain the minimum (r_{min}) and maximum (r_{max}) values of anisotropy. These values were used to determine the fraction of labelled peptide bound to the protein (fraction ligand bound, L_b, Eq. 4), and fitted (Eq. 5) in OriginPro 8.5 to determine the dissociation constant, K_d.

$$I = 2PG + S \qquad \textit{Equation 1}$$

Where *I* is the total intensity, *G* is an instrument factor which was set to 1, and *r* is the anisotropy.

$$r = \frac{S - PG}{I} \qquad \textit{Equation 2}$$

$$y = r_{max} + \frac{r_{min} - r_{max}}{1 + \left({}^{x}/_{x_0}\right)^{p}} \qquad \textit{Equation 3}$$

$$Lb = \frac{r - r_{min}}{(\lambda(r_{max} - r) + r - r_{min})} \qquad \textit{Equation 4}$$

Where x_0 is the midpoint, *p* is the power and λ is the ratio of $I_{bound}/I_{unbound}$ and is equal to 1.

$$y = \frac{(K_d + x + [FL]) - \sqrt{((K_d + x + [FL])^2 - 4x[FL])}}{2} \qquad \textit{Equation 5}$$

Where *Lb* is the fraction ligand bound, [FL] is the concentration of fluorescent ligand, y is Lb*[FL], x = [added titrant].

5.3.7.2 MCL-1/FITC-NOXA-B direct titration

Titration of MCL-1 into NOXA-B was performed in a 384 well plate in Tris Buffer (50 mM Tris, 150 mM NaCl, 0.01% Triton X-100, pH 7.4) with the concentration of MCL-1 starting at 10 µM, diluted over 24 points in a 1/2 regime with [FITC-NOXA-B] fixed at 25 nM. Plates were read after 1 hour and 20 hours incubation. Assays were run in triplicate (both test wells and control wells).

5.3.7.3 BCL-X_L/FITC-BID direct titration

Titration of BCL-X_L into FITC-BID was performed in a 384 well plate in Tris Buffer (50 mM, Tris, 150 mM NaCl, 0.01% Triton-X-100, pH 7.4) + 0.02 mg/mL BSA with the concentration of BCL-X_L starting from 10 µM, diluted over 24 points in a 1/2 regime with [FITC-BID] fixed at 25 nM. Plates were read after 1 hour incubation and 20 hours incubation, data shown is from 1 hour incubation (assay died at 20 hour time point). Assays were run in triplicate (both test wells and control wells).

5.3.8 Competition assays

5.3.8.1 General remarks

All competition assays were performed in 384 well plates with the concentration of peptide competitor serially diluted over 16 or 24 points in a 1/2 regime with [tracer] fixed at 25 nM. The assays consisted of three test rows (containing protein, tracer and competitor), and three control rows (tracer peptide was replaced with buffer). As in *section 5.3.7.1*, the intensity, *I*, and anisotropy, *r*, were calculated using equations 1 and 2, respectively. A plot of anisotropy against competitor concentration was plotted in OriginPro 8.5 and fitted using Eq. 4 (*section 5.3.7.1*) to determine an IC_{50} value.

5.3.8.2 Competition of MCL-1/FITC-NOXA-B interaction by Ru(II)(bpy)$_3$NOXA-B

MCL-1/FITC-NOXA-B competition assays were performed in Tris Buffer (50 mM, Tris, 150 mM NaCl, pH 7.4) + 0.01% Triton-X-100 with [MCL-1] fixed at 75 nM. Assays were run in triplicate (both test wells and control wells). Plates were read after 1 h and 20 h incubation, data shown is from 20 h incubation.

5.3.8.3 Competition of BCL-X_L/FITC-BID interaction by Ru(II)(bpy)$_3$NOXA-B

BCL-X_L/FITC-BID competition assays were performed in Tris Buffer (50 mM, Tris, 150 mM NaCl, 0.01% Triton-X-100, pH 7.4) + 0.02 mg/mL BSA with [BCL-X_L] fixed at 50 nM. Both control wells and test wells were run in triplicate. Plates were read after 1 h and 20 h incubation, data shown is from 20 h incubation.

5.3.9 Ligand-directed photolabelling experiments

5.3.9.1 Labelling of recombinant MCL-1 with TAMRA-RTA

To a solution of MCL-1 (final concentration 5 µM) in ammonium acetate buffer (50 mM, pH 7.5) was added Ru(II)(bpy)$_3$-NOXA-B (final concentration 1 µM), TAMRA-RTA (final concentration 10 µM) and APS (final concentration 10 µM), and the mixture was incubated at r.t. for 5 min. For the competition experiment, Ac-BID (final concentration 0-1000 µM) was added and the mixture was incubated for 10 min prior to incubation with Ru(II)(bpy)$_3$-NOXA-B. The mixture was irradiated for 1 min at r.t., 5 cm from the light source (Kessil H150W-BLUE LED, 32W, 2 × lamps). The reaction was immediately quenched by the addition of DTT (final concentration 1 mM) and analysed using ESI-MS (*section 5.3.5.1*) and/or SDS-PAGE (*section 5.3.3*). Fluorescently modified peptides were analysed using a Molecular Imager ChemiDoc

XRS System (Bio-Rad, CA). After obtaining the fluorescence image, the same gel was stained with Coomassie Brilliant Blue (CBB) and visualised on a Molecular Imager ChemiDoc XRS System. ImageJ was used to quantify bands on the gel.

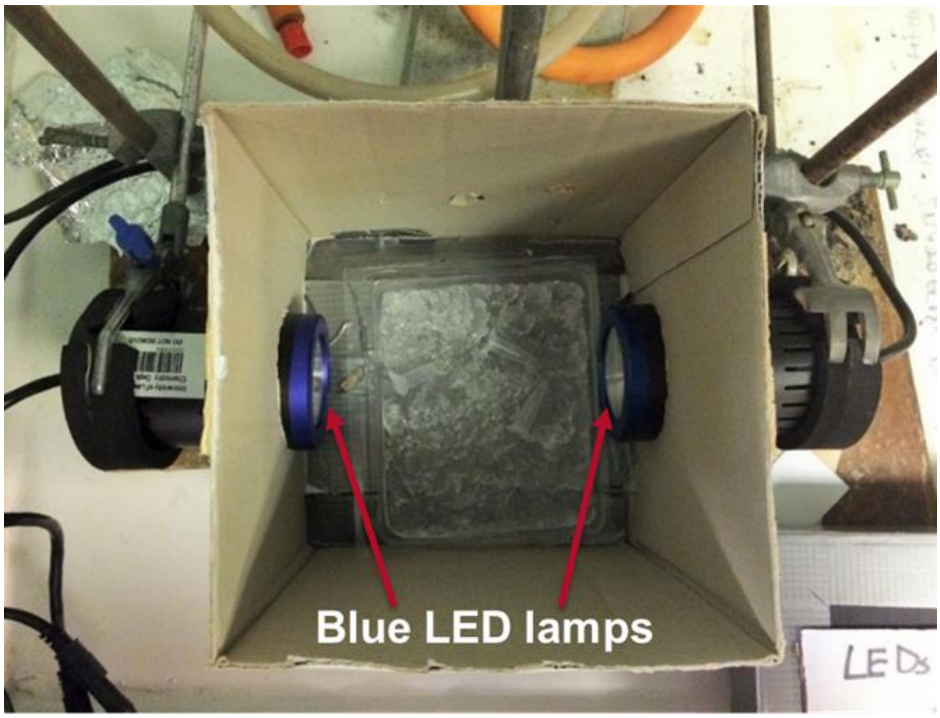


Figure 52: *Photolabelling set-up: two Kessil H150W-BLUE LED lamps facing each other, parallel to base of the box containing the samples in Eppendorf tubes.*

5.3.10 Labelling of recombinant MCL-1 with biotin-RTA

To a solution of MCL-1 (5 μM) in ammonium acetate buffer (50 mM, pH 7.5) was added Ru(II)$(bpy)_3$-NOXA-B (final concentration 1 μM), Biotin-RTA (final concentration 5 μM) and APS (final concentration 10 μM) and the mixture was incubated at r.t. for 5 min. The mixture was irradiated for 1 min at r.t., 5 cm from light source (Kessil H150W-BLUE LED, 32W, 2 × lamps). The reaction was immediately quenched by the addition of dithiothreitol (final concentration 1 mM) and analysed using ESI-MS (*section 5.3.5.1*) and/or SDS-PAGE (*section 5.3.5.2*). Gels (15% acrylamide) were run at a constant voltage of 180 V for 45 min, then the gel was stained with Coomassie Brilliant Blue (CBB) and the image obtained on a Molecular Imager ChemiDoc XRS.

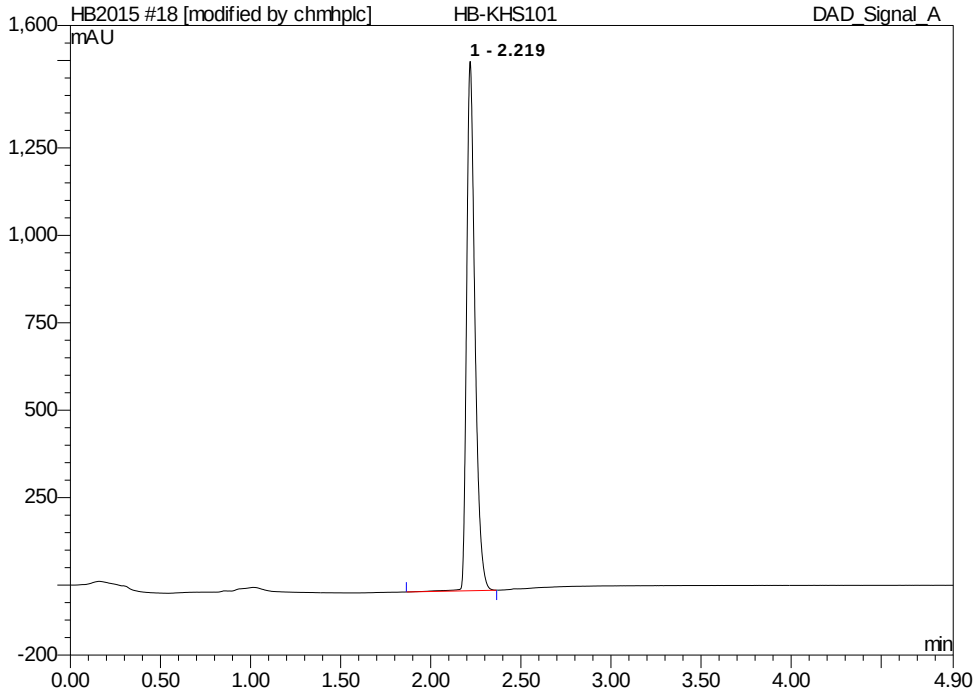
5.3.11 Peptide mapping of RTA-labelled MCL-1

To a solution of MCL-1 (50 µM) in ammonium acetate buffer (50 mM, pH 7.5) was added Ru(II)$(bpy)_3$-NOXA-B (final concentration 10 µM), RTA **120** (final concentration 10 µM) and APS (final concentration 10 µM) and the mixture was incubated at r.t. for 5 min. The mixture was irradiated for 1 min at r.t., 5 cm from light source (Kessil H150W-BLUE LED, 32W, 2 × lamps) and immediately quenched by the addition of dithiothreitol (final concentration 10 mM). The small molecules were removed using a centrifugal filter (MWCO: 10,000, Amicon Ultra, 0.5 mL). In-solution digestion and MS/MS analysis was performed by Dr Rachel George. Glu-C and trypsin solutions (20 ng μL^{-1} in 25 mM ammonium bicarbonate) were added in a 1:50 ratio (protease:total protein content). Samples were incubated at 37 ºC with shaking for 18 h. The digest reaction was stopped by adding 5 µL of 1% HCOOH, then subjected to purification using a Sep-pak 18 column (see *section 5.2.6.3*). The sample was subjected to RP-HPLC (see *section 5.2.6.4*) and the labelled peptides were analysed using a Xevo G2-XS Q-TOF mass spectrometer. Data processing was performed using the MassLynX v4.1 suite of software. Peptide MS/MS data were processed with PEAKS Studio (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada).

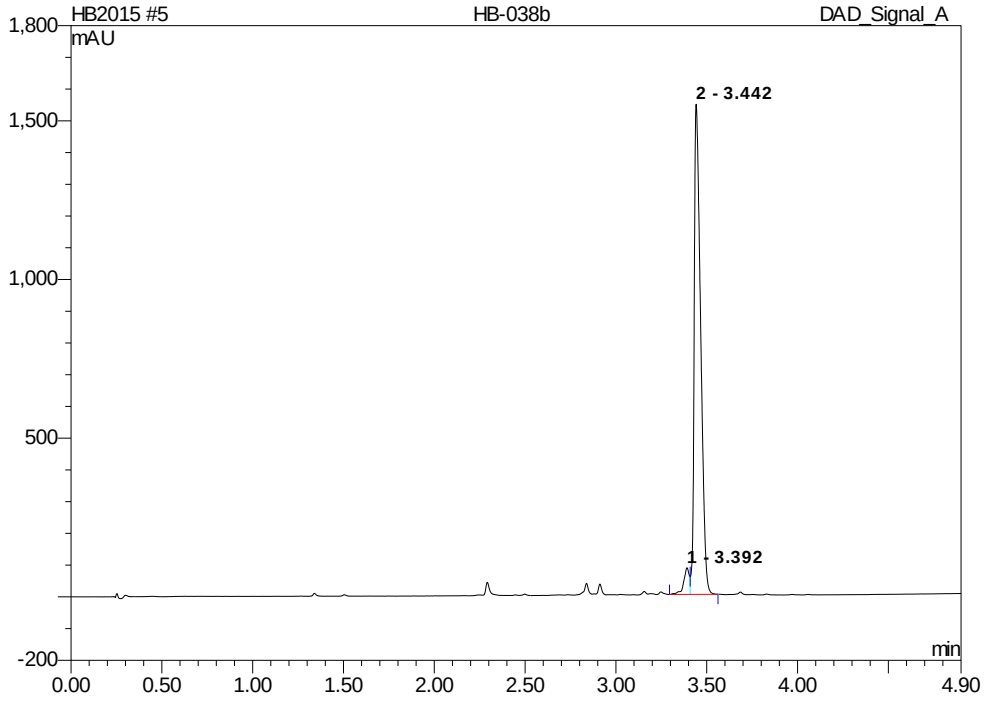
Appendix I

Analytical HPLC data for KHS101 analogues

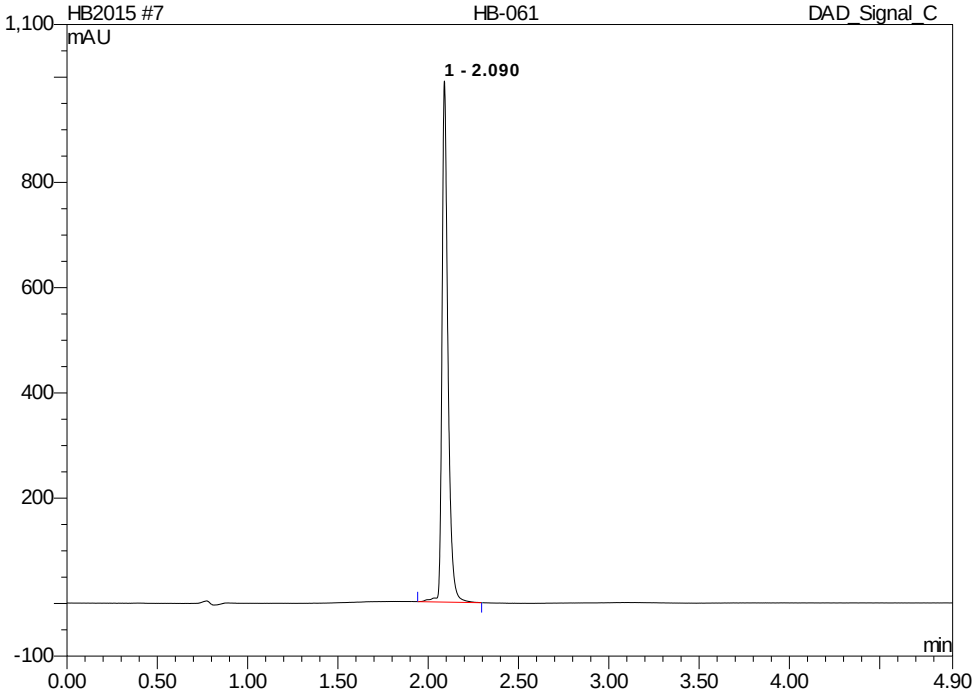
Analytical HPLC data for compound **41** (**KHS101**)



