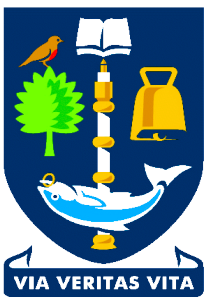


University of Glasgow

Synthesis of Activity-Based Probes for the Metalloprotease CSN5

Siddique Amin, BSc (Hons)

Submitted in fulfilment of the requirements for the degree of Master of Science by Research

School of Chemistry

College of Science & Engineering

University of Glasgow

December 2020

Abstract

The Cullin-RING E3 ligase (CRL) family consists of over 200 enzymes that are responsible for the ubiquitylation and regulation of a huge variety of proteins involved in the immune response, cell cycle control and cell proliferation. CRLs are activated by covalent attachment to the *C*-terminus of the ubiquitin-like modifier NEDD8 by forming an isopeptide bond in a process called NEDDylation. Inhibition of CRLs is mediated by the COP9 signalosome complex (CSN) which cleaves the isopeptide bond binding to NEDD8 and removes it. CRL homeostasis is disrupted by CSN overactivity and is associated with the pathogenesis of ~50% of all cancers in humans. The means by which the activity of the catalytic subunit of CSN (CSN5) is regulated is not fully understood and thus new tools are required to study its activity. CSN5 contains a JAMM domain with a Zn^{2+} atom facilitating its catalytic activity. The synthesis of several activity-based probes for CSN5 was attempted, each containing a zinc-binding group (ZBG) for attachment onto the catalytic site, a NEDD8 peptide to enable selectivity to CSN5 and a biotin tag for subsequent biochemical assays. Three zinc-binding groups were considered including hydroxamic acid, imidazole and a precursor of a potent small molecule CSN5 inhibitor, CSN5i-3. The multi-step synthesis of the CSN5i-3 precursor was attempted but not completed due to time constraints. In an endeavour to mimic the entire substrate of CSN5, the synthesis of NEDD8 was attempted but not successful. In a separate approach, a peptide of the NEDD8 *C*-terminal tail was synthesised and successfully biotinylated, however conjugation to the zinc-binding groups hydroxamic acid and imidazole were unsuccessful due to the low solubility of the peptide.

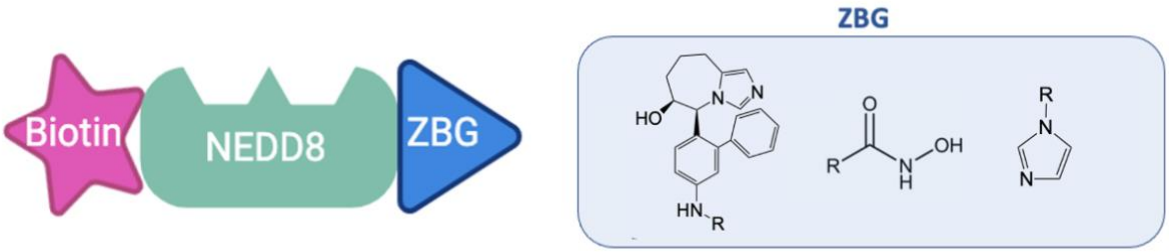


Table of Contents

Acknowledgements

I am hugely grateful to Dr Andrew Jamieson for hosting me in his group for my project. I thank him not only for his academic support but also for his mentorship and showing me nothing but compassion and patience throughout my endeavour to become a chemist.

Likewise it has been a pleasure working within the Jamieson group and I thank everyone for their support. I am grateful to Amit Mahindra, Kopano Mapesa and Ross Gillespie for helping me learn organic synthesis and to Caroline Morris, Andrea González Muñoz, Danielle Morgan and Lewis Archibald for showing me the ropes in peptide chemistry. I am especially thankful to Amit for dedicating so much time to helping me with my project and for encouraging me when things didn't go as planned.

Thank you to my colleagues in the Chemical Biology lab for their kindness and friendship.

I could not have completed this project without the incredible support of my colleagues, friends and family. Thank you to the people in my life who celebrate my successes with me and pick me up when things get tough (e.g. a pandemic).

Declaration

I certify that the thesis presented here for examination for a MSc(Res) degree of the University of Glasgow is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice.

I declare that the thesis does not include work forming part of a thesis presented successfully for another degree.

I declare that this thesis has been produced in accordance with the University of Glasgow's Code of Good Practice in Research.

I acknowledge that if any issues are raised regarding good research practice based on review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Siddique Amin

01/12/2020

Abbreviations

AIBN – Azobisisobutyronitrile
ATP – Adenosine triphosphate
BOC – *tert*-butoxycarbonyl
cat. – Catalytic
CCl_4 – Carbon tetrachloride
DNA – Deoxyribonucleic Acid
cDNA – Complementary DNA
DCC – *N*,*N*'-dicyclohexylcarbodiimide
DCM – Dichloromethane
DIEA – *N*,*N*-Diisopropylethylamine
DMAP – 4-dimethylaminopyridine
Dmb – Dimethoxybenzyl
DMF – Dimethylformamide
DMSO – Dimethylsulfoxide
DPPA – Diphenylphosphoryl azide
dppf – 1,1'-bis(diphenylphosphino)ferrocene
DUB – Deubiquitinase
EtOAc – Ethyl acetate
Et_3N – Triethylamine
Eq – Equivalents
ESI – Electrospray ionisation
Fmoc – Fluorenylmethoxycarbonyl
HCTU – *O*-(1*H*-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HECT – Homologous to the E6-AP Carboxyl Terminus
Hmb – Hydroxymethylbutyrate
HPLC – High performance liquid chromatography
HRMS – High resolution mass spectrometry
kDa – Kilodalton
LC-MS – Liquid-chromatography mass spectrometry
MIU – Motif Interacting Ubiquitin
mw – Microwave
NMP – *N*-methyl-2-pyrrolidone
NMR – Nuclear magnetic resonance
RING – Really Interesting New Gene
Rt – Retention time
rt – Room temperature
SPPS – Solid-phase peptide synthesis
TAMRA – Tetramethylrhodamine
TFA – Trifluoroacetic acid
TGT – Tentagel 4-carboxytrityl
THF – Tetrahydrofuran
Ub – Ubiquitin
UBD – Ubiquitin Binding Domain
UBL – Ubiquitin-Like Protein
UIM – Ubiquitin Interacting Motif

Chapter 1 – Introduction

1.1 - Ubiquitylation

Ubiquitin (Ub) is an essential small protein of ~8.5 kDa and 76 residues which is highly conserved and ubiquitously expressed in eukaryotic cells. Ubiquitylation refers to a reversible post-translational modification whereby ubiquitin is covalently transferred onto a substrate protein by forming an isopeptide bond between the carboxyl group of the ubiquitin C-terminal glycine residue (Gly76) and the ε-amino group of a lysine residue on the substrate.[1] The substrate protein can be subject to the attachment of a single ubiquitin (monoubiquitylation) or multiple ubiquitin molecules on different residues (multi-monoubiquitylation).[2] Ubiquitin itself also contains 7 lysine residues that can act as sites of ubiquitylation to form extended polyubiquitin (polyUb) chains. There is diversity in ubiquitin chain formation such that polyUb chains can consist of linkages of the same type (homotypic polyubiquitylation) or linkages built up of different lysine residues (heterotypic polyubiquitylation).[3] These ubiquitin molecules and chains act as signals that assert an array of cellular actions upon a substrate protein, which is dictated by the type of ubiquitin linkages. For example, polyubiquitin chains linked by Lys48 residues target a substrate protein for proteasomal degradation, a major process in protein homeostasis, whereas Lys63-linked polyubiquitin chains enable endocytosis of the substrate protein.[4] Furthermore, the *N*-terminal methionine (Met1) amino group of ubiquitin also acts a site of ubiquitylation to form Met1-linked linear polyubiquitin chains which play an important role in the NF-κB signalling pathway.[5]

The process of ubiquitylation involves a sequential cascade of three enzymes.[3] The E1 ubiquitin-activating enzyme utilises a molecule of ATP to activate the C-terminal glycine of ubiquitin, which subsequently forms a thioester linkage with a catalytic cysteine residue on the E1 enzyme. The ubiquitin is transferred from the E1 enzyme onto the active site cysteine on an E2 ubiquitin-conjugating enzyme. Finally, the E2-Ub binds onto an E3 ubiquitin ligase which catalyses the transfer of ubiquitin onto the substrate protein to form an isopeptide bond (Figure 1).[6]

The reverse process of ubiquitylation (ubiquitin deconjugation) is facilitated by ubiquitin-specific proteases referred to as deubiquitylating enzymes or deubiquitinases (DUBs). These act to break the isopeptide bond between ubiquitin and its target protein, releasing free ubiquitin in the process.[7]

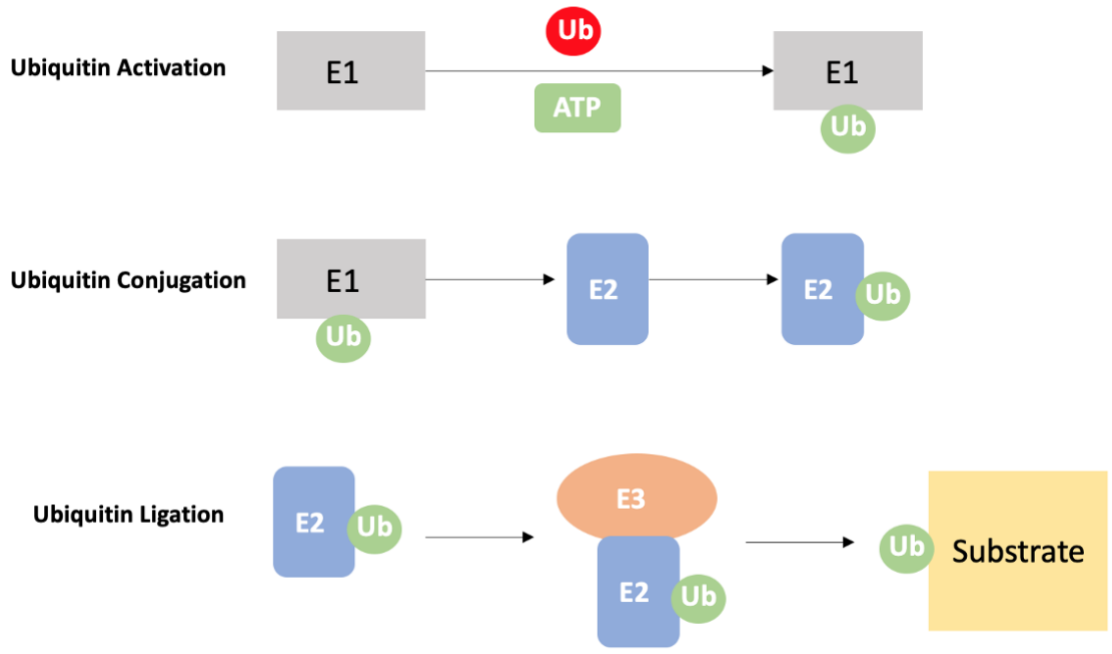


Figure 1: The ubiquitin conjugation system consisting of the E1, E2, E3 enzymatic cascade.

1.1.1 – Structural Features of Ubiquitin

Ubiquitin exhibits a compact globular structure with a 5-stranded antiparallel β-sheet which "grasps" a 3.5-turn α-helix and a 3_{10} helix (Figure 2A).[8] This structure is collectively referred to as the β-grasp superfold.[9] Ubiquitin also possesses a flexible C-terminal tail that protrudes away from the globular fold and ends with a di-glycine motif, with the distal glycine involved in the covalent attachment to a substrate lysine. Many proteins interact with ubiquitin non-covalently to fulfil a cellular role and this interaction is facilitated by ubiquitin-binding domains (UBDs) on those proteins.[10] Unlike other post-translational modifications such as phosphorylation, recognition of a ubiquitylated protein is not dictated by the substrate but is instead largely determined by regions on the surface of ubiquitin and the conformation of

polyUb chains.[11,12] The ubiquitin surface is mostly polar with the exception of a hydrophobic patch located within the β-sheet consisting of residues Ile44, Leu8 and Val70. This hydrophobic patch is also commonly referred to as the Ile44 patch as this residue is described as being the centre of the area.[12] The Ile44 patch is the major site of protein-protein interactions enabled by UBDs such as Ubiquitin Interacting Motif (UIM) and Motif Interacting Ubiquitin (MIU) which specifically recognise and bind to this region.[10,12] For example, the MIU domain was first described in the interaction between ubiquitin and Rpn10, a subunit of the 26S proteasome.[13]

Whilst a major hotspot for protein interactions, the Ile44 patch is not the exclusive site of ubiquitin recognition and such interactions can also occur by other surface regions such as the Ile36 patch. This region is centred upon the eponymous Ile36 and consists of residues Leu8, Leu71 and Leu73.[12,14] Likewise, other surface patches include the Phe4 patch (Phe4/Gln2/Thr14) and the Asp58 patch (Asp58/Arg54/Glu51) (Figure 2).[15]

Certain UBDs may recognise and bind to multiple regions on ubiquitin. For example, the A20 zinc finger (A20_ZnF) domain recognises both the hydrophobic Ile44 patch and the polar Asp58 patch. Proteins may contain multiple UBDs such as Rabex-5 which has both a UIM and A20_ZnF domain, thus two molecules of Rabex-5 can interact with a single molecule of ubiquitin by binding on two distinct regions.[16,17] Furthermore, having multiple UBDs can enable a protein to bind polyUb chains.

Depending on the linkage type, a polyUb chain may adopt a distinct conformation. For example, Lys48-linked polyUb chains form a compact structure whilst Lys63-linked chains adopt a wider, more open conformation.[4,18] Likewise, the conformation of the polyUb chain dictates which surface regions of the ubiquitin molecules are exposed for binding. Proteins may contain multiple UBDs that are spaced in a way that only facilitate binding to a certain polyUb chain conformation. For example, Rap80 is a protein involved in DNA repair that contains two MIU domains that both recognise Ile44 patches. However due to spacing between the domains they can only bind Lys63-linked polyUb chains, thus allowing for selectivity in their binding.[10,19]

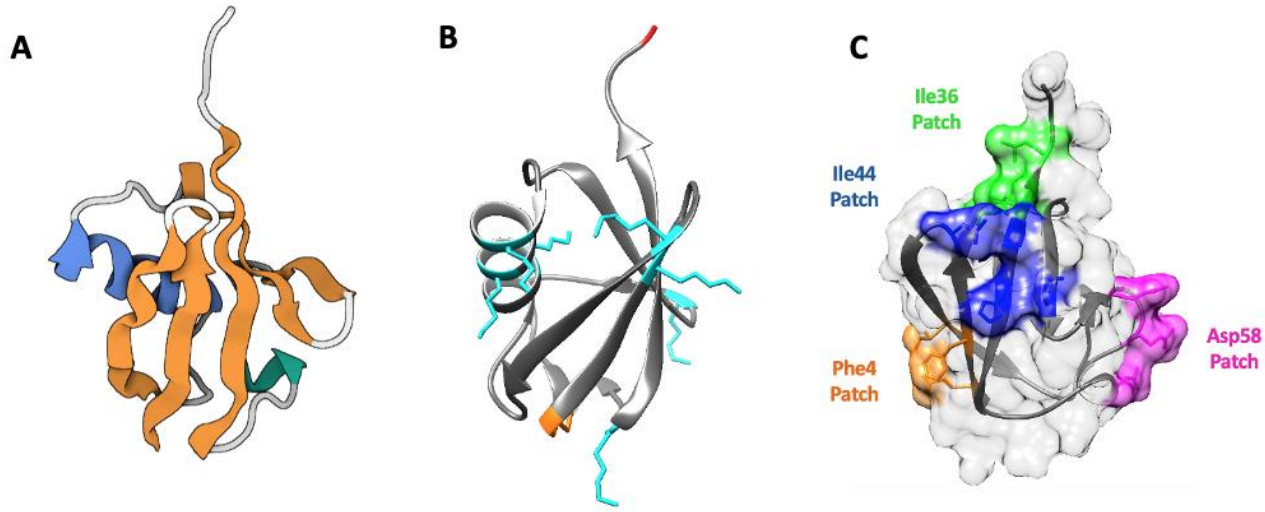


Figure 2: (A) Cartoon representation of the ubiquitin structure with secondary structures highlighted: 3.5 turn alpha helix in blue, β-sheet in orange and 3_{10} helix in teal. (B) Structure of ubiquitin with its 7 lysine residues coloured in cyan, Met1 in orange and Gly76 in red. (C) Surface patches on ubiquitin for non-covalent protein-protein interactions. [PDB: 1UBQ] (Figure 2A generated in Biorender and Figures 2B-C in UCSF Chimera)

1.1.2 – Ubiquitin-Like Modifiers

The characteristic β-grasp superfold structure is not exclusively found in ubiquitin. This structure, along with ubiquitin's conserved amino acid sequence, has enabled the identification of proteins which share structural and sequential similarities to ubiquitin, referred to as ubiquitin-like modifiers or ubiquitin-like proteins (UBLs).[9,12] Functionally, UBLs are distinguished into two groups. Type I UBLs are conjugated to proteins by an ATP-dependent enzymatic cascade consisting of E1, E2, and E3 enzymes. Therefore the mechanism of Type I UBL conjugation closely resembles that of ubiquitylation.[20] Likewise, Type I UBLs possess a C-terminal tail ending in a glycine motif facilitating the covalent attachment onto a substrate protein. Members of this group include, but are not limited to, Interferon-Stimulating Gene 15 (ISG-15), Small Ubiquitin-Like Modifier (SUMO) and Neural Precursor Cell Expressed Developmentally Down-Regulated Protein 8 (NEDD8). In contrast, Type II UBLs do not engage in substrate conjugation and are instead normally found as part of multi-domain proteins and can facilitate non-covalent protein-protein interactions.[20,21] For example, the E3 ligase UHRF1 contains a Type II UBL domain which, through its hydrophobic

patch, enables binding to an E2 enzyme UBE2D1 to form an E3-E2 complex required to ubiquitylate histone H3 in newly-replicated chromatin. This plays a key role in epigenetic regulation as the methyltransferase DNMT1 is recruited to the ubiquitylated histone to perform DNA methylation, thus ensuring chromatin homeostasis.[22]

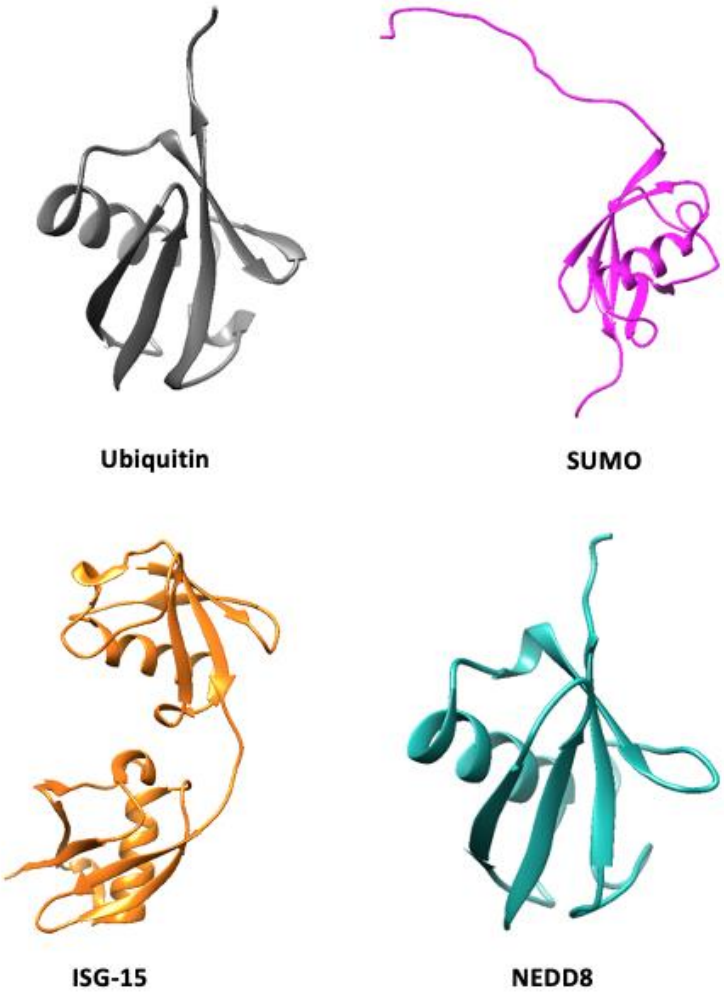


Figure 3: Structures of ubiquitin and examples of ubiquitin-like modifiers. [PDB: 1UBQ, 2N1V, 1Z2M, 1NDD] (Figure generated in UCSF Chimera).

ISG-15

ISG-15 was the first protein to be identified as a UBL and was later sub-classified as a Type I UBL. Although first being discovered within interferon-activated cells in 1979, it wasn't until 1987 that it was found as a UBL when it was shown to cross-react with ubiquitin antibodies.[23–25] With a size of 17 kDa and 165 residues, ISG-15 consists of two ubiquitin-like β-grasp folds in tandem and structurally resembles di-Ub.[26] As its name suggests, the expression of ISG-15

is dependent on interferon (IFN)-β and is therefore implicated in the innate immune response. Like ubiquitin, its conjugation to lysine residues on proteins (ISGylation) is catalysed be the E1-E2-E3 enzyme cascade however it occurs predominantly within the cell ribosome where the principal ISG-15 E3 ligase, Herc5, is localised.[27] Therefore ISGylation usually occurs on newly-translated proteins, namely those involved in type 1 interferon (IFN-1) signalling which are responsible for eliciting an anti-viral response by activating immune cells such as macrophages and natural killer cells.[23,26,27]

Furthermore, unlike ubiquitin, ISG-15 is conjugated exclusively as a monomer on its target protein and does not form polyISG-15 chains.[23] ISGylation plays several key roles in the regulation of innate immunity. For example, the pattern recognition receptor RIG-I is a promoter of IFN-1 signalling and its ISGylation results in degradation by autophagy, consequently inhibiting IFN-1 expression which is key in preventing auto-immunity.[23,28] ISG-15 also acts to stimulate the antiviral response by prolonging the activation of proteins such as Interferon-Regulatory Factor 3 (IRF3) whose ISGylation prevents its ubiquitylation and subsequent proteasomal degradation.[29]

Alongside stimulation of the antiviral response by conjugation onto host proteins, ISGylation can also occur on viral proteins, *I.e*. proteins produced by an invading virus itself[23]. This can interfere with the ability of viral proteins interact with each other or with host proteins. For example, the influenza A protein NS1 (NS1A) interacts with and supresses the host enzyme Protein Kinase R (PKR) which plays a role in antiviral defence. ISGylation of NS1A on Lys-41 blocks this interaction with PKR and thus prevents the suppression of PKR-mediated host defence.[26,30,31]

Interestingly, to combat this response, viruses have evolved to encode for DUBs which act to deconjugate ISG15 from their proteins. For example, Severe Acute Respiratory Syndrome-Related Coronavirus (SARS-CoV) was amongst the first viruses discovered to contain DUBs which deconjugate ISGylated proteins.[23,32] SARS-CoV papain-like protease (PLP) shows selectivity towards ISG15 and it has been demonstrated that PLP deconjugates ISG15 from IRF3 to supress IFN1 signalling.[33]

SUMO

Although ISG-15 was the first UBL to be discovered, it is still not relatively well-studied. In contrast, SUMO is the most extensively researched and characterised UBL.[12] The SUMO family of proteins consists of 4 isoforms in mammalian cells, three of which are physiologically active (SUMO-1, -2 and -3) and a fourth (SUMO-4) which is the least studied isoform and is not known to be functional under physiological conditions.[34] At 11 kDa and 100 residues, SUMO contains the characteristic structural β-grasp superfold which resembles that of ubiquitin and has a *C*-terminal di-glycine motif to facilitate substrate conjugation. However SUMO has an extended flexible *N*-terminal region, making it larger than and structurally and functionally distinct from ubiquitin.[35,36]

SUMO exhibits a highly diverse role in cellular homeostatic processes, including cytosolic transport, transcriptional regulation and DNA repair. However, unlike ubiquitin, it is not directly involved in the targeting of proteins for degradation.[34,37] SUMOylation refers to the reversible *C*-terminal conjugation of SUMO onto a lysine residue of a substrate protein by forming an isopeptide bond. Just like other Type I UBLs, SUMOlyation utilises the ATP-dependent E1-E2-E3 conjugation pathway.[34] The removal of SUMO from its substrate, deSUMOylation, is performed by SUMO-specific proteases (SENPs) which target specific SUMO isoforms and cleaves the isopeptide bond with the substrate.[38] A well-researched example of SUMOylation is that of the DNA-repair response enzyme Thymine DNA Glycosylase (TDG) which binds to and excises mismatched G-U and G-T base pairs within damaged DNA. Following excision, the enzyme binds with high affinity to the repaired DNA and cannot continue its repair cycle. SUMOlyation of TDG assets a conformational change which releases it from the DNA and, following deSUMOylation, the enzyme can continue to bind to mismatched bases and repair DNA mutations.[34,35,37]

NEDD8

At 8.5 kDa and 76 residues, NEDD8 is the closest relative to ubiquitin and, amongst other UBLs, shares the greatest structural similarity to ubiquitin. With its C-terminal di-glycine motif, NEDD8 can be conjugated to substrate lysine residues in a process referred to as

NEDDylation.[39,40] The gene *NEDD8* encodes for an 81-residue precursor protein that is deemed inactive. Processing of NEDD8 involves cleavage of 5 residues from the C-terminus to generate the di-glycine motif. This proteolytic step is performed by the cysteine protease deneddylase 1 (DEN1) which forms the mature 76-residue NEDD8 protein.[41,42] The conjugation of NEDD8 onto a substrate lysine is facilitated by the E1-E2-E3 enzymatic cascade analogous to ubiquitin and other Type I UBLs. NEDD8-Activating Enzyme (NAE), a heterodimer consisting of amyloid-β precursor protein-binding protein 1 (APPBP1 or NAE1) and ubiquitin-activating enzyme 3 (UBA3), is the only known NEDD8-specific E1.[43,44] Following activation by NAE, NEDD8 is transferred onto one of two NEDD8-conjugating E2 enzymes, UBE2M (also known as UBC12) or UBE2F.[43] Subsequently, NEDD8 is transferred onto an E3 ligase which is responsible for conjugation of NEDD8 onto a substrate protein. Whilst there is a variety of E3s which can target both ubiquitin and NEDD8, there is only one E3 specific for NEDD8, a complex consisting of the ligases Defective in Cullin Neddylation Protein 1 (Dcn1) and RING box protein 1 (Rbx1).[45,46]

NEDD8 shares a high percentage of amino acid sequence homology to ubiquitin, with ~58% sequence identity and ~77% sequence similarity.[43] A sequence comparison between NEDD8 and ubiquitin reveals that many residues essential for ubiquitin function are conserved within NEDD8 (Figure 4). For example, lysine residues are largely conserved between the two proteins and allow for NEDD8 to form polymeric chains.[47] Although the biological role of these poly-NEDD8 chains has yet to be ascertained. Similarly, the residues Ile44, Leu8, His68 and Val70 are conserved in NEDD8 and these make up the canonical Ile44 hydrophobic patch seen in ubiquitin.[43] Being the major site of non-covalent interactions, the Ile44 patch also serves as a key-binding site for protein-protein interactions in NEDD8. Likewise, the Ile36 patch consisting of residues Ile36, Leu8, Leu71 and Leu73 is also conserved in NEDD8 and serves as another means of forming non-covalent protein-protein interactions.[14,43]

On the contrary, sequential differences between the two proteins allow for differentiation between binding partners. Lys63 is a key site for ubiquitin-chain formation with significant biological function, however in NEDD8 this position is occupied with a glycine residue and therefore lacks this important conjugation site.[43] Like ubiquitin, the surface of NEDD8 is largely polar with the exception of the conserved hydrophobic patches. The main regions of

divergence between NEDD8 and ubiquitin are observed as clustered areas on the surface of the molecules.[48] NEDD8 contains a region of charged residues (Lys22, Arg25, Glu28, Glu31) whereas in ubiquitin this region is uncharged (Thr22, Asn25, Ala28, Gln31).