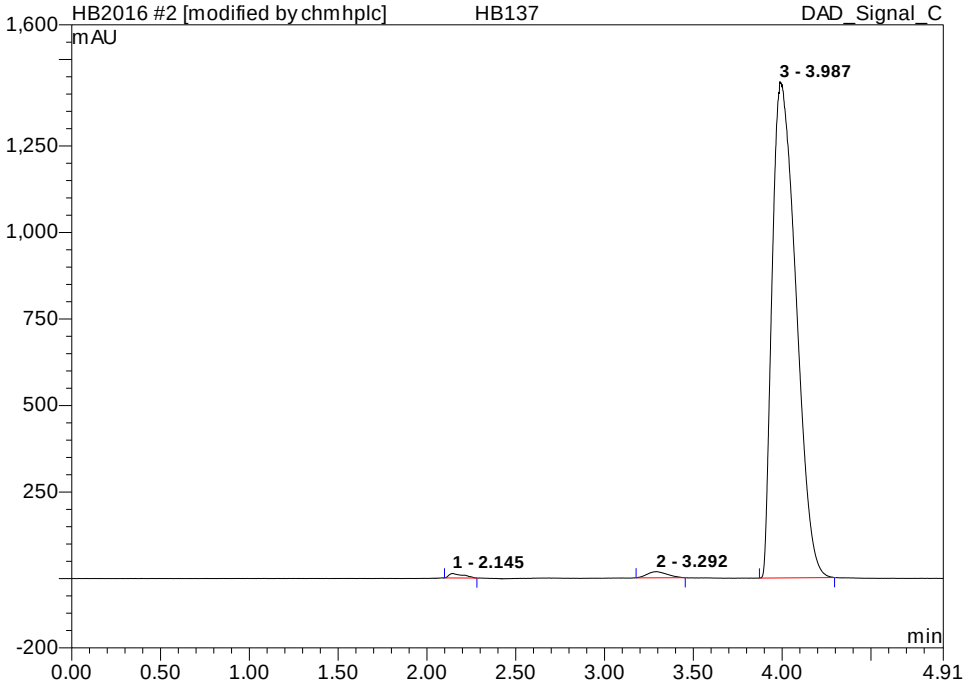
Analytical HPLC data for compound **51a** (**KHS101-side product**)



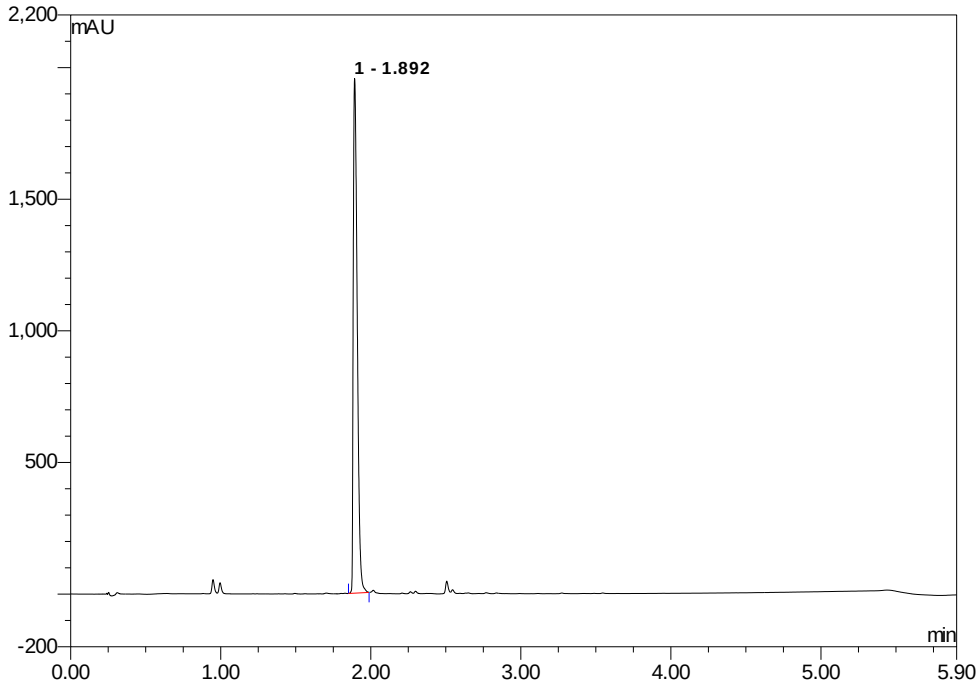
Analytical HPLC data for compound **61** (**KHS101-BP**)



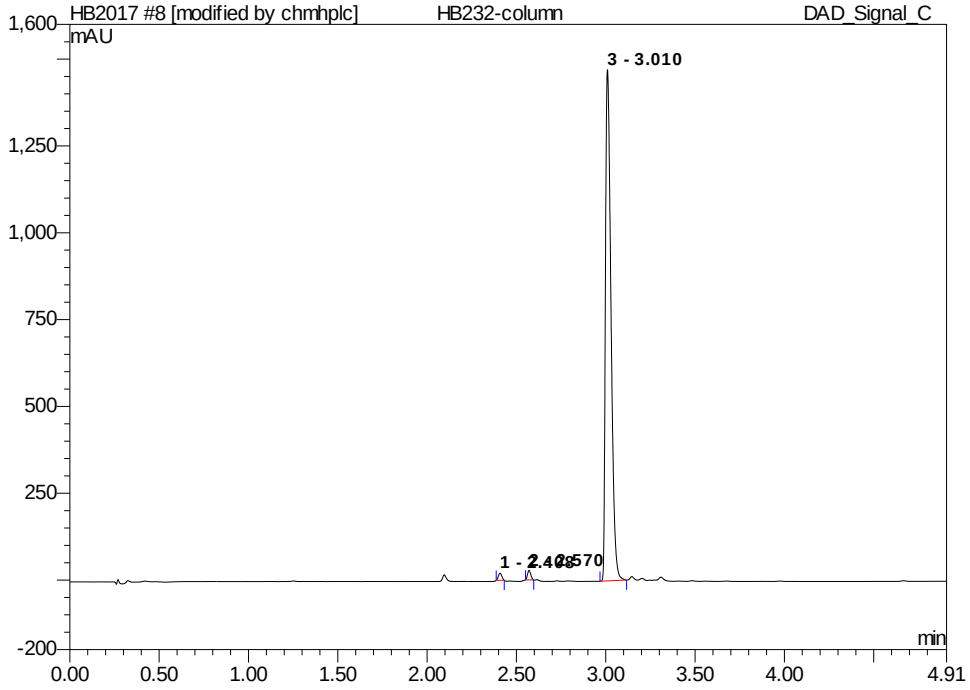
Analytical HPLC data for compound **62** (**KHS101-Biotin**)



Analytical HPLC data for compound **60** (**HB072**)



Analytical HPLC data for compound **89** (**KHS101-Diazirine**)



HIF-1α/p300 fluorescence anisotropy assay

The structure of regioisomer **51a** resembles alpha-helix mimics of HIF-1α (hypoxia inducible factor, alpha subunit, see *Figure 53*).[142,143] This could be important to investigate further in light of recent literature suggesting that KHS101 plays a role in regulating the interaction between HIF-1β (also termed ARNT) and its co-activator TACC3.[86] The transcription factor HIF-1 is a heterodimer, composed of one alpha (HIF-1α) and one beta subunit (HIF-1β/ARNT). Gardner and co-workers revealed an indirect mode of function for KHS101 on HIF complex formation; instead of binding directly to the interacting domains of HIF-1β or TACC3, KHS101 promoted turnover of both TACC3 and HIF-1α in NPCs. HIF-1 plays a major role in regulating the hypoxic response, and high levels of HIF-1 are known to correlate with tumour metastasis and angiogenesis - indicating the potential for HIF-1 inhibitors to reduce tumour formation and progression.[144] Compound **51a** was tested in a fluorescence anisotropy competitive HIF-1α/p300 assay (p300 is a transcriptional co-activator of HIF-1α), using a 42-residue FITC-labelled HIF-1α peptide (derived from the *C*-terminal transactivation domain) and the p300 CH1 domain (residues 330-420).[142] However, no binding of **51a** to p300 was observed.

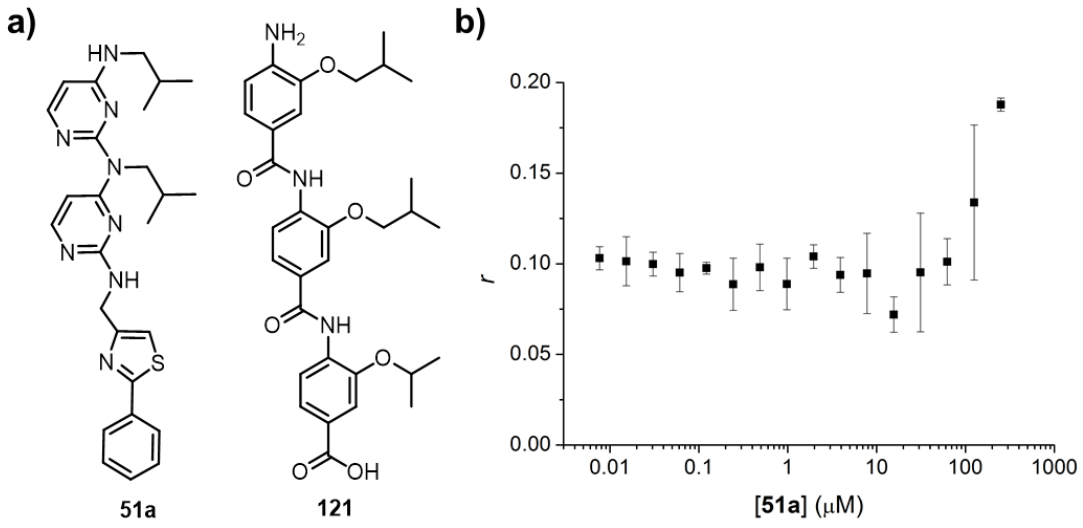


Figure 53: *Investigation of **51a** as an alpha helix mimic of HIF-1α. a) Chemical structures of a HIF-1α mimic known to inhibit the HIF-1α/p300 interaction (oligoamide **121**, IC_{50} 9.2 μM) and side product from KHS101 synthesis (**51a**). b) Fluorescence anisotropy competition assay for the inhibition of FITC-HIF-1α/p300 by **51a**. Error bars represent the standard deviation of three replicates.*

Fluorescence anisotropy assays were run in 384 well Optiplates and scanned using a Perkin Elmer EnVision™ 2103 MultiLabel plate reader, with each well run in triplicate. Fluorescein labelled peptides used an excitation and emission wavelength of 480 nm and 535 nm (5 nm bandwidths). Assays were run in Tris buffer (20 mM Tris, 100 mM NaCl, 1 mM DTT, pH 7.46). The fluorescent tracer used in the fluorescence anisotropy assay was a 42-residue HIF-1α peptide derived from the *C*-terminal transactivation domain (C-TAD), purchased from Proteogenix, France. The sequence is given below:

FITC-Ahx-SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN-NH_2

For the direct titration, protein was serially diluted into a solution of FITC-HIF-1α C-TAD, with the concentration of p300 starting at 15 µM, diluted over a ½ regime with [FITC-HIF-1α C-TAD] fixed at 50 nM. Plates were read after 10 minutes of incubation.

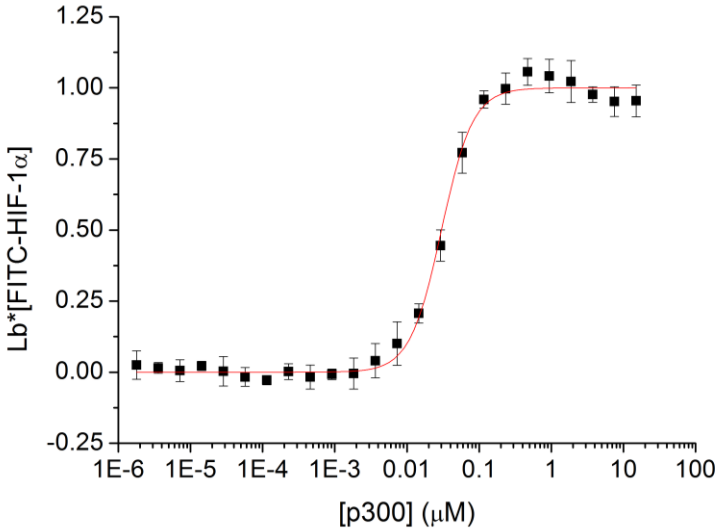


Figure 54: *Direct titration of p300 into a constant concentration of FITC-HIF-1α. Error bars represent the standard deviation of 3 replicates. Plot of bound labelled peptide against protein concentration to give a K_d of 30.1 ± 1.9 nM (which corresponds to that obtained by Kristina Paraschiv, unpublished results).*

For the competition assay, **51a** was serially diluted over 16 points in a ½ regime with [FITC-HIF-1α C-TAD] fixed at 25 nM. The assays consisted of three test rows (containing protein, tracer and competitor), and three control rows (where the tracer peptide was replaced with buffer). As in *section 5.2.5.1*, the intensity, *I*, and anisotropy, *r*, were calculated using equations 1 and 2, respectively. A plot of anisotropy against competitor concentration was plotted in OriginPro 8.5 (***figure 6***).

Immobilisation of KHS101-biotin for SPR studies

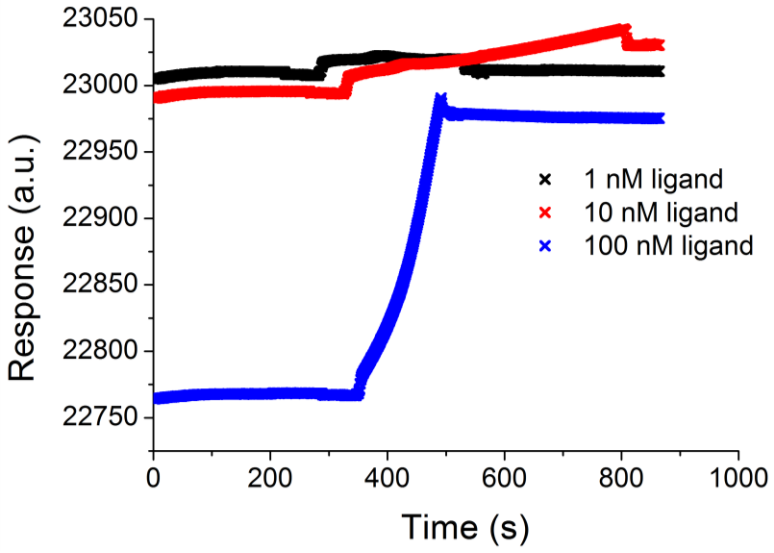


Figure 55: *Immobilisation of flow cells with KHS101-biotin at concentrations of 1 nM, 10 nM and 100 nM, respectively. Sensorgrams demonstrate successful derivatisation of each flow cell with their respective KHS101-biotin concentrations.*

MS/MS analysis of unmodified HSPD1

(continued on page 140)

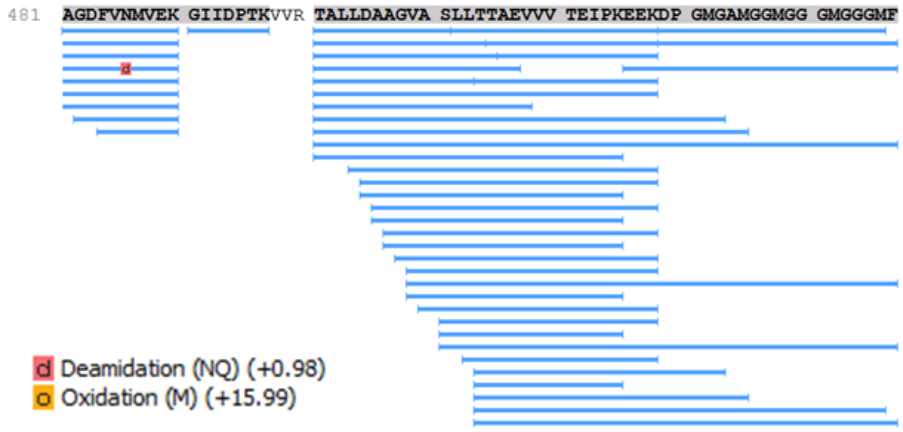


Figure 56: *Results from peptide mapping of unmodified HSPD1. Sample contained 10 µM HSPD1 in 50 mM $(NH_4)HCO_3$, 10 mM $MgCl_2$, pH 7.0. Protease used: trypsin. Sequence coverage (75%). Peptides detected in MS are shown in blue, with post-translational modifications deamidation and oxidation indicated by a red 'd' and orange 'o'.*

MS/MS analysis of HSPD1 crosslinked with KHS101-diazirine

(Continued on page 142)

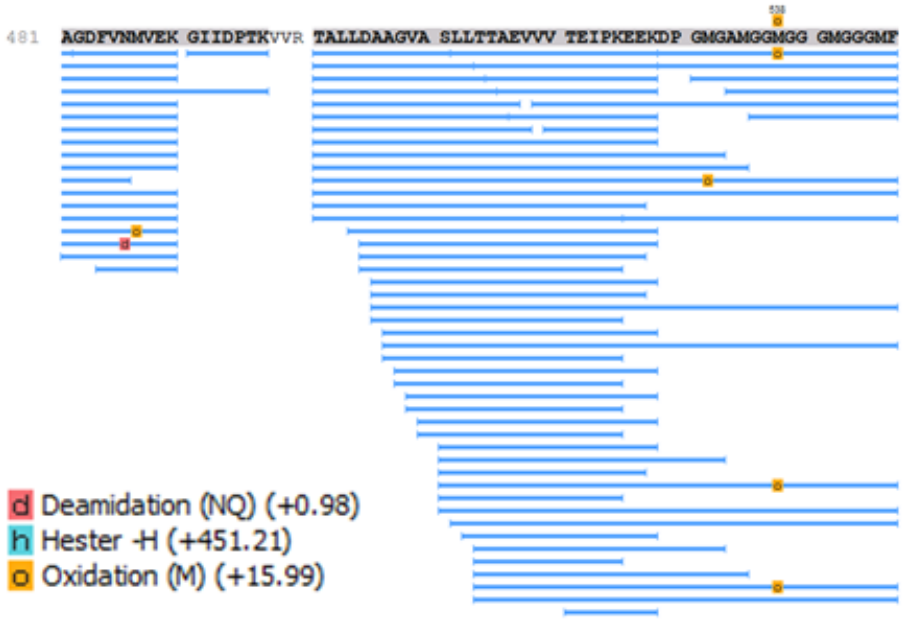


Figure 57: *Results from peptide mapping of HSPD1 modified with KHS101-diazirine. Conditions for photocrosslinking reaction: 10 µM HSPD1, 30 µM KHS101-diazirine, 30 min irradiation at 365 nm; buffer: 50 mM $(NH_4)HCO_3$, 10 mM $MgCl_2$, pH 7.0. Protease used: trypsin. Sequence coverage (80%). Peptides detected in MS are shown in blue, with post-translational modifications deamidation and oxidation indicated by a red 'd' and orange 'o'. Residues modified by KHS101-diazirine are denoted by a blue 'h' (+451.21 Da), however the exact site of modification cannot be confirmed.*

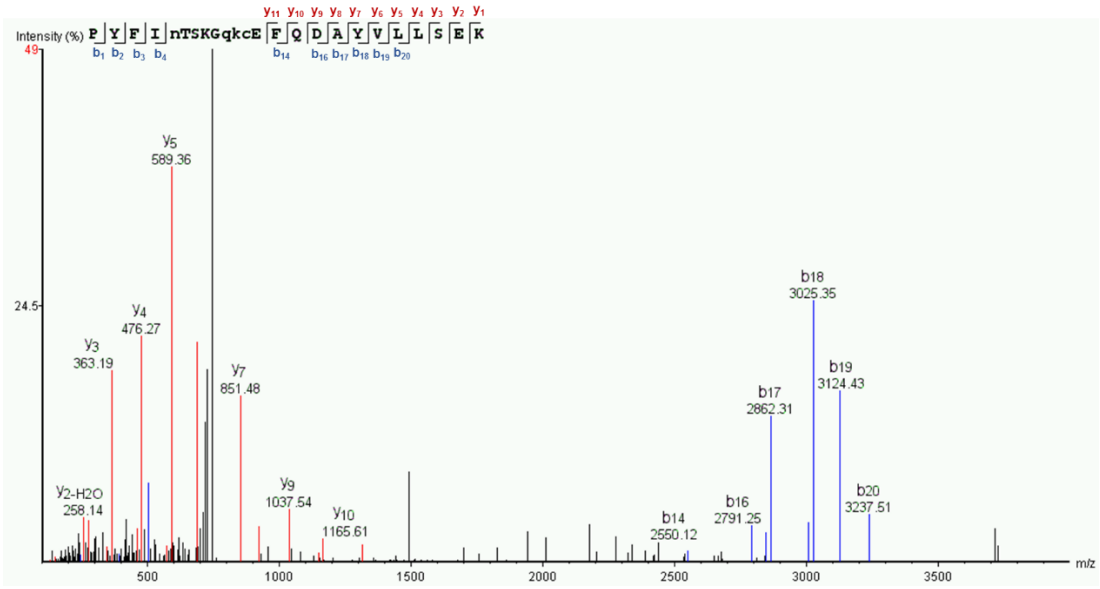


Figure 58: *Q-TOF MS/MS spectrum for 'peptide 1' modified with KHS101-diazirine. Observed y and b ions are shown in red and blue, respectively.*

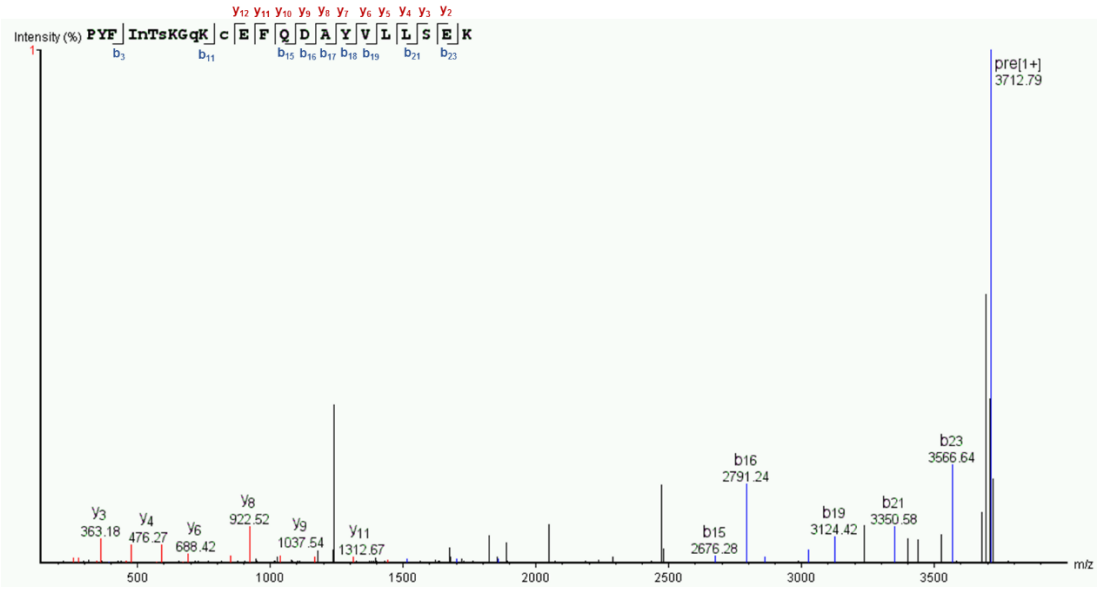


Figure 59: *Q-TOF MS/MS spectrum for 'peptide 2' modified with KHS101-diazirine. Observed y and b ions are shown in red and blue, respectively.*

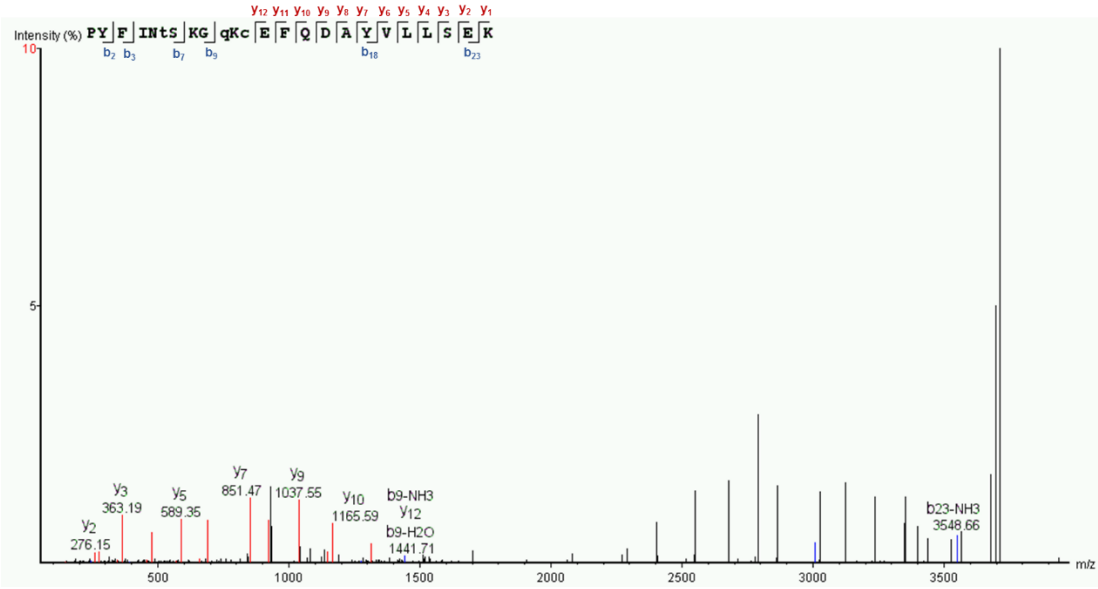


Figure 60: *Q-TOF MS/MS spectrum for 'peptide 3' modified with KHS101-diazirine. Observed y and b ions are shown in red and blue, respectively.*

Appendix II

Attempted synthesis of acyl-imidazole-based Tag Transfer Reagent

Initial efforts focussed on the synthesis of modular chemical tools that could be rapidly assembled from a key building block (**122**, *Figure 61*). This would enable TTRs based on any binding peptide to be synthesised, allowing the labelling of a range of different proteins. To this end, a "click-able" 2-cyano-1,3-benzothiazole handle was incorporated into the design of the modular TTRs, allowing peptides bearing an *N*-terminal cysteine residue to be attached in a rapid and reagent-free manner. The 1,2-aminothiol-cyanobenzothiazole (CBT) ligation between 2-cyano-1,3-benzothiazole with β-aminothiols is highly selective, does not require a copper catalyst (the use of which has shown to be toxic to cells)[145] and is 3 orders of magnitude faster than the Staudinger ligation.[146] Acyl-imidazole chemistry (see *section 1.4.3.*, Chapter 1) was chosen as the reactive entity, permitting traceless labelling of a protein of interest.

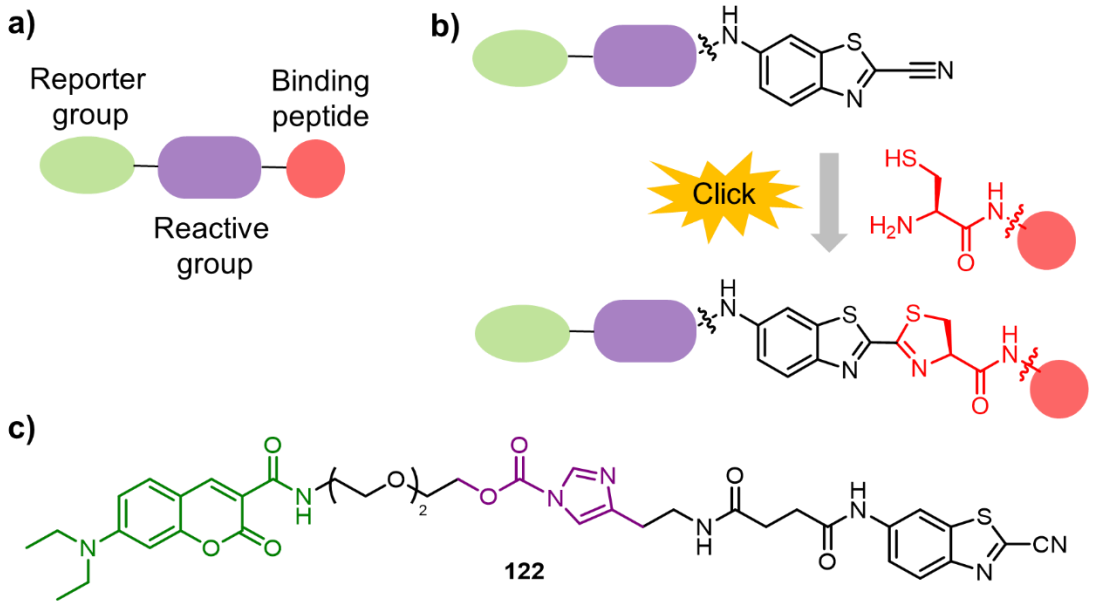
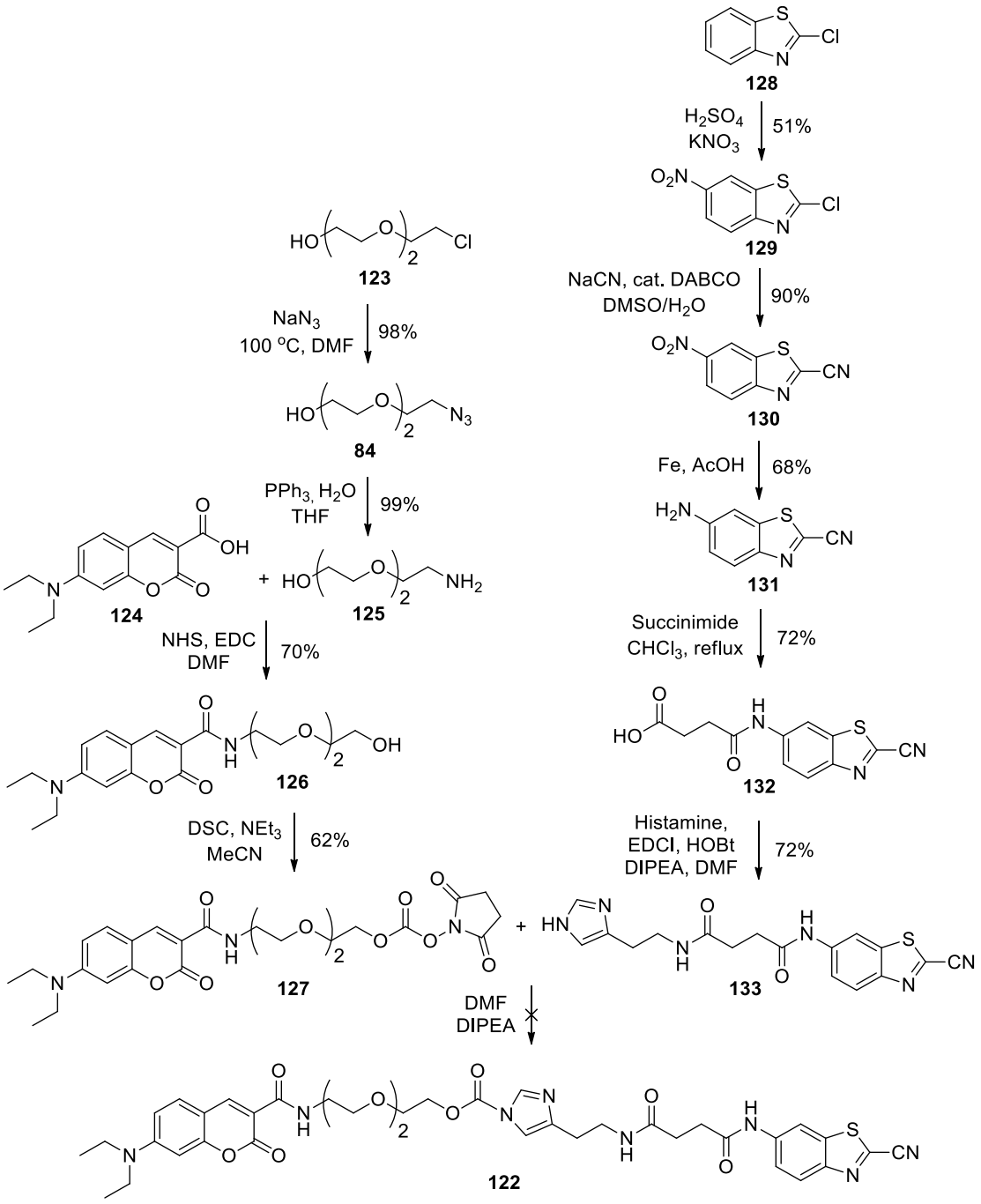


Figure 61: *Design of modular TTRs based on acyl-imidazole chemistry. a) General structure of fully assembled Tag Transfer Reagent. b) Schematic representing coupling of CBT-functionalised building block with binding peptide via a bioorthogonal click reaction. c) Chemical structure of CBT-functionalised building block* ***122****.*

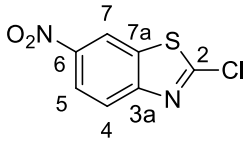
A convergent synthesis of building block **122** was followed (outlined in *Scheme 21*). The route to CBT intermediate **131** was optimised within the Bon group.[122] Due to its sensitivity towards nucleophilic attack, the carbamate bond connecting the imidazole to the oligo ethylene glycol (OEG) chain was planned as

the final synthetic step. Unfortunately, the coupling of NHS carbonate **127** and histamine **133** failed to afford any product, possibly due to hydrolysis of carbonate **127** or carbamate **122**. Further efforts to obtain **122** were not pursued, while investigations into the use of Ru(II)(bpy)$_3$-mediated photolabelling commenced.



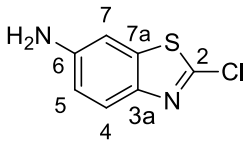
Scheme 21: *Synthetic route to building block **122** based on acyl-imidazole chemistry. Final synthetic step failed to afford any product. Synthetic route to ACBT intermediate **131** was optimised, and the procedure was published (Hauser et. al.).*[122]

2-Chloro-6-nitro-1,3-benzothiazole 129[147]



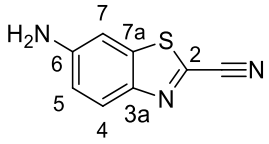
2-Chloro-1,3-benzothiazole (5.0 g, 29.5 mmol) was added portionwise to concentrated H_2SO_4 (30 mL) in a cooled round-bottomed flask (ice bath), followed by the portionwise addition of potassium nitrate (3.3 g, 32.4 mmol). The mixture was stirred at 0 °C (ice bath) for 30 min, then for 1 hour at ambient temperature. The solution was poured onto ice (150 mL) and filtered; the collected solid was washed with ice-cold water until the filtrate was acid-free. The crude product was dried *in vacuo* to remove water, then purified by recrystallisation from EtOH to afford the *title compound* as a yellow solid (3.2 g, 51%); m.p. 190.1-190.6 °C (lit.[148] 190-191 °C); δ_H (300 MHz, $CDCl_3$) 8.75 (1H, d, *J* 2.3, 7-C*H*), 8.39 (1H, dd, *J* 9.0, 2.3, 5-C*H*), 8.08 (1H, d, *J* 9.0, 4-C*H*); δ_C (125 MHz, $CDCl_3$) 158.9 (*C*-3a), 154.9 (*C*-2), 145.6 (*C*-6), 136.6 (*C*-7a), 123.5 (*C*H-4), 122.4 (*C*H-5), 117.8 (*C*H-7); IR (neat, ν_{max}/cm^{-1}) 1326, 748; HRMS (ESI): calcd. for $C_7H_3ClN_2O_2S$ $[M+Na]^+$ 236.9496, found 236.1646.