Another main region of divergence is a cluster containing nonpolar and charged residues on NEDD8 (Lys4, Glu12, Glu14, Gly63 and Gly64) whilst in ubiquitin these positions contain residues Phe4, Thr12, Thr14, Lys63 and Glu64. A small site of divergence is an area containing Asn51 and Glu53 in NEDD8 which are polar and charged, respectively, whereas in ubiquitin these sites are occupied by Glu51 and Gly53 which are, respectively, charged and nonpolar. Lastly, Arg72 in the ubiquitin C-terminal tail is occupied by an alanine residue in NEDD8 (Figure 5).[43] Collectively, whilst possessing highly similar structures, these sites of divergence allow for the discrimination between NEDD8 and ubiquitin by their respective conjugation and deconjugation enzymes.[43] These differences are important as, whilst ubiquitylation occurs on most eukaryotic proteins, NEDD8 is significantly less promiscuous and is conjugated onto a significantly smaller number of substrates which are mostly members of the Cullin family of scaffolding proteins.[49]

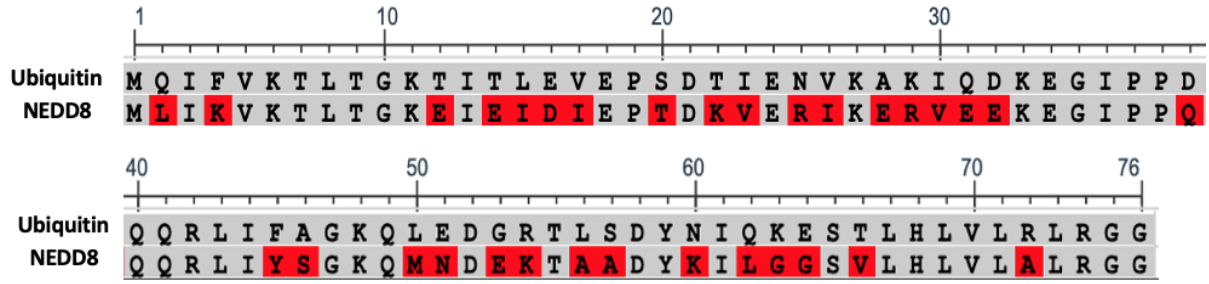


Figure 4: Alignment of human ubiquitin and NEDD8 protein sequences with divergences highlighted in red within the NEDD8 sequence. Performed using NCBI BLAST.[50,51]

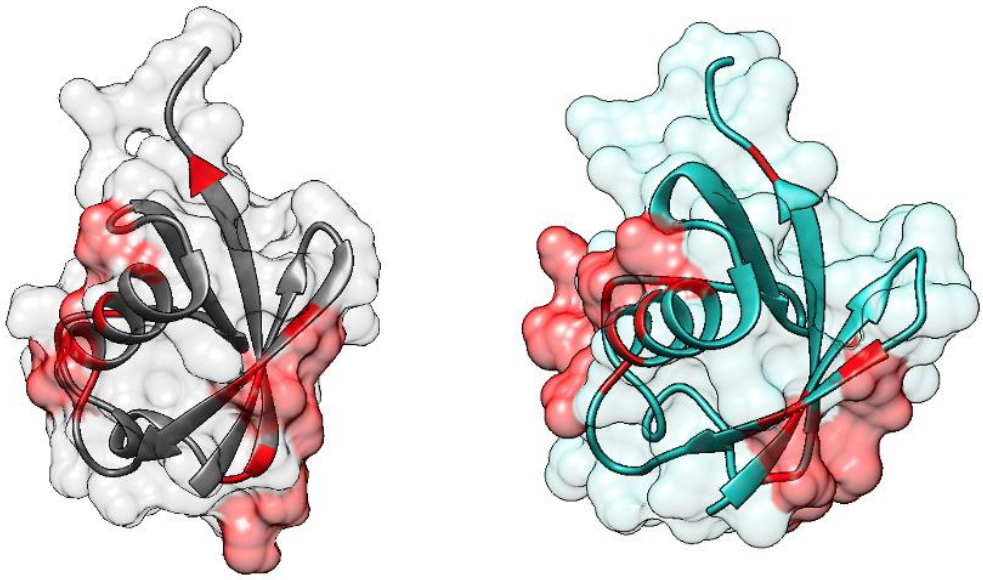


Figure 5: Structures of ubiquitin and NEDD8 with clusters of divergence coloured in red. [PDB: 1UBQ, 1NDD] (Figure generated in UCSF Chimera)

For example, the E1 enzyme NAE needs to distinguish between NEDD8 and ubiquitin for its activation. This is achieved by recognition of residues in the flexible *C*-terminal tail.[44] Upon docking to NAE subunit UBA3, a conformational change in the NEDD8 *C*-terminal tail allows for the residues Leu71 and Ala72 to bind with Leu206 and Tyr207 within UBA3.[43,52] The main difference between the NEDD8 and ubiquitin *C*-terminal tail is a single residue in position 72 which is essential in the specific recognition of NEDD8. In this connection, it has been demonstrated that a mutation of Ala72 in NEDD8 to Arg impedes this specificity for UBA3 binding to NEDD8.

Furthermore, repulsion of Arg72 is mediated by Arg190 in the UBA3 binding site.[52] Therefore, even small divergences between the NEDD8 and ubiquitin sequences can lead to key differences in enzyme recognition. In contrast, the similarity between the two proteins can lead to crosstalk between their pathways and shared recognition by enzymes. The ubiquitin-activating E1 enzyme UBE1, for example, is able to recognise and activate NEDD8 under conditions of cellular stress to help alleviate the relevant stress response.[53] However this is kinetically slow and unfavourable, therefore NEDDylation by ubiquitin-specific enzymes is uncommon but not impossible.[43]

Whilst conjugation of NEDD8 by ubiquitylation machinery is unfavourable, binding of NEDD8 to UBDs is very common. Due to the presence of the Ile44 hydrophobic patch on NEDD8 which is conserved from ubiquitin, UBDs can bind to both and many show little preference between the two proteins.[54] Whilst the UBA ubiquitin-binding domain in RAD23, a ubiquitin carrier, has a greater affinity for ubiquitin over NEDD8, the protein ubiquilin-1 has no preference for ubiquitin over NEDD8.[43,55] Likewise, in addition to ubiquitin binding, Epidermal growth factor receptor substrate 15 (EPS15) also binds NEDD8 *via* its UIM motif.[56] The ability for a protein UBD to distinguish between ubiquitin and NEDD8 is not yet understood.

Recently, phage lambda display of a human brain cDNA library has revealed a novel binding domain in the protein KHNYN named CUBAN (Cullin-Binding domain Associating with NEDD8).[54] This domain is unique due to its preference for NEDD8 binding over ubiquitin, a trait not seen in other UBDs. Furthermore, it has been demonstrated that binding of NEDD8 to the CUBAN domain requires the NEDD8 to be conjugated onto a cullin-family protein. Structural insights into the CUBAN-NEDD8 interaction were determined by NMR spectroscopy and it was revealed that positively charged His and Arg residues in CUBAN bind to NEDD8 at residues Ile13 and Glu14 in the 2nd β-strand and polar residues Glu31, Glu32, Lys33, Glu34 in the α-helix.[54] Therefore, rather than a hydrophobic interaction, the CUBAN domain interacts with NEDD8 in an electrostatic manner.[43]

1.1.3 – Ubiquitin Conjugation: E3 ligases

There exists over 600 E3 ligases in the human proteome which perform the conjugation of ubiquitin and UBLs to their target.[57] These E3s are functionally classified based on their mechanism of conjugation and the presence of characteristic structural domains. There are two major types of E3s, HECT and RING, which have distinct ubiquitylation machinery.

HECT E3s

HECT E3s were the first class of E3 ligases to be discovered and there are 28 known HECT E3 ligases in humans.[58] The enzyme consists of two components, the *N*-terminal substrate binding domain (*N*-lobe) which is attached by a hinge to the *C*-terminal HECT domain (*C*-lobe) which contains a catalytic cysteine residue.[59] Ubiquitylation by HECT E3s is carried out in a two-step process. The *N*-lobe acts as docking site for the E2-Ub thioester and binds onto the E2 enzyme. Once docked, the Ub is transferred from the E2 onto the catalytic cysteine on the *C*-lobe by transthiolation. In a second step, ubiquitin is transferred from the HECT domain onto the substrate protein where it forms an isopeptide bond with a lysine residue on the substrate.[60] The flexible hinge region allows the lobes to orient themselves in a manner that ubiquitin transfer onto the substrate is optimal.[60,61]

RING E3s

With ~600 members, the RING class of E3 ligases are the most abundant.[59] These E3s are characterised by the presence of a RING finger domain, which is made up of 8 cysteine and histidine residues coordinating two zinc ions in a cross-braced structure.[62] Unlike HECT domains, the RING domain is not inherently catalytic and lacks a cysteine residue to mediate transthiolation with ubiquitin. Instead, RING E3s act as an adaptor which binds the E2-Ub thioester and juxtaposes it with respect to the substrate to mediate the direct transfer of ubiquitin from the E2 onto the substrate protein.[63]

Cullin-RING E3s (CRLs)

Within the RING E3 family exists sub-families, with the largest consisting of over 200 members known as the Cullin-RING E3 ligases (CRLs).[64] CRLs exist as multiprotein complexes consisting of four main components, a RING finger domain for E2-Ub thioester binding, a substrate recognition domain, adaptor proteins and a cullin which acts as an overall scaffold for the CRL complex. Each component of the complex is variable and can form many different combinations, leading to a large number of CRLs being expressed and consequently a high degree of substrate specificity.[65]

There are seven members of the cullin family of proteins expressed in mammals (CUL1, CUL2, CUL3, CUL4A, CUL4B, CUL5 and CUL7), named after their involvement in selecting or "culling" proteins for ubiquitin-mediated degradation by the proteasome.[66] The *C*-terminal winged helix B (WHB) domain of the cullin binds to one of two RING domains, RING Box Protein (Rbx)-1 or Rbx2. The *N*-terminal region of cullin interacts with an adaptor protein used to link the substrate recognition domain with the rest of the CRL complex. The adaptor protein can be one of Skp1, Elongin-C, BTB and DNA damage binding protein 1 (DDB1), amongst others.[64] Substrate specificity of the CRL is determined by the specific substrate recognition protein in the complex, such as F-box proteins, VHL-box, SOCS-box, DCAF etc.[65]

The best characterised CRL is CRL1, which is also commonly referred to as the SCF (Skp1-Cul1-F-box) complex.[65] Cul1 acts as the scaffold for the complex, with RING domain Rbx1 bound to its *C*-terminal region and the adaptor Skp1 bound on the *N*-terminus.[64] Skp1 is used to facilitate the attachment of the substrate receptor F-box with the rest of the complex. The specific F-box protein of the SCF is variable, with the three most well-studied being Skp2, FBXW7 (F-box WD40 repeat-containing protein-7) and β-TrCP (β-transducin repeat-containing protein).[65] For example, $SCF^{\beta\text{-}TrCP}$ plays a key role in the DNA damage response by targeting cell-cycle proteins Cdc25 and Wee1 for ubiquitylation and proteasomal degradation. Degradation of these proteins results in inhibition of G1/S-phase progression and G2/M entry to mitosis to ultimately halt the cell cycle during DNA damage.[67]

Activation of CRLs to allow for ubiquitylation of a substrate requires the E3 itself to undergo a post-translational modification, whereby the *C*-terminal domain of cullin (Lys720 in Cul1, for example) must undergo NEDDylation.[64,68] NEDDylation of cullin induces a conformational change between the cullin WHB-Rbx interface whereby the Rbx protein, which is interacting with E2-Ub thioester, dissociates from cullin and positions itself towards the protein substrate, allowing close interaction with E2-Ub and subsequent ubiquitylation of the substrate to occur.[69] Inhibition of CRL E3 ligase activity is performed by deneddylation of the cullin by the COP9 signalosome, which is discussed in greater detail in chapter 1.2.[70]

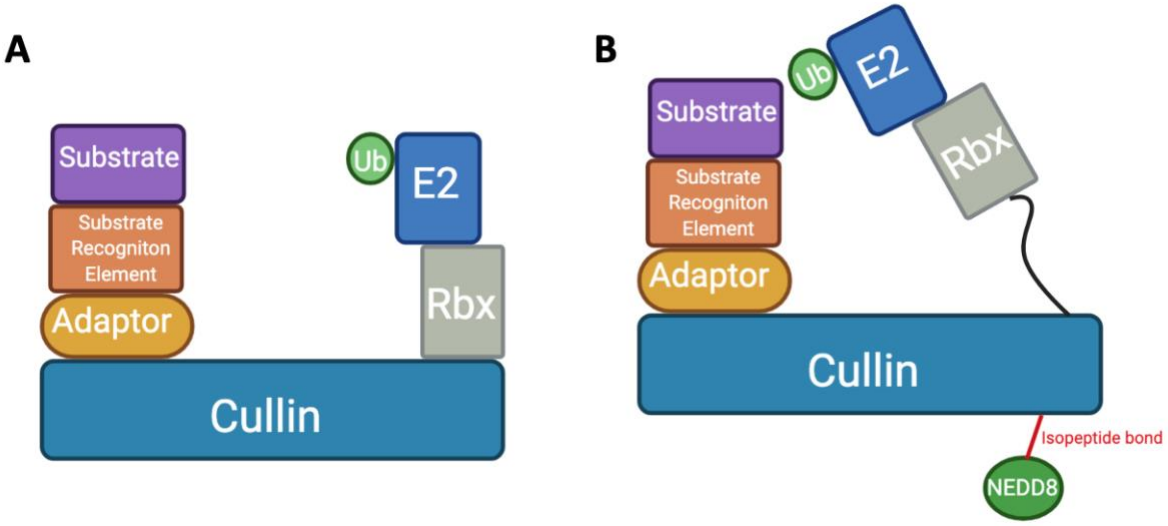


Figure 6: (A) Cullin Ring Ligase (CRL) complex in its inactive state. (B) Neddylation of CRL on the *C*-terminal WHB domain of cullin enables a conformational change in the complex allowing E3 activity and ubiquitylation.

1.1.4 – Ubiquitin Deconjugation: DUBs

Protein ubiquitylation (and covalent attachment of UBLs) is a reversible process. The removal of ubiquitin from a substrate by cleavage of the isopeptide bond (deubiquitylation) is catalysed by a group of isopeptidases known as deubiquitinases (DUBs). Deubiquitylation has a diverse set of functions, including the editing of polyUb chains, maturation of newly synthesised ubiquitin and regulation of cellular processes.[71] Whilst proteins are targeted for proteasomal degradation by being polyubiquitylated on Lys48 residues, the 26S proteasome itself requires the target protein to be cleared of all ubiquitin molecules before degradation can occur.[72] Importantly, deubiquitylation releases free ubiquitin which can be recycled for further use, thus ensuring that the cell retains an abundance of ubiquitin.[73] There are two main classes of DUBs differentiated based on their mechanism of action, cysteine/thiol DUBs and metalloDUBs. There are currently six known families of thiol DUBs, Ubiquitin *C*-terminal Hydrolase (UCH), ovarian tumour proteases (OTU), Ubiquitin-Specific Proteases (USP), Josephin, MIU-containing Novel DUB family (MINDY) and, most recently discovered, ZUFSP/ZUP1. Meanwhile there exists only a single family of metalloDUB, JAB1/MPN/MOV34 metalloenzymes (JAMM).[74]

The ubiquitin recognition/binding site on a DUB is referred to as the S1 site which, like UBDs, binds to hydrophobic patches on ubiquitin.[75] Since NEDD8 also contains these surface patches, DUBs can distinguish it from ubiquitin in a similar manner to E1, E2 and E3 enzymes, *I.e*. due to Ala72 in the *C*-terminus. DUBs which cleave polyUb chains contain another binding site called the S1' site. In the case of diUb, the distal ubiquitin binds onto the S1 site and the proximal ubiquitin to the S1' site.[76,77] DUBs can contain further interaction sites such as S2, S2', S3 etc which mediate binding to polyUb chains.[76] The S1' site is key in determining linkage specificity in the cleavage of polyUb chains as the exact orientation of the proximal ubiquitin affects the lysine residue presented to the catalytic site. The S1' site can also be located on binding partners of the DUB.[75] For example, the JAMM-family DUBs AMSH and AMSH-LP have low intrinsic DUB activity until they bind with Signal Transducing Adaptor Molecule (STAM).[78] STAM contains a UIM domain which interacts with the proximal ubiquitin of Lys63-linked diUb and thus the AMSH-STAM complex is specific for this chain type.[75,79]

Thiol DUBs

Thiol DUBs are named as such since they contain a cysteine residue in their catalytic site. The catalytic site is made up of a triad of three residues, cysteine, histidine and an acidic residue (e.g. aspartic acid in USP family DUBs). In the triad, the thiol group of cysteine is deprotonated by histidine. Following binding of ubiquitin to the S1 site, the deprotonated thiol performs a nucleophilic attack upon the carbonyl group of the isopeptide bond. This forms a negatively charged tetrahedral intermediate which is stabilised by an oxyanion hole at the carbon. The intermediate subsequently collapses and the isopeptide bond is hydrolysed. At this point, if the DUB is cleaving a diUb, the proximal ubiquitin is released from the S1' site. An acyl intermediate is formed on the (distal) ubiquitin at the S1 site. Deacylation by water is performed on this site and another oxyanion hole is formed, which hydrolyses the bond with cysteine and releases the ubiquitin from the S1 site, subsequently allowing the regeneration of the catalytic triad (Figure 7).[75]

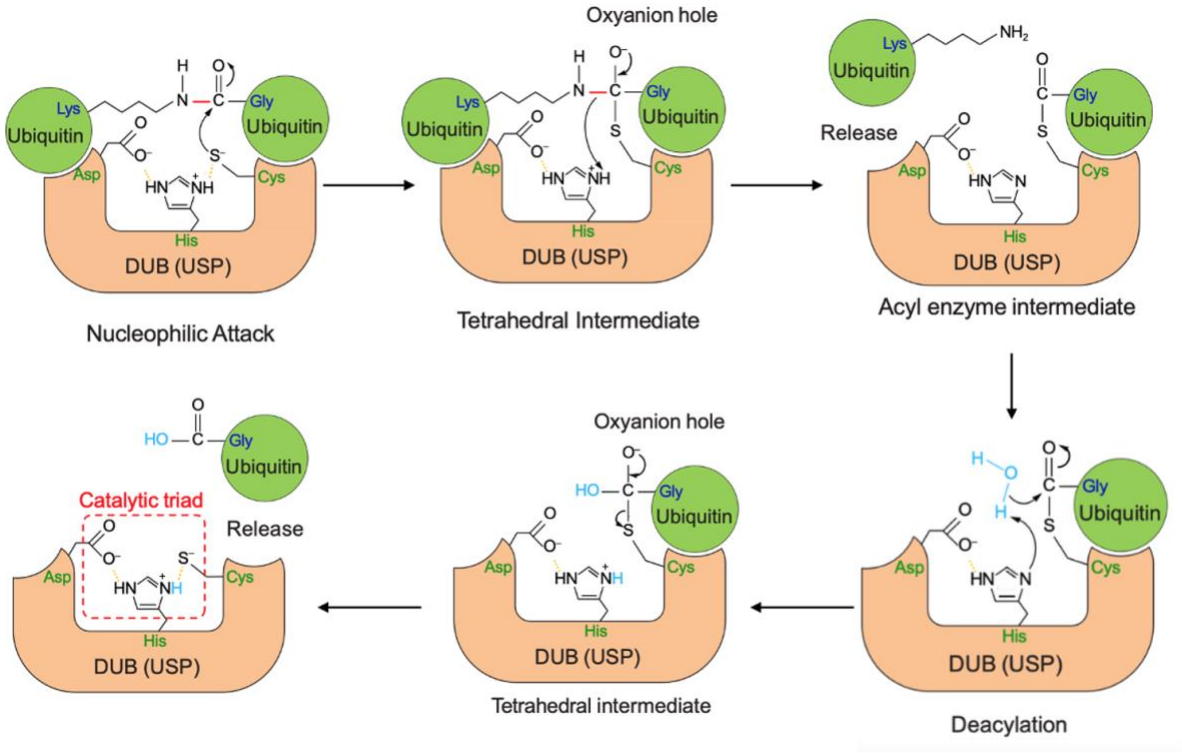


Figure 7: Catalytic mechanism of thiol DUBs, with USP as an example. Hydrogen bonds are represented by yellow dotted lines. Reproduced from Mevissen *et al*.[75]

MetalloDUBs

JAMM family DUBs do not contain a cysteine in the catalytic site. Instead, the catalytic mechanism is dependent on a zinc atom in the active site, which is coordinated by two histidines, an acidic residue and a water molecule. Following binding of ubiquitin to the S1 site and, in the case of polyUb, the S1' site, the hydroxyl group of the water molecule, following activation, performs a nucleophilic attack on the carbonyl carbon of the isopeptide bond and forms a tetrahedral intermediate. Following collapse of the intermediate and hydrolysis, the isopeptide bond is broken and both ubiquitins are released at the same time. Subsequently, the catalytic site is regenerated (Figure 8).[75]

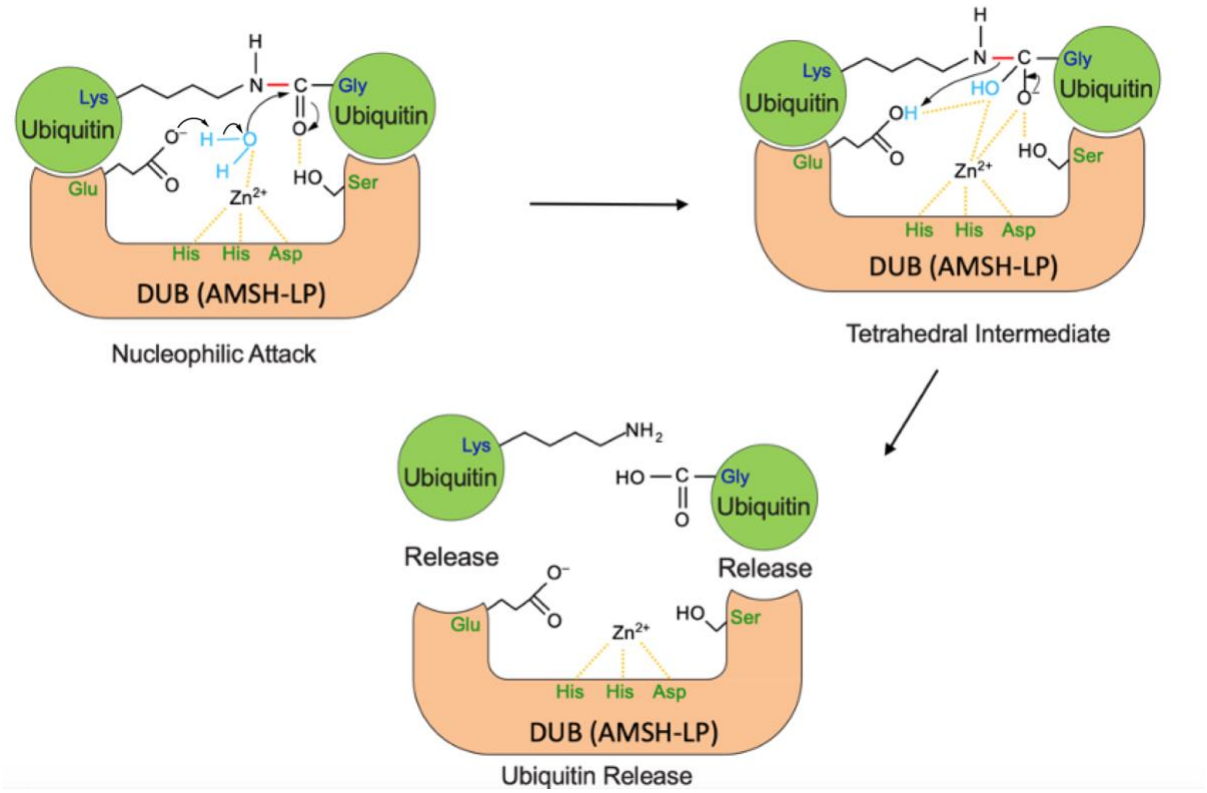


Figure 8: Catalytic mechanism of JAMM family metallo DUBs, with AMSH-LP as an example. Hydrogen bonds are represented by yellow dotted lines. Reproduced from Mevissen *et al*.[75]

1.2 – The COP9 Signalosome

The COP9 Signalosome (CSN) is a vital eukaryotic multiprotein complex that plays a key role in ubiquitin homeostasis by acting as a deneddylase. The CSN isopeptidase activity is specifically directed towards the cleavage of NEDD8 from cullins within CRL E3 ligases, which subsequently deactivates them. Unlike other isopeptidases, CSN activity is specific for NEDDylated cullins and does not cleave NEDD8 conjugations to other proteins nor ubiquitin/UBL conjugations.

1.2.1 – CSN Structure

The human CSN complex is composed of eight subunits (CSN1-8), all of which are required for full *in vitro* activity of the enzyme (Figure 9).[80] Two of these subunits contain JAMM domains (CSN5 and CSN6) and the other six subunits have PCI (proteasome lid-CSN-initiation factor) domains. X-ray crystallography has revealed that the 6 PCI domain subunits assemble into a horseshoe-like ring while CSN5 and CSN6 form a heterodimeric complex.[81] The main function of the PCI subunits is to provide a structural scaffold for the complex while the CSN5-CSN6 heterodimer is involved in catalytic activity.[81,82] The *C*-terminal α-helices of each subunit are united to form a helical bundle over the PCI ring, upon which the CSN5-CSN6 heterodimer sits and is slightly juxtaposed from the rest of the complex. Therefore the overall structure of CSN is a three-tier assembly consisting of the PCI ring, helical bundle and JAMM dimer.[81–83]

The structure of CSN is similar those of the 19S proteasome lid and the eukaryotic translation initiation factor (eIF3) which also contain JAMM and PCI domains and are arranged in a three-tier complex.[80,84] Furthermore, CSN2 and CSN4 recruit NEDDylated cullins of CRLs to the CSN complex to allow for the cleavage of NEDD8.[81] Whilst both CSN5 and CSN6 contain JAMM domains, only CSN5 has an active domain with a zinc atom required for metalloprotease activity, therefore CSN5 acts as the catalytic site within the CSN.[81] However, deletion of CSN6 severely impairs CSN5 catalytic activity, suggesting that CSN6 plays a role in stabilising the CSN5 JAMM domain to allow it to function.[81,82]

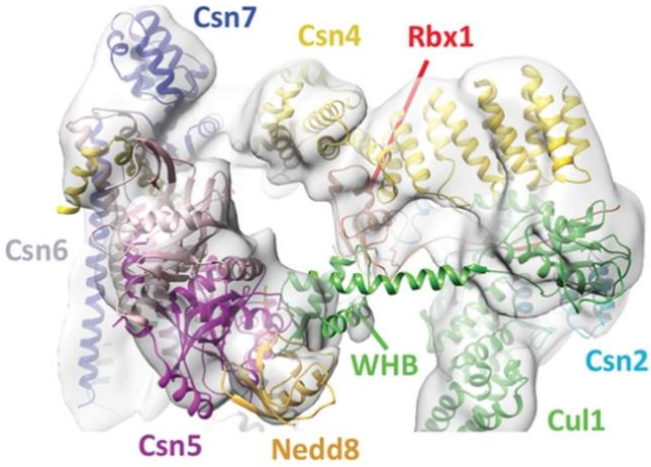


Figure 9: Molecular model of the human COP9 signalosome complex bound to NEDDylated cullin, based on cryo-electron microscopy (cryo-EM). Figure taken from Mosadeghi *et al*.[85]

1.2.2 – CSN5 Activation

Within inactive CSN, the JAMM motif of CSN5 consists of a Zn^{2+} ion coordinated by two histidine residues (His138 and His140), Asp151 and a Glu104.[81] An active JAMM domain requires a water molecule coordinating the zinc atom, however within inactive CSN5, this water is replaced with Glu104 of the insertion-1 (Ins-1) loop segment of CSN5.[81,86] In this conformation, the Ins-1 loop covers the active site and prevents it from binding to substrate isopeptide bonds and thus the enzyme is in an autoinhibited state.[81] Furthermore the active site water is coordinated and polarised by an acidic Glu76 residue in CSN5.[81,87] Therefore, in order for CSN5 to become active, Glu104 needs to be removed from the catalytic site, the Ins-1 loop must undergo a conformational shift to open up the active site and Glu76 must activate the catalytic water molecule (Figure 10).[81]

It has been evidenced that activation of CSN5 requires binding of a NEDDylated CRL to the CSN complex. Once binding has occurred, a significant conformational change in CSN5 causes the Ins-1 loop to swing out to reveal the active site. Conformational shift of the Ins-1 loop subsequently removes Glu104 from the catalytic cleft and allows water to be coordinated within the active site to facilitate isopeptidase activity.[81,87]

Studies using electron microscopy have revealed that, following binding of CSN2 to the *C*-terminal WHB of NEDDylated cullin, CSN4 undergoes a conformational rotation in its *N*-terminus.[81,88] Within the CSN complex, CSN4 interacts with the JAMM domain of CSN6 through a conserved β-hairpin loop in CSN6 called the Ins-2 loop. A conformational change in CSN4 proceeding cullin binding induces a change in the Ins-2 loop of the CSN6 JAMM domain, which enhances the CSN5-CSN6 dimer interaction and allows the CSN5 JAMM domain to be stabilised.[81] A CSN6 mutant which lacks Ins-2 loop residues was used to elucidate the specific role of the CSN4-CSN6 interaction. It was revealed that in this mutant, the enzyme had a higher activity than wild type CSN. Furthermore, the CSN6 mutant was used to study specificity of the CSN complex towards NEDD8 by surveying the ability of the mutant to cleave an artificial DUB substrate, rhodamine-ubiquitin. Whilst wildtype CSN does not cleave the rhodamine-ubiquitin, the CSN6 mutant complex lacking the Ins-2 loop was able to cleave the isopeptide bond with a K_m of 1.8 μM.[81] Therefore it is proposed that the specificity of CSN isopeptidase activity towards NEDDylated CRLs is due to the CSN4-CSN6 interaction and, once this interaction is mutated, CSN5 is able to act as a more promiscuous isopeptidase.[81]

Due to the fast rate of isopeptidase activity, an X-Ray structure of the CSN5-NEDD8 complex has not been solved. This is due to NEDD8 being rapidly released from the complex following cleavage of the NEDD8-cullin isopeptide bond.[85,89] Therefore, whilst the major sites of protein-protein interactions between NEDD8 and CSN5 are the *C*-terminus and, most likely, the Ile36 patch of NEDD8 due to its close proximity to the *C*-terminus, there is little structural information to analyse these interactions and any other potential interactions between the complex. Cryo-EM structures of the complex have been solved however these are at a relatively low resolution, thus a high-resolution X-Ray structure would enable further structural analysis of the complex.