2-Chloro-6-amino-1,3-benzothiazole 130[149]



Tin(II) chloride dihydrate (23.1 g, 102.5 mmol) was added to a stirred solution of **129** (4.3 g, 19.8 mmol) in EtOH (50 mL), and the mixture was heated to reflux for approximately 1 hour. After cooling to room temperature, the reaction mixture was poured onto ice, neutralised to pH 7 with saturated $NaHCO_3$ solution, then adjusted to pH 9 with 4M aqueous NaOH solution. The resulting mixture was stirred for 1 hour, extracted with EtOAc (4 × 20 mL), then washed with brine (1 × 20 mL). Charcoal was added to the organic layer, which was stirred for 10 minutes. The charcoal was removed by filtration through celite; the filtrate was concentrated to afford the *title compound* as a pale pink solid (40%, 1.47 g). m.p. 141.2-143.1 °C; R_f 0.29 (EtOAc-Hexane 1:3); δ_H (300 MHz, $CDCl_3$) 7.70 (1H, d, *J* 8.7, 4-C*H*), 6.99 (1H, d, *J* 2.3, 7-C*H*), 6.81 (1H, dd, *J* 8.7, 2.3, 5-C*H*), 3.84 (2H, s, NH_2); δ_C (125 MHz, $CDCl_3$) 148.9 (*C*-2), 145.3 (*C*-6), 144.5 (*C*-3a), 138.2 (*C*-7a), 123.8 (*C*H-4), 116.1 (*C*H-5), 105.1 (*C*H-7); IR (neat, ν_{max}/cm^{-1}) 3310, 1326, 809; HRMS (ESI): calcd. for $C_7H_5ClN_2S$ $[M+H]^+$ 184.9940, found 184.9948.

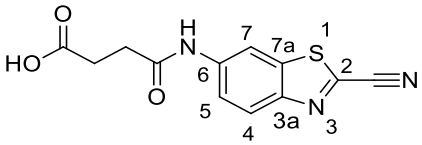
2-Cyano-6-amino-1,3-benzothiazole 131[150]



Potassium cyanide (2.92 g, 44.8 mmol) was dissolved in DMSO (225 mL) by heating to 140 °C with stirring for 16 hours. The mixture was cooled to 115 °C, then **130** (1.65 g, 8.9 mmol) was added as a solution in DMSO (65 mL), using a syringe pump (1 mL every 5 minutes). The reaction mixture was stirred at 115 °C for 12 hours. The reaction mixture was cooled to room temperature and then poured into a solution of KH_2PO_4 (0.2 M, 200 mL). The aqueous solution was extracted with EtOAc (6 x 100 mL). The combined organics were washed with brine (1 × 100 mL), dried (Na_2SO_4) and concentrated to yield the crude product as a brown solid. The crude material was purified by recrystallisation from ethanol to give the *title compound* as yellow needles (546 mg, 46%). m.p. 194.2 °C (decomposed); R_f 0.21 (DCM); δ_H (300 MHz, DMSO-d_6) 7.86 (1H, d, *J* 9.0, 4-C*H*), 7.15 (1H, d, *J* 2.2, 7-C*H*), 6.96 (1H, dd, *J* 9.0, 2.2, 5-C*H*), 6.14 (2H, s, NH_2); δ_C (125 MHz, DMSO-d_6) 150.5 (*C*-2), 143.2 (*C*-6), 138.4 (*C*-3a), 127.8 (*C*-7a), 125.2 (*C*H-4), 114.3 (*C*H-5), 117.6 (*C*≡N), 102.3 (*C*H-7); IR (neat, ν_{max}/cm^{-1}) 3349, 2219 (C≡N), 1150, 809; HRMS (ESI): calcd. for $C_8H_5N_3S$ $[M+H]^+$ 176.0278, found 176.0282.

Additional safety precautions taken: reaction was covered by a blast shield, medical O_2 breathing apparatus was available nearby along with a first aider trained in its administration. The reaction was poured was into KH_2PO_4 to quench the DMSO/KCN solution after the reaction had terminated, and all glassware and needles were quenched in a bleach bath. All aqueous washes were quenched with bleach. Three pairs of gloves were worn during manipulation of any point in the reaction. The COSHH form was signed off by the deputy head of the safety committee, Bruce Turnbull.

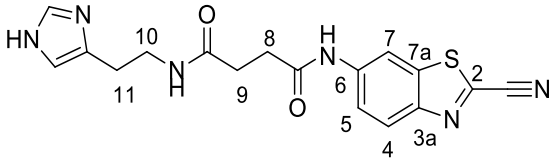
4-[(2-Cyano-6-benzothiazolyl)amino]-4-oxobutanoic acid 132



131 (450 mg, 2.57 mmol) and succinic anhydride (385 mg, 3.86 mmol) were suspended in chloroform (60 mL) and heated to reflux for 20 hours. The reaction mixture was cooled to room temperature and concentrated to afford a cream solid. The crude material was purified by recrystallisation from ethanol to give the *title compound* as off-white needles (510 mg, 72%). m.p. 194.4-195.2 °C; R_f 0.2 (DCM–

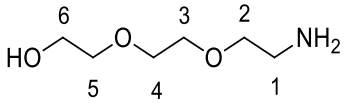
MeOH 98:2); δ_H (300 MHz, DMSO-d_6) 8.75 (1H, d, *J* 2.1, 7-C*H*), 8.19 (1H, d, *J* 9.0, 4-C*H*), 7.72 (1H, dd, *J* 9.0, 2.1, 5-C*H*), 2.73 – 2.56 (4H, m, alkyl-CH_2); δ_C (125 MHz, DMSO-d_6) 176.2 (*C*-amide), 173.2 (*C*-acid), 149.6 (*C*-3a), 141.0 (*C*-6), 138.1 (*C*-2), 136.4 (*C*-7a), 125.9 (*C*H-4), 122.0 (*C*H-5), 114.1 (*C*≡N), 112.4 (*C*H-7), 32.5 (*C*H_2-alkyl), 29.8 (*C*H_2-alkyl); IR (neat, ν_{max}/cm^{-1}) 3322, 2233 (C≡N), 1673; HRMS (ESI): calcd. for 297.0184 $[M+Na]^+$, found 297.0179.

N-(2-cyano-1,3-benzothiazol-6-yl)-*N*'-[2-(1H-imidazol-4-yl)ethyl]butane diamide 133



To a solution of **132** (100 mg, 0.36 mmol) in anhydrous DMF (10 mL) was added histamine dichloride (67 mg, 0.36 mmol), HOBt•H_2O (74 mg, 0.55 mmol), EDCI•HCl (104 mg, 0.55 mmol) and DIPEA (190 µL, 1.09 mmol). The resulting mixture was stirred for 16 hours at ambient temperature. The solvent was evaporated to give an orange residue, which was purified by reverse-phase column chromatography (C_{18}, eluting with H_2O (0.1% HCOOH)-MeCN (0.1% HCOOH) 95:5→60:40 to give the *title compound* as an orange solid (96 mg, 72%). δ_H (500 MHz, D_2O) 8.51 (1H, s, imidazole-C*H*), 8.09 (1H, d, *J* 2.1, 7-C*H*), 7.78 (1H, d, *J* 9.0, 4-C*H*), 7.36 (1H, d, *J* 9.0, 2.1, 5-C*H*), 7.24 (1H, s, imidazole-C*H*), 3.51 (2H, t, *J* 6.6, 11-CH_2), 2.92 (2H, t, *J* 6.6, 10-CH_2), 2.71 (2H, t, *J* 6.9, 9-CH_2), 2.59 (2H, t, *J* 6.9, 8-CH_2); δ_C (125 MHz, D_2O) 174.8 (*C*-amide), 172.9 (*C*-amide), 147.1 (*C*-3a), 138.1 (*C*-2), 136.2 (*C*-6), 135.5 (*C*H-histamine), 133.1 (*C*-7a), 130.9 (*C*-histamine), 124.1 (*C*H-4), 120.9 (*C*H-5), 116.2 (*C*H-histamine), 112.9 (*C*≡N), 111.2 (*C*H-7), 38.1 (*C*H_2-alkyl), 31.8 (*C*H_2-alkyl), 30.5 (*C*H_2-alkyl), 24.2 (*C*H_2-11); IR (neat, ν_{max}/cm^{-1}) 3198, 1629, 1568; HRMS (ESI): calcd. for $C_{17}H_{17}N_6O_2S$ $[M+H]^+$ 369.1128, found 369.1128.

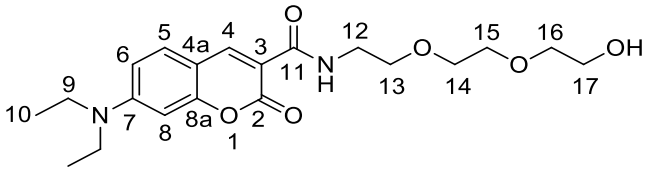
2-(2-(2-Aminoethoxy)ethoxy)ethanol 125[151]



Azide **84** (2.5 g, 14.19 mmol) was dissolved in THF (30 mL) and the solution was cooled to 0°C (ice bath). To this solution was added triphenylphosphine (4.1 g, 15.62 mmol) and the mixture was stirred for 6 hours at ambient temperature. The reaction mixture was then diluted into water (30 mL) and washed with toluene (2 × 20 mL). The aqueous layer was concentrated to give the *title compound* as a clear yellow oil (2.1 g, 99%). δ_H (500 MHz, $CDCl_3$) 3.58–3.55 (2H, m, 6-CH_2), 3.52–

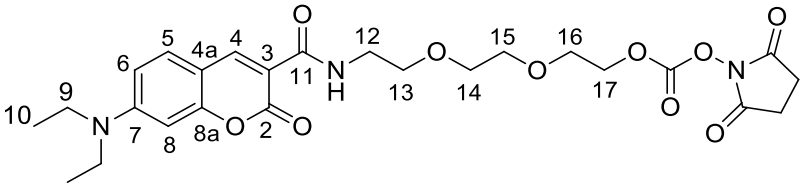
3.48 (6H, m, 3-C*H* ,4-C*H*, 5-CH_2), 3.44 (2H, t, *J* 5.0, 2-CH_2), 2.73 (2H, t, *J* 5.0, 1-CH_2); δ_C (125 MHz, $CDCl_3$) 72.6 (*C*H_2-4), 72.2 (*C*H_2-5), 70.1 (*C*H_2-3), 70.0 (*C*H_2-2), 61.0 (*C*H_2-6), 41.0 (*C*H_2-1); IR (neat, ν_{max}/cm^{-1}) 3357, 1066; HRMS (ESI): calcd. for $C_6H_{15}NO_3$ $[M+H]^+$, 150.1130 found 150.1135.

7-(diethylamino)-*N*-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)-2-oxo-2H-chromene-3-carboxamide 126



To a solution of 7-diethylaminocoumarin-3-carboxylic acid (1 g, 3.83 mmol) in DMF (50 mL), was added *N*-hydroxysuccinimide (563 mg, 4.9 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (969 mg, 5.05 mmol). The resulting mixture was stirred for 5 minutes at ambient temperature. After stirring, **125** (742 mg, 4.98 mmol) was added, and the mixture was stirred at ambient temperature for 16 hours. The solvent was co-evaporated with toluene to give a residue that was dissolved in $CHCl_3$ (50 mL). The solution was washed with saturated aqueous $NaHCO_3$ (3 × 20 mL) and the organic layer was dried over anhydrous Na_2SO_4 and then concentrated *in vacuo*. The crude orange oil was purified by flash column chromatography (SiO_2; DCM-MeOH 99:1→98:2) to give the *title compound* as an orange oil (1.37 g, 70%). R_f 0.25 (DCM-MeOH 98:2); δ_H (500 MHz, $CDCl_3$) 9.10 (1H, d, *J* 5.8, amide-N*H*), 8.70 (1H, s, 4-C*H*), 7.42 (1H, d, *J* 8.9, 5-C*H*), 6.64 (1H, dd, *J* 8.9, 2.5, 6-C*H*), 6.50 (1H, d, *J* 2.5, 8-C*H*), 3.77 (2H, t, *J* 4.4, 12-CH_2), 3.73–3.59 (10H, m, 13-CH_2, 14-CH_2, 15-CH_2, 16-CH_2, 17-CH_2), 3.45 (4H, q, *J* 7.1, 9-CH_2), 1.64 (1H, s, O*H*), 1.23 (6H, t, *J* 7.2, 10-CH_3); δ_C (125 MHz, $CDCl_3$) 163.2 (*C*-2), 162.8 (*C*-11), 157.6 (*C*-8a), 152.5 (*C*-3), 148.1 (*C*-7), 131.1 (*C*-5), 110.3 (*C*-4), 109.9 (*C*-6), 108.4 (*C*-4a), 96.6 (*C*-8), 72.8-69.6 (*C*H_2-13, *C*H_2-14, *C*H_2-15, *C*H_2-16), 61.9 (*C*H_2-17), 45.1 (*C*H_2-9), 39.4 (*C*H_2-12), 12.4 (*C*H_3-10); IR (neat, ν_{max}/cm^{-1}) 3253, 2968, 1695, 1509; HRMS (ESI): calcd. for $C_{40}H_{56}N_4NaO_{12}$ $[2M+Na]^+$ 807.3787, found 807.3810.

2-[2-(2-{[7-(diethylamino)-2-oxo-2H-chromen-3-yl]formamido}ethoxy) ethoxy]ethyl 2,5-dioxopyrrolidin-1-yl carbonate 127



To a solution of **126** (350 mg, 0.89 mmol) in MeCN (15 mL) was added *N,N'*-disuccinimidyl carbonate (686 mg, 2.67 mmol) and NEt_3 (0.37 mL, 2.67 mmol). The resulting mixture was stirred for 1 hour at ambient temperature. The mixture was diluted with water (100 mL) and extracted with EtOAc (3 × 75 mL). The organics were washed with saturated aqueous $NaHCO_3$ (2 × 50 mL), dried over anhydrous Na_2SO_4, filtered and concentrated to give the *title compound* as an orange oil (293 mg, 62%). δ_H (500 MHz, $CDCl_3$) 9.06 (1H, s, amide-N*H*), 8.70 (1H, s, 4-C*H*), 7.44 (1H, d, *J* 9.0, 8-C*H*), 6.64 (1H, dd, *J* 9.0, 2.5, 6-C*H*), 6.49 (1H, d, *J* 2.5, 5-C*H*), 4.58–4.41 (2H, m, OEG-CH_2), 3.87–3.80 (2H, m, OEG-CH_2) 3.70 – 3.65 (8H, m, OEG-CH_2), 3.45 (4H, q, *J* 7.2, 9-CH_2), 2.69 (4H, t, succinimide-CH_2), 1.23 (6H, t, *J* 7.1, 10-CH_3); δ_C (125 MHz, $CDCl_3$) 197.9 (*C*-succinimide), 172.3 (*C*-2), 168.7 (*C*-8a), 165.3 (*C*-3), 162.6 (*C*H-4), 157.7 (*C*-carbonate), 152.7 (*C*-11), 148.4 (*C*-7), 131.4 (*C*H-5), 110.0 (*C*H-6), 108.4 (*C*-4a), 96.6 (*C*H-8), 70.8 – 68.5 (*C*H_2-13, *C*H_2-14, *C*H_2-15, *C*H_2-16, *C*H_2-17), 45.1 (*C*H_2-9), 39.6 (*C*H_2-12), 25.4 (*C*H_2-succinimide), 12.4 (*C*H_3-10); IR (neat, ν_{max}/cm^{-1}) 1785, 1697, 1615, 1206, 1073; HRMS (ESI): calcd. for $C_{25}H_{32}N_3O_{10}$ $[M+H]^+$ 534.2082, found 534.2092.

Appendix III

Analytical data for *wt* NOXA-B peptide

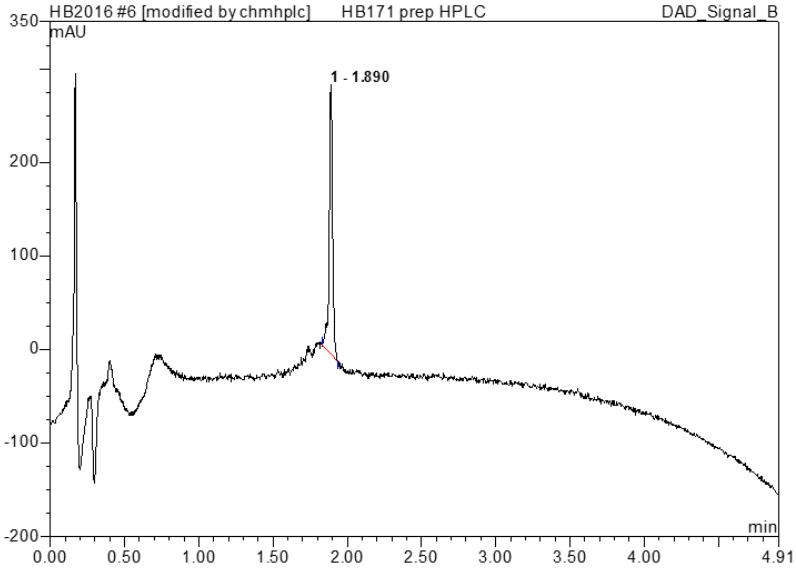


Figure 62: *Analytical HPLC data for wt NOXA-B peptide*.

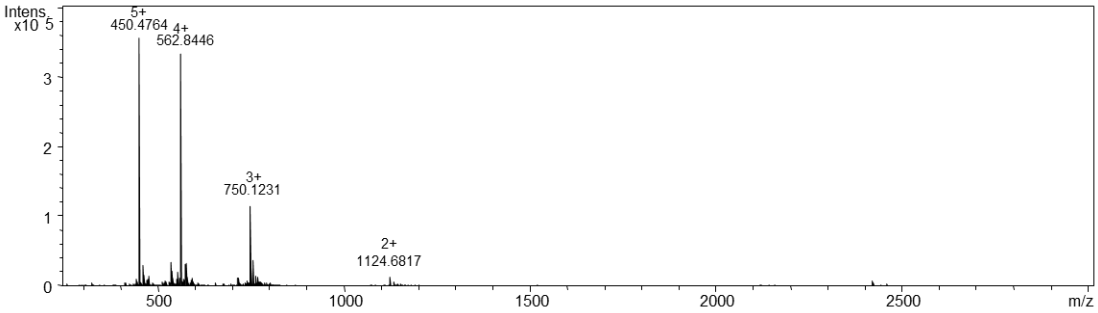


Figure 63: *HRMS (ESI) spectrum for wt NOXA-B peptide.*

Table 7: *Tabulated HRMS data for wt NOXA-B peptide showing the expected (Exp[d]) and observed (Obs[d]) masses for the multiple charge states.*

Peptide	$[M+3H]^{3+}$ Obs[d]	$[M+3H]^{3+}$ Exp[d]	$[M+4H]^{4+}$ Obs[d]	$[M+4H]^{4+}$ Exp[d]
wt NOXA-B	749.7887	749.7889	562.8939	562.2762

MCL-1/FITC-NOXA-B direct titration

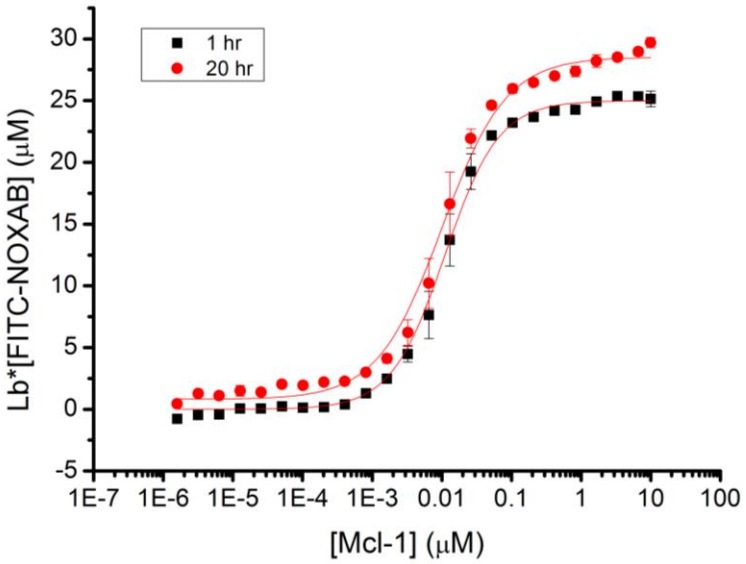


Figure 64: *Direct titration of MCL-1 into FITC-NOXA-B to give a K_d of 11 ± 2.7 nM that corresponds to that described previously (10 ± 8.9 nM, Dr Katherine Horner, unpublished results). Error bars represent the standard deviation of three replicates.*

BCL-X_L/FITC-BID direct titration

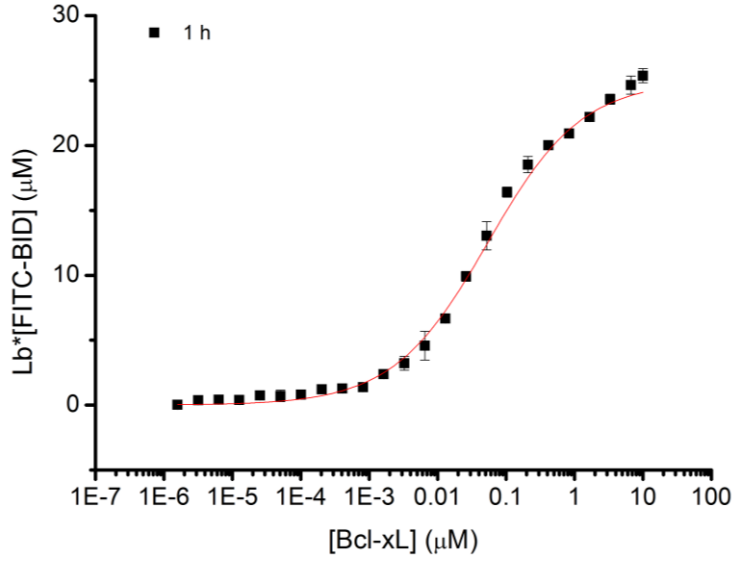


Figure 65: *Direct titration of BCL-X_L into a constant concentration of FITC-BID to give a K_d of 54 ± 8.3 nM that corresponds to that described previously (79 ± 6.0 nM).*[131] *Error bars represent the standard deviation of three replicates.*

Labelling of recombinant MCL-1 with TAMRA-RTA

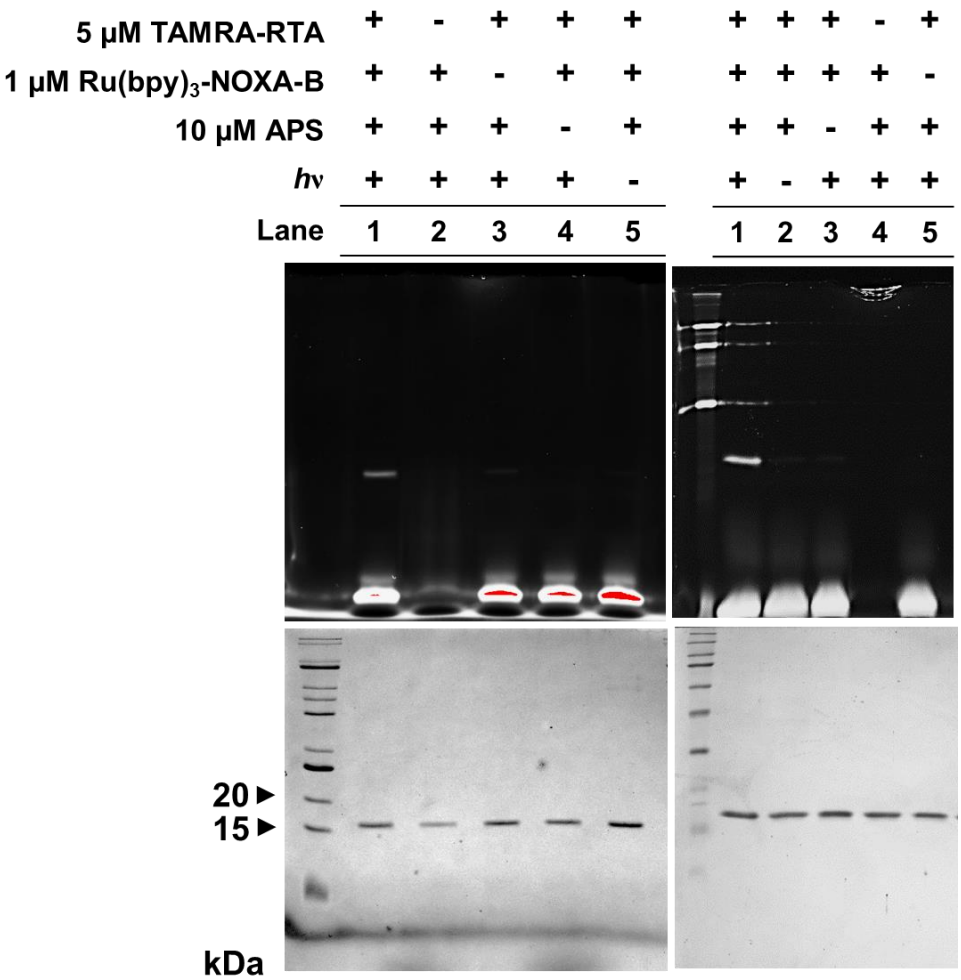


Figure 66: *Photolabelling of MCL-1 with TAMRA-RTA* ***111****. Fluorescence image (top) and Coomassie Brilliant Blue (CBB, bottom) stained SDS-PAGE gels; fluorescence image shows MCL-1 only labelled when 1 min hν, 5 µM TAMRA-RTA and 1 µM Ru(II)(bpy)$_3$-NOXA-B are present (5 µM MCL-1 and 10 µM APS used), CBB stain shows MCL-1 for all conditions.*

Selective labelling of MCL-1 with TAMRA-RTA over BCL-X_L and *h*DM2

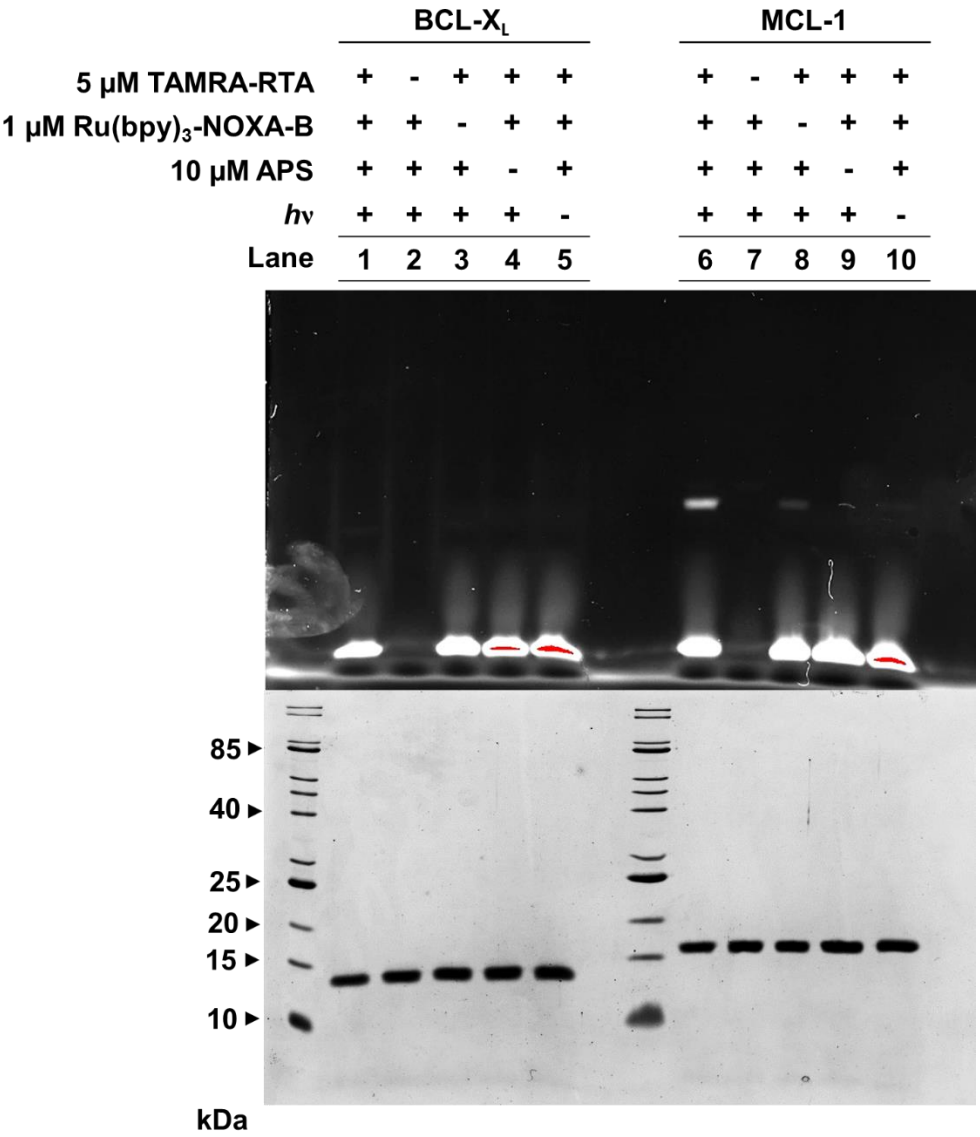


Figure 67: *Photolabelling of BCL-X_L and MCL-1 with TAMRA-RTA* ***111****. Fluorescence image (top) and Coomassie Brilliant Blue (CBB, bottom) stained SDS-PAGE gels; fluorescence image shows MCL-1 only labelled when 1 min hν, 5 µM TAMRA-RTA and 1 µM Ru(II)(bpy)$_3$-NOXA-B are present (5 µM MCL-1 and 10 µM APS used), CBB stain shows BCL-X_L and MCL-1 for all conditions.*

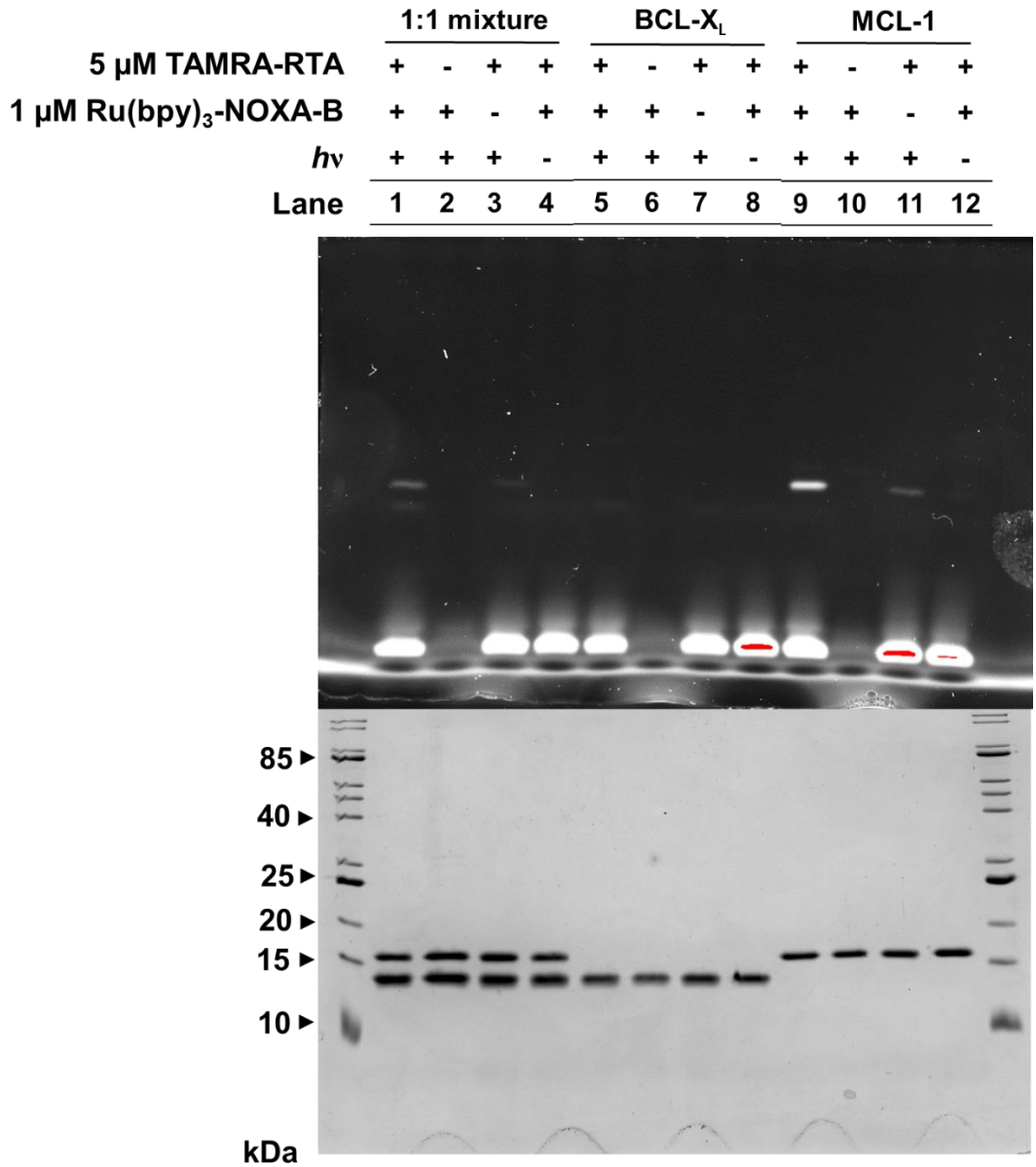


Figure 68: *MCl-1 is selectively labelled with TAMRA-RTA* ***111*** *over BCL-X_L in a 1:1 mixture of proteins, while BCL-X_L is not labelled in solution on its own. Conditions: 5 µM protein, 5 µM TAMRA-RTA and 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min hν.*