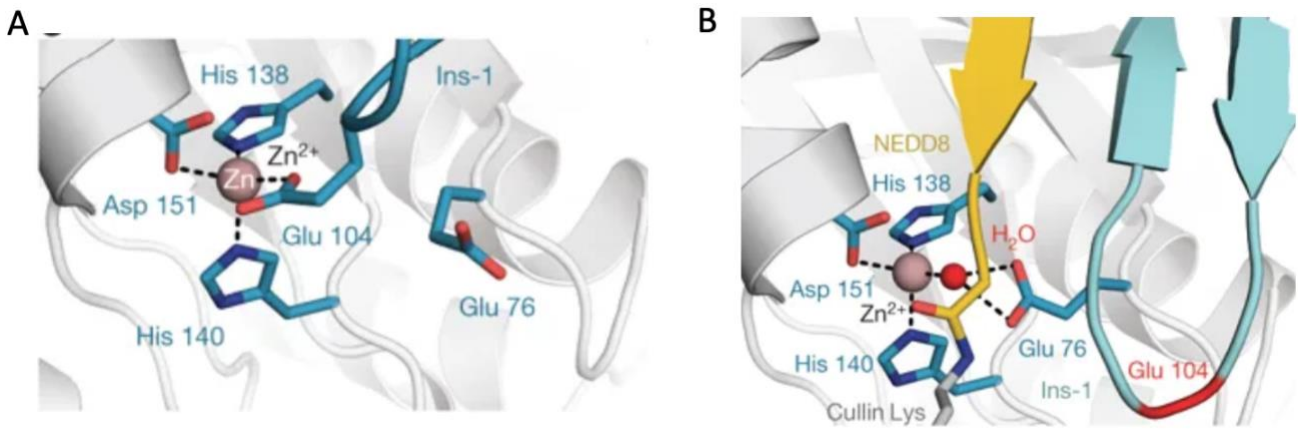


Figure 10: The catalytic JAMM site of CSN5. (A) CSN5 is autoinhibited with the Ins-1 loop blocking access of the site to water for activity. (B) In its catalytically active conformation, the Ins-1 loop (yellow) is shifted to allow coordination of water to the JAMM site. Images taken from Mosadeghi *et al.* based on cryo-EM structures.[85]

1.2.2 – Cellular Roles of CSN Activity

CRLs perform the ubiquitylation of proteins involved in a huge variety of key cellular functions such as immune response, cell cycle control, cell proliferation and stress pathways.[90] The regulation of CRLs is vital to ensure the adequate homeostasis of these cellular pathways and maintain a healthy cell. By deneddylating CRLs, CSN acts to inhibit these E3s and negatively regulate their activity. Therefore, CSN plays a vital role in maintaining cellular and proteome homeostasis.

For example, the aforementioned SCF complex ubiquitylates key players in the cell cycle for their degradation and thus halts cell cycle progression during DNA damage.[91] However, overactivity of the SCF can prevent cell cycle progression and impede cell division.[65] Likewise, other substrates of the SCF include oncoproteins such as c-Jun and β-catenin and tumour suppressors including p21, p27, and NF1.[65,92,93] Abnormal overactivation of SCF has been found in several cancers in humans.[94] Similarly, CRL4 also targets DNA damage response proteins such as the DNA replicating factor CDT1 and XPC alongside regulators of histone methylation such as the Histone-Lysine *N*-methyltransferase PR-Set7.[65,95] Thus CRL4 is involved in regulating DNA replication and chromatin modifications. Just like SCF, in human

cancers abnormal activation of CRL4 has been observed.[96] Therefore, it is vital that the CSN negatively regulates CRL activity to prevent cancer progression and maintain cell homeostasis.

Moreover, CSN asserts cellular roles indirect of its isopeptidase activity. For example immunoprecipitation analysis has shown that CSN3 within the CSN complex acts as a docking site for protein kinases such as Protein Kinase D (PKD) and casein kinase 2 (CK2) which phosphorylate tumour suppressor p53, transcription factor c-Jun and also CSN7.[97] Furthermore, CSN also associates with the DUB USP15 to deubiquitylate IκBα within the NF-κB signalling pathway, a vital regulator of the innate and adaptive immune response.[98] IκBα acts as an inhibitor of NF-κB by tethering it to the cytoplasm of the cell where it is inactive. Following phosphorylation by IkB Kinase (IKK) complexes, IκBα is ubiquitylated by $SCF^{\beta-TrCP}$ for its degradation, allowing NF-κB to translocate to the nucleus and become active as a transcriptional factor to mediate inflammation during an immune response. Not only does CSN deneddylate $SCF^{\beta-TrCP}$ and prevent further ubiquitylation, it binds with USP15 which cleaves ubiquitin from IκBα to allow its rapid replenishment and consequently inhibit NF-κB translocation.[98]

Interestingly, it has emerged that CSN5 also directly interacts with IKK complexes. Alongside phosphorylating IkBα for its subsequent ubiquitylation, IKK also phosphorylates CSN5 on Ser201 and Thr205 and this has been shown to attenuate CSN5 deneddylase activity, allowing for IkBα degradation and NF-κB translocation.[99,100] This was determined by generating phospho-mutants of CSN5 and determining its deneddylation activity for NEDD8-Cul1. It was shown that the mutant generates less free Cul1 than the wildtype enzyme, suggesting that the mutant CSN5 exhibits less deneddylation activity, however some activity still persists.[99] It has also been shown that treatment with IKK activator TNF-α results in increased interaction of CSN with CRL5 and proteasomal DUBs, but the effects of these interactions downstream of TNF-α activation have not yet been studied.[99]

1.2.3 – CSN in Disease Pathogenesis

Dysregulation of CSN has been associated within the pathology of several diseases. In particular, CSN5 overexpression is seen in a variety of cancers as it has been evidenced to positively regulate oncogenes and negatively regulate tumour suppressors.[101] Most notably, low levels of the cyclin-dependent kinase inhibitor p27, which is a tumour suppressor, are seen in ~50% of cancers and this downregulation is correlated with CSN5 activity.[102] CSN5 is implicated in the translocation of p27 from the nucleus to the cytoplasm and thus inhibits its antiproliferative effects, resulting in cell proliferation and tumour progression.[103]

Moreover, the tumour suppressor p53 is activated in response to DNA damage and acts as a transcription factor to express downstream genes such as p21 which inhibit the cell cycle and halt cell proliferation, thus preventing cancer progression.[104] The RING E3 ligase Murine Double Minute 2 (MDM2) regulates p53 activity by targeting it for ubiquitylation and proteasomal degradation.[105] It is extremely important that correct levels of p53 are maintained in the cell as p53 accumulation can induce apoptosis whereas under expression leads to tumour progression.[106,107]

Phosphorylation of p53 by ataxia telangiectasia mutated (ATM) hinders its binding to MDM2 and prevents its ubiquitylation, while MDM2 phosphorylation by ATM destabilises its RING domain and inhibits its activity.[108,109] Both of these phosphorylation events ultimately allow for p53 accumulation in the DNA damage response. However, CSN5 has been implicated in the stabilisation of MDM2 and subsequent degradation of p53, thereby allowing tumour progression and cancer pathophysiology.[110] The exact mechanism for this MDM2 activation is unclear but may be associated with the stabilisation of its upstream enzymes by means of protecting them from ubiquitylation by CRLs.

Vascular smooth muscle cells (VSMCs) are a major component of blood vessels and, in healthy vessels, exhibit low levels of proliferation and differentiation.[111] However under pathological conditions there is increased proliferation of these VMSCs which is implicated in the development of vascular diseases such as atherosclerosis and pulmonary/systemic hypertension, and also in cancer.[90] SCF ubiquitylates cell division cycle 25A (CDC25A) and

targets it for degradation.[112] CDC25A is known to positively regulate the cell cycle and enable increased cell division.[113] Therefore increased CSN activity leads to accumulation of CDC25A and it has been hypothesised that this leads to increased proliferation of VSMCs.[90]

The *CSN5* gene only rarely undergoes mutations.[101] Therefore it is not clear how CSN5 overexpression occurs and whether there is an upstream enzyme regulating CSN5 activity and expression. Whilst CSN interacts with proteins such as Myc, USP15 and protein kinases, it is not currently known whether these proteins modulate CSN5 activity.[114] Therefore further investigations into CSN5 activation and expression are essential for pharmacological intervention and drug discovery.

1.2.4 – CSN5i-3

As the catalytic centre of CSN activity, CSN5 is a viable drug target for inhibition of the CSN complex. High-throughput screening of compounds against CSN5 and subsequent optimisation allowed for the development of the first small-molecule inhibiter of CSN5, named CSN5i-3 (Figure 11).[115] This molecule acts as a zinc chelator and binds to the zinc atom within the JAMM domain of CSN5 in a monodentate manner. This zinc chelating property is achieved through the imidazole ring in the molecule. Specificity towards CSN5 is accomplished due to an amide-linked region of the molecule that extends to the active site cleft of CSN5, as revealed by the co-crystal structure determined by X-Ray diffraction (Figure 11). Moreover, CSN5i-3 is the first example of a JAMM-family DUB inhibitor. CSN5i-3 potently inhibits CSN5 and has an IC_{50} of 5.8 nM. The small molecule is also specific for CSN5 inhibition and does not show affinity towards other tested JAMM family proteases, such as AMSH-LP which is structurally the most similar DUB to CSN5.[115] However, the selectivity of CSN5i-3 for CSN5 has not been tested against the JAMM-family DUBs BRCC36 and MYSM1.

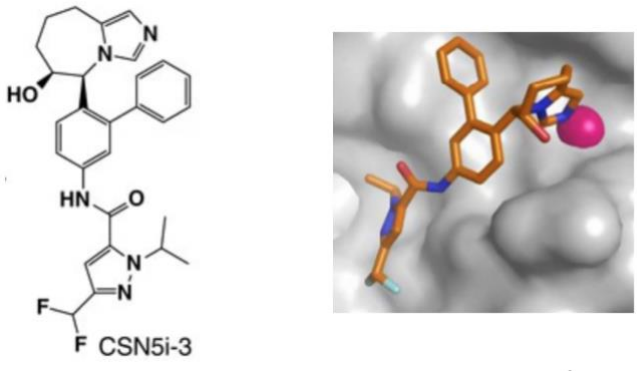


Figure 11: CSN5i-3 with a model of it bound to the catalytic Zn^{2+} atom of CSN5. Figure taken from Schlierf *et al*.[115]

1.3 – Activity Based Probes

1.3.1 – The Need for Tools to Study Enzyme Activity

Many enzymes in the cell are members of larger signalling pathways, leading to their tight regulation that is necessary to maintain proteome homeostasis. The dysregulation of enzymes is a major contributor to disease pathology and therefore many drugs act to restore this homeostasis by activating or inhibiting these enzymes.[116] Therefore it is important to have tools and methods to study enzyme activity, however this can be difficult, especially in enzymes that are subject to posttranslational modifications for regulation of their activity.[117]

A large number of protein studies are performed by means of measurement of mRNA levels or protein abundance, however this is not a feasible way of studying protein activity. Many proteases, for example, are found in an autoinhibited state and must undergo activation before they are enzymatically functional.[117] So, whilst an enzyme may be abundant in a cell, its expression does not correlate with its activity and therefore these aforementioned studies will not provide any quantitative information on enzyme activity. Many studies instead focus on the abundance of an enzyme's target substrate. However, this is a highly indirect approach as proteins can be subject to regulation by a plethora of different enzymes, therefore these studies can provide highly ambiguous data and non-robust conclusions.[117]

Activity-based probes (ABPs) have emerged as powerful tools in the study of enzyme activity. ABPs irreversibly bind onto the catalytic site of an enzyme only in its active state and their binding can be monitored to quantify enzyme activity. ABPs covalently label enzymes of interest in an activity-dependent manner. There are 3 components of an ABP: a reactive group referred to as a warhead; a recognition element; a reporter tag.[118]

1.3.2 – Components of an Activity Based Probe

Warhead

Labelling of an ABP to the enzyme active site is achieved through the warhead which targets the catalytic component of the enzyme. This is normally an electrophile on the probe which reacts with an active site nucleophile on the target.[119] Where this reactivity is not possible, a photocrosslinking approach can by utilised whereby a UV-sensitive linker on the probe forms a radical intermediate when stimulated with UV light and can form covalent bonds with the target enzyme.[120] Enzymes which contain metals in their active site can also be targeted with a metal-binding moiety or photocrosslinker.[118] It is important that the warhead is able to reach the active site in close proximity to facilitate binding. As such, it is key that the other components of the ABP do not interfere with the ability of the warhead to closely interact with the target with appropriate geometry. For example, a bulky recognition element may hinder the ability for the warhead to bind to the target.

Recognition Element

Whilst the warhead attaches the probe onto a catalytic site, it does not provide selectivity for a certain enzyme. Therefore, a second component, the recognition element, acts to direct the ABP towards a specific enzyme or class of enzyme by possessing affinity for the enzyme.[117] This is usually accomplished with the use of peptides which mimic a portion of the substrate, a protein domain or a whole protein. A major caveat with the recognition element is the poor cell permeability of many peptides and proteins.[121] Accordingly, many ABP studies are limited to cell lysates unless the issue of cell permeability is addressed. The incorporation of a cell-

penetrating peptide (CPP), a short peptide consisting of basic residues (lysine and arginine) onto the recognition element has been used for ABPs in proteomic studies.[122,123]

Furthermore, it is important when designing an ABP to consider both the reactivity of the warhead and selectivity of the recognition element. A highly reactive warhead coupled with a low affinity recognition element may not provide significant selectivity due to promiscuity of the warhead. However, if coupled with a high affinity recognition element, then the reactivity of the warhead likely will not hinder probe selectivity. Likewise, a low reactive warhead combined with a high affinity recognition element will provide high selectivity but the efficiency of probe labelling will be reduced.[124]

Reporter Tag

A key feature of the ABP which differentiates it from an enzyme inhibitor is the attachment of a reporter tag which facilitates detection of probe labelling and quantitative enzyme activity assays. The choice of reporter tag is dictated by the desired probe labelling assay or study, *i.e*. ABP target identification or ABP target validation.[125] Affinity labels such as epitope tags, for example haemagglutinin (HA) and poly-histidine tags, can allow for immunoblotting and enrichment of a probe-labelled enzyme from a cell sample to determine labelling to target proteins. Biotin tags can also be used to facilitate proteomic analysis of probe labelling within a whole-cell lysate to efficiently identify ABP targets. Biotin is particularly useful for proteomics since treatment of a cell sample with streptavidin/avidin will induce high-affinity binding of biotin to avidin and significantly reduce background labelling to non-targets.[126] The use of fluorophores such as tetramethylrhodamine (TAMRA) and fluorescein isothiocyanate (FITC) can be utilised in experiments to validate probe-labelling to a specific target. For example, if a particular protein is being targeted by the ABP, then fluorescent labelling can allow for the rapid validation of probe labelling to the target protein.[125] Fluorescent labelling also facilitates studies of labelling kinetics such as fluorescence polarisation (FP) to quantify the association of the fluorophore with the target protein which allows determination of binding affinity and rate of binding. This is also useful for inhibitor screening as it can provide information on inhibitor activity, as probe labelling decreases with inhibitor binding.[127] Therefore, whilst useful if a specific ABP protein target is sought out, the use of fluorescent

reporter tags, unlike epitope/biotin tags, provide little purpose for the identification of yet-unknown ABP targets.

1.3.3 – Activity Based Probes in Ubiquitylation

E3 ABPs

ABPs have been used to elucidate key findings in the activation of enzymes. Due to the multi-step enzymatic cascade underlying the activation of its components, ABPs have shown particular promise within the ubiquitin-proteasome system. In particular, the first probes for E3 ligases were developed in 2016 and has allowed for studies on interesting pathologically-important proteins such as the E3 ligase Parkin.[128] Parkin plays a major role in mitochondrial homeostasis by ubiquitylating substrates such as mitofusin (Mfn)-1 and -2 which normally act as fusion proteins to facilitate oxidative phosphorylation within the mitochondria.[129] However during mitochondrial damage it is vital that these substrates are degraded to prevent oxidative stress which would otherwise lead to neurodegeneration.[130] Autosomal recessive mutations in Parkin are a major cause of Parkinson's Disease, which is the second most common human neurodegenerative disease.[131] Mutations in Parkin lead to its loss of activity and results in accumulation of Parkin substrates, leading to Parkinson's Disease pathophysiology.[130] The mechanism of Parkin activation was unclear as it is autoinhibited in a resting state and therefore its abundance does not correlate with substrate ubiquitylation. An ABP for Parkin was developed which has an activated vinylsulfide (AVS) conjugated to *C*-terminus of ubiquitin acting as an electrophilic warhead for catalytic cysteine within the Parkin active site.[128] Selectivity of the ABP to Parkin was achieved by conjugating the AVS-Ub warhead to the E2 enzyme UBE2L3 which acts as a substrate for Parkin. A TAMRA fluorophore was used as the reporter tag which allowed for probe labelling studies for Parkin (Figure 12). Activity-based protein profiling with the ABP showed that activation of Parkin requires its phosphorylation by PTEN-induced kinase 1 (PINK1) on Ser65 in response to mitochondrial depolarisation. Therefore the ABP has shown promise as a novel biomarker for Parkin activity, which is clinically useful as there is currently no test for the early diagnosis of Parkinson's Disease.[128] Furthermore, an ABP is a highly powerful tool for drug discovery as it allows for screening of drugs which modulate target activity.

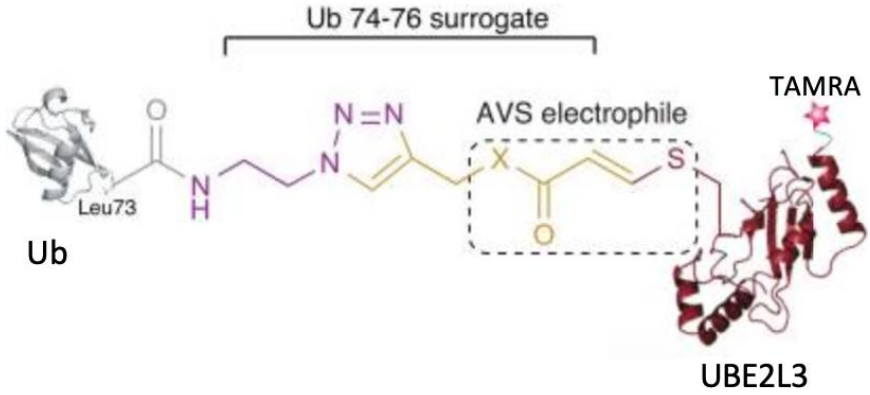


Figure 12: Structure of ABP for Parkin. The warhead consists of Ub-AVS. The E2 enzyme UBE2L3 is the recognition element, which is attached to a TAMRA reporter tag. Figure taken from Pao *et al.*[128]

DUB ABPs

The development of ABPs targeting DUBs is also an area of particular interest and has led to many significant findings in the field. Up until recently, only thiol DUB ABPs have been developed. The first ABP for DUBs was designed in 2001 using ubiquitin as a recognition element and vinyl sulfone for the electrophilic warhead to target the catalytic cysteine in a range of thiol DUBs.[132] Similarly, a vinyl sulfone (VS) DUB probe containing ubiquitin-aldehyde as the recognition element was utilised to solve the first structure of a DUB bound in complex with ubiquitin.[133,134] These probes were also used to discover novel DUBs such as the proteasomal DUB USP14 and the OTU family of DUBs.[135,136] ABPs have also been utilised to determine the effects of a stimulus on DUB activity, so as to elucidate cellular function. For example, in an *in vitro* adipogenesis model, the use of the Ub-VS probe showed an increase in USP7 activity, revealing its role in the regulation of adipogenesis (the production of fat cells, adipocytes). It was then proved that USP7 does in fact play a key role in adipogenesis through the deubiquitylation of the adipocyte differentiation regulator Tip60.[137] Similarly, this ABP was employed in Salmonella-infected macrophages to look at DUB activity within the innate immune response. It was identified that UCHL3 activity decreased whereas proteasomal DUB UCHL5 activity increased which induces the secretion of the cytokine interleukin-1β (IL-1β).[136,138]

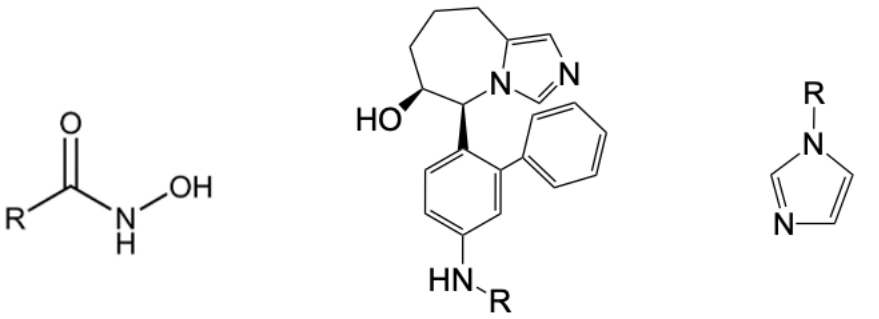
A widely-used application of ABPs is in the screening of inhibitor selectivity. Namely, the Ub-VS probe was used to determine selectivity during the development of the only DUB inhibitor to reach clinical trials, VLX1570. Assays with the ABP showed that VLX1570 selectively inhibits proteasomal DUBs USP14 and UCHL5, thus behaving as a proteasome inhibitor.[139]

Although there have been many scientific advances using thiol DUB ABPs, probes for JAMM-family DUBs have only recently been developed. Within the former, a nucleophilic thiol group forms an intermediate with the carbonyl group in the isopeptide bonds of ubiquitin conjugates. This is relatively straightforward to mimic by attaching an electrophilic warhead such as VS, AVS or dehydroalanine (Dha) onto ubiquitin.[128,132,140] However, JAMM DUBs do not directly bind to Ub conjugates, the active site Zn^{2+} ion coordinates water to attack the isopeptide bond. Therefore, it has been difficult to design a probe that can mimic this substrate. Recently, Hameed *et al.* developed the first ABP for a JAMM DUB, incorporating ubiquitin as the recognition element and, on the *C*-terminus of ubiquitin, a derivative of the zinc chelator 8-mercaptoquinoline (8MQ) as the warhead.[141] The Ub-8MQ probe was shown to label an array of JAMM-family DUBs including Rpn11/Rpn8, AMSH and AMSH-LP. Applications of this probe within cellular and inhibitor studies have not yet been described.

1.4 – Project Aims

This project aimed to synthesise an activity-based probe for the Cop9 signalosome subunit deneddylase CSN5. The warhead consists of a zinc chelator, which binds to the Zn^{2+} in the active site JAMM domain of CSN5. Selectivity of the ABP is facilitated by the incorporation of NEDD8 as the recognition element. A biotin reporter tag allows for *in vitro* testing of probe labelling.

Three different zinc binding groups were set out to be incorporated onto the *C*-terminus of NEDD8 to act as the warhead: hydroxamic acid, which is a commonly used zinc chelator due to its strong affinity for zinc binding; a precursor of the selective CSN5 small molecule inhibitor CSN5i-3; and imidazole, which provides the zinc chelating functionality of CSN5i-3 (Figure 13).



Hydroxamic Acid CSN5i-3 Precursor Imidazole

Figure 13: Warheads to be installed on the activity-based probe.

As NEDD8 is the enzymatic substrate for CSN5, the full NEDD8 protein was to be used as the recognition element. The *C*-terminus of NEDD8 in particular binds to the active site of CSN5, therefore a 12-residue peptide of the NEDD8 *C*-terminus was utilised in a second probe.
To allow for *in vitro* probe labelling of cell lysates, biotin was conjugated onto the *N*-terminus of the NEDD8 recognition element.

Figure 14: Schematic of the proposed CSN5-targetting activity-based probe. Biotin is conjugated to the *N*-terminus of NEDD8 and the zinc-binding group (ZBG) to the *C*-terminus.

Following assembly of the probe, co-immunoprecipitation assays will be performed to validate selective labelling of the probes to CSN5. Subsequent work can be performed to study the activity of CSN5, such as inhibitor screening and experiments to ascertain the biochemical mechanism of CSN5 activation. Rapid isopeptidase activity has prevented the acquirement of an X-Ray structure of the CSN5-NEDD8 complex. Therefore, labelling with the full-NEDD8 probe allows for the complex to be trapped in a bound state and can facilitate the use of X-Ray diffraction to solve a high-resolution crystal structure of the complex. Due to only the recent development of JAMM ABPs, these DUBs have not undergone the rigorous

studies of their thiol-based DUB counterparts. However, whilst untested, the ubiquitin-8MQ ABP does not likely target CSN5, which is selective for NEDD8. Therefore a NEDD8-based ABP for CSN5 can enable research into the unknown properties of, not only CSN5, but the entire CSN complex. This holds significant pharmacological interest due to the implications of CSN5 overactivity and overexpression in human cancers.

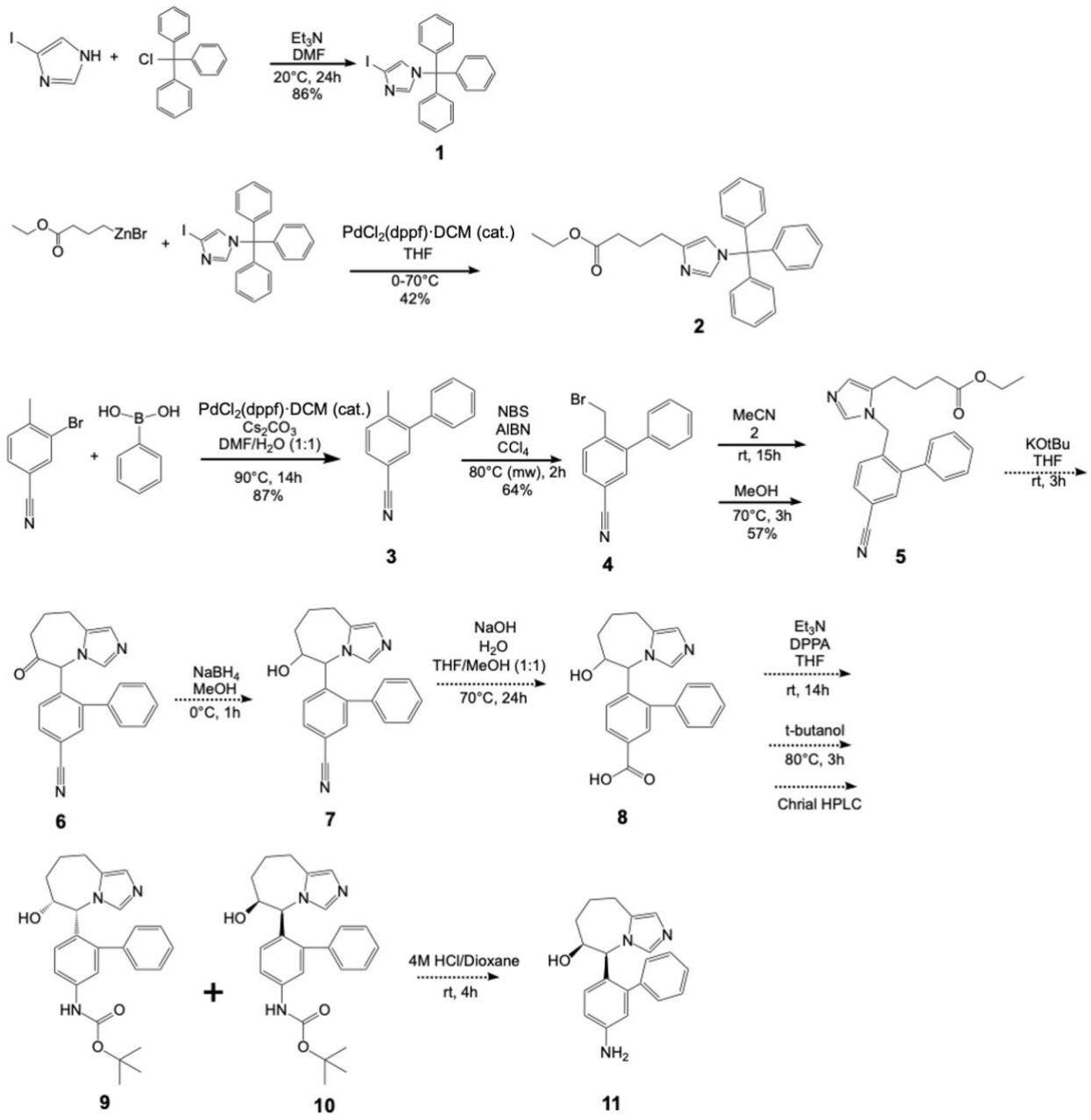
Chapter 2 – Results and Discussion

2.1 – CSN5i-3 Precursor Warhead Synthesis

2.1.1 – Synthesis

In the discovery of CSN5i-3, a precursor molecule was synthesised and reported in the literature, which was used to derive CSN5i-3. It was decided that this molecule would be used as a warhead in the CSN5 ABP. This molecule contains a zinc-chelating imidazole for attachment to the CSN5 JAMM site and has a free amine to conveniently facilitate the conjugation to the NEDD8 *C*-terminus. Schlierf *et al.*[115] previously described the synthesis of the target molecule which was reproduced in this project, with alterations in the synthesis highlighted in this chapter. The overall synthetic route of the molecule is highlighted in Scheme 1. All successful reactions produced moderate to high yields.