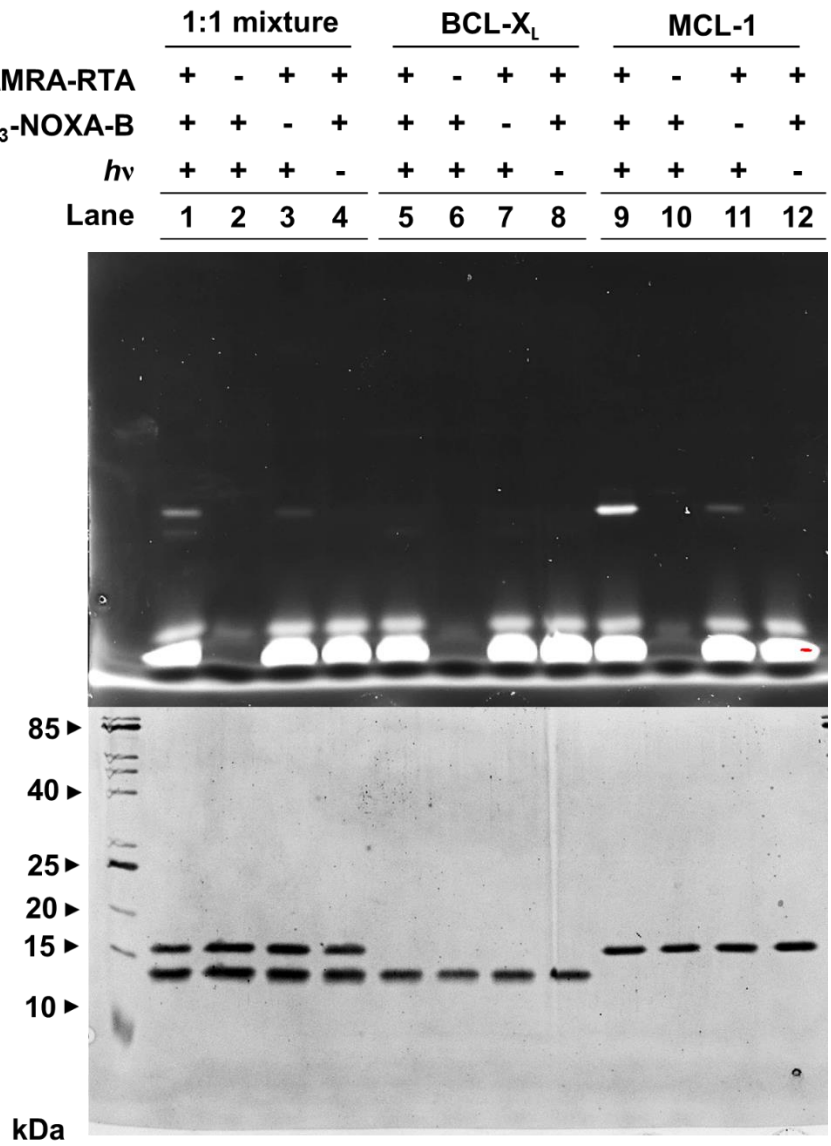


Figure 69: *MCl-1 is selectively labelled with TAMRA-RTA* ***111*** *over BCL-X_L in a 1:1 mixture of proteins, while BCL-X_L is not labelled in solution on its own. Conditions: 5 µM protein, 5 µM TAMRA-RTA and 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min hν.*

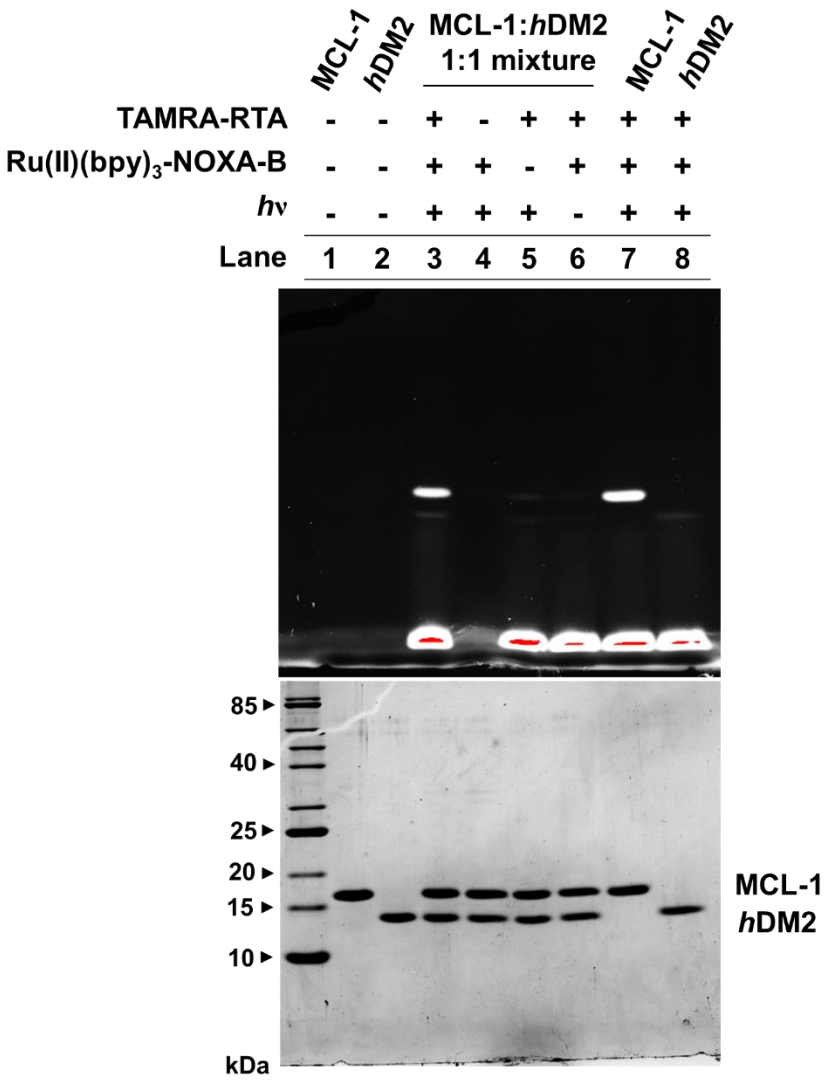


Figure 70: *MCL-1 is selectively labelled over hDM2 in a mixture of proteins (lane 1). Conditions: 5 µM protein, 5 µM TAMRA-RTA and 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min hν.*

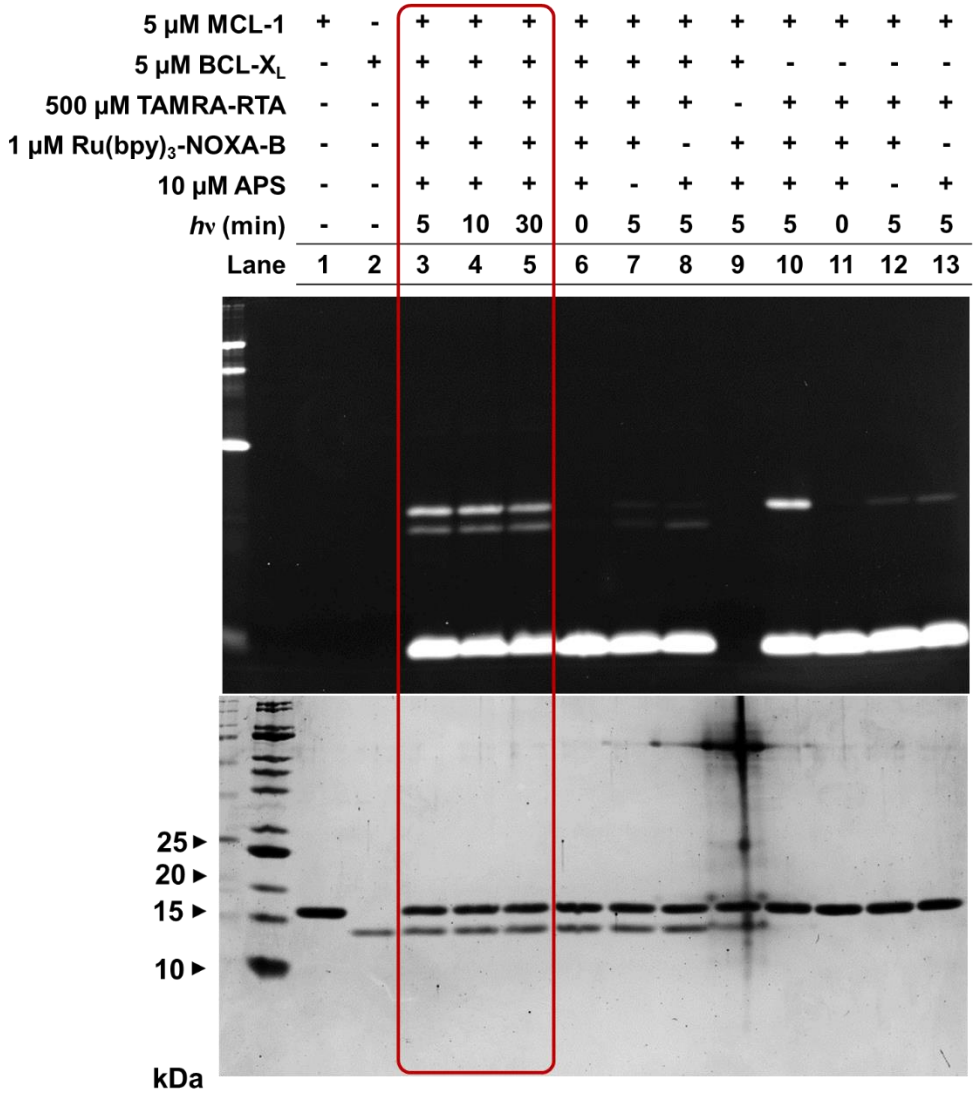


Figure 71: *Both MCL-1 and BCL-X_L are labelled in a 1:1 mixture of proteins (lanes 3, 4 and 5). Conditions: 5 µM protein, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 500 µM TAMRA-RTA, 10 µM APS, 50 mM $(NH_4)HCO_3$ (pH 7.4), 1 min hν.*

Competition experiments using Ac-NOXA-B and Ac-BID

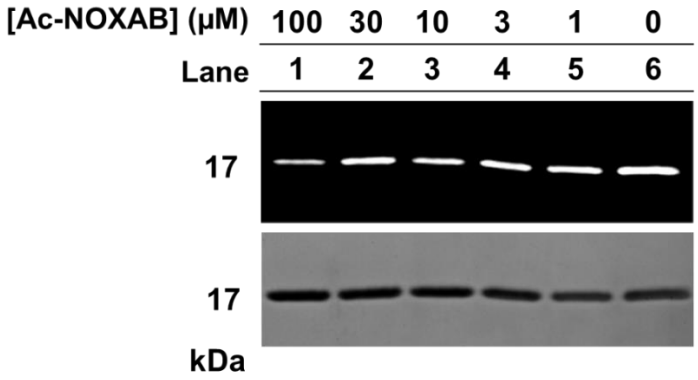


Figure 72: *Competition experiment using Ac-NOXA-B. Conditions: 5 µM MCL-1, 5 µM TAMRA-RTA, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min hν, 0-100 µM Ac-NOXA-B.*

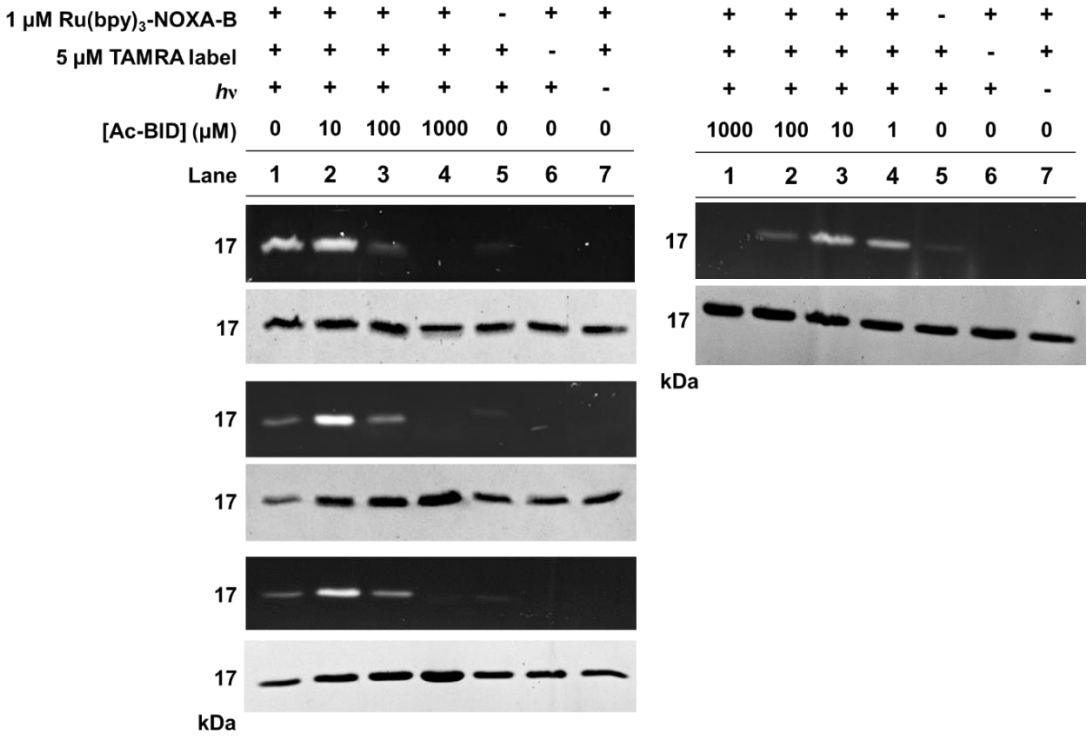


Figure 73: *Competition experiments with Ac-BID. Conditions: 5 µM MCL-1, 5 µM TAMRA-RTA, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 10 µM APS, 1 min hν, 0-1000 µM Ac-NOXA-B.*

ESI-MS analysis of MCL-1 photolabelling reactions

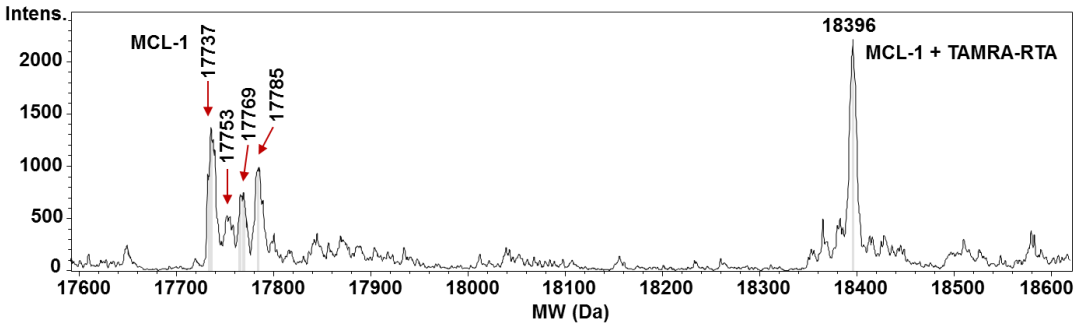


Figure 74: *ESI-MS data showing photolabelling of MCL-1 with TAMRA-RTA after 1 minute of irradiation. Conditions: 5 µM protein, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 5 µM TAMRA-RTA, 10 µM APS, 50 mM (NH$_4$)HCO$_3$ (pH 7.4). Masses indicative of oxidation of amino acid residues on unlabelled MCL-1 are present in the mass spectrum (+16 Da).*

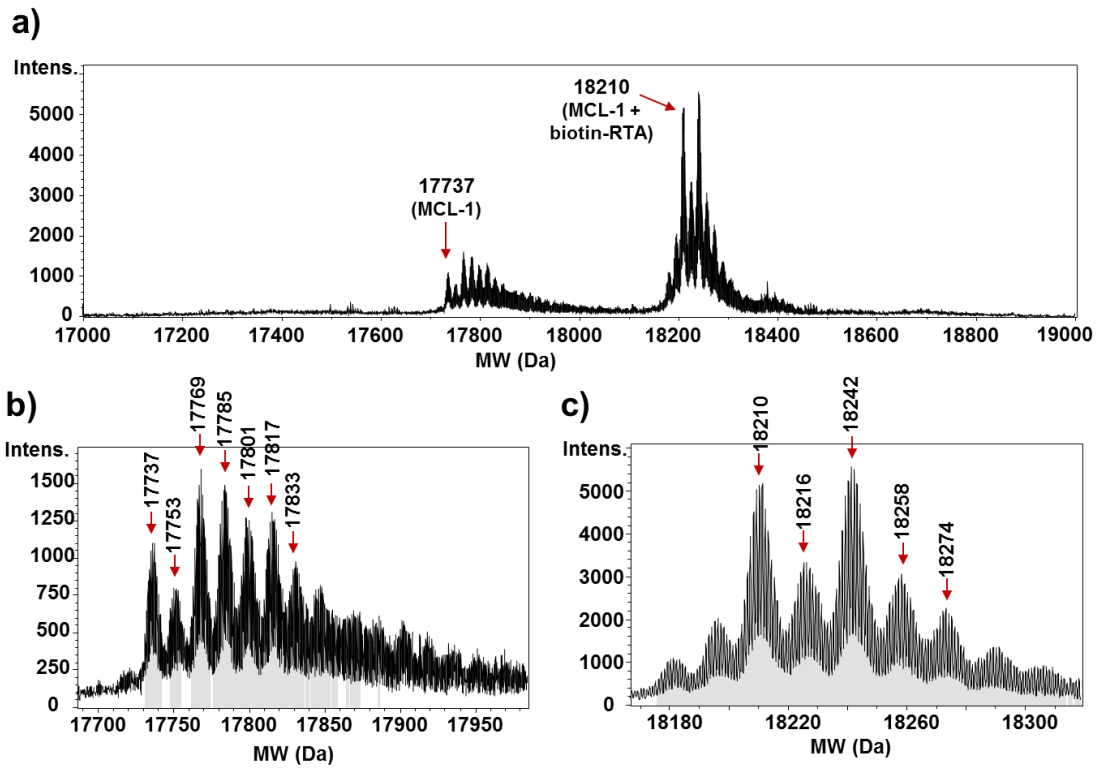


Figure 75: *a) ESI-MS data showing photolabelling of MCL-1 with biotin-RTA after 15 mins irradiation. Conditions: 5 µM protein, 1 µM Ru(II)(bpy)$_3$-NOXA-B, 5 µM TAMRA-RTA, 10 µM APS, 50 mM (NH$_4$)HCO$_3$ (pH 7.4). b) Zoom in on MCL-1 peak, showing masses +16 Da, indicating oxidation of amino acid residues. b) Zoom in on labelled MCL-1 peak, showing masses +16 Da, indicating oxidation of amino acid residues.*