The synthesis begins from commercially available 4-iodoimidazole which underwent protection with a trityl group by an S_N1 reaction to give **1** in 86% yield. This was then followed by Negishi coupling with an alkyl zinc bromide to give **2** in 42% yield. In parallel, Suzuki coupling of 3-bromo-4-methylbenzonitrile to phenylboronic acid was performed to give **3** in 87% yield. Compound **3** then underwent bromination by a Wohl-Ziegler reaction producing **4** in 64% yield. The Negishi product **2** and the aryl bromide nitrile **4** were conjugated together by S_N2 and the trityl group was sequentially deprotected to give **5** in 57% yield. This would then further undergo cyclisation and subsequent reduction of the ketone to a hydroxide. The nitrile on the product can then be hydrolysed into a carboxylic acid. This carboxylic acid undergoes Curtius rearrangement to produce a BOC-protected amide, which can finally undergo deprotection to yield the final product.

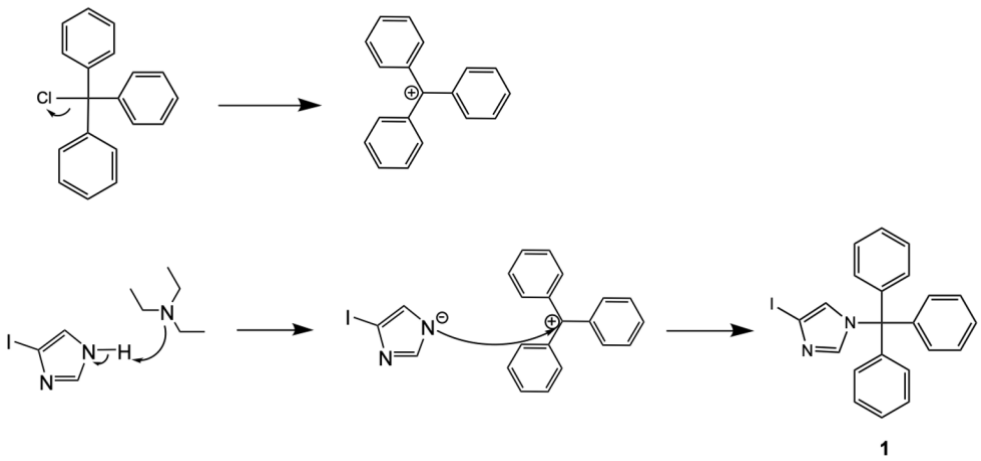


Scheme 1: Synthetic route for the synthesis of the CSN5i-3 precursor, based on Schlierf *et al.*[115]

2.1.2 – Trityl Protection of Iodoimidazole

The first step in the synthesis consists of the protection of the 4-iodoimidazole amine by nucleophilic substitution (S_N1) reaction to form 4-iodo-tritylimidazole (**1**) . This step is key in ensuring that coupling of the Negishi product (**2**) occurs on the correct nitrogen of the imidazole and also prevents the zinc bromide Negishi reagent from being quenched. Under basic conditions, the trityl chloride undergoes a kinetically slow reaction with the chloride acting as a leaving group to produce a stable tertiary carbocation. This is the rate limiting step of the S_N1 reaction and the rate is increased by the stability of the carbocation. Therefore S_N1 reactions with unstable carbocations are less favourable kinetically. Furthermore, the choice of leaving group in the reactant is important in determining the rate of reaction. A good leaving group such as chloride will improve the reaction rate. In a rapid second step, following deprotonation of the imidazole amine by Et_3N, the nitrogen anion acts as a nucleophile attacking the tertiary carbocation. Alternatively, the lone pair on the amine can attack the carbocation first and is then deprotonated by the Et_3N. The reaction proceeded readily and provided compound **1** in an 86% yield.

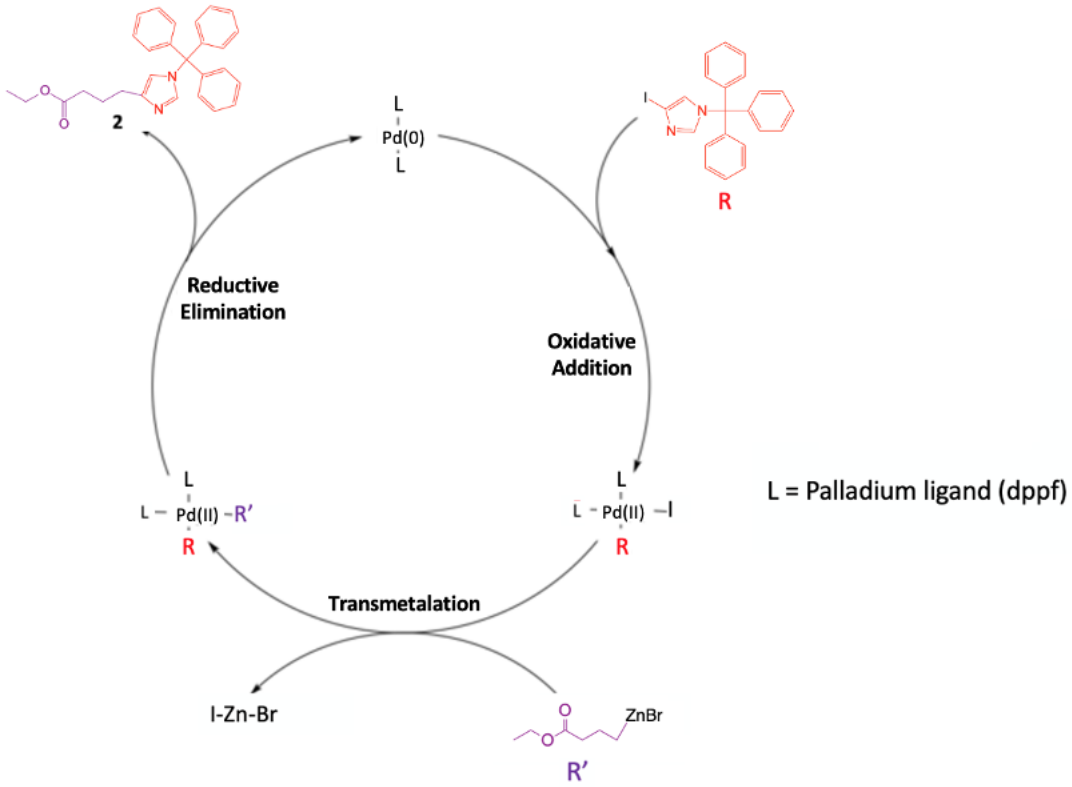
S_N1 reactions typically prefer polar protic solvents, however trityl chloride is not soluble in these solvents. Therefore, the reaction must be performed using an aprotic solvent such as DMF, DCM or THF. As there are no hydrogen-bonding interactions that can occur, the reactivity of trityl chloride largely relies on the electron withdrawing effects of the benzyl groups.



Scheme 2: Mechanism for S_N1 reaction of 4-iodoimidazole and trityl chloride.

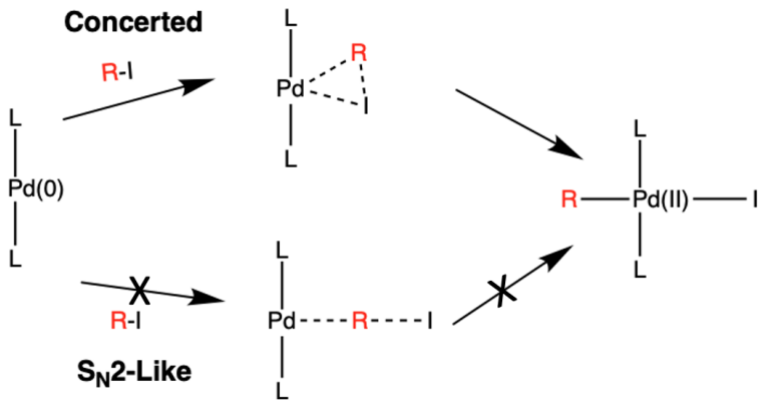
2.1.3 – Negishi Coupling

The 4-iodo-tritylimidazole (**1**) underwent Negishi coupling to 4-ethoxy-4-oxobutylzinc bromide to form Ethyl 4-(1-trityl-1*H*-imidazol-4-yl)butanoate (**2**). Negishi couplings are palladium- (and less-commonly nickel-) catalysed cross-coupling reactions which couple organozinc reagents as nucleophiles to organohalides or sulfonate esters as electrophiles by forming carbon-carbon bonds. This type of coupling was discovered by Ei-ichi Negishi who first described it in 1977 for the coupling of asymmetrical biaryls.[142] However, the reaction is not limited to the synthesis of biaryls. Whilst the reaction can be catalysed by either palladium or nickel, palladium is typically preferred as it tends to produce higher yields and has greater functional group tolerance. The Negishi coupling is distinct from most other palladium-catalysed coupling reactions as it enables the coupling of sp^3, sp^2 and sp carbons as opposed to Heck reactions, for example, which only allow for sp^2 coupling. Whilst Stille coupling also permits coupling of these carbons, its use of organostannes is highly toxic and therefore Negishi coupling is preferable due to its relatively low toxicity. Furthermore, unlike Suzuki and Stille coupling, Negishi coupling does not require a base due to the high reactivity of the organozinc reagent. A major stipulation of Negishi couplings is that organozinc compounds are air and moisture sensitive and must therefore be performed in a dry and inert environment to avoid reactions with water and oxygen.



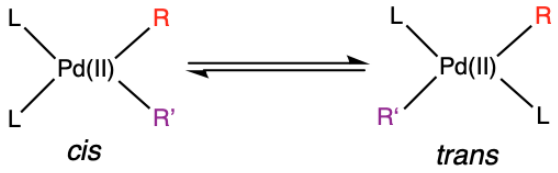
Scheme 3: Catalytic cycle for palladium-catalysed Negishi coupling to synthesise **2**

In the first step of the reaction, Pd(0) undergoes oxidative addition with the 4-iodo-tritylimidazole, oxidising the palladium to Pd(II) in the process. Whilst the exact mechanism of oxidative addition in Negishi couplings is not known, there are two likely possibilities for this step. The first is a concerted reaction and the second possible way is by an S_N2-like reaction. Due to the steric hindrance caused by the iodoimidazole, the latter is unlikely. In this step, the bond between the iodine and tritylimidazole is broken and palladium is subsequently bound to both components.



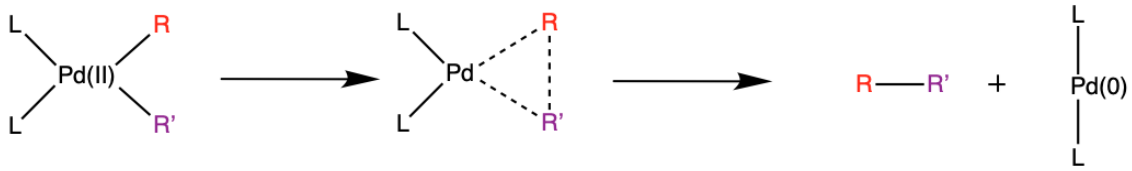
Scheme 4: Oxidative addition step of Negishi coupling. Can only progress by concerted route due to steric hindrance of the iodoimidazole.

The second step is transmetalation, which refers to the transfer of an organic substituent from a more electropositive metal to a less electropositive metal. In the case of this Negishi coupling, the zinc in 4-ethoxy-4-oxobutylzinc bromide exchanges its organic substrate for the iodine in the Pd(II) complex. This results in the palladium forming a complex with both the tritylimidazole and 4-ethoxy-4-oxobutyl substituents and also the formation of a zinc halide salt. The resulting palladium complex exists in equilibrium between its *cis* and *trans* isomers, however only the *cis* form is active in the coupling.



Scheme 5: *cis*/*trans* isomerisation of the palladium complex following transmetalation.

The third step of the reaction is reductive elimination, which is the reverse of oxidative addition, whereby the Pd(II) eliminates the product by forming a C-C σ-bond between R and R'. This is a concerted process and therefore the organic substituents must be *cis* to each other. In the process, Pd(II) is reduced to regenerate catalytic Pd(0).



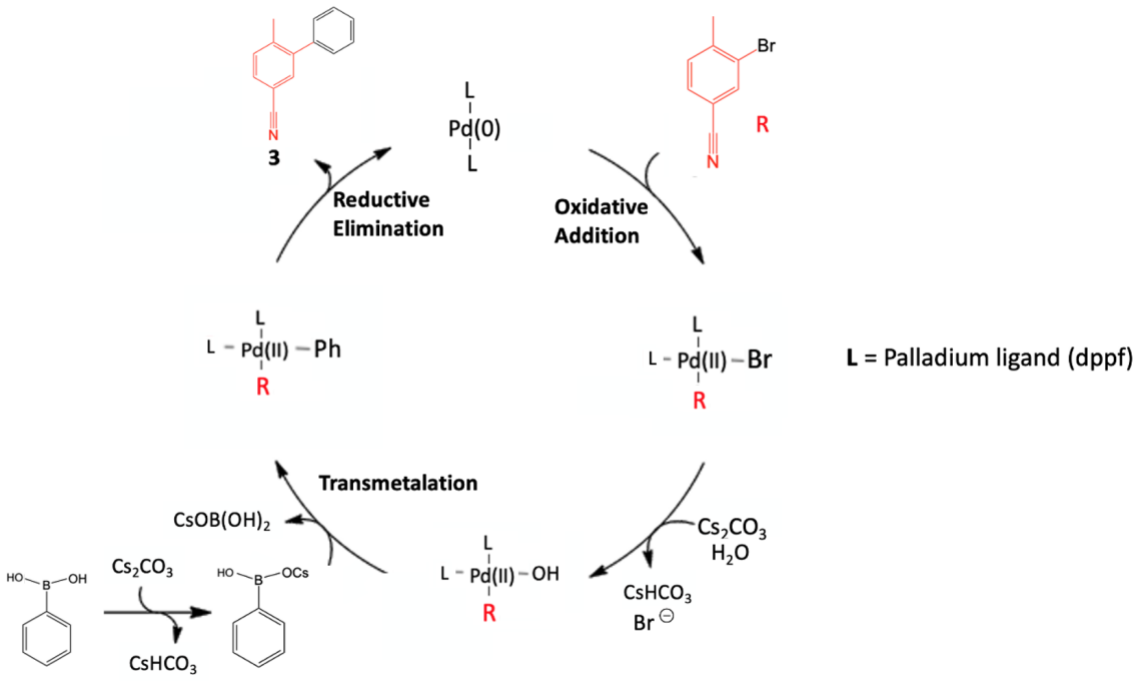
Scheme 6: Reductive elimination of the *cis* palladium complex by a concerted route.

Compound 2 was successfully synthesised, as determined by NMR and HRMS with a moderate yield of 42%. Whilst it is known that Negishi couplings occur with lower yields than other cross-coupling reactions, the yield achieved was still considerably lower than the reaction published by Schlierf *et al.* (72%).[115] The first likely reason for this is loss of product during purification. Furthermore, the reaction was attempted with an extended reaction time (14 h) however, this did not improve yield. Additionally, it was observed by TLC that full conversion of the limiting reagent was not accomplished. After multiple attempts of the synthesis, it was realised that conversion of starting material is dependent on the dryness and inertness of the reaction environment. Initially the reaction was performed by purging the flask with nitrogen gas, which is routinely used in the lab, however under this condition only low yields of ~15-20% were reached. Upon switching to argon gas, higher yields of up to 42% were achieved. This is due to argon being considerably drier and heavier than nitrogen and therefore adds a more robust blanket of protection on the organozinc reagent against air and moisture. This emphasises the necessity of stringent reaction conditions to prevent moisture and air from reacting with the organozinc reagent and interfering with the reaction. It can be stipulated that further adaptations and improved technical proficiency in the experimental setup to ensure that no moisture or air contamination occurs may improve the yield.

2.1.4 – Suzuki Coupling

Suzuki coupling of 3-bromo-4-methylbenzonitrile and phenylboronic acid to 6-methyl-[1,1'-biphenyl]-3-carbonitrile (**3**) was performed. The Suzuki coupling, like the Negishi reaction, is a palladium-catalysed cross-coupling reaction and was first described by Akira Suzuki in 1979.[143] It is commonly used for the formation of carbon-carbon single bonds between an organoboron species and a halide to produce a conjugated system of alkenes, styrenes and

biaryls. Unlike Negishi coupling, Suzuki reactions only permit the coupling of Sp^2 and Sp^3 carbons. Furthermore, Suzuki coupling requires the use of a base. Benefits of Suzuki coupling over Negishi include higher functional group tolerance, greater yields and the reaction is not air or moisture sensitive and can thus proceed in less stringent conditions. Moreover, the reaction is significantly less toxic for the environment than other coupling methods due to the use of boronic acids which are safer than organozinc or organotin reagents, for example. Likewise, the use of H_2O as a solvent in the reaction makes it more economical and can be widely used with water-soluble reagents. Mechanistically, Suzuki coupling proceeds similarly to the previously described Negishi coupling, with the main difference being the use of base to produce a boronate intermediate prior to transmetalation.



Scheme 7: Catalytic cycle of coupling of Suzuki coupling for the synthesis of **3.**

The first step involves the oxidative addition of 3-bromo-4-methylbenzonitrile to Pd(0), forming an organopalladium complex and oxidising Pd(0) to Pd(II) in the process. In this step, the carbon-bromide bond is broken and allows for both the bromide and the organic substituent to bind to the palladium.

The second step is transmetalation which is mechanistically distinct from Negishi couplings. A base (Cs_2CO_3) is added which activates water to undergo a ligand exchange from the palladium complex, resulting in a palladium-hydroxide complex and free bromide. In parallel,

phenylboronic acid also reacts with Cs_2CO_3 to form a boronate complex. The hydroxide on the palladium complex is exchanged for the phenyl group on the boronate complex, resulting in a palladium complex with both organic substituents (Scheme 7).[144]

As in Negishi coupling, transmetalation results in *cis*/*trans* isomerisation of the palladium complex, with only the *cis* conformation able to undergo reductive elimination. In this step, the organic substituents form a carbon-carbon bond and are released from the palladium complex. Pd(II) is reduced to Pd(0) in the process and thus the catalyst is regenerated.

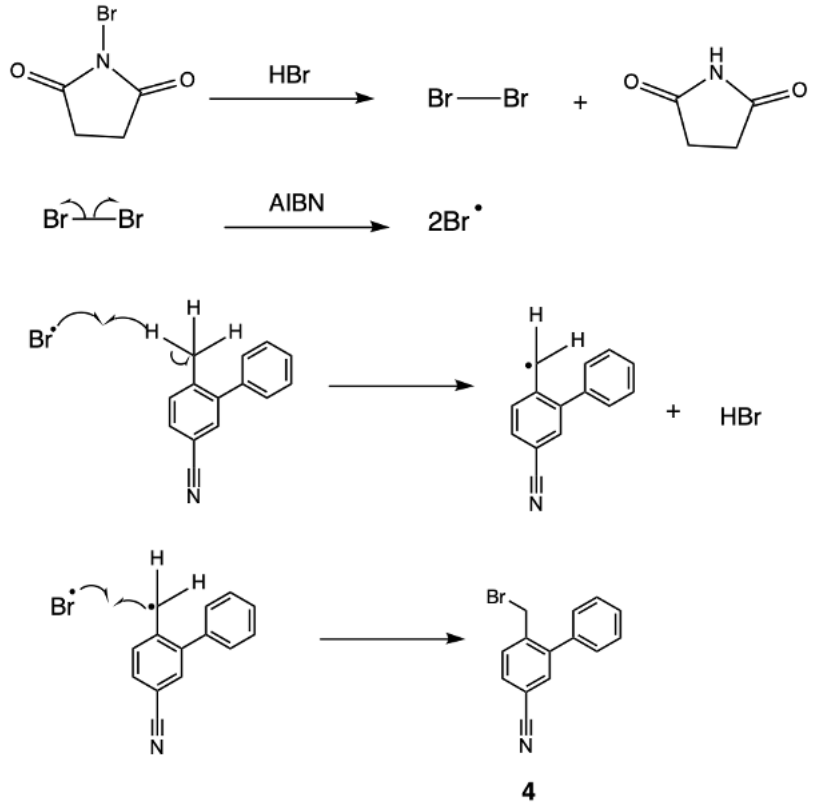
The product **3** was successfully synthesised with a yield of 87%, as determined by NMR and HRMS. As expected, the yield of this coupling reaction was greater than that of the Negishi coupling. This is likely due to the lack of need for rigorous reaction conditions. It was hypothesised that an excess of phenylboronic acid is unnecessary for the reaction to proceed, and was thus lowered. However, it was found that the reaction does not proceed without an excess of boronic acid.

2.1.5 – Wohl-Ziegler Reaction

Bromination of the methyl group on **3** was accomplished by Wohl-Ziegler substitution to result in 6-bromomethyl-[1,1'-biphenyl]-3-carbonitrile (**4**). Wohl-Ziegler substitution refers to the bromination of an allylic carbon using *N*-bromosuccinamide (NBS) and a radical initiator. The reaction was first discovered in 1919 by Alfred Wohl[145] with the use of *N*-bromoacetamide but was later modified in 1942 by Karl Ziegler[146] to include NBS after studying its value in the reaction.

HBr is usually present in trace amounts with NBS. Under heat, NBS reacts with HBr to produce Br_2 and succinimide. With the use of the radical initiator AIBN, an initiation reaction occurs whereby Br_2 undergoes homolytic cleavage to produce bromine radicals. A free bromine radical then abstracts a hydrogen from the methyl group of compound **3**, resulting in a radical on the allylic carbon and the reformation of HBr. This HBr is recycled and reacts with NBS to further produce Br_2 to repeat the process. The second free Br radical reacts with the allylic carbon radical, resulting in bromination. Due to the formation of HBr in the previous step, the

reaction is repeated until the limiting reagent **3** or NBS are depleted or upon quenching the reaction with H_2O.



Scheme 8: Mechanism of Wohl-Ziegler bromination using NBS for the synthesis of **4**.

The use of NBS allows for the Br_2 to be kept at low concentrations during the reaction. This is important as the desired substitution reaction is competing with an addition reaction. High amounts of Br_2 favours addition reactions with bromine to the phenyl group. Whilst phenyl groups are typically resistant to addition reaction, the presence of heat and AIBN can drive this forward and result in the formation of an undesired product. Therefore, by keeping the amount of Br_2 low, the use of NBS favours the substitution reaction.

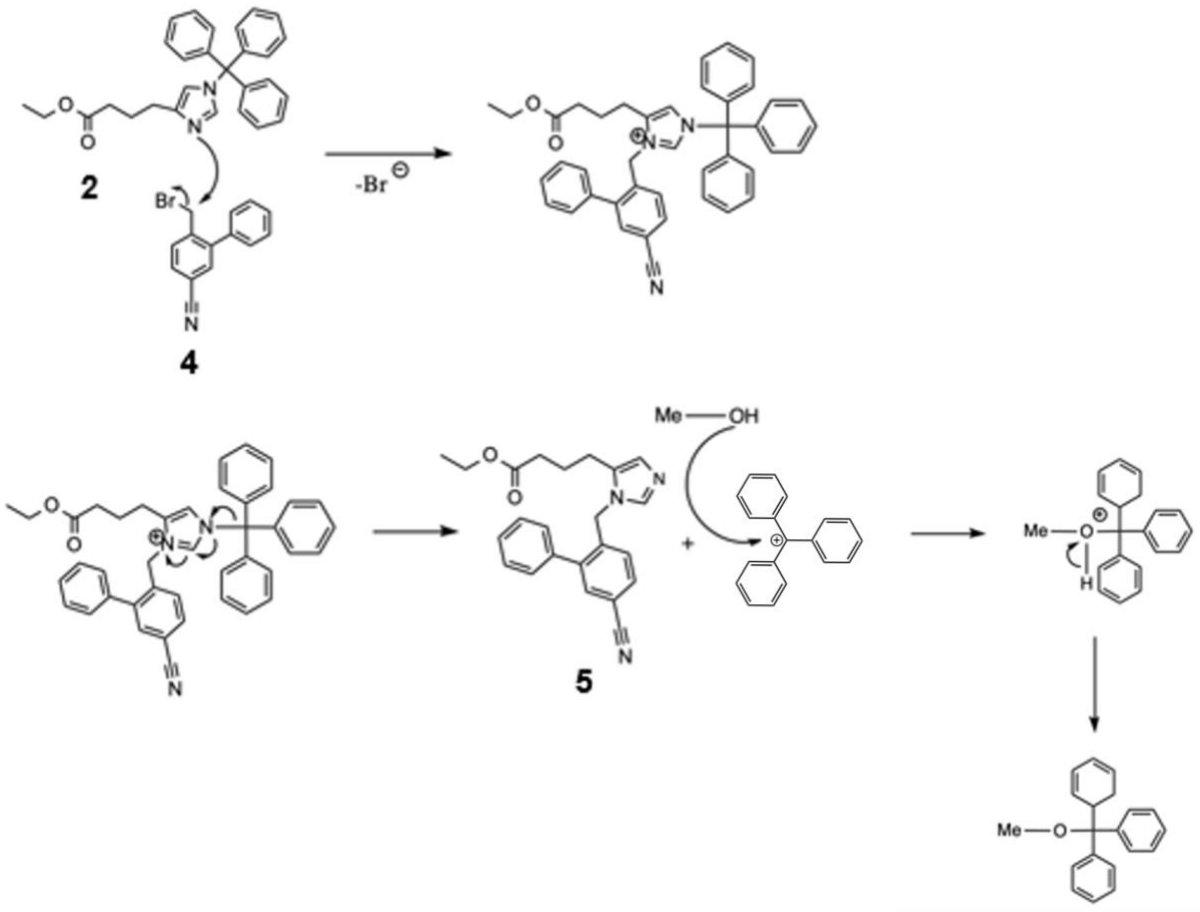
Typically, a consideration for allylic radical initiation reactions is that the allyl radical is resonance stabilised and therefore the substitution can occur on the undesired carbon. However, as the alkyl used is bound to a phenyl group, the aromaticity of the ring prevents substitution at the undesired carbon atoms.

The reaction is performed in carbon tetrachloride (CCl_4) which limits the solubility of NBS and therefore the amount of Br_2 generated is kept to a minimum. Reaction completion is also indicated by the precipitation of the resulting succinimide. A major issue with the use of CCl_4 as a solvent is its high toxicity and ozone-depleting causing environmental concerns. Going forward, environmentally safe reagents should be explored to limit the amount of hazardous waste generated. Such a solvent includes trifluorotoluene whose use has been reported in a similar reaction.[147]

Initial reactions were attempted by stirring at reflux for 20h, however it was observed by TLC that a large amount of the limiting reagent **3** had not been consumed. In an endeavour to increase conversion, it was hoped that an increase in reaction pressure would improve conversion to product. The reaction was further performed under microwave conditions with heating for 2h. It was found by TLC that the use of microwave-assisted synthesis greatly improved consumption of the starting material with an observed yield of 64% of the product **4**. This moderate yield can be attributed to product loss during purification.

2.1.6 – Conjugation of Negishi and Wohl-Ziegler Products

The products **2** and **4** underwent coupling and subsequent deprotection in a one pot synthesis reaction, producing dialkyl imidazole (**5**). At room temperature, the imidazole nitrogen on **2** acted as a nucleophile and attacks the bromo-methyl group on **4** in an S_N2 reaction. Following this conjugation, acetonitrile was removed and methanol was added to the reaction and the reaction mixture was then refluxed. The trityl-imidazole bond is broken and the electrons in the imidazole ring are transferred onto the nitrogen cation, resulting in product **5**. The carbocation formed on the trityl group is attacked by methanol forming a stable methoxy substituted trityl side product. Because acetonitrile does not possess an acidic hydrogen, it can add to the trityl group but the adduct is unstable and so can reattach to the imidazole in a reverse reaction.



Scheme 9: Mechanism of S_N2 reaction and subsequent deprotection of trityl group.

Despite difficulties in the purification of the product whereby the polarities of the product and an impurity were similar and thus were difficult to separate. Compound **5** was isolated in a yield of 57%.

As S_N2 reactions favour polar aprotic solvents, acetonitrile was used in the reaction for this step. However, the yield may be improved with the use of anhydrous acetonitrile, ensuring that no water hinders the progression of the reaction. Furthermore, the reaction was only performed at a small scale with only 50 mg (0.18 mmol) of limiting reagent. Therefore scaling up the reaction would likely make the purification less technically challenging, resulting in larger yields.

2.1.7 – Conclusion

The synthesis of a precursor molecule to the CSN5 inhibitor CSN5i-3 was attempted by the route highlighted in chapter 2.1.1 (scheme 1). Of the proposed synthetic route, five steps were successfully accomplished to produce compound **5** at an overall yield of 11.5%. Due to time constraints, the synthesis of the target molecule was not completed, with a further five steps remaining. As this molecule was intended for use as a warhead in the CSN5 ABP, alternative zinc binding domains were explored and the overall design of the probe was modified accordingly. This is discussed in detail in chapter 2.2.

2.2 – NEDD8 Synthesis

2.2.1 – Computational Design of NEDD8 Synthesis

Molecular Docking

As mentioned in chapter 1.3.2, the recognition element needs to be designed in a way such that it does not disrupt the binding of the warhead to the target enzyme at its catalytic site. The full NEDD8 protein is the native substrate for CSN5 and thus readily occupies its binding site. Whilst this will unlikely interfere with the binding of a small warhead such as hydroxamic acid or imidazole, the relatively large size of NEDD8 will likely cause hindrance in the binding of a larger warhead such as the CSN5i-3 precursor to the catalytic Zn^{2+} atom in the CSN5 active site. Therefore, in designing the synthesis of NEDD8 for conjugation to the CSN5i-3 precursor, the size of the warhead must be taken into account and the sequence of NEDD8 must be truncated to facilitate optimal binding of the probe into the CSN5 catalytic site. Truncation of the NEDD8 sequence must occur at the *C*-terminus as this is the region of the protein which binds onto the CSN5 active site. Based on this, truncation at this end allows the CSN5i-3 precursor warhead to act as a surrogate for the deleted *C*-terminal NEDD8 residues and thus facilitate binding onto the CSN5 active site. In order to determine the required extent of truncation, molecular docking was performed of both the *C*-terminus of NEDD8 (residues 73-76) and the CSN5i-3 precursor warhead (Figure 15). The CSN5i-3 precursor was *tert*-butoxycarbonyl (BOC)-protected to prevent interference of the free amine in the docking and thus generating false-positive scores.

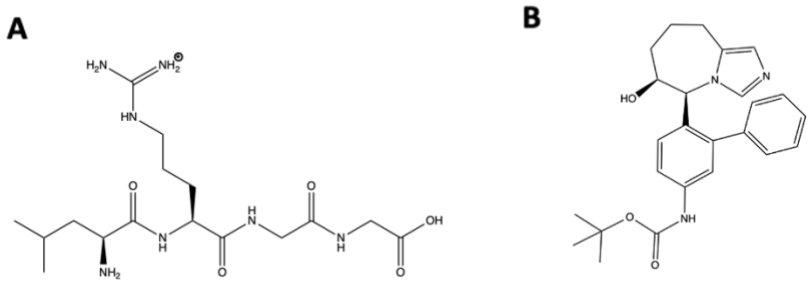


Figure 15: Ligands used for docking onto CSN5. (A) *C*-terminal residues 73-76 of NEDD8 (LRGG). (B) CSN5i-3 precursor with BOC protection of its free amine.

The generated ligands were individually docked onto a model of active CSN5 (PDB: 5JOG) on the Zn^{2+} residue in the JAMM domain. The docking results were scored by GlideScore[2] and models showing the strongest docking (highest scoring) of the ligands to CSN5 were selected. The docked 3D models were subsequently analysed to determine how the ligands bind to and interact with CSN5. This was further determined by generation of a 2D ligand interaction plot, showing which atoms of the ligand interact with CSN5 residues (Figure 16).

The interactions of CSN5 with the *C*-terminal peptide and the CSN5i-3 precursor were compared to look for any correlations in their binding. It can be seen in Figure 16 that the *C*-terminal Gly76 of NEDD8 interacts with the catalytic Zn^{2+} atom of the CSN5 JAMM domain, as expected. Likewise, the imidazole of the CSN5i-3 precursor also interacts with the catalytic Zn^{2+} atom, showing that its predicted coordination to zinc was successful in the docking. Furthermore, it was observed that the amide bond of the CSN5i-3 precursor binds to Thr154 on CSN5 by hydrogen bonding, mimicking the interaction of the Arg74-Leu73 amide bond of the NEDD8 *C*-terminus (Figure 16). It was concluded that the CSN5i-3 precursor acts as a surrogate for residues 74-76 of the NEDD8 *C*-terminus, and so a probe of NEDD8 incorporating this warhead must have a truncation of 3 residues at the *C*-terminus to produce a 73 residue protein. Ideally a model would be created with the warhead conjugated to Leu73 of the NEDD8 protein, however software limitations prevented docking with such a ligand. Alternatively, a model of this probe was generated with the warhead conjugated onto a leucine residue acetylated at its free amine to prevent interference in the docking (Figure 17).

This probe was used as a ligand for docking onto the active CSN5 model, and the binding interactions were analysed. It was observed that binding occurs as expected, with the warhead mimicking the *C*-terminal residues 74-76 (Figure 18).

In conclusion, assembly of an ABP utilising NEDD8 and the CSN5i-3 precursor requires synthesis of NEDD8 (1-73) instead of the full protein (1-76) to allow for binding to the CSN5 active site.

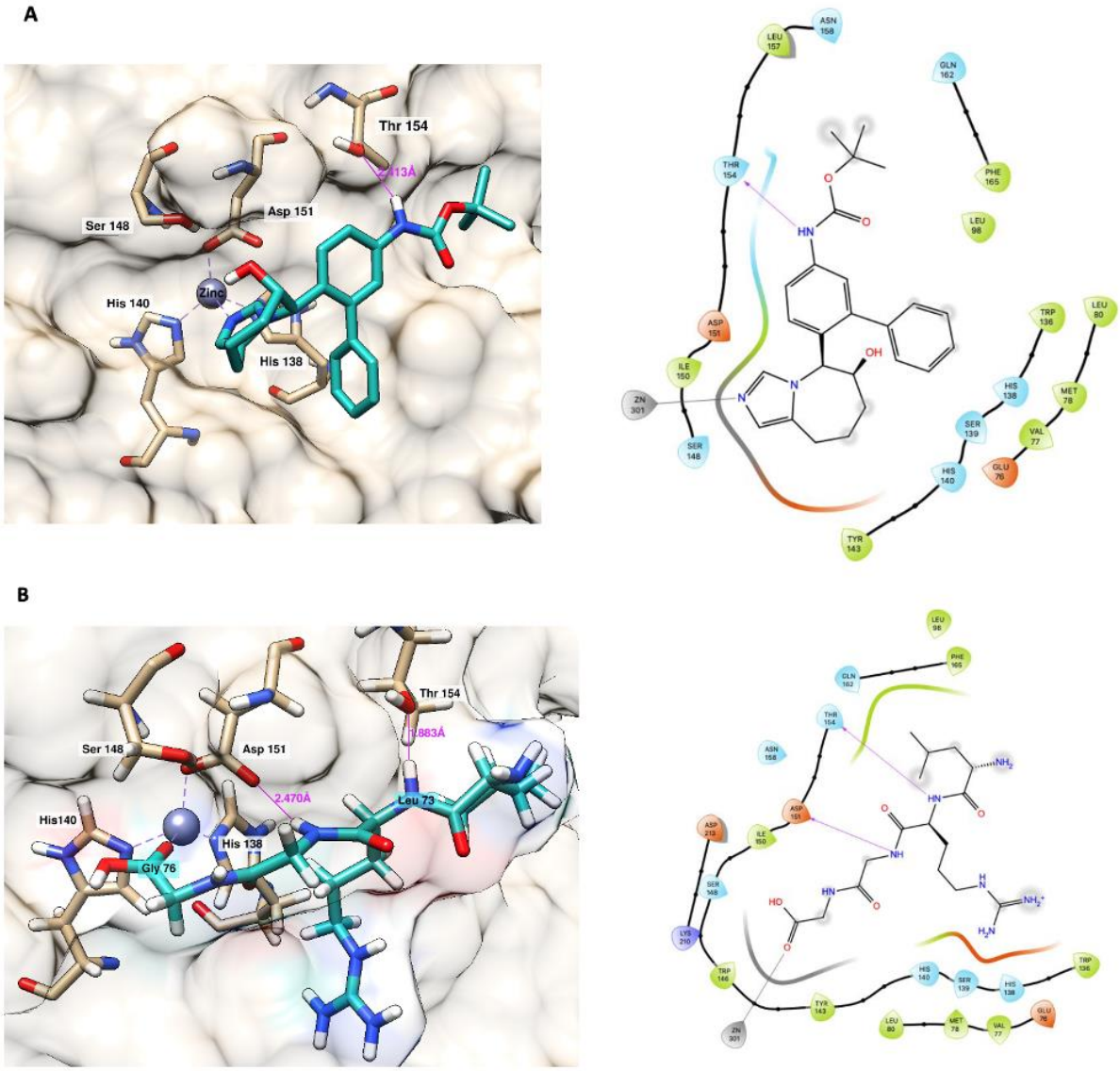


Figure 16: 3D models and 2D ligand interaction plots of ligands docked onto a model of active CSN5. Hydrogen bonding is indicated by pink lines. **(A)** Docking of CSN5i-3 precursor showing hydrogen bonding between amine and Thr154 on CSN5. **(B)** Docking of NEDD8 *C*-terminal peptide LRGG (73-76) revealed hydrogen bonding between the Leu73-Arg74 amide bond of the ligand and Thr154 on CSN5, showing a correlation between the docked models.

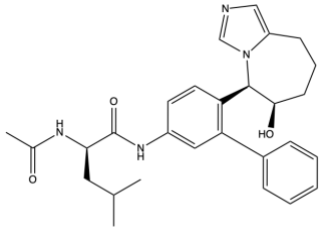


Figure 17: CSN5i-3 precursor warhead conjugated onto an acetylated leucine residue.

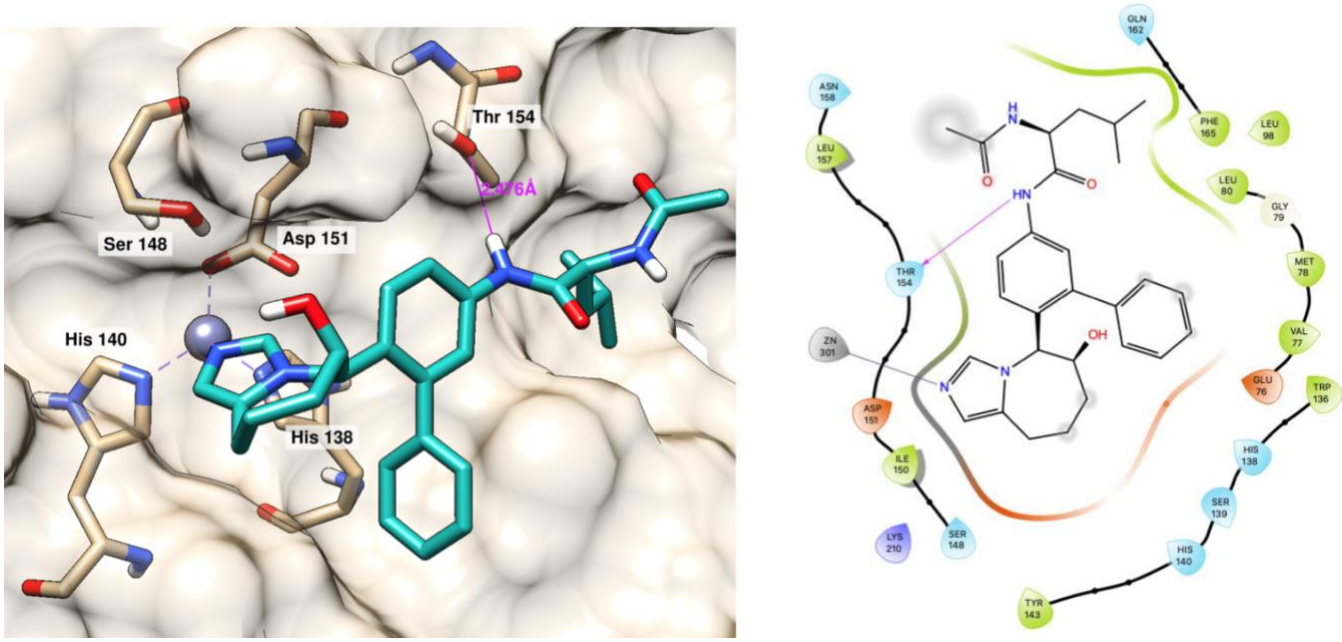


Figure 18: 3D model and 2D ligand interaction diagram of the CSN5i-3 precursor conjugated to leucine, docked onto a model of active CSN5. Hydrogen bonding is indicated by pink lines. The imidazole of the ligand coordinates zinc and, at the same time, the amide bond between the small molecule and leucine forms a hydrogen bond with Thr154.

Structure Prediction

To predict any issues that may occur during the synthesis of NEDD8, the sequence was submitted to the protein structure prediction server I-TASSER. This service utilises the PDB to look for structural templates and uses these in iterative fragment assembly simulations to generate atomic models from an amino acid sequence.[148] This was performed to check if the service could correctly predict the secondary structures of NEDD8 and highlight any problem areas which could occur during protein synthesis. If the service incorrectly predicted the

secondary structure, it may be an indication of potential issues that may arise during synthesis and would allow for adequate planning of the synthesis design to address these problems. This predicted structure was compared to the published crystal structure of NEDD8. The server was able to accurately predict the secondary structures, and no major deviations from the published structure were observed. Therefore, it was deemed unlikely that issues in secondary structure formation would arise during NEDD8 synthesis.

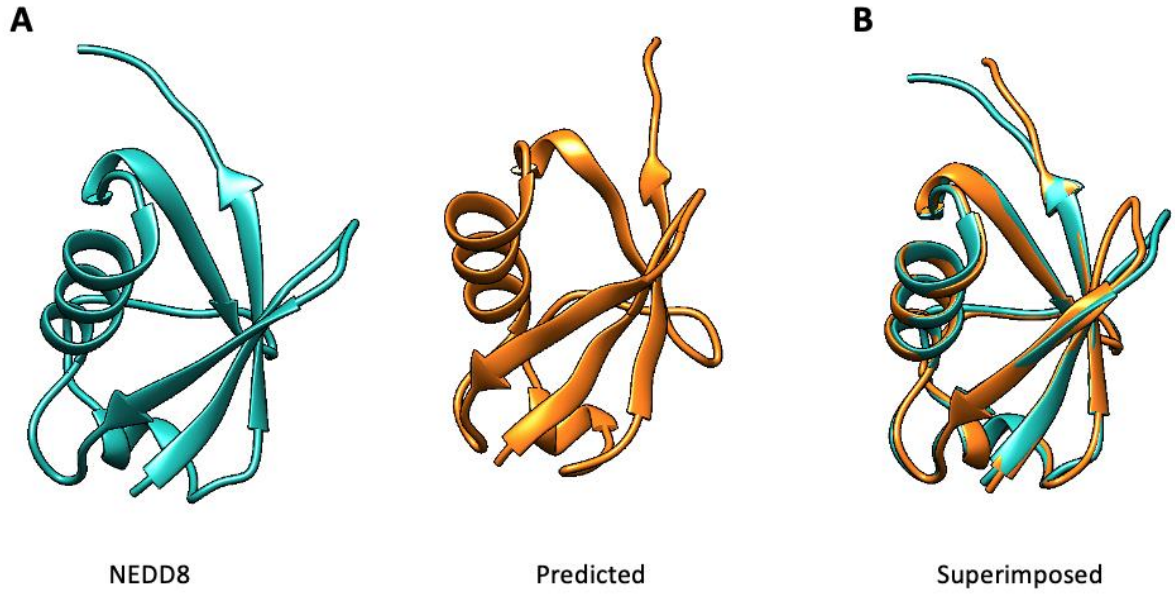


Figure 19: (A) NEDD8 structure (left) [PDB: 1NDD] and output model of structural NEDD8 prediction by I-TASSER (right). **(B)** Superimposition of models to display alignment of secondary structures.

2.2.2 – Solid-Phase Peptide Synthesis of NEDD8 Protein

In the first instance, synthesis of the full NEDD8 protein was attempted. The chemical synthesis of NEDD8 has not been reported but the synthesis of ubiquitin has been described by Oualid *et* al.[149] This provided precedence for the total synthesis of NEDD8 using a solid phase approach.

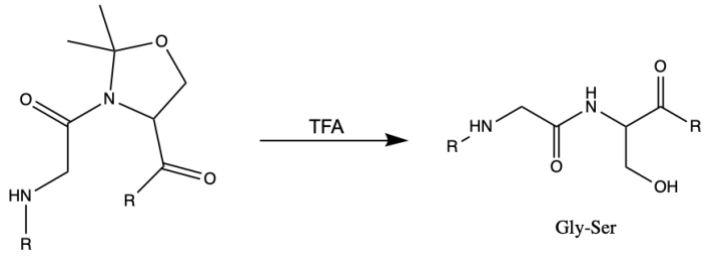
A major benefit of using the full protein as a recognition element in the ABP is that it acts as a complete native substrate for CSN5, and is thus able to replicate all of the protein-protein interactions between CSN5 and NEDD8. Furthermore, successful labelling of active CSN5 with

this ABP will produce a stable NEDD8-CSN5 complex which could be used to obtain the first X-Ray crystal structure of this complex. Previously, it has not been possible to obtain an X-Ray structure due to the isopeptidase activity of CSN5 rapidly cleaving the isopeptide bond between the proteins. Because ABP labelling is irreversible, a stably-bound CSN5-NEDD8 complex will be produced which could allow for extensive structural analysis by crystallography.

Due to the incomplete synthesis of the CSN5i-3 precursor, as described in Chapter 2.1, development of the ABPs were proceeded by utilising hydroxamic acid and imidazole as the respective zinc binding groups (Figure 13). As these warheads are significantly less bulky than the CSN5i-3 precursor, it does not require attachment to a truncated NEDD8 protein and must instead be conjugated to the full NEDD8 protein (1-76). It was therefore unnecessary to synthesise a truncated NEDD8 protein and synthesis of the full NEDD8 protein (1-76) was instead proceeded by Fmoc-based solid-phase peptide synthesis (SPPS).

Whilst computational prediction (Chapter 2.2.1) showed that there was unlikely to be any major issues in secondary structure formation, the large 5-stranded β-sheet was still likely to be a major problem area in the synthesis. β-structure formation is known to cause lowered solubility and solvation and, in turn, aggregation of a peptide during synthesis[150]. Due to the relatively large sequence to be synthesised, there was a high possibility of protein aggregation during the formation of this β-sheet. Pseudoproline dipeptides were utilised in an attempt to minimise the risk of forming these secondary structures, and thus avoid aggregation during the synthesis.[150] Pseudoprolines consist of serine- or threonine-derived oxazolidines or cysteine-derived thiazolodines which form heterocycles akin to proline. This creates a kink in the peptide and prevents β-sheet formation and aggregation during synthesis.[150] Following complete synthesis of the protein, pseudoprolines can be cleaved with TFA to generate the intended peptide sequence and allow for protein folding (Scheme 10).[150,151] Pseudoproline dipeptide building blocks were inserted in the NEDD8 sequence where serine or threonine residues were present. The building blocks were placed at regions where they are spaced apart by at least 5 residues to allow for optimum formation of kinks in the protein backbone. Proline residues in the sequence are also taken into account and thus pseudoprolines do not

need to be incorporated in a region of the sequence where there is a native proline (Figure 20).



Scheme 10: Gly-Ser pseudoproline dipeptide heterocycle can undergo cleavage by TFA to result in native peptide sequence.

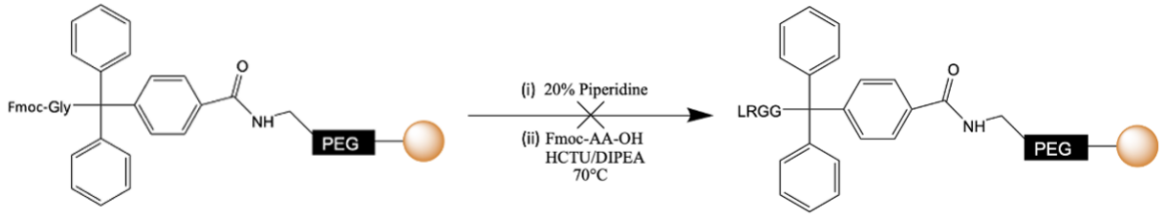
MLIKVKTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGG

Figure 20: NEDD8 sequence with pseudoproline dipeptide building block insertions highlighted in yellow. Native prolines in the sequence are highlighted in green. The building blocks were incorporated across the whole sequence and spaced by at least 5 residues.

Synthesis of the protein was performed on a 0.05mmol scale using NovaSyn tentagel 4-carboxytrityl (TGT) resin pre-loaded with Fmoc-glycine at 0.21mmol/g loading. This resin consists of polystyrene beads conjugated to a polyethylene glycol (PEG) linker, and subsequently an amide-carboxytrityl linker where Fmoc-Gly is coupled. The PEG linker increases the polarity and swelling ability of the resin, making it useful for the synthesis of long peptides or proteins as it reduces the risk of aggregation.[152,153] NovaSyn TGT also has a low-loading capacity which minimises entanglement of peptide chains during elongation and is therefore preferred for long peptide synthesis as it reduces the risk of issues such as aggregation.[154,155] This resin generates a peptide with carboxyl functionalisation at the *C*-terminus. Moreover, the resin is highly acid labile and can therefore be cleaved under weak acidic conditions. This was necessary to avoid removal of protecting groups on the peptide to allow for selective modifications of the *C*-terminus.

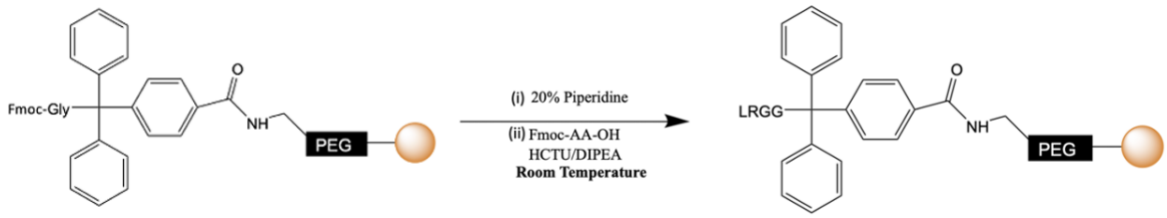
To determine the optimum conditions for protein synthesis, an initial microwave-assisted SPPS of the NEDD8 *C*-terminal peptide LRGG (residues 73-76) was attempted. At a 0.05mmol scale, the Fmoc-Gly NovaSyn TGT resin was swollen in DCM before deprotection of the Fmoc in 20% piperidine in DMF. Coupling of the first residue (glycine) was performed using HCTU and DIEA in DMF at 70°C for 5 minutes under microwave irradiation. Following coupling of

the first residue, the peptide was washed with DMF and the deprotection/coupling process was repeated for each residue in the elongation of the peptide to result in Fmoc-LRGG (Scheme 11). This was cleaved from the resin with 95% TFA and analysed using LC-MS. However, no peptide mass was observed.

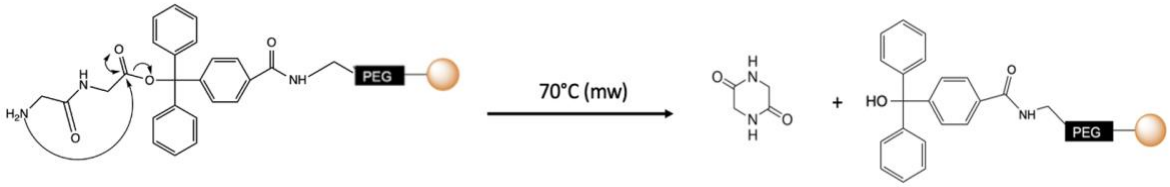


Scheme 11: Scheme of unsuccessful microwave-assisted SPPS of the NEDD8 *C*-terminal peptide LRGG.

To determine if the unsuccessful synthesis was due to temperature, the synthesis was repeated without microwave assistance and the coupling was instead performed at room temperature for 1 hour (Scheme 12). It was confirmed by LC-MS that the synthesis was successful under these conditions. Therefore, the synthesis of the NEDD *C*-terminus was possible at room temperature but not under microwave conditions. It was suspected that diketopiperazine formation of the di-glycine residues was occurring, resulting in low peptide yield. Here, the free amine of the distal glycine, following Fmoc deprotection, attacks the carbonyl group linking the proximal glycine to the resin. This results in cyclisation of the di-glycine peptide and its detachment from the resin (Scheme 13). To confirm this, a di-glycine dipeptide was synthesised on Fmoc-Gly NovaSyn TGT resin using HCTU/DIEA at 70°C under microwave for 5 minutes. Following deprotection of the distal glycine, the piperidine/DMF eluent was collected without prior resin cleavage and analysed by ESI high resolution mass spectrometry. The mass of cyclic di-glycine was observed (expected mass = 113 Da, observed mass = 114 Da), confirming diketopiperazine formation. To further confirm that no peptide was bound to the resin, cleavage with 95% TFA was performed prior to analysis and no peptide mass was observed. Therefore, microwave-assisted SPPS of the di-glycine residues catalyses diketopiperazine formation, which does not occur at room temperature. This is highly unusual as the bulkiness of the NovaSyn TGT resin would be expected to provide steric hindrance to the free amine nucleophile and prevent cyclisation.



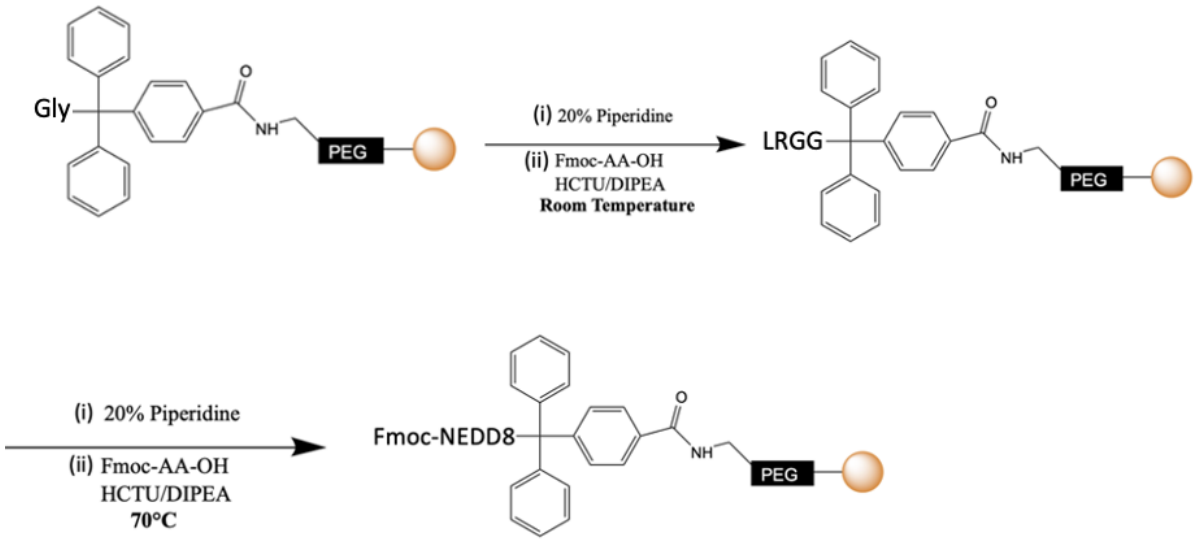
Scheme 12: Scheme of SPPS of NEDD8 *C*-terminus at room temperature



Scheme 13: Mechanism of cyclic di-glycine (diketopiperazine) formation under microwave conditions.