MS/MS analysis of MCL-1 modified with minimal RTA

Figure 76: *MS/MS results from unmodified MCL-1. Sequence coverage 97%, using Glu-C and trypsin proteases.*

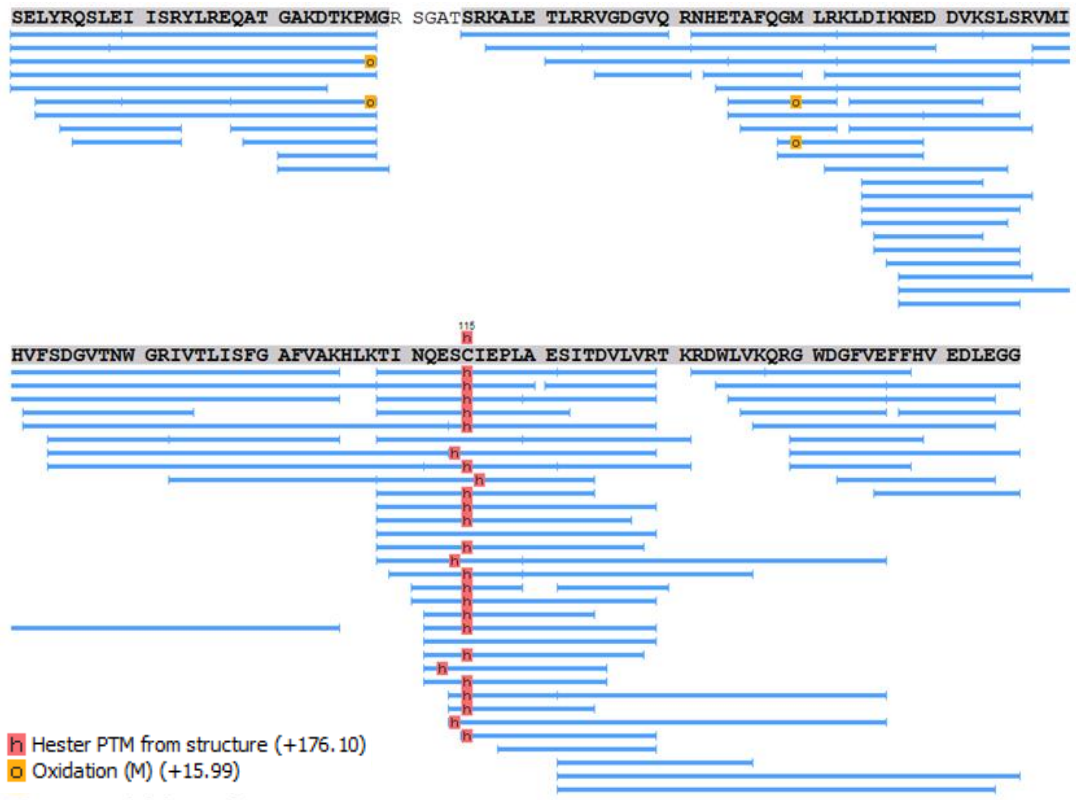


Figure 77: *MS/MS results from MCL-1 modified with RTA* ***120****. Sequence coverage 94%, using Glu-C and trypsin proteases.*

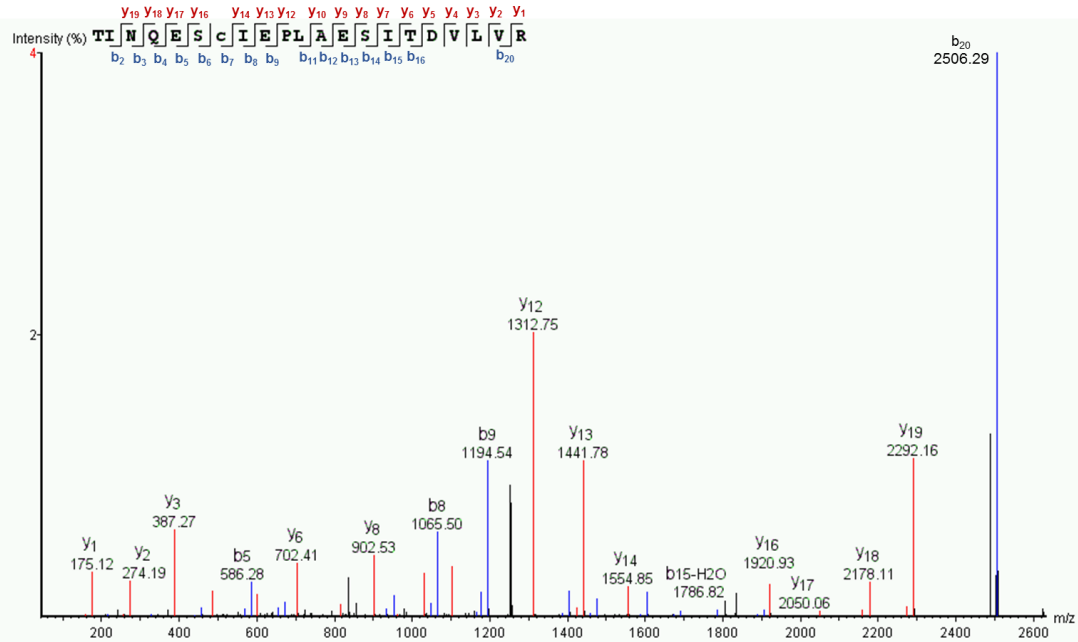


Figure 78: *Q-TOF MS/MS spectrum for a selected peptide modified with minimal RTA* ***120****. Observed y and b ions are shown in red and blue, respectively.*

Bibliography

1. Prescher, J. A. & Bertozzi, C. R. Chemistry in living systems. *Nat. Chem. Biol.* **1,** 13–21 (2005).

2. Giepmans, B. N. G., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. The Fluorescent Toolbox for Assessing Protein Location and Function. *Science.* **312,** 217–224 (2006).

3. Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat. Methods* **2,** 905–909 (2005).

4. Takaoka, Y., Ojida, A. & Hamachi, I. Protein organic chemistry and applications for labeling and engineering in live-cell systems. *Angew. Chemie Int. Ed.* **52,** 4088–4106 (2013).

5. Kim, Y. *et al.* Efficient Site-Specific Labeling of Proteins via Cysteines. *Bioconjug Chem.* **19,** 786–791 (2008).

6. Yuan, Y. *et al.* Labeling thiols on proteins, living cells, and tissues with enhanced emission induced by FRET. *Sci. Rep.* **3,** 1–7 (2013).

7. Koniev, O. & Wagner, A. Developments and recent advancements in the field of endogenous amino acid selective bond forming reactions for bioconjugation. *Chem. Soc. Rev.* **44,** 5495–5551 (2015).

8. Sletten, E. M. & Bertozzi, C. R. From mechanism to mouse: a tale of two bioorthogonal reactions. *Acc. Chem. Res.* **44,** 666–676 (2011).

9. Bertozzi, C. R. & Saxon, E. Cell Surface Engineering by a Modified Staudinger Reaction. *Science.* **287,** 2007–2010 (2000).

10. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective ‘Ligation’ of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed.* **41,** 2596–2599 (2002).

11. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67,** 3057–3064 (2002).

12. Agard, N. J., Prescher, J. & Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126,** 15046–15047 (2004).

13. Baskin, J. M. *et al.* Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci.* **104,** 16793–16797 (2007).

14. De Araújo, A. D. *et al.* Diels-alder ligation of peptides and proteins. *Chem. Eur. J.* **12,** 6095–6109 (2006).

15. Suchanek, M., Radzikowska, A. & Thiele, C. Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat. Methods* **2,** 261–267 (2005).

16. Lauehlin, S. T. & Bertozzi, C. R. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via staudinger ligation. *Nat. Protoc.* **2,** 2930–2944 (2007).

17. Jao, C. Y., Roth, M., Welti, R. & Salic, A. Metabolic labeling and direct imaging of choline phospholipids in vivo. *Proc. Natl. Acad. Sci.* **106,** 15332–15337 (2009).

18. Wofsy, L., Metzger, H. & Singer, S. J. Affinity Labeling - A General Method for Labeling the Active Sites of Antibody and Enzyme Molecules. *Biochemistry* **1,** 1031–1039 (1962).

19. Mills, J. S. *et al.* Identification of a Ligand Binding Site in the Human Neutrophil Formyl Peptide Receptor Using a Site-specific Fluorescent Photoaffinity Label and Mass Spectrometry. *J. Biol. Chem.* **273,** 10428–10435 (1998).

20. Robinette, D., Neamati, N., Tomer, K. B. & Borchers, C. H. Photoaffinity labeling combined with mass spectrometric approaches as a tool for structural proteomics. *Expert Rev Proteomics* **3,** 399–408 (2006).

21. Dormán, G. & Prestwich, G. D. Benzophenone photophores in biochemistry. *Biochemistry* **33,** 5661–5673 (1994).

22. Nielsen, P. E. & Buchardt, O. Aryl azides as photoaffinity labels: a photochemical study of some 4-substituted aryl azides. *Photochem. Photobiol.* **35,** 317–323 (1982).

23. Smith, E. & Collins, I. Photoaffinity labeling in target- and binding-site identification. *Future Med. Chem.* **7,** 159–183 (2015).

24. Wright, M. H. & Sieber, S. A. Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat. Prod. Rep.* **33,** 681–708 (2016).

25. Murale, D. P., Hong, S. C., Haque, M. M. & Lee, J. S. Photo-affinity labeling (PAL) in chemical proteomics: A handy tool to investigate protein-protein interactions (PPIs). *Proteome Sci.* **15,** 1–34 (2017).

26. Preston, G. W. & Wilson, A. J. Photo-induced covalent cross-linking for the analysis of biomolecular interactions. *Chem. Soc. Rev.* **42,** 3289–3301 (2013).

27. Brunner, J. New Photolabeling and Crosslinking Methods. *Annu. Rev. Biochem.* **62,** 483–514 (1993).

28. Das, M. & Fox, C. F. Chemical cross-linking in biology. *Annu. Rev. Biophys. Bioeng.* **8,** 165–193 (1979).

29. Bayley, H. & Knowles, J. R. Photoaffinity labeling. *Methods Enzymol.* **46,** 69–114 (1977).

30. Staros, J. Aryl Azide Photolabels in Biochemistry. *Trends Biochem. Sci.* **5,** 320–322 (1980).

31. Bayley, H. in *Laboratory Techniques in Biochemistry and Molecular Biology* 25–65 (Elsevier, 1983).

32. Shields, C. J. *et al.* Photochemistry of Aryl Azides: Detection and Characterization of a Dehydroazepine by Time-Resolved Infrared Spectroscopy and Flash Photolysis at Room Temperature. *J. Am. Chem. Soc.* **109,** 4723–4726 (1987).

33. Poe, R., Schnapp, K., Young, M. J. T., Grayzar, J. & Platz, M. S. Chemistry and Kinetics of Singlet (Pentafluorophenyl)nitrene. *J. Am. Chem. Soc.* **114,** 5054–5067 (1992).

34. Albini, A., Bettinetti, G. & Minoli, G. Reactivity of singlet and triplet arylnitrenes: Temperature-dependent photodecomposition of 1-(2-azidophenyl)-3,5-dimethylpyrazole. *J. Am. Chem. Soc.* **119,** 7308–7315 (1997).

35. Mohrig, J. R. & Keegstra, K. Alkyldiazonium Cations. I. Direct Observation of

the 2,2,2-Trifluoroethyldiazonium Ion. *J. Am. Chem. Soc.* **89,** 5492–5493 (1967).

36. Brunner, J., Senn, H. & Richards, F. M. 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *J. Biol. Chem.* **255,** 3313–3318 (1980).

37. Hashimoto, M. & Hatanaka, Y. Practical conditions for photoaffinity labeling with 3-trifluoromethyl-3- phenyldiazirine photophore. *Anal. Biochem.* **348,** 154–156 (2006).

38. Sumranjit, J. & Chung, S. J. Recent advances in target characterization and identification by photoaffinity probes. *Molecules* **18,** 10425–10451 (2013).

39. Galardy, E., Craig, C., Jamieson, J. D. & Printz, P. Photoaffinity Labeling of Peptide Hormone Binding Sites. *J. Biol. Chem.* **249,** 3510–3518 (1974).

40. Preston, G. W. & Wilson, A. J. Photo-induced covalent cross-linking for the analysis of biomolecular interactions. *Chem Soc Rev.* **42,** 3289–3301 (2013).

41. MacKinnon, A. L. & Taunton, J. Target Identification by Diazirine Photocrosslinking and Click Chemistry. *Curr. Protoc. Chem. Biol.* **1,** 55–73 (2009).

42. Young, T. S. *et al.* Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci.* **108,** 11052–11056 (2011).

43. Hayashi, T. & Hamachi, I. Traceless affinity labeling of endogenous proteins for functional analysis in living cells. *Acc. Chem. Res.* **45,** 1460–1469 (2012).

44. Hamachi, I., Nagase, T. & Shinkai, S. A general semisynthetic method for fluorescent saccharide-biosensors based on a lectin. *J. Am. Chem. Soc.* **122,** 12065–12066 (2000).

45. Takaoka, Y., Tsutsumi, H., Kasagi, N., Nakata, E. & Hamachi, I. One-pot and sequential organic chemistry on an enzyme surface to tether a fluorescent probe at the proximity of the active site with restoring enzyme activity. *J. Am. Chem. Soc.* **128,** 3273–3280 (2006).

46. Tsukiji, S., Miyagawa, M., Takaoka, Y., Tamura, T. & Hamachi, I. Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol.* **5,** 341–

343 (2009).

47. Tsukiji, S. *et al.* Quenched ligand-directed tosylate reagents for one-step construction of turn-on fluorescent biosensors. *J. Am. Chem. Soc.* **131,** 9046–54 (2009).

48. Tamura, T., Tsukiji, S. & Hamachi, I. Native FKBP12 engineering by ligand-directed tosyl chemistry: Labeling properties and application to photo-cross-linking of protein complexes in vitro and in living cells. *J. Am. Chem. Soc.* **134,** 2216–2226 (2012).

49. Fujishima, S., Yasui, R., Miki, T., Ojida, A. & Hamachi, I. Ligand-directed acyl imidazole chemistry for labeling of membrane-bound proteins on live cells. *J. Am. Chem. Soc.* **134,** 3961–3964 (2012).

50. Matsuo, K. *et al.* One-step construction of caged carbonic anhydrase I using a ligand-directed acyl imidazole-based protein labeling method. *Chem. Sci.* **4,** 2573 (2013).

51. Takaoka, Y., Nishikawa, Y., Hashimoto, Y., Sasaki, K. & Hamachi, I. Ligand-directed dibromophenyl benzoate chemistry for rapid and selective acylation of intracellular natural proteins. *Chem. Sci.* **6,** 3217–3224 (2015).

52. Matsuo, K., Nishikawa, Y., Masuda, M. & Hamachi, I. Live-Cell Protein Sulfonylation Based on Proximity-driven *N* -Sulfonyl Pyridone Chemistry. *Angew. Chemie Int. Ed.* **57,** 659–662 (2018).

53. Koshi, Y. *et al.* Target-specific chemical acylation of lectins by ligand-tethered DMAP catalysts. *J. Am. Chem. Soc.* **130,** 245–251 (2008).

54. Song, Z. *et al.* Extended Affinity-guided DMAP Chemistry with a Finely Tuned Acyl Donor for Intracellular FKBP12 Labeling. *Chem. Lett.* **44,** 333–335 (2015).

55. Wang, H. *et al.* Chemical cell-surface receptor engineering using affinity-guided, multivalent organocatalysts. *J. Am. Chem. Soc.* **133,** 12220–12228 (2011).

56. Tamura, T. *et al.* Affinity-Guided Oxime Chemistry for Selective Protein Acylation in Live Tissue Systems. *J. Am. Chem. Soc.* **139,** 14181–14191 (2017).

57. Amaike, K., Tamura, T. & Hamachi, I. Recognition-driven chemical labeling of endogenous proteins in multi-molecular crowding in live cells. *Chem. Commun.* **53,** 11972–11983 (2017).

58. Yamaura, K. *et al.* Live Cell Off-target Identification of Lapatinib Using Ligand-Directed Tosyl Chemistry. *Chem. Commun.* 2–5 (2014).

59. Gerson, F. *et al.* 4-Dialkylaminopyridines as Highly Active Acylation Catalysts. *Angew. Chem. Int. Ed. Engl.* **17,** 569–583 (1978).