Upon this realisation, SPPS of the full NEDD8 protein was attempted by first synthesising the *C*-terminal region at room temperature before proceeding to elongate the peptide by microwave-assisted synthesis at 70°C. This approach avoids the issue of di-glycine cyclisation whilst also avoiding complete protein synthesis at room temperature, which would be a highly time-consuming process.

The synthesis was performed on a 0.05 mmol scale, again using Fmoc-Gly NovaSyn TGT resin. The glycine, arginine and leucine residues were sequentially coupled onto deprotected glycine-loaded resin using HCTU/DIEA in DMF at room temperature for 1 hour. Following the synthesis of LRGG at room temperature, the peptide was elongated using HCTU/DIEA coupling in DMF at 70°C under microwave for 5 minutes (Scheme 14). Due to the large β-sheet secondary structure of the protein there was a risk of decreased solvation of the growing peptide chain and thus inefficient amino acid coupling later in the synthesis. Therefore, after coupling of 32 residues (NEDD8 45-76), the subsequent 44 residues (NEDD8 1-44) underwent double coupling to increase the probability of successful amino acid coupling.



Scheme 14: Scheme of attempted synthesis of the full NEDD8 sequence. Coupling of the initial 3 residues performed at room temperature to avoid diketopiperazine formation followed by microwave-assisted synthesis at 70°C for elongation of the peptide.

To monitor the synthesis of the protein, a test cleave was performed at intervals during the synthesis whereby a small amount of resin was collected and the peptide cleaved with 95% TFA for analysis by LC-MS and HPLC (Figure 21).

MLIKVKTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGG
3 2 1

Figure 21: NEDD8 sequence with sites of test cleavage highlighted in cyan and numbered chronologically.

The first test cleave was performed after coupling of 24 residues and was deemed successful with the correct mass observed on LC-MS (expected mass = 2819 Da, observed mass = 2820 Da) (Figure 22) and a significant analytical HPLC peak of the crude product was seen.
The second test cleave was carried out after coupling of 51 residues. Whilst the correct mass was observed on LC-MS (expected = 5835 Da, observed = 5835 Da) the analytical HPLC trace only showed a very small product peak, suggesting that only trace amounts of the product was synthesised.

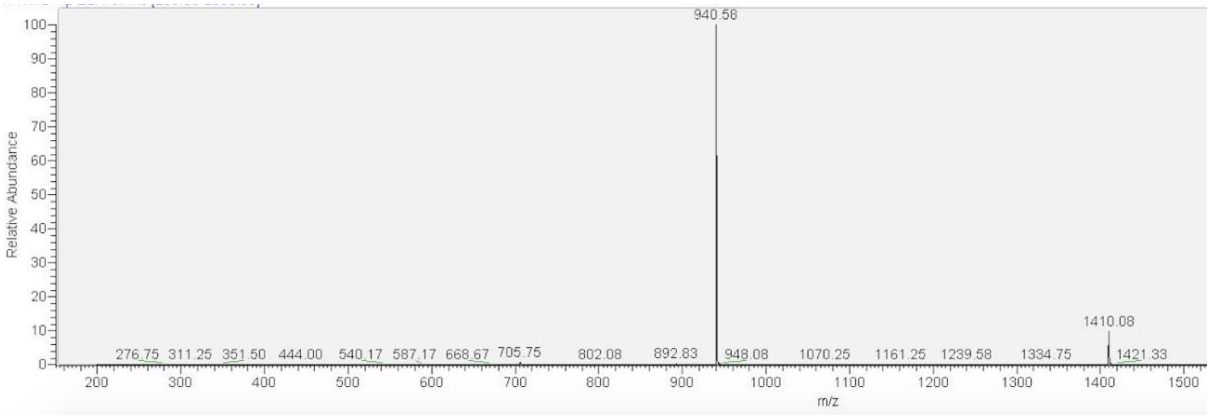


Figure 22: LC-MS of NEDD8 residues 52-76 + Fmoc. m/z expected [M/3] = 940, found 940. m/z expected [M/2] = 1410, found 1410.

The final test cleave was performed upon completion of the protein synthesis. The mass of the desired product was not observed on LC-MS (expected = 8782 Da) and thus the synthesis was unsuccessful. However, a clear peak was observed on the analytical HPLC and, upon analysis on LC-MS, the mass appeared to be of 7224 Da. This mass does not correspond with the product nor any amino acid deletions from the *N*-terminus of the sequence. Therefore, the observed product was unable to be identified. A likely explanation for the failed synthesis is aggregation of the protein upon elongation. Whilst not unexpected given the large β-sheet component, the risk of aggregation should have been minimised with the use of low-loading resin and the insertion of pseudoprolines, as in the reported synthesis of ubiquitin.[149] Potential next steps for synthesis of the protein may consist of increasing the amino acid concentration and coupling time to allow greater chance of coupling, or by conducting the entire synthesis at room temperature. An additional strategy to prevent aggregation is by the incorporation of Hmb or Dmb amino acids.[155] These act similarly to pseudoprolines whereby the backbone amide bond of the residue is protected by a cyclic ring which creates a kink in the sequence and prevents secondary structure formation until cleaved. However, whilst pseudoproline dipeptides can only be placed where there is a serine or threonine residue in the sequence, Hmb/Dmb amino acids can be inserted at any point in the sequence. Therefore incorporation of the amino acids would likely alleviate the risk of protein aggregation.

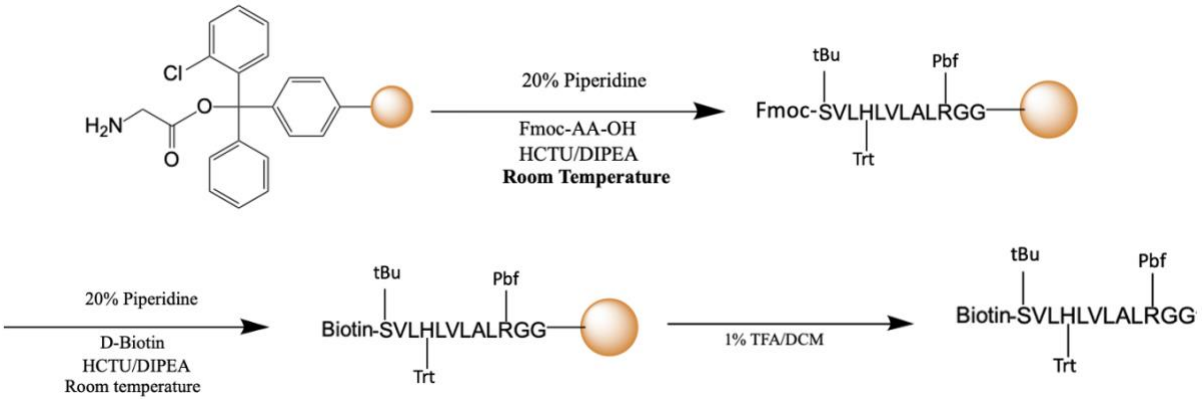
2.2.3 – Solid-Phase Peptide Synthesis of NEDD8 *C*-terminal Peptide

The *C*-terminus of NEDD8 is the site of isopeptidase activity of CSN5, and as such is a major protein-protein interaction site within the complex. Therefore, a *C*-terminal peptide of NEDD8 was deemed an appropriate recognition element in the development of a CSN5 ABP. This retains the ability of the probe warhead to interact directly with the catalytic zinc atom of the CSN5 JAMM site whilst still acting as a substrate of CSN5, thereby enabling efficient probe binding. Previous studies on ubiquitin and Ubl *C*-terminal ABPs with vinyl sulfone warheads have demonstrated that a 12-mer *C*-terminal peptide is sufficient to achieve probe labelling to DUBs and Ubl-specific proteases.[156] This approach was applied here with the synthesis of a 12-mer from the *C*-terminus of NEDD8 (SVLHLVLALRGG).

As the sequence to be synthesised was relatively short, and had little risk of aggregation, a low-loading resin such as NovaSyn TGT was not necessary for this synthesis. Alternatively, a less costly, widely available, high-loading resin was used. The synthesis was carried out on a 0.3 mmol scale using 2-chlorotrityl resin preloaded with glycine at a loading of 0.78 mmol/g. Similarly to NovaSyn TGT resin, this resin is highly acid labile and also produces carboxyl functionalisation of the *C*-terminus. The resin was swollen with DCM and the residues were sequentially coupled on using HCTU and DIEA in DMF at room temperature for 1 hour. Fmoc deprotection was performed using 20% piperidine in DMF.

As biotin is a highly versatile label which allows for a diverse array of probe-labelling assays such as proteomics and immunoprecipitation, it was decided that biotin would be used as the reporter tag in the ABP. Following synthesis of the peptide, the *N*-terminal serine underwent Fmoc deprotection before coupling of D-biotin using HCTU and DIEA in DMF for 1 hour. A test cleave of the biotinylated peptide was performed and analysed by ESI high resolution mass spectrometry (expected mass = 1460 Da, observed = 1461 Da) and analytical HPLC. The HPLC chromatogram showed a large significant peak, therefore the peptide synthesis and subsequent biotinylation was deemed successful.

Since the warhead was to be conjugated onto the *C*-terminus of the peptide, it needed to be ensured that, where available, the residues on the peptide remained protected with Pbf, Trt and *t*Bu groups on serine, histidine and arginine, respectively, so that any modifications occurred only at the *C*-terminus and that these residues did not interfere with the conjugation reactions, such as by cyclisation. Cleavage of the peptide from resin with 95% TFA also deprotects the peptide and removes the protecting groups throughout the chain. Therefore, the acid labile 2-chlorotrityl resin was used in the synthesis as it allows for cleavage of the peptide in low concentrations of TFA (1% TFA in DCM) whilst still ensuring that the protecting groups are not removed (Scheme 15).



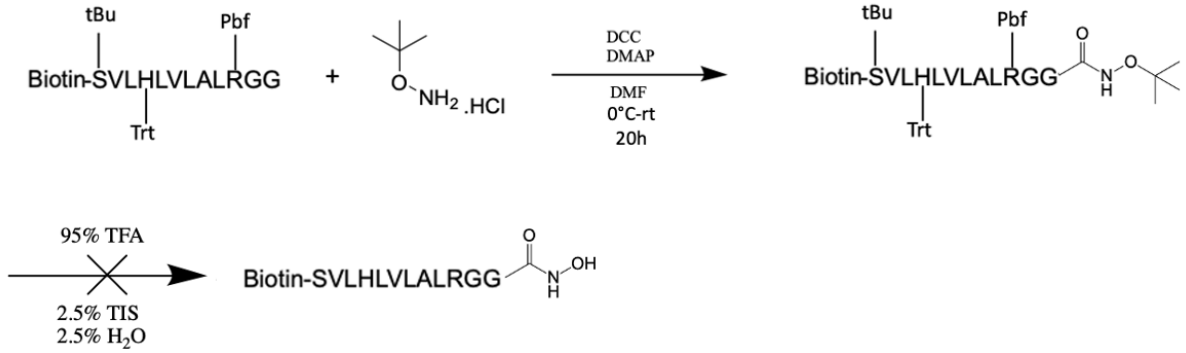
Scheme 15: Synthesis of fully-protected biotinylated NEDD8 *C*-terminal peptide using H-Gly-2-ClTrt resin.

2.2.4 – Conjugation of Hydroxylamine to NEDD8 *C*-terminal Peptide

Having successfully synthesised the biotinylated NEDD8 *C*-terminal 12-mer peptide with protecting groups on arginine, histidine and serine residues, the *C*-terminal glycine residue underwent modification to yield a hydroxamic acid. This is a zinc binding group for use as a warhead in the ABP to target the catalytic JAMM site of CSN5.

The crude protected biotinylated peptide was dissolved in DMF before the coupling of *O*-*tert*-butyl hydroxyl amine hydrochloride using DCC and DMAP. Following the reaction, the peptide was treated with 95% TFA to cleave the protecting groups (Scheme 16). When analysed with LC-MS the mass of the product was not seen (expected mass = 1476 Da, observed = 1461 Da) and the analytical HPLC chromatogram shows a significant peak. The observed mass of 1461

Da correlates to that of the unprotected biotinylated peptide. Therefore the conjugation reaction was unsuccessful and the peptide remained unmodified.



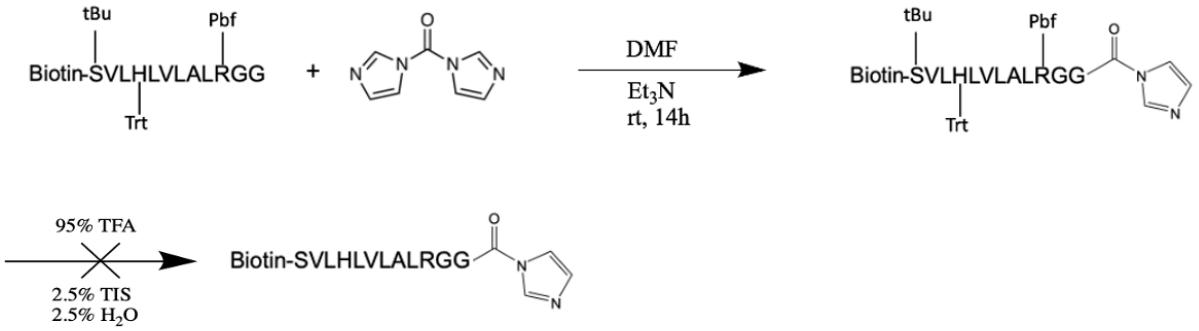
Scheme 16: Conjugation of hydroxyl amine to biotinylated NEDD8 *C*-terminal peptide.

The hydrophobicity of the peptide caused issues in this reaction, as attempting to dissolve the peptide in DMF resulted in a thick jelly-like solid. This created a significant problem as the peptide needed to be fully dissolved in the solvent before any further reactions can occur. A relatively large volume of DMF was required for the peptide to dissolve, resulting in a highly diluted solution. It is therefore likely that the reaction concentration was too low and thus did not proceed. The addition of a greater concentration of *O*-*tert*-butyl hydroxyl amine hydrochloride, DCC and DMAP may improve the reaction and allow it to progress. To improve the solubility of the peptide, the sequence could be elongated to include more polar residues. However, the *C*-terminal end of NEDD8 is largely non-polar and must be extended significantly before generating a soluble peptide. Alternatively, a polar peptide sequence, for example tetra-arginine (RRRR), can be included on the *N*-terminus to improve solubility without the need of synthesising a large NEDD8 peptide.

2.2.5 – Conjugation of Imidazole to NEDD8 *C*-terminal Peptide

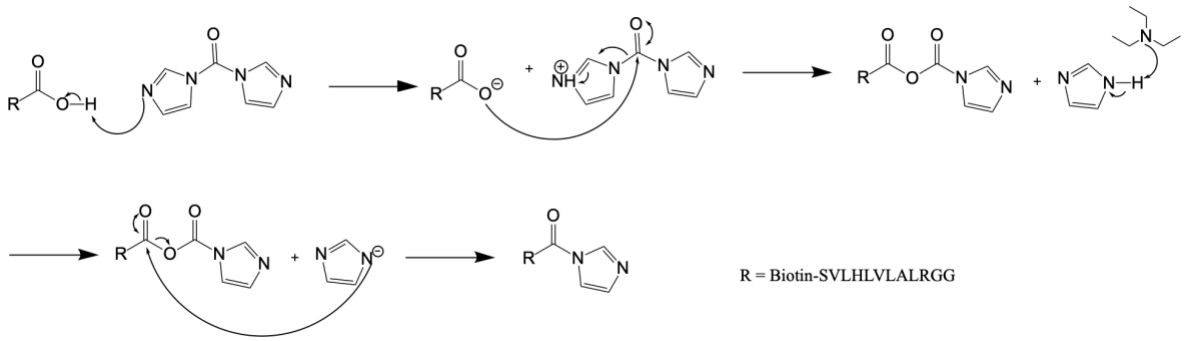
In a parallel approach, the biotinylated NEDD8 12-mer *C*-terminal peptide underwent conjugation of imidazole to the *C*-terminus. Imidazole behaves as the zinc binding group in the inhibitor CSN5i-3 and so was selected as a warhead within the CSN5 ABP. Stability studies of N-acylimidazole have shown that this system remains stable under acidic conditions.[157] Therefore there was not a risk of the imidazole group being removed under the TFA

conditions used for peptide cleavage. The peptide was dissolved in DMF and reacted with carbonyldiimidazole (CDI) and Et_3N at room temperature. Following the reaction, the protecting groups on the peptide were cleaved using 95% TFA (Scheme 17). The product was subsequently analysed by LC-MS, however the correct mass was not observed (expected mass = 1512, observed mass = 1516). A significant peak corresponding to the observed mass was seen on the analytical HPLC chromatogram. With an additional mass of 4 Da it is evident that the desired product was not formed, however it is unclear what the observed product of the reaction is. As in chapter 2.2.4, there was difficulty in dissolving the peptide in DMF as it solidified into a jelly-like substance. A large volume of DMF was required to dissolve the peptide and thus the reaction concentration would have again been lowered. In the first instance, the concentration of the reaction would be increased in an endeavour to drive the desired reaction forward.



Scheme 17: Conjugation of imidazole to the *C*-terminus of the biotinylated NEDD8 12-mer *C*-terminal peptide.

The mechanism of the conjugation reaction is highlighted in Scheme 18. The CDI first deprotonates the *C*-terminal hydroxyl group of the peptide. The oxyanion then attacks the carbonyl group of CDI and releases free imidazole. Et_3N deprotonates the free imidazole which acts as a nucleophile to attack the proximal carbonyl group on the peptide *C*-terminus. The distal carbonyl group is hindered by the attached imidazole and thus nucleophilic addition does not occur there. Upon addition by imidazole, the bond between the carbonyl group and oxygen is broken, yielding the product.



Scheme 18: Mechanism of imidazole conjugation to the biotinylated peptide *C*-terminus using carbonyldiimidazole and Et_3N.

Chapter 3 – Conclusions

In an aim to investigate the biochemical mechanisms which modulate the activity of the COP9 Signalosome (CSN), this project endeavoured to generate an activity-based probe (ABP) specific for the catalytic subunit CSN5. In the first instance, the synthesis of a precursor molecule of the potent small molecule inhibitor CSN5i-3 was attempted. This molecule was to be used as a zinc-binding reactive group attached to the *C*-terminus of NEDD8, which acts as the recognition element in the ABP. By means of molecular docking, it was determined that NEDD8 must be truncated to 73 residues to facilitate docking of the ABP to the CSN5 active site. The synthesis route was derived from the published synthesis of CSN5i-3 by Schlierf *et al.*[115] and was followed closely. Due to time constraints and challenges arising during the synthesis, namely with optimising the organometallic couplings and subsequent purifications, the desired final product was not achieved.

Due to these time limitations, priority was shifted to the generation of ABPs utilising alternative warheads for binding onto the catalytic Zn^{2+} residue of CSN5. These warheads included hydroxamic acid, a widely used zinc-binding group, and imidazole, which acted as the zinc-binding group in CSN5i-3. Two different NEDD8-based recognition elements were proposed in the generation of ABPs with these warheads, the first of which was the full NEDD8 protein. Synthesis of NEDD8 (1-76) was attempted using an Fmoc-based solid-phase approach, however the desired product was not successfully formed. Test cleaves during the synthesis allowed for the successful characterisation of the peptide at 24 residues (Asp52). However, after the synthesis of 51 residues, a second test cleave showed that the product was only produced in trace amounts, as determined by analytical HPLC. Upon completion of the synthesis protocol, the desired product mass was not observed. Modifications in the synthetic approach such as inclusion of DMB dipeptides to prevent protein aggregation may improve the synthesis and should be explored as a next step. The second recognition element synthesised for these ABPs was a 12-mer *C*-terminal peptide of NEDD8 consisting of residues 65-76. This peptide was successfully synthesised by an Fmoc-based solid-phase approach. Biotin was coupled onto the *N*-terminus of the peptide for use as a versatile recognition element for subsequent biological assays.

Conjugation of the imidazole and hydroxamic acid warheads to the NEDD8 12-mer peptide *C*-terminus was attempted, however in both instances the reaction was unsuccessful. A major issue with these reactions was the hydrophobicity of the 12-mer peptide preventing sufficient mixing with the reagents and thus likely hindering the reaction. In the first instance, the solubility of the peptide must be improved before conjugation to these warheads can occur to produce the desired ABPs.

Following the generation of these probes, they can be used within *in vitro* assays to assess labelling to CSN5. If probe labelling proves successful, the ABP can be utilised within biochemical studies to monitor CSN5 activity within various conditions to determine how it is altered. This is a highly powerful approach to elucidate the unknown properties of the enzyme CSN5 and, by association, the full CSN complex and its role within the pathogenesis of human cancer and cardiovascular diseases.

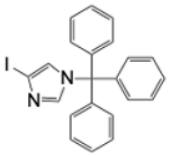
Chapter 4 – Experimental

3.1 – Organic Synthesis

3.1.1 – General Information

All reagents were purchased from commercial sources and used without further purification. Anhydrous THF was obtained by passage through a PureSolv™ solvent purification system and stored with molecular sieves under nitrogen. Reactions were performed using round bottomed flasks on Radleys Pro hotplates. All glassware was oven dried and cooled in a desiccator before use. Microwave-assisted synthesis was performed in a CEM® Discover microwave reactor using a 35 mL reactor vial. All flash chromatography was performed using a Biotage® Isolera™ One automated flash column chromatography system using Biotage® SNAP KP-Sil normal phase pre-packed silica columns. Solvent evaporation was performed under reduced pressure on a Heidolph rotary evaporator. All reactions were monitored using Thin Layer Chromatography (TLC). TLC was also used to determine the optimal solvent system for flash chromatography. This was carried out using aluminium plates precoated with silica gel (0.25 mm, 60 Å pore-size) impregnated with a 254 nm fluorescent indicator. Visualization of TLC was performed using UV light at a wavelength of 254 nm. Proton Nuclear Magnetic Resonance spectroscopy (^{1}H NMR) was performed on an AVANCE III 400 Bruker (400 MHz) system. Proton chemical shifts are expressed in parts per million (ppm, δ scale) and are referenced to the appropriate deuterated solvent signal ($CDCl_3$ δ = 7.26 ppm and DMSO-d_6 δ = 2.50 ppm). Coupling constants (J) are quoted in Hz. Peak patterns are described as singlet (s), doublet (d), triplet (t), quartet (q) and multiplet (m). Carbon Nuclear Magnetic Resonance spectroscopy (^{13}C NMR) was performed on an AVANCE III 400 Bruker (101 MHz) system. Chemical shifts are reported in ppm referenced to residual protium in the deuterated solvent ($CDCl_3$ δ = 77.0 ppm and DMSO-d_6 δ = 39.5 ppm). High-resolution mass spectrometry (HRMS) was performed on a Bruker microTOF-Q II (ESI+).

3.1.2 – 4-iodo-1-trityl-1*H*-imidazole (1)

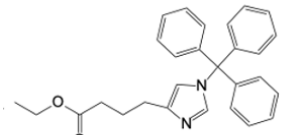


1

To a solution of 4-iodoimidazole (500 mg, 2.6 mmol) in DMF (10 mL) was added trityl chloride (800 mg, 2.9 mmol) and Et_3N (0.429 mL, 3.12 mmol). This was stirred at room temperature for 24 h before being poured into water and filtered. The filtrate was dried *in vacuo* and the crude residue was purified by flash chromatography, eluting at a gradient of 10-50% EtOAc in petroleum ether. The eluent was concentrated *in vacuo* to afford **1** as a white powder at 710 mg (86%).

R_F = 0.5 (EtOAc/petroleum ether 2:9). ^{1}H NMR (400 MHz, CDCl3) δ 7.39 – 7.30 (m, 10H), 7.15 – 7.08 (m, 6H), 6.91 (d, *J* = 1.4 Hz, 1H). ^{13}C NMR (101 MHz, CDCl3) δ 142.0, 140.7, 129.9, 128.5, 128.3, 127.1. HRMS: *m/z* expected $[M+Na]^+$ = 459.0334 found 459.0334. Characterisation data is in alignment with the literature.[158]

3.1.3 – Ethyl 4-(1-trityl-1*H*-imidazol-4-yl)butanoate (2)

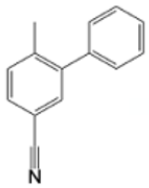


2

To a solution of **1** (1.0 g, 2.3 mmol) in anhydrous THF (5 mL) was added $PdCl_2$(dppf)·DCM under an argon atmosphere and cooled to 0 °C. A solution of 4-ethoxy-4-oxobutylzinc bromide (6.9 mL, 3.45 mmol) was added to the reaction mixture dropwise. The reaction was then heated to 70 °C and stirred for 3 h. The reaction was cooled to room temperature and diluted with EtOAc before filtering through celite. The filtrate was poured into saturated $NaHCO_3$ solution and extracted with EtOAc (3 X). The organic layers were combined, washed with brine and dried by passing through a hydrophobic frit before being concentrated *in vacuo*. The crude residue was purified by flash chromatography, eluting at a gradient of 0-100% EtOAc in hexanes. The eluent was concentrated *in vacuo* to afford a brown oil at 608 mg (42%).

R_F = 0.45 (EtOAc/hexanes 2:9). ^{1}H NMR (400 MHz, DMSO) δ 7.47 – 7.33 (m, 9H), 7.26 (d, *J* = 1.4 Hz, 1H), 7.15 – 7.05 (m, 6H), 6.62 – 6.57 (s, 1H), 4.02 (q, *J* = 7.1 Hz, 2H), 2.43 (t, *J* = 7.4 Hz, 2H), 2.25 (t, *J* = 7.4 Hz, 2H), 1.77 (m, 2H), 1.15 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, DMSO) δ 172.8, 142.4, 140.3, 137.7, 129.2, 128.1, 127.9, 117.6, 74.3, 59.6, 32.8, 26.9, 24.2, 14.1. HRMS: m/z expected $[M+Na]^+$ = 447.2049 found 447.1980. Characterisation data is in alignment with the literature.[115]

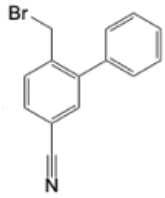
3.1.4 – 6-methyl-[1,1'-biphenyl]-3-carbonitrile (3)



3

3-bromo-4-methylbenzonitrile (1.0 g, 5.1 mmol) was dissolved in DMF/H_2O (1:1, 17 mL). Under a nitrogen atmosphere was added phenylboronic acid (933 mg, 7.65 mmol), Cs_2CO_3 (2.49 g, 7.65 mmol) and $PdCl_2$(dppf)·DCM (416 mg, 0.51 mmol) to the reaction mixture. This was stirred at 90 °C for 14 h. The reaction was cooled to room temperature and diluted with EtOAc before filtering through celite to remove the Pd catalyst. The filtrate was poured into H_2O and extracted with EtOAc (3 X). The organic layers were combined, washed with brine and dried by passing through a hydrophobic frit filter before being concentrated *in vacuo*. The crude residue was purified by flash chromatography, eluting at a gradient of 0-50% EtOAc in hexanes. The eluent was concentrated *in vacuo* to afford a white powder at 860 mg (87%).
R_F = 0.45 (EtOAc/hexanes 1:9). ^{1}H NMR (400 MHz, DMSO) δ 7.71 (dd, *J* = 7.9, 1.8 Hz, 1H), 7.61 (d, *J* = 1.8 Hz, 1H), 7.56 – 7.31 (m, 6H), 2.28 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 142.4, 141.2, 139.0, 132.7, 131.5, 130.8, 128.9, 128.4, 127.7, 118.8, 108.93, 20.4. HRMS: m/z expected $[M+Na]^+$ = 216.0788 found 216.0744. Characterisation data is in alignment with the literature.[115]

3.1.5 – 6-bromomethyl-[1,1'-biphenyl]-3-carbonitrile (4)

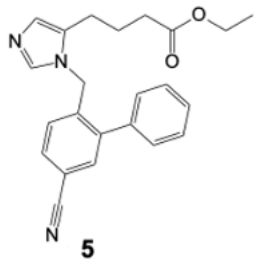


4

To a solution of **3** (800 mg, 4.1 mmol) in CCl_4 (14 mL) was added NBS (873 mg, 4.72 mmol) and AIBN (67.3 mg, 0.41 mmol). This underwent microwave-assisted synthesis at 80 °C, 40W for 2 h with constant stirring. The reaction was cooled to room temperature and poured into H_2O where it was extracted with DCM (3 X). The combined organic layers were washed with brine and passed through a hydrophobic frit to dry before being concentrated *in vacuo*. The crude residue was purified by flash chromatography, eluting at a gradient of 0-50% EtOAc in petroleum ether. The eluent was concentrated *in vacuo* to afford a viscous brown oil at 536 mg (64%).

R_F = 0.58 (EtOAc/hexanes 1:9). ^{1}H NMR (400 MHz, DMSO) δ 7.88 (dd, *J* = 8.0, 1.8 Hz, 1H), 7.80 (d, *J* = 8.1 Hz, 1H), 7.74 (d, *J* = 1.7 Hz, 1H), 7.58 – 7.38 (m, 5H), 4.62 (s, 2H). ^{13}C NMR (101 MHz, DMSO) δ 142.5, 140.6, 137.7, 133.8, 132.1, 131.6, 128.7, 128.6, 128.2, 118.3, 111.4, 31.2. HRMS: m/z expected $[M+Na]^+$ = 295.9895 found 295.9853. Characterisation data is in alignment with the literature.[115]

3.1.6 – Ethyl-4-(1-((5-cyano-[1,1'-biphenyl]-2-yl)methyl)-1H-imidazol-5-yl) butanoate (5)



5

Ethyl 4-(1-trityl-1H-imidazol-4-yl)butanoate (**2**) (89 mg, 0.21 mmol) was dissolved in MeCN (2.2 mL) before the subsequent addition of 6-Bromomethyl-[1,1'-biphenyl]-3-carbonitrile (**4**) (50 mg, 1.8 mmol). The reaction mixture was stirred at room temperature for 15 h. The reaction mixture was concentrated *in vacuo* and the residue was dissolved in MeOH (1.5 mL). The reaction was heated to 70 °C and stirred for 3 h. The reaction mixture was cooled to room temperature and then concentrated *in vacuo* before dissolving in DCM and washing in saturated $NaHCO_3$ solution (3 X). The combined organic layers were washed with brine and passed through a hydrophobic frit filter to dry. The crude residue was concentrated *in vacuo* and purified by flash chromatography, eluting at a gradient of 0-10% MeOH in DCM. The eluent was concentrated to afford a yellow powder at 45 mg (57%).

R_F = 0.6 (MeOH/DCM 0.5:9.5). ^{1}H NMR (400 MHz, DMSO) δ 7.82 (dd, *J* = 8.1, 1.8 Hz, 1H), 7.76 (d, *J* = 1.8 Hz, 1H), 7.56 – 7.38 (m, 5H), 7.08 – 6.59 (m, 2H), 5.15 (s, 2H), 3.98 (q, *J* = 7.1 Hz, 2H), 2.16 (dt, *J* = 12.1, 7.5 Hz, 4H), 1.60 – 1.48 (m, 2H), 1.12 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, DMSO) δ 172.4, 141.5, 140.4, 137.6, 133.4, 131.7, 128.9, 128.7, 128.4, 128.3, 128.2, 127.7, 126.1, 118.4, 110.6, 59.7, 45.4, 32.7, 23.0, 22.3, 14.1. HRMS: m/z expected $[M+H]^+$ = 374.1790 found 374.1817. Characterisation data is in alignment with the literature.[115]

4.2 – Molecular Docking

Peptide models were generated using Schrödinger Bioluminate.[159–161] 3D models of peptides conjugated to CSN5i-3 were created within UCSF Chimera.[162] All other 3D ligand models were generated in Schrödinger Maestro.[159–161] All ligands were energy minimised within Schrödinger Maestro using the LigPrep tool. The active site model was achieved by importing the structure of CSN5 bound to the ligand CSNi-3 (PDB: 5JOG) to Schrödinger Maestro and removing the ligand to produce an unbound model of CSN5 in its active form. The protein was further prepared by adding hydrogens and removing waters using the Protein Preparation wizard in Schrödinger Maestro. Molecular docking of the ligands to active-form CSN5 was performed in Schrödinger Glide using a constraint for the interaction with the active site Zn^{2+} atom. Following docking, the resulting poses were scored using Glidescore[2] which ranks the poses in order of binding affinity and the top scoring poses were selected for visualisation. 3D figures of the poses were generated in UCSF Chimera. 2D ligand interaction analysis diagrams were produced in Schrödinger Maestro.

4.3 – Peptide Synthesis

4.3.1 – General Information

All Fmoc-protected amino acids were purchased from either CEM Corporation or Pepceuticals Ltd. Fmoc-pseudoproline dipeptides, Fmoc-Gly NovaSyn TGT resin and H-Gly 2-chlorotrityl resin were purchased from Sigma-Aldrich. Peptide-grade DMF was purchased from Cambridge Reagents Limited and HCTU from Pepceuticals Ltd. All other reagents were purchased from Sigma-Aldrich and used without any further purification. All glassware was dried at 110°C for at least 14 hours and cooled in a desiccator before use. Unless otherwise stated, all DMF used in this section is of peptide grade.

All peptide synthesis was performed on a Biotage® Initiator+ Alstra™ automated microwave peptide synthesiser using Biotage® 10mL or 30mL reactor vials fitted with polytetrafluoroethylene (PTFE) frit filters. Analytical HPLC of peptides was performed on a Shimadzu reverse-phase HPLC with UV-vis detection (monitoring at 214 nm and 280 nm) using

a Gemini™ C18 (5µm, 110Å, 150 mm × 4.6 mm) column at a flow rate of 1 mL/min. A solvent system of H_2O + 0.1% TFA (buffer A) and 95% acetonitrile + 0.1% TFA (buffer B) was utilised for analytical HPLC at a linear gradient of 0-100% buffer B over 50 minutes. Column retention time (Rt) is reported in minutes. Liquid chromatography mass spectrometry (LC-MS) was carried out using a Dionex Ultimate 3000 LC fitted to a Thermo Scientific LCQ Fleet quadrupole mass spectrometer using electrospray ionisation in positive mode (ESI+) with a C18 column (2.1 x 50 mm). A solvent system of H_2O + 0.1% TFA (buffer A) and 95% acetonitrile + 0.1% TFA (buffer B) was utilised with a linear gradient of 5-100% buffer B over 20 minutes at a flow rate of 0.6 mL/min. High resolution mass spectrometry was performed on a Bruker microTOF-Q II using ESI+. Solvent evaporation was performed under reduced pressure on a Heidolph rotary evaporator. Peptide lyophilisation was carried out using a Christ Alpha 2-4 LDplus freeze dryer.

Resin swelling was performed in DCM for 20 mins at room temperature. Room temperature coupling of standard Fmoc-protected amino acids was carried out for 60 mins in an automated peptide synthesiser. Microwave-assisted coupling of standard amino acids was performed at 70°C for 10 mins. Double coupling refers to the coupling of a residue proceeded by DMF washing 4 times before repeating the coupling procedure with fresh reagents. All arginine residues were double-coupled at room temperature for 60 mins and then at 70°C for a further 5 mins. All histidine residues were coupled at room temperature for 5 mins and subsequently at 50°C for a further 5 mins. HCTU was dissolved in DMF at a 0.5 M concentration and DIEA was dissolved in NMP at either a 1 M or 2 M concentration. Fmoc deprotection was performed using 20% piperidine in DMF + 5% formic acid at room temperature for 5 mins followed by an additional 3 mins at 70°C.

Test cleavage was carried out using a cleavage cocktail of 95% TFA, 2.5% TIS and 2.5% H_2O. 1 mL of the cleavage cocktail was added to a small sample of resin and shaken for 1 hour at room temperature. The solvent was filtered through a PTFE frit filter, collected into a Falcon tube and the TFA was evaporated using a stream of nitrogen gas. The peptide was subsequently precipitated in ice-cold Et_2O and centrifuged at 3500 x g for 5 mins. The supernatant was discarded and the pellet was washed with Et_2O three times. The pellet was resuspended in a mixture of acetonitrile and water before subsequent analysis.

4.3.2 – Attempted NEDD8 Synthesis

Synthesis of the full NEDD8 protein was performed on Fmoc-Gly NovaSyn TGT resin at a 0.05 mmol scale using a Biotage® 10 mL reactor vial. All amino acids were dissolved in DMF at a 0.12 M concentration. Coupling of standard Fmoc-protected amino acids was carried out using 5 Eq and pseudoproline dipeptide coupling using 2.5 Eq. 5 Eq HCTU and 10 Eq 1 M DIEA were used for the coupling. Due to the relatively small scale of the reaction, an excess of reagents were used to increase the reaction volume to accommodate the 10 mL reactor vial. Coupling of the first 3 residues (LRG) was performed at room temperature and all subsequent couplings were carried out with microwave assistance, as described in Chapter 4.3.1. After coupling of 32 residues, all subsequent residues were double coupled. Fmoc deprotection was performed using 20% piperidine in DMF + 5% formic acid, as described in Chapter 4.3.1.

Test cleavage was performed using a cleavage cocktail of 95% TFA, 2.5% TIS and 2.5% H_2O, as described in Chapter 4.3.1.

4.3.3 – Diketopiperazine Test

Synthesis of the diglycine dipeptide was performed on Fmoc-Gly NovaSyn TGT resin at a 0.05 mmol scale using a Biotage® 10 mL reactor vial. The Fmoc-protected glycine was dissolved in DMF at a 0.12 M concentration. Coupling and deprotection was performed as described in Chapter 4.3.2. After deprotection of the *N*-terminal glycine, a sample of the 20% piperidine/DMF solution was collected and filtered through a PTFE frit filter before analysis by HRMS.

4.3.4 – NEDD8 *C*-terminal 12-mer Peptide Synthesis

The 12-mer peptide was synthesised using H-Gly 2-chlorotrityl resin at a 0.3 mmol scale using a Biotage® 30 mL reactor vial. All amino acids were dissolved in DMF at a 0.2 M concentration and coupled using 4 Eq. 4 Eq HCTU and 8 Eq 2M DIEA were used for coupling. All residues were coupled at room temperature or as specified in Chapter 4.3.1. Fmoc deprotection was

performed using 20% piperidine in DMF + 5% formic acid, as described in Chapter 4.3.1. D-Biotin was dissolved in DMSO at a 0.2 M concentration and 4 Eq was coupled onto the *N*-terminus of the peptide using 4 Eq HCTU and 8 Eq 2M DIEA. The biotin coupling was performed at room temperature for 60 minutes. Upon completion of the synthesis, the resin-bound peptide was washed with DMF 4 times for 45s and then washed again with DCM 4 times for 45s.

Test cleavage was carried out using a cleavage cocktail of 95% TFA, 2.5% TIS and 2.5% H_2O, as described in Chapter 4.3.1. Full cleavage of the peptide from resin whilst ensuring it remains fully protected was performed using 10mL 1% TFA in DCM. This was stirred for 3 hours before the solvent was filtered through a PTFE frit filter and collected in a round-bottom flask. The peptide was concentrated by rotary evaporation and dissolved in a solution of acetic acid, acetonitrile and H_2O. The peptide was subsequently lyophilised to yield a light-yellow powder.

4.3.5 – Attempted Peptide Conjugation to Hydroxylamine

NEDD8 *C*-terminal 12-mer peptide (50 mg, 0.35 mmol) was dissolved in 5 mL DMF and cooled to 0°C on ice. *O*-*tert*-butyl hydroxyl amine hydrochloride (67 mg, 0.53 mmol, 1.5 equiv) was added to the solution under nitrogen atmosphere followed by DMAP (9 mg, 0.07 mmol) and was stirred for 5 mins at 0°C. DCC (110 mg, 0.54 mmol) was added to the mixture and allowed to stir at room temperature for 20h. The solvent was removed *in vacuo* by rotary evaporation and the crude product was dissolved in acetonitrile/H_2O before analysis by LC-MS and analytical HPLC.

4.3.6 – Attempted Peptide Conjugation to Imidazole

To a solution of NEDD8 *C*-terminal 12-mer peptide (30 mg, 0.21 mmol) in 4 mL DMF was added *N,N'*- Carbonyldiimidazole (44 mg, 0.27 mmol). The mixture was stirred at room temperature for 3h before the addition of Et_3N (0.038 mL, 0.27 mmol). The mixture was subsequently stirred for a further 14h at room temperature. The solvent was removed *in vacuo* by rotary evaporation and the crude product was dissolved in acetonitrile/H_2O before analysis by LC-MS and analytical HPLC.

References

1 O. Kerscher, R. Felberbaum and M. Hochstrasser, *Annu. Rev. Cell Dev. Biol.*, 2006, 22, 159–180.
2 W. Li and Y. Ye, *Cell. Mol. Life Sci.*, 2008, 65, 2397–2406.
3 M. Akutsu, I. Dikic and A. Bremm, *J. Cell Sci.*, 2016, **129**, 875–880.
4 D. Komander and M. Rape, *Annu. Rev. Biochem.*, 2012, **81**, 203–229.
5 M. Hrdinka and M. Gyrd-Hansen, *Mol. Cell*, 2017, 68, 265–280.
6 M. J. Clague, C. Heride and S. Urbé, *Trends Cell Biol.*, 2015, 25, 417–426.
7 M. J. Clague, I. Barsukov, J. M. Coulson, H. Liu, D. J. Rigden and S. Urbé, *Physiol. Rev.*, 2013, **93**, 1289–1315.
8 S. Vijay-kumar, C. E. Bugg and W. J. Cook, *J. Mol. Biol.*, 1987, **194**, 531–544.
9 C. Grabbe and I. Dikic, *Chem. Rev.*, 2009, **109**, 1481–1494.
10 L. Randles and K. J. Walters, *Front. Biosci.*, 2012, **17**, 2140–2157.
11 B. T. Seet, I. Dikic, M. M. Zhou and T. Pawson, *Nat. Rev. Mol. Cell Biol.*, 2006, 7, 473–483.
12 J. M. Winget and T. Mayor, *Mol. Cell*, 2010, 38, 627–635.
13 P. Young, Q. Deveraux, R. E. Beal, C. M. Pickart and M. Rechsteiner, *J. Biol. Chem.*, 1998, **273**, 5461–5467.
14 H. B. Kamadurai, J. Souphron, D. C. Scott, D. M. Duda, D. J. Miller, D. Stringer, R. C. Piper and B. A. Schulman, *Mol. Cell*, 2009, **36**, 1095–1102.
15 J. A. Ronau, J. F. Beckmann and M. Hochstrasser, *Cell Res.*, 2016, 26, 441–456.
16 S. Lee, Y. C. Tsai, R. Mattera, W. J. Smith, M. S. Kostelansky, A. M. Weissman, J. S. Bonifacino and J. H. Hurley, *Nat. Struct. Mol. Biol.*, 2006, **13**, 264–271.
17 C. Raiborg, T. Slagsvold and H. Stenmark, *Trends Biochem. Sci.*, 2006, **31**, 541–544.
18 T. Tenno, K. Fujiwara, H. Tochio, K. Iwai, E. H. Morita, H. Hayashi, S. Murata, H. Hiroaki, M. Sato, K. Tanaka and M. Shirakawa, *Genes to Cells*, 2004, **9**, 865–875.
19 Y. Sato, A. Yoshikawa, H. Mimura, M. Yamashita, A. Yamagata and S. Fukai, *EMBO J.*, 2009, **28**, 2461–2468.
20 L. Cappadocia and C. D. Lima, *Chem. Rev.*, 2018, 118, 889–918.
21 S. Jentsch and G. Pyrowolakis, *Trends Cell Biol.*, 2000, 10, 335–342.
22 B. M. Foster, P. Stolz, C. B. Mulholland, A. Montoya, H. Kramer, S. Bultmann and T. Bartke, *Mol. Cell*, 2018, **72**, 739-752.e9.
23 J. V. Dzimianski, F. E. M. Scholte, É. Bergeron and S. D. Pegan, *J. Mol. Biol.*, 2019, 431, 4203–4216.
24 P. J. Farrell, R. J. Broeze and P. Lengyel, *Nature*, 1979, 279, 523–525.
25 D. S. Kessler, D. E. Levy and J. E. Darnell, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 8521–8525.
26 D. Zhang and D. E. Zhang, *J. Interf. Cytokine Res.*, 2011, 31, 119–130.
27 L. A. Durfee, N. Lyon, K. Seo and J. M. Huibregtse, *Mol. Cell*, 2010, **38**, 722–732.
28 M.-J. Kim, S.-Y. Hwang, T. Imaizumi and J.-Y. Yoo, *J. Virol.*, 2008, **82**, 1474–1483.
29 H.-X. Shi, K. Yang, X. Liu, X.-Y. Liu, B. Wei, Y.-F. Shan, L.-H. Zhu and C. Wang, *Mol. Cell. Biol.*, 2010, **30**, 2424–2436.
30 W. Yuan and R. M. Krug, *EMBO J.*, 2001, **20**, 362–371.
31 S. Li, J.-Y. Min, R. M. Krug and G. C. Sen, *Virology*, 2006, **349**, 13–21.
32 H. A. Lindner, N. Fotouhi-Ardakani, V. Lytvyn, P. Lachance, T. Sulea and R. Ménard, *J.*

Virol., 2005, **79**, 15199–15208.
33 D. Shin, R. Mukherjee, D. Grewe, D. Bojkova, K. Baek, A. Bhattacharya, L. Schulz, M. Widera, A. R. Mehdipour, G. Tascher, P. P. Geurink, A. Wilhelm, G. J. van der Heden van Noort, H. Ovaa, S. Müller, K. P. Knobeloch, K. Rajalingam, B. A. Schulman, J. Cinatl, G. Hummer, S. Ciesek and I. Dikic, *Nature*, , DOI:10.1038/s41586-020-2601-5.
34 K. A. Wilkinson and J. M. Henley, *Biochem. J.*, 2010, 428, 133–145.
35 Y. Yang, Y. He, X. Wang, Z. Liang, G. He, P. Zhang, H. Zhu, N. Xu and S. Liang, *Open Biol.*, 2017, 7.
36 G. Gill, *Genes Dev.*, 2004, 18, 2046–2059.
37 R. Geiss-Friedlander and F. Melchior, *Nat. Rev. Mol. Cell Biol.*, 2007, 8, 947–956.
38 C. Guo and J. M. Henley, *IUBMB Life*, 2014, 66, 71–77.
39 L. Zhou, Y. Jiang, Q. Luo, L. Li and L. Jia, *Mol. Cancer*, 2019, 18, 77.
40 T. Kamitani, K. Kito, H. P. Nguyen and E. T. H. Yeh, *J. Biol. Chem.*, 1997, **272**, 28557–28562.
41 T. A. Soucy, L. R. Dick, P. G. Smith, M. A. Milhollen and J. E. Brownell, *Genes and Cancer*, 2010, **1**, 708–716.
42 T. Gan-Erdene, K. Nagamalleswari, L. Yin, K. Wu, Z. Q. Pan and K. D. Wilkinson, *J. Biol. Chem.*, 2003, **278**, 28892–28900.
43 E. Santonico, in *Ubiquitin Proteasome System - Current Insights into Mechanism Cellular Regulation and Disease*, IntechOpen, 2019.
44 B. A. Schulman and J. Wade Harper, *Nat. Rev. Mol. Cell Biol.*, 2009, 10, 319–331.
45 R. Hjerpe, Y. Thomas, J. Chen, A. Zemla, S. Curran, N. Shpiro, L. R. Dick and T. Kurz, *Biochem. J.*, 2012, **441**, 927–936.
46 T. Kurz, Y. C. Chou, A. R. Willems, N. Meyer-Schaller, M. L. Hecht, M. Tyers, M. Peter and F. Sicheri, *Mol. Cell*, 2008, **29**, 23–35.
47 D. Girdwood, D. P. Xirodimas and C. Gordon, , DOI:10.1371/journal.pone.0020089.
48 F. G. Whitby, G. Xia, C. M. Pickart and C. P. Hill, *J. Biol. Chem.*, 1998, **273**, 34983–34991.
49 A. G. Van Der Veen and H. L. Ploegh, *Annu. Rev. Biochem.*, 2012, **81**, 323–357.
50 S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Res.*, 1997, 25, 3389–3402.
51 S. F. Altschul, J. C. Wootton, E. M. Gertz, R. Agarwala, A. Morgulis, A. A. Schäffer and Y. K. Yu, *FEBS J.*, 2005, 272, 5101–5109.
52 H. Walden, M. S. Podgorski, D. T. Huang, D. W. Miller, R. J. Howard, D. L. Minor, J. M. Holton and B. A. Schulman, *Mol. Cell*, 2003, **12**, 1427–1437.
53 O. Leidecker, I. Matic, B. Mahata, E. Pion and D. P. Xirodimas, *Cell Cycle*, 2012, **11**, 1142–1150.
54 L. Castagnoli, W. Mandaliti, R. Nepravishta, E. Valentini, A. Mattioni, R. Procopio, M. Iannuccelli, S. Polo, M. Paci, G. Cesareni and E. Santonico, *FEBS J.*, 2019, **286**, 653–677.
55 R. K. Singh, S. Zerath, O. Kleifeld, M. Scheffner, M. H. Glickman and D. Fushman, *Mol. Cell. Proteomics*, 2012, **11**, 1595–1611.
56 S. Oved, Y. Mosesson, Y. Zwang, E. Santonico, K. Shtiegman, M. D. Marmor, B. S. Kochupurakkal, M. Katz, S. Lavi, G. Cesareni and Y. Yarden, *J. Biol. Chem.*, 2006, **281**, 21640–21651.
57 A. J. George, Y. C. Hoffiz, A. J. Charles, Y. Zhu and A. M. Mabb, *Front. Genet.*, 2018, 9, 29.

58 D. Rotin and S. Kumar, *Nat. Rev. Mol. Cell Biol.*, 2009, 10, 398–409.
59 N. Zheng and N. Shabek, *Annu. Rev. Biochem.*, 2017, 86, 129–157.
60 L. Buetow and D. T. Huang, *Nat. Rev. Mol. Cell Biol.*, 2016, 17, 626–642.
61 M. A. Verdecia, C. A. P. Joazeiro, N. J. Wells, J. L. Ferrer, M. E. Bowman, T. Hunter and J. P. Noel, *Mol. Cell*, 2003, **11**, 249–259.
62 C. Chasapis and G. Spyroulias, *Curr. Pharm. Des.*, 2009, **15**, 3716–3731.
63 R. J. Deshaies and C. A. P. Joazeiro, *Annu. Rev. Biochem.*, 2009, **78**, 399–434.
64 H. C. Nguyen, W. Wang and Y. Xiong, *Subcell. Biochem.*, 2017, **83**, 323–347.
65 Y. Zhao and Y. Sun, *Curr. Pharm. Des.*, 2013, **19**, 3215–3225.
66 A. Sarikas, T. Hartmann and Z. Q. Pan, *Genome Biol.*, 2011, 12, 1–12.
67 N. Watanabe, H. Arai, Y. Nishihara, M. Taniguchi, N. Watanabe, T. Hunter and H. Osada, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 4419–4424.
68 T. Hori, F. Osaka, T. Chiba, C. Miyamoto, K. Okabayashi, N. Shimbara, S. Kato and K. Tanaka, *Oncogene*, 1999, **18**, 6829–6834.
69 D. M. Duda, L. A. Borg, D. C. Scott, H. W. Hunt, M. Hammel and B. A. Schulman, *Cell*, 2008, **134**, 995–1006.
70 S. Lyapina, G. Cope, A. Shevchenko, G. Serino, T. Tsuge, C. Zhou, D. A. Wolf, N. Wei, A. Shevchenko and R. J. Deshaies, *Science (80-.).*, 2001, **292**, 1382–1385.
71 D. Komander, M. J. Clague and S. Urbé, *Nat. Rev. Mol. Cell Biol.*, 2009, 10, 550–563.
72 S. A. H. de Poot, G. Tian and D. Finley, *J. Mol. Biol.*, 2017, 429, 3525–3545.
73 A. Y. Amerik, S. Swaminathan, B. A. Krantz, K. D. Wilkinson and M. Hochstrasser, *EMBO J.*, 1997, **16**, 4826–4838.
74 D. Kwasna, S. A. Abdul Rehman, J. Natarajan, S. Matthews, R. Madden, V. De Cesare, S. Weidlich, S. Virdee, I. Ahel, I. Gibbs-Seymour and Y. Kulathu, *Mol. Cell*, 2018, **70**, 150-164.e6.
75 T. E. T. Mevissen and D. Komander, *Annu. Rev. Biochem.*, 2017, **86**, 159–192.
76 Y. Kulathu, *Cell Chem. Biol.*, 2016, 23, 432–434.
77 T. E. T. Mevissen, M. K. Hospenthal, P. P. Geurink, P. R. Elliott, M. Akutsu, N. Arnaudo, R. Ekkebus, Y. Kulathu, T. Wauer, F. El Oualid, S. M. V. Freund, H. Ovaa and D. Komander, *Cell*, 2013, **154**, 169.
78 J. McCullough, P. E. Row, Ó. Lorenzo, M. Doherty, R. Beynon, M. J. Clague and S. Urbé, *Curr. Biol.*, 2006, **16**, 160–165.
79 C. W. Davies, L. N. Paul and C. Das, *Biochemistry*, 2013, **52**, 7818–7829.
80 N. Wei, G. Serino and X. W. Deng, *Trends Biochem. Sci.*, 2008, 33, 592–600.
81 G. M. Lingaraju, R. D. Bunker, S. Cavadini, D. Hess, U. Hassiepen, M. Renatus, E. S. Fischer and N. H. Thomä, *Nature*, 2014, **512**, 161–165.
82 C. Gutierrez, I. E. Chemmama, H. Mao, C. Yu, I. Echeverria, S. A. Block, S. D. Rychnovsky, N. Zheng, A. Sali and L. Huang, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 4088–4098.
83 E. S. Fischer, K. Böhm, J. R. Lydeard, H. Yang, M. B. Stadler, S. Cavadini, J. Nagel, F. Serluca, V. Acker, G. M. Lingaraju, R. B. Tichkule, M. Schebesta, W. C. Forrester, M. Schirle, U. Hassiepen, J. Ottl, M. Hild, R. E. J. Beckwith, J. W. Harper, J. L. Jenkins and N. H. Thomä, *Nature*, 2014, **512**, 49–53.
84 N. Wei and X. W. Deng, in *Annual Review of Cell and Developmental Biology*, Annu Rev Cell Dev Biol, 2003, vol. 19, pp. 261–286.
85 R. Mosadeghi, K. M. Reichermeier, M. Winkler, A. Schreiber, J. M. Reitsma, Y. Zhang, F. Stengel, J. Cao, M. Kim, M. J. Sweredoski, S. Hess, A. Leitner, R. Aebersold, M.