60. Ladbury, J. E. *et al.* Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reappraisal. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 3199–3203 (1995).

61. Hayashi, T. *et al.* Semisynthetic Lectin−4-Dimethylaminopyridine Conjugates for Labeling and Pro fi ling Glycoproteins on Live Cell Surfaces. *J. Am. Chem. Soc.* **135,** 12252–12258 (2013).

62. Matsuo, K. *et al.* One-step construction of caged carbonic anhydrase I using a ligand-directed acyl imidazole-based protein labeling method. *Chem. Sci.* **4,** 2573 (2013).

63. Wakayama, S. *et al.* Chemical labelling for visualizing native AMPA receptors in live neurons. *Nat. Commun.* **8,** 1–14 (2017).

64. Kiyonaka, S. *et al.* Ligand-directed chemistry of AMPA receptors confers live-cell fluorescent biosensors. *ACS Chem. Biol.* Article ASAP (2018). doi:10.1021/acschembio.7b01042

65. Takaoka, Y., Nishikawa, Y., Hashimoto, Y., Sasaki, K. & Hamachi, I. Ligand-directed dibromophenyl benzoate chemistry for rapid and selective acylation of intracellular natural proteins. *Chem. Sci.* **6,** 3217–3224 (2015).

66. Popp, B. V. & Ball, Z. T. Structure-selective modification of aromatic side chains with dirhodium metallopeptide catalysts. *J. Am. Chem. Soc.* **132,** 6660–6662 (2010).

67. Chen, Z. *et al.* Catalytic protein modification with dirhodium metallopeptides: Specificity in designed and natural systems. *J. Am. Chem. Soc.* **134,** 10138–10145 (2012).

68. Vohidov, F., Coughlin, J. M. & Ball, Z. T. Rhodium(II) Metallopeptide Catalyst Design Enables Fine Control in Selective Functionalization of Natural SH3 Domains. *Angew. Chem. Int. Ed.* **54,** 4587–4591 (2015).

69. Chen, Z., Popp, B. V., Bovet, C. L. & Ball, Z. T. Site-specific protein modification with a dirhodium metallopeptide catalyst. *ACS Chem. Biol.* **6,** 920–925 (2011).

70. Kim, K., Fancy, D. A., Carney, D. & Kodadek, T. Photoinduced Protein Cross-Linking Mediated by Palladium Porphyrins. *J. Am. Chem. Soc.* **121,** 11896–11897 (1999).

71. Fancy, D. A. & Kodadek, T. Chemistry for the analysis of protein–protein interactions: Rapid and efficient cross-linking triggered by long wavelength light. *Proc. Natl. Acad. Sci.* **96,** 6020–6024 (1999).

72. Meunier, S., Strable, E. & Finn, M. G. Crosslinking of and Coupling to Viral Capsid Proteins by Tyrosine Oxidation. *Chem. Biol.* **11,** 319–326 (2004).

73. Fancy, D. A. *et al.* Scope , limitations and mechanistic aspects of the photo-induced cross-linking of proteins by water-soluble metal complexes. *Chem. Biol.* **7,** 697–708 (1999).

74. Sato, S. & Nakamura, H. Ligand-directed selective protein modification based on local single-electron-transfer catalysis. *Angew. Chemie - Int. Ed.* **52,** 8681–8684 (2013).

75. Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* **402,** 47–52 (1999).

76. Sato, S., Morita, K. & Nakamura, H. Regulation of Target Protein Knockdown and Labeling Using Ligand-Directed Ru(bpy)3 Photocatalyst. *Bioconjug. Chem.* **26,** 250–256 (2015).

77. Tsushima, M., Sato, S. & Nakamura, H. Selective purification and chemical labeling of a target protein on ruthenium photocatalyst-functionalized affinity beads. *Chem. Commun.* **53,** 4838–4841 (2017).

78. Preusser, M. *et al.* Current concepts and management of glioblastoma. *Ann. Neurol.* **70,** 9–21 (2011).

79. Carlsson, S. K., Brothers, S. P. & Wahlestedt, C. Emerging treatment strategies for glioblastoma multiforme. *EMBO Mol. Med.* **6,** 1359–1370 (2014).

80. Niewald, M. *et al.* Toxicity after radiochemotherapy for glioblastoma using temozolomide - a retrospective evaluation. *Radiat. Oncol.* **6,** 1–7 (2011).

81. Combs, S. E., Schmid, T. E., Vaupel, P. & Multhoff, G. Stress Response Leading to Resistance in Glioblastoma — The Need for Innovative Radiotherapy (iRT) Concepts. *Cancers (Basel).* **8,** 1–14 (2016).

82. Bonavia, R., Inda, M. D. M., Cavenee, W. K. & Furnari, F. B. Heterogeneity maintenance in glioblastoma: A social network. *Cancer Res.* **71,** 4055–4060 (2011).

83. Dirks, P. B. Brain tumor stem cells: The cancer stem cell hypothesis writ large. *Mol. Oncol.* **4,** 420–430 (2010).

84. Wurdak, H. *et al.* A small molecule accelerates neuronal differentiation in the adult rat. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 16542–22360 (2010).

85. Gergely, F. *et al.* The TACC domain identifies a family of centrosomal proteins that can interact with microtubules. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 14352–14357 (2000).

86. Guo, Y. *et al.* Regulating the ARNT/TACC3 axis: Multiple approaches to manipulating protein/protein interactions with small molecules. *ACS Chem. Biol.* **8,** 626–635 (2013).

87. Duncan, C. G. *et al.* Integrated genomic analyses identify ERRFI1 and TACC3 as glioblastoma-targeted genes. *Oncotarget* **1,** 265–277 (2010).

88. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science.* **344,** 1396–1401 (2014).

89. Allen, M., Bjerke, M., Edlund, H., Nelander, S. & Westermark, B. Origin of the U87MG glioma cell line: Good news and bad news. *Sci. Transl. Med.* **8,** 1946–6242 (2016).

90. Awoussong, P. K. *et al.* Heterocycles 26: synthesis, characterisation, and anticancer activity of some thiazolic chalcones. *Med. Chem. Rev.* **24,** 131–141 (2015).

91. Hantzsch, A. & Weber, J. H. Ueber Verbindungen des Thiazols (Pyridins der Thiophenreihe). *Ber. Dtsch. Chem. Ges.* **20,** 3118–3132 (1887).

92. Delépine, M. *Bull.Soc.Chim.Fr.* **13,** 352–361 (1895).

93. Klionsky, D. J. & Emr, S. D. Autophagy as a Regulated Pathway of Cellular Degradation Daniel. *Science.* **290,** 1717–1721 (2009).

94. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102,** 15545–15550 (2005).

95. Salisbury, C. M. & Cravatt, B. F. Optimization of activity-based probes for proteomic profiling of histone deacetylase complexes. *J. Am. Chem. Soc.* **130,** 2184–2194 (2008).

96. Bukau, B. & Horwich, A. L. The Hsp70 and Hsp60 Chaperone Machines. *Cell* **92,** 351–366 (1998).

97. Lianos, G. D. *et al.* The role of heat shock proteins in cancer. *Cancer Lett.* **360,** 114–118 (2015).

98. Tang, H. *et al.* Down-regulation of HSP60 Suppresses the Proliferation of Glioblastoma Cells via the ROS/AMPK/mTOR Pathway. *Sci. Rep.* **6,** 28388 (2016).

99. Wiechmann, K. *et al.* Mitochondrial Chaperonin HSP60 Is the Apoptosis-Related Target for Myrtucommulone. *Cell Chem. Biol.* **24,** 614–623 (2017).

100. Nielsen, K. L., Lennan, N. M. C., Masters, M. & Cowan, N. J. A Single-Ring Mitochondrial Chaperonin (Hsp60-Hsp10) Can Substitute for GroEL-GroES In Vivo. *J. Bacteriol.* **181,** 5871–5875 (1999).

101. Itoh, H., Komatsuda, A., Wakui, H., Miura, A. B. & Tashima, Y. Mammalian HSP60 Is a Major Target for an Immunosuppressant Mizoribine. *J. Biol. Chem.* **274,** 35147–35151 (1999).

102. Chen, H. *et al.* Identification of Human T Cell Receptor gammadelta-recognized Epitopes/Proteins via CDR3delta Peptide-based Immunobiochemical Strategy. *J. Biol. Chem.* **283,** 12528–12537 (2008).

103. Kvach, M. V. *et al.* Practical Synthesis of Isomerically Pure 5- and 6-

Carboxytetramethylrhodamines, Useful Dyes for DNA Probes. *Bioconjug. Chem.* **20,** 1673–1682 (2009).

104. Wallrabe, H. & Periasamy, A. Imaging protein molecules using FRET and FLIM microscopy. *Curr. Opin. Biotechnol.* **16,** 19–27 (2005).

105. Yokoyama, M., Yoshida, S. & Imamoto, T. Organic Reactions Using Trimethylsilyl Polyphosphate (PPSE): A Convenient Synthesis of Nitriles from Carboxamides. *Synthesis (Stuttg).* **7,** 591–592 (1982).

106. Brunet, A., Aslam, T. & Bradley, M. Bioorganic & Medicinal Chemistry Letters Separating the isomers — Efficient synthesis of the N -hydroxysuccinimide esters of 5 and 6-carboxyfluorescein diacetate and 5 and 6-carboxyrhodamine B. *Bioorg. Med. Chem. Lett.* **24,** 3186–3188 (2014).

107. Liu, Y., Zhang, Z., Zhang, Q., Baker, G. L. & Worden, R. M. Biomembrane disruption by silica-core nanoparticles: Effect of surface functional group measured using a tethered bilayer lipid membrane. *Biochim. Biophys. Acta - Biomembr.* **1838,** 429–437 (2014).

108. Neises, B. & Steglich, W. Simple Method for the Esterification of Carboxylic Acids. *Angew. Chemie Int. Ed.* **17,** 522–524 (1978).

109. Hahn, S. & Kim, D. Transient protein-protein interaction of the SH3-peptide complex via closely located multiple binding sites. *PLoS One* **7,** 32804 (2012).

110. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. *Structure* **18,** 1233–1243 (2010).

111. Phizicky, E. M. & Fields, S. Protein-protein interactions: methods for detection and analysis. *Microbiol. Rev.* **59,** 94–123 (1995).

112. Westermarck, J., Ivaska, J. & Corthals, G. L. Identification of Protein Interactions Involved in Cellular Signaling. *Mol. Cell. Proteomics* **12,** 1752–1763 (2013).

113. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340,** 245–246 (1989).

114. Bartel, P. L. & Fields, S. Analyzing protein-protein interactions using two-

hybrid system. *Methods Enzymol.* **254,** 241–263 (1995).

115. Dziembowski, A. & Séraphin, B. Recent developments in the analysis of protein complexes. *FEBS Lett.* **556,** 1–6 (2004).

116. Wetie, A. G. N. *et al.* Investigation of stable and transient protein-protein interactions: past, present and future. *Proteomics* **13,** 1–33 (2013).

117. Stagljar, I., Korostensky, C., Johnsson, N. & te Heesen, S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5187–92 (1998).

118. Sinz, A., Arlt, C., Chorev, D. & Sharon, M. Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. *Protein Sci.* **24,** 1193–1209 (2015).

119. Kao, A. *et al.* Mapping the Structural Topology of the Yeast 19S Proteasomal Regulatory Particle Using Chemical Cross-linking and Probabilistic Modeling. *Mol. Cell. Proteomics* **11,** 1566–1577 (2012).

120. Amini, F., Denison, C., Lin, H., Kuo, L. & Kodadek, T. Using Oxidative Crosslinking and Proximity Labeling to Quantitatively Characterize Protein-Protein and Protein-Peptide Complexes. *Chem. Biol.* **10,** 1115–1127 (2003).

121. Denison, C. & Kodadek, T. Toward a General Chemical Method for Rapidly Mapping Multi-Protein Complexes. *J. Proteome Res.* **3,** 417–425 (2004).

122. Hauser, J. R. *et al.* Economical and scalable synthesis of 6-amino-2-cyanobenzothiazole. *Beilstein J. Org. Chem.* **12,** 2019–2025 (2016).

123. Stewart, M. L., Fire, E., Keating, A. E. & Walensky, L. D. The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nat. Chem. Biol.* **6,** 595–601 (2010).

124. Oltersdorf, T. *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **435,** 677–681 (2005).

125. Muchmore, S. W. *et al.* X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. *Nature* **381,** 335–341 (1996).

126. Sattler, M. *et al.* Structure of Bcl-xL-Bak Peptide Complex: Recognition Between Regulators of Apoptosis. *Science.* **275,** 983–987 (1997).

127. Chen, L. *et al.* Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Mol. Cell* **17,** 393–403 (2005).

128. Czabotar, P. E. *et al.* Structural insights into the degradation of Mcl-1 induced by BH3 domains. *Proc. Natl. Acad. Sci.* **104,** 6217–6222 (2007).

129. Barnard, A. *et al.* Orthogonal functionalisation of α-helix mimetics. *Org. Biomol. Chem.* **12,** 6794–6799 (2014).

130. Sullivan, B. P., Salmon, D. J. & Meyer, T. J. Mixed Phosphine 2,2'-Bipyridine Complexes of Ruthenium. *Inorg. Chem.* **17,** 3334–3341 (1978).

131. Miles, J. A. *et al.* Hydrocarbon constrained peptides – understanding preorganisation and binding affinity. *Chem. Sci.* **7,** 3694–3702 (2016).

132. Castranova, V., Asgharian, B., Sayre, P., Virginia, W. & Carolina, N. Fluorescence anisotropy (polarization): from drug screening to precision medicine. *Expert Opin Drug Discov.* **10,** 1145–1161 (2015).

133. Van Houten, J. & Watts, R. J. Temperature Dependence of the Photophysical and Photochemical Properties of the Tris(2,2′-bipyridyl)ruthenium(II) Ion in Aqueous Solution. *J. Am. Chem. Soc.* **98,** 4853–4858 (1976).

134. Kellogg, R. E. & Bennett, R. G. Radiationless intermolecular energy transfer. III. Determination of phosphorescence efficiencies. *J. Chem. Phys.* **41,** 3042–3045 (1964).

135. Ghosh, I., Marzo, L., Das, A., Shaikh, R. & König, B. Visible Light Mediated Photoredox Catalytic Arylation Reactions. *Acc. Chem. Res.* **49,** 1566–1577 (2016).

136. Yoshioka, E. *et al.* Aqueous-Medium Carbon-Carbon Bond-Forming Radical Reactions Catalyzed by Excited Rhodamine B as a Metal-Free Organic Dye under Visible Light Irradiation. *J. Org. Chem.* **81,** 7217–7229 (2016).

137. Lee, J., Udugamasooriya, D. G., Lim, H. & Kodadek, T. Potent and selective photo-inactivation of proteins with peptoid-ruthenium conjugates. *Nat. Chem. Biol.* **6,** 258–260 (2010).

138. Plante, J. P. *et al.* Oligobenzamide proteomimetic inhibitors of the p53–hDM2

protein–protein interaction. *Chem. Commun.* **34,** 5091–5093 (2009).

139. Still, W. C., Kahn, M. & Mitra, A. Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution. *J. Org. Chem.* **43,** 2923–2925 (1978).

140. Bur, D. *et al.* US8674111 B2: Oxazole and thiazole derivatives as ALX receptor agonists. (2010).

141. Zhang, R. *et al.* Development of a Ruthenium(II) Complex-Based Luminescent Probe for Hypochlorous Acid in Living Cells. *Inorg. Chem.* **52,** 10325−10331 (2013).

142. Burslem, G. M. *et al.* Small-Molecule Proteomimetic Inhibitors of the HIF-1 – p300 Protein–Protein Interaction. *ChemBioChem* **15,** 1083–1087 (2014).

143. Burslem, G. M. *et al.* Biomolecular Chemistry Synthesis of highly functionalized oligobenzamide proteomimetic foldamers by late stage introduction of sensitive groups. *Org. Biomol. Chem.* **14,** 3782–3786 (2016).

144. Semenza, G. L. Hypoxia-Inducible Factors in Physiology and Medicine. *Cell* **148,** 399–408 (2012).

145. Kennedy, D. C. *et al.* Cellular consequences of copper complexes used to catalyze bioorthogonal click reactions. *J. Am. Chem. Soc.* **133,** 17993–18001 (2011).

146. Godinat, A. *et al.* A biocompatible in vivo ligation reaction and its application for noninvasive bioluminescent imaging of protease activity in living mice. *ACS Chem. Biol.* **8,** 987–999 (2013).

147. Wang, P. *et al.* Site-specific immobilization of biomolecules by a biocompatible reaction between terminal cysteine and 2-cyanobenzothiazole. *Chem. Commun.* **49,** 8644–8646 (2013).

148. Katz, L. Antituberculous Compounds. II. 2-Benzalhydrazinobenzothiazoles. *J. Am. Chem. Soc.* **73,** 4007–4010 (1951).

149. Bellamy, F. D. & Ou, K. Selective reduction of aromatic nitro compounds with stannous chloride in non acidic and non aqueous medium. *Tetrahedron Lett.* **25,** 839–842 (1984).

150. McCutcheon, D. C., Porterfield, W. B. & Prescher, J. A. Rapid and scalable assembly of firefly luciferase substrates. *Org. Biomol. Chem.* **3,** 5012–5016 (2014).

151. Oudet, P. & Mioskowskia, C. Synthesis of New Phospholipids Linked to Steroid-Hormone Derivatives. *Helv. Chim. Acta* **74,** 1697–1706 (1991).