Peter, R. J. Deshaies and R. I. Enchev, *Elife*, , DOI:10.7554/eLife.12102.
86 Y. Sato, A. Yoshikawa, A. Yamagata, H. Mimura, M. Yamashita, K. Ookata, O. Nureki, K. Iwai, M. Komada and S. Fukai, *Nature*, 2008, **455**, 358–362.
87 V. Kumar, M. Naumann and M. Stein, *Front. Chem.*, 2018, **6**, 480.
88 R. I. Enchev, D. C. Scott, P. C. A. da Fonseca, A. Schreiber, J. K. Monda, B. A. Schulman, M. Peter and E. P. Morris, *Cell Rep.*, 2012, **2**, 616–627.
89 S. V. Faull, A. M. C. Lau, C. Martens, Z. Ahdash, K. Hansen, H. Yebenes, C. Schmidt, F. Beuron, N. B. Cronin, E. P. Morris and A. Politis, *Nat. Commun.*, 2019, **10**, 1–13.
90 J. Milic, Y. Tian and J. Bernhagen, *Biomolecules*, 2019, 9, 217.
91 R. I. Enchev, B. A. Schulman and M. Peter, *Nat. Rev. Mol. Cell Biol.*, 2015, 16, 30–44.
92 M. Tan, Y. Zhao, S. J. Kim, M. Liu, L. Jia, T. L. Saunders, Y. Zhu and Y. Sun, *Dev. Cell*, 2011, **21**, 1062–1076.
93 J. R. Skaar, V. D'Angiolella, J. K. Pagan and M. Pagano, *Cell*, 2009, 137, 1358.e1-1358.e2.
94 L. Jia and Y. Sun, *Curr. Cancer Drug Targets*, 2011, **11**, 347–356.
95 T. Abbas, E. Shibata, J. Park, S. Jha, N. Karnani and A. Dutta, *Mol. Cell*, 2010, **40**, 9–21.
96 H. Nakade, K. Migita, S. Matsumoto, K. Wakatsuki, T. Kunishige, S. Miyao and M. Sho, *Int. J. Clin. Oncol.*, 2020, **25**, 446–455.
97 S. Uhle, O. Medalia, R. Waldron, R. Dumdey, P. Henklein, D. Bech-Otschir, X. Huang, M. Berse, J. Sperling, R. Schade and W. Dubiel, *EMBO J.*, 2003, **22**, 1302–1312.
98 K. Schweitzer, P. M. Bozko, W. Dubiel and M. Naumann, *EMBO J.*, 2007, **26**, 1532–1541.
99 J. Zhang, R. Zhao, C. Yu, C. L. N. Bryant, K. Wu, Z. Liu, Y. Ding, Y. Zhao, B. Xue, Z. Q. Pan, C. Li, L. Huang and L. Fang, *J. Proteome Res.*, 2020, **19**, 1119–1130.
100 L. Orel, H. Neumeier, K. Hochrainer, B. R. Binder and J. A. Schmid, *J. Cell. Mol. Med.*, 2010, **14**, 1555–1568.
101 M. H. Lee, R. Zhao, L. Phan and S. C. J. Yeung, *Cell Cycle*, 2011, 10, 3057–3066.
102 F. J. Esteva, A. A. Sahin, G. Z. Rassidakis, L. X. H. Yuan, T. L. Smith, Y. Yang, M. Z. Gilcrease, M. Cristofanilli, R. Nahta, L. Pusztai and F.-X. Claret, *Jun Activation Domain Binding Protein 1 Expression Is Associated with Low p27 Kip1 Levels in Node-Negative Breast Cancer*, 2003.
103 K. Tomoda, Y. Kubota and J. Y. Kato, *Nature*, 1999, **398**, 160–165.
104 A. G. Georgakilas, O. A. Martin and W. M. Bonner, *Trends Mol. Med.*, 2017, 23, 310–319.
105 G. Lozano and R. Montes De Oca Luna, *Biochim. Biophys. Acta - Rev. Cancer*, 1998, 1377.
106 K. Karakostis and R. Fåhraeus, *BMC Cancer*, 2019, 19, 915.
107 B. J. Aubrey, G. L. Kelly, A. Janic, M. J. Herold and A. Strasser, *Cell Death Differ.*, 2018, 25, 104–113.
108 S. Y. Shieh, M. Ikeda, Y. Taya and C. Prives, *Cell*, 1997, **91**, 325–334.
109 Q. Cheng and J. Chen, *Cell Cycle*, 2010, 9, 472–478.
110 K. Tomoda, N. Yoneda-Kato, A. Fukumoto, S. Yamanaka and J. Y. Kato, *J. Biol. Chem.*, 2004, **279**, 43013–43018.
111 R. Mazurek, J. M. Dave, R. R. Chandran, A. Misra, A. Q. Sheikh and D. M. Greif, in *Advances in Pharmacology*, Academic Press Inc., 2017, vol. 78, pp. 323–350.
112 J. Jin, T. Shirogane, L. Xu, G. Nalepa, J. Qin, S. J. Elledge and J. W. Harper, *Genes Dev.*, 2003, **17**, 3062–3074.

113 T. Shen and S. Huang, *Anticancer. Agents Med. Chem.*, 2012, **12**, 631–639.
114 A. S. Adler, M. Lin, H. Horlings, D. S. A. Nuyten, M. J. Van De Vijver and H. Y. Chang, *Nat. Genet.*, 2006, **38**, 421–430.
115 A. Schlierf, E. Altmann, J. Quancard, A. B. Jefferson, R. Assenberg, M. Renatus, M. Jones, U. Hassiepen, M. Schaefer, M. Kiffe, A. Weiss, C. Wiesmann, R. Sedrani, J. Eder and B. Martoglio, *Nat. Commun.*, , DOI:10.1038/ncomms13166.
116 J. P. C. Le Quesne, K. A. Spriggs, M. Bushell and A. E. Willis, *J. Pathol.*, 2010, 220, 140–151.
117 L. E. Sanman and M. Bogyo, *Annu. Rev. Biochem.*, 2014, **83**, 249–273.
118 A. Saghatelian, N. Jessani, A. Joseph, M. Humphrey and B. F. Cravatt, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 10000–10005.
119 P. Yang and K. Liu, *ChemBioChem*, 2015, 16, 712–724.
120 E. W. S. Chan, S. Chattopadhaya, R. C. Panicker, X. Huang and S. Q. Yao, *J. Am. Chem. Soc.*, 2004, **126**, 14435–14446.
121 N. J. Yang and M. J. Hinner, *Methods Mol. Biol.*, 2015, 1266, 29–53.
122 X. Antoniou and T. Borsello, *Pharmaceuticals*, 2010, 3, 379–392.
123 F. Fan, S. Nie, E. B. Dammer, D. M. Duong, D. Pan, L. Ping, L. Zhai, J. Wu, X. Hong, L. Qin, P. Xu and Y. H. Zhang, *J. Proteome Res.*, 2012, **11**, 5763–5772.
124 S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb and R. Aebersold, *Nat. Biotechnol.*, 1999, **17**, 994–999.
125 S. Wang, Y. Tian, M. Wang, M. Wang, G. Sun and X. Sun, *Front. Pharmacol.*, 2018, **9**, 353.
126 C. E. Chivers, A. L. Koner, E. D. Lowe and M. Howarth, *Biochem. J.*, 2011, **435**, 55–63.
127 D. Lahav, B. Liu, R. J. B. H. N. Van Den Berg, A. M. C. H. Van Den Nieuwendijk, T. Wennekes, A. T. Ghisaidoobe, I. Breen, M. J. Ferraz, C. L. Kuo, L. Wu, P. P. Geurink, H. Ovaa, G. A. Van Der Marel, M. Van Der Stelt, R. G. Boot, G. J. Davies, J. M. F. G. Aerts and H. S. Overkleeft, *J. Am. Chem. Soc.*, 2017, **139**, 14192–14197.
128 K. C. Pao, M. Stanley, C. Han, Y. C. Lai, P. Murphy, K. Balk, N. T. Wood, O. Corti, J. C. Corvol, M. M. K. Muqit and S. Virdee, *Nat. Chem. Biol.*, 2016, **12**, 324–331.
129 M. E. Gegg, J. M. Cooper, K. Y. Chau, M. Rojo, A. H. V. Schapira and J. W. Taanman, *Hum. Mol. Genet.*, 2010, **19**, 4861–4870.
130 T. M. Dawson and V. L. Dawson, *Mov. Disord.*, 2010, **25**, S32.
131 T. R. Mhyre, J. T. Boyd, R. W. Hamill and K. A. Maguire-Zeiss, *Subcell. Biochem.*, 2012, **65**, 389–455.
132 A. Borodovsky, B. M. Kessler, R. Casagrande, H. S. Overkleeft, K. D. Wilkinson and H. L. Ploegh, *EMBO J.*, 2001, **20**, 5187–5196.
133 M. Hu, P. Li, M. Li, W. Li, T. Yao, J. W. Wu, W. Gu, R. E. Cohen and Y. Shi, *Cell*, 2002, **111**, 1041–1054.
134 S. C. Johnston, S. M. Riddle, R. E. Cohen and C. P. Hill, *EMBO J.*, 1999, **18**, 3877–3887.
135 M. Y. Balakirev, S. O. Tcherniuk, M. Jaquinod and J. Chroboczek, *EMBO Rep.*, 2003, **4**, 517–522.
136 D. S. Hewings, J. A. Flygare, M. Bogyo and I. E. Wertz, *FEBS J.*, 2017, **284**, 1555–1576.
137 Y. Gao, A. Koppen, M. Rakhshandehroo, I. Tasdelen, S. F. Van De Graaf, J. Van Loosdregt, O. Van Beekum, N. Hamers, D. Van Leenen, C. R. Berkers, R. Berger, F. C. P. Holstege, P. J. Coffer, A. B. Brenkman, H. Ovaa and E. Kalkhoven, *Nat. Commun.*, 2013, **4**, 1–10.
138 E. Kummari, N. Alugubelly, C.-Y. Hsu, B. Dong, B. Nanduri and M. J. Edelmann, *PLoS*

One, 2015, **10**, e0135531.
139 X. Wang, M. Mazurkiewicz, E. K. Hillert, M. H. Olofsson, S. Pierrou, P. Hillertz, J. Gullbo, K. Selvaraju, A. Paulus, S. Akhtar, F. Bossler, A. C. Khan, S. Linder and P. D'Arcy, *Sci. Rep.*, , DOI:10.1038/srep26979.
140 S. D. Whedon, N. Markandeya, A. S. J. B. Rana, N. A. Senger, C. E. Weller, F. Tureček, E. R. Strieter and C. Chatterjee, *J. Am. Chem. Soc.*, 2016, **138**, 13774–13777.
141 D. S. Hameed, A. Sapmaz, L. Burggraaff, A. Amore, C. J. Slingerland, G. J. P. Westen and H. Ovaa, *Angew. Chemie Int. Ed.*, 2019, **58**, 14477–14482.
142 S. Baba and E. ichi Negishi, *J. Am. Chem. Soc.*, 1976, 98, 6729–6731.
143 N. Miyaura, K. Yamada and A. Suzuki, *Tetrahedron Lett.*, , DOI:10.1016/S0040-4039(01)95429-2.
144 K. Matos and J. A. Soderquist, *J. Org. Chem.*, 1998, **63**, 461–470.
145 A. Wohl, *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.*, 1919, **52**, 51–63.
146 K. Ziegler, G. Schenck, E. W. Krockow, A. Siebert, A. Wenz and H. Weber, *Justus Liebig's Ann. der Chemie*, 1942, **551**, 1–79.
147 D. Suarez, G. Laval, S. M. Tu, D. Jiang, C. L. Robinson, R. Scott and B. T. Golding, *Synthesis (Stuttg).*, 2009, **2009**, 1807–1810.
148 J. Yang and Y. Zhang, *Web Serv. issue Publ. online*, , DOI:10.1093/nar/gkv342.
149 F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. De Jong, H. Hilkmann, T. K. Sixma and H. Ovaa, *Angew. Chemie - Int. Ed.*, 2010, **49**, 10149–10153.
150 T. Haack and M. Mutter, *Tetrahedron Lett.*, 1992, **33**, 1589–1592.
151 T. Wöhr and M. Mutter, *Tetrahedron Lett.*, 1995, **36**, 3847–3848.
152 Z. Wang, R. Yang, J. Zhu and X. Zhu, in *Science China Chemistry*, Springer, 2010, vol. 53, pp. 1844–1852.
153 B. G. De La Torre, A. Jakab and D. Andreu, *Int. J. Pept. Res. Ther.*, 2007, **13**, 265–270.
154 K. C. PUGH, E. J. YORK and J. M. STEWART, *Int. J. Pept. Protein Res.*, 2009, **40**, 208–213.
155 M. Paradís-Bas, J. Tulla-Puche and F. Albericio, *Chem. Soc. Rev.*, 2016, 45, 631–654.
156 A. Borodovsky, H. Ovaa, W. J. N. Meester, E. S. Venanzi, M. S. Bogyo, B. G. Hekking, H. L. Ploegh, B. M. Kessler and H. S. Overkleeft, *ChemBioChem*, 2005, **6**, 287–291.
157 S. Zaramella, R. Strömberg and E. Yeheskiely, *European J. Org. Chem.*, 2002, **2002**, 2633.
158 K. L. Kirk, *J. Heterocycl. Chem.*, 1985, 22, 57–59.
159 N. K. Salam, M. Adzhigirey, W. Sherman and D. A. Pearlman, *Protein Eng. Des. Sel.*, 2014, **27**, 365–374.
160 H. Beard, A. Cholleti, D. Pearlman, W. Sherman and K. A. Loving, *PLoS One*, 2013, **8**, e82849.
161 K. Zhu, T. Day, D. Warshaviak, C. Murrett, R. Friesner and D. Pearlman, *Proteins Struct. Funct. Bioinforma.*, 2014, **82**, 1646–1655.
162 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

Appendix

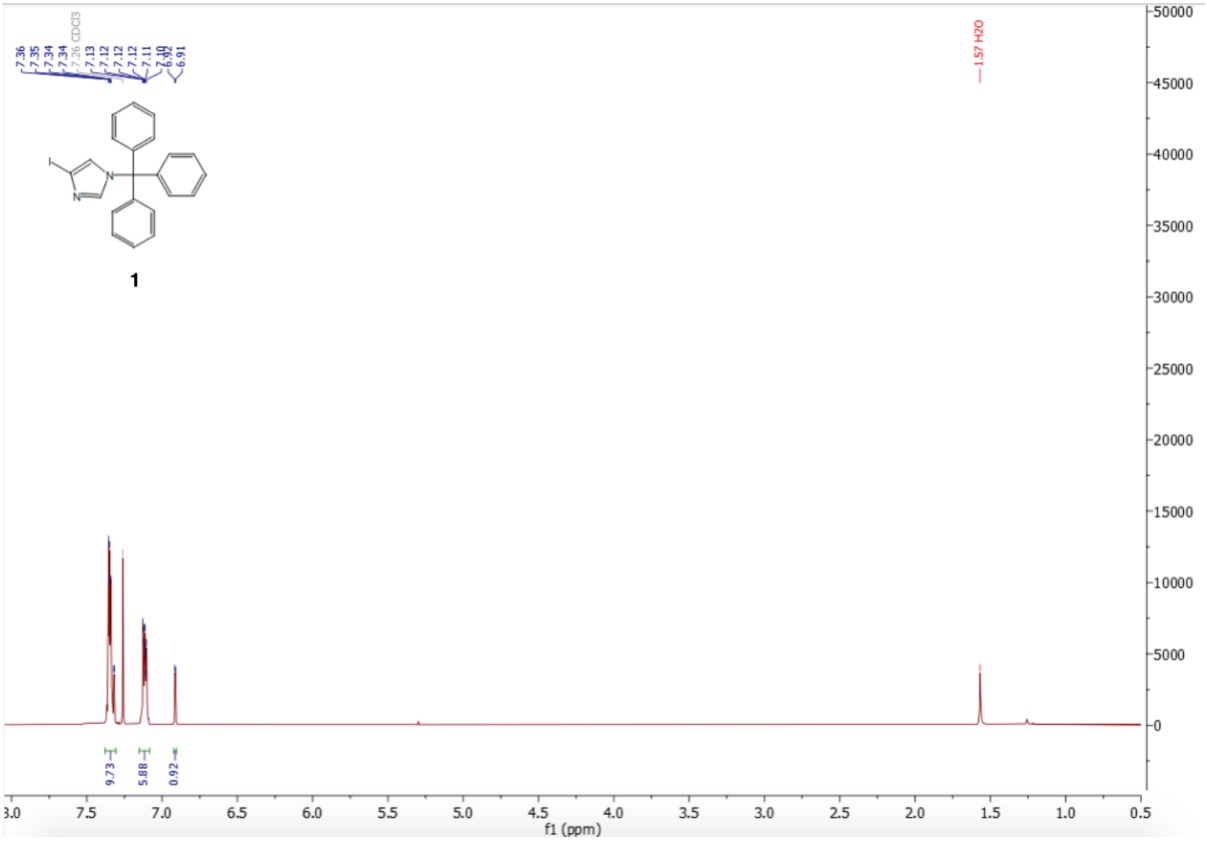


Figure S1: ^{1}H NMR (400 MHz, CDCl3) of 4-iodo-1-trityl-1H-imidazole (**1**)

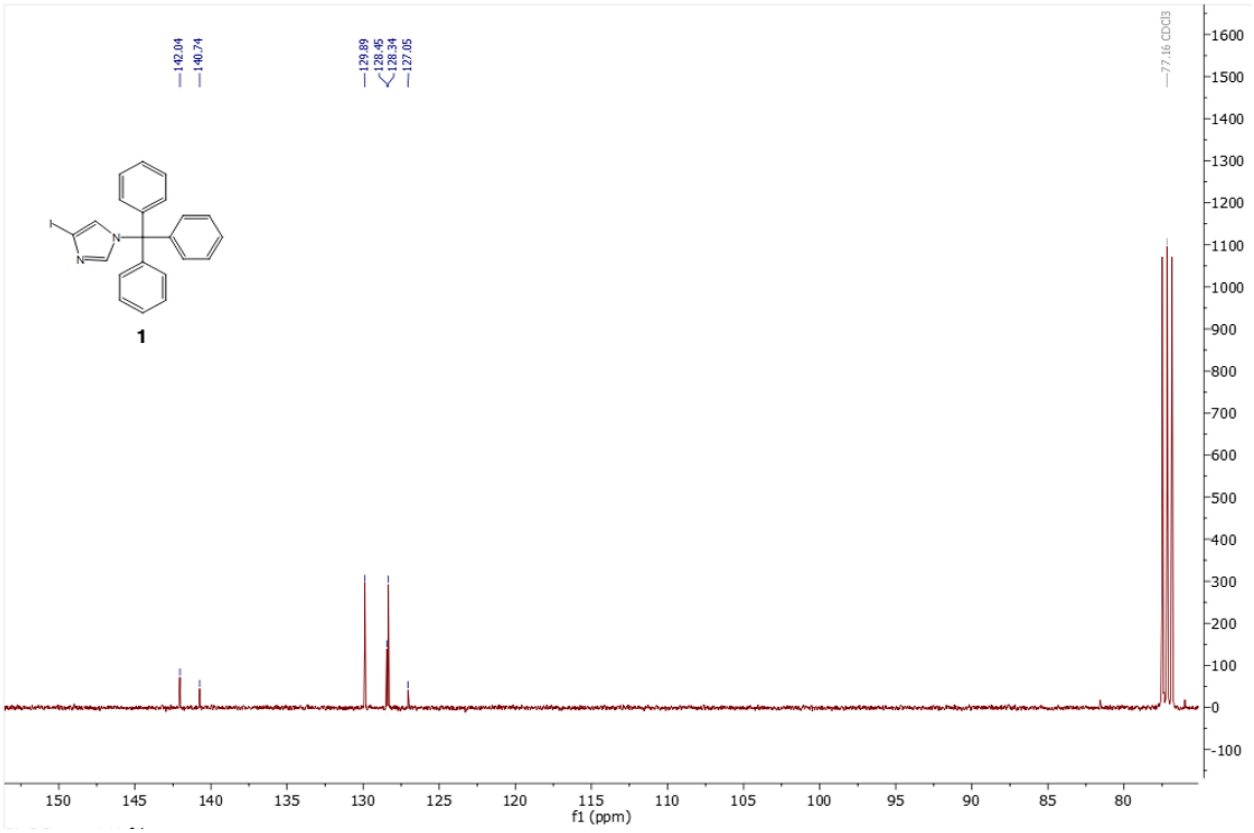


Figure S2: ^{13}C NMR (101 MHz, CDCl3) of 4-iodo-1-trityl-1H-imidazole (**1**)

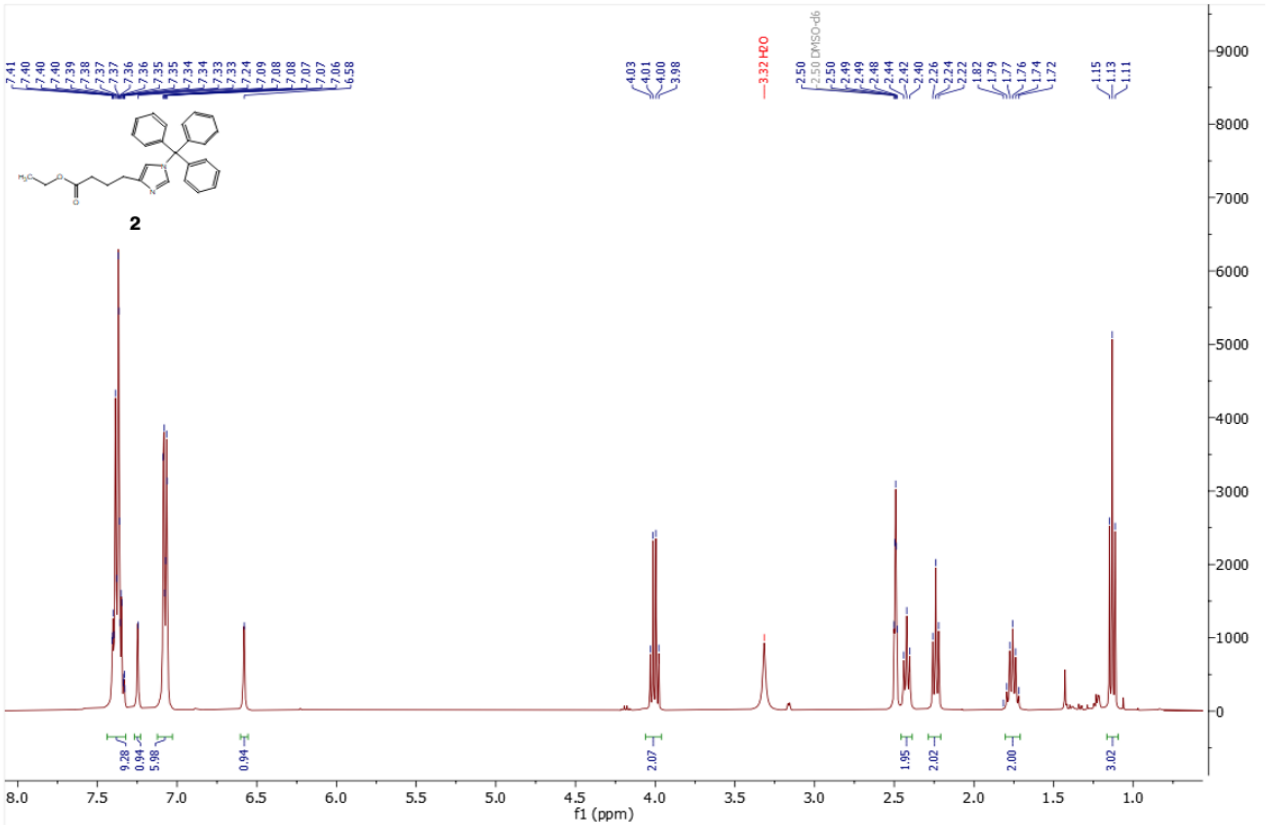


Figure S3: ^{1}H NMR (400 MHz, DMSO) of Ethyl 4-(1-trityl-1H-imidazol-4-yl)butanoate (**2**)

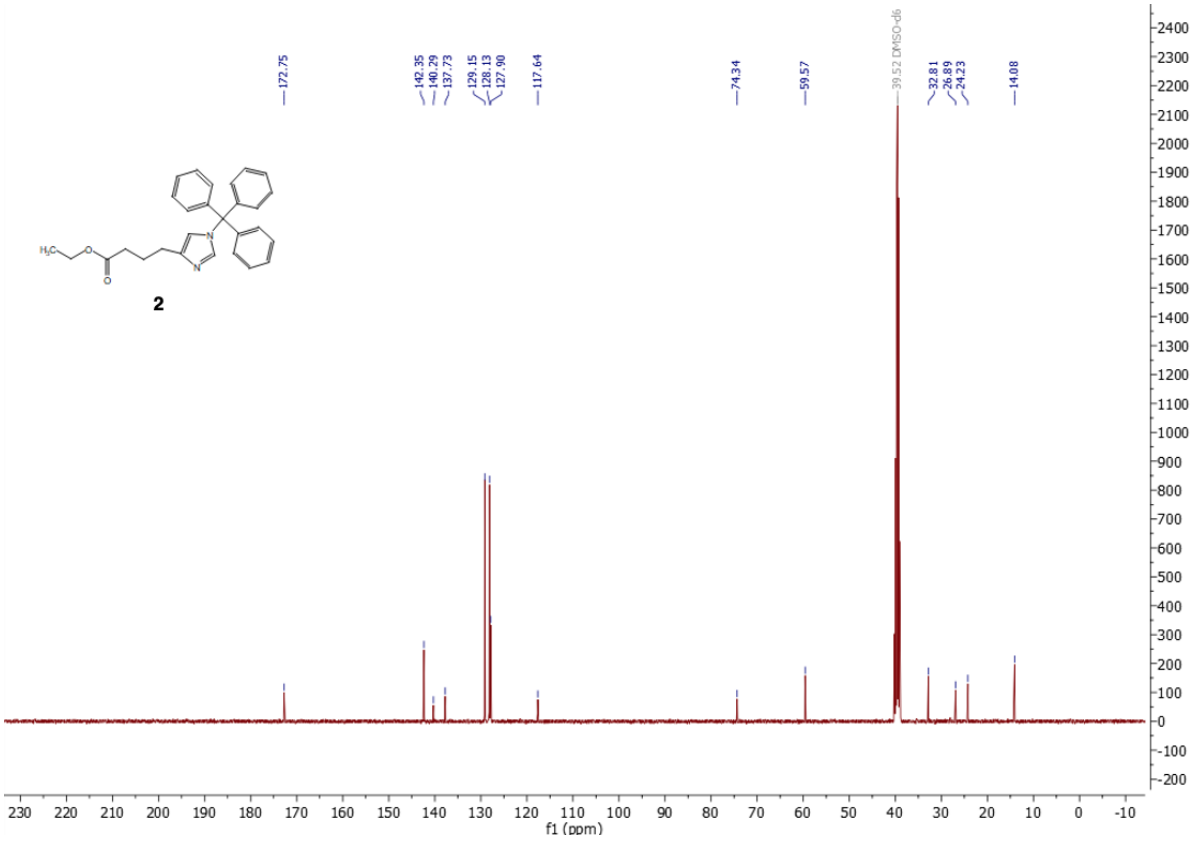


Figure S4: ^{13}C NMR (101 MHz, DMSO) of Ethyl 4-(1-trityl-1H-imidazol-4-yl)butanoate (**2**)

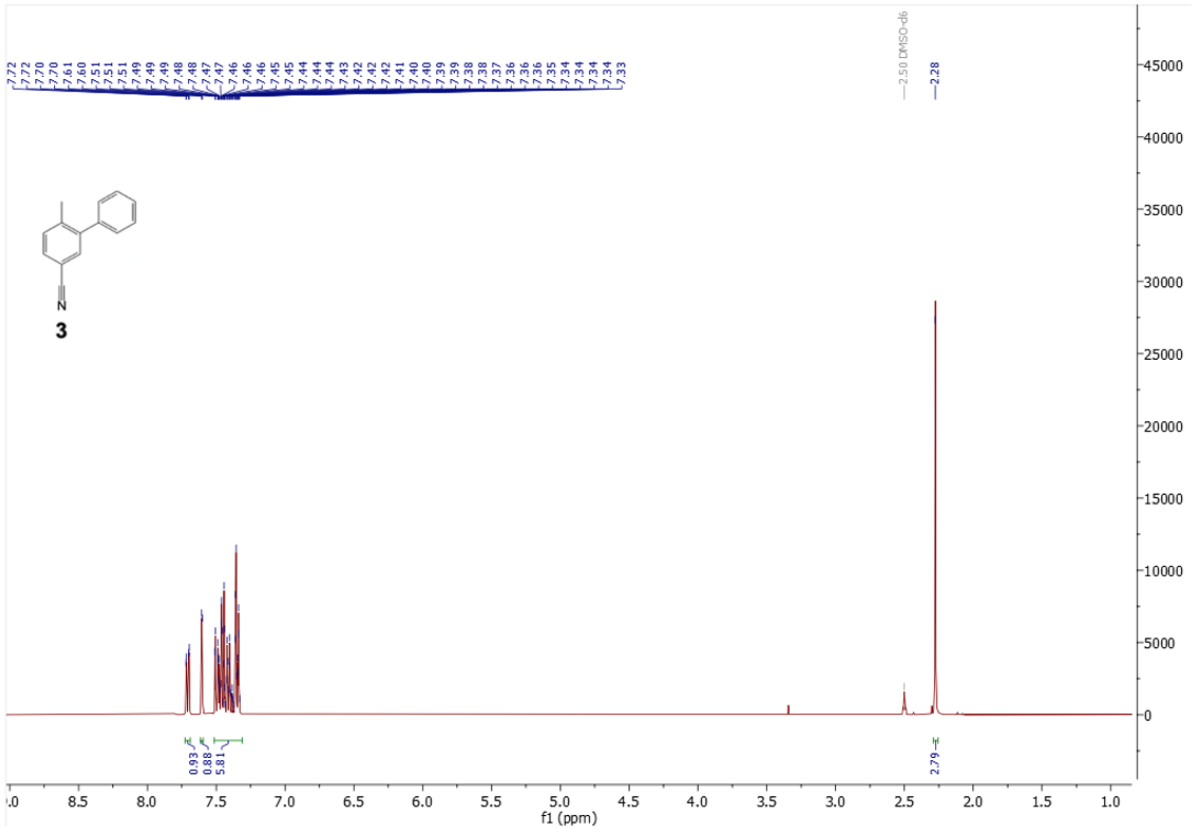


Figure S5: ^{1}H NMR (400 MHz, DMSO) of 6-Methyl-[1,1'-biphenyl]-3-carbonitrile (**3**)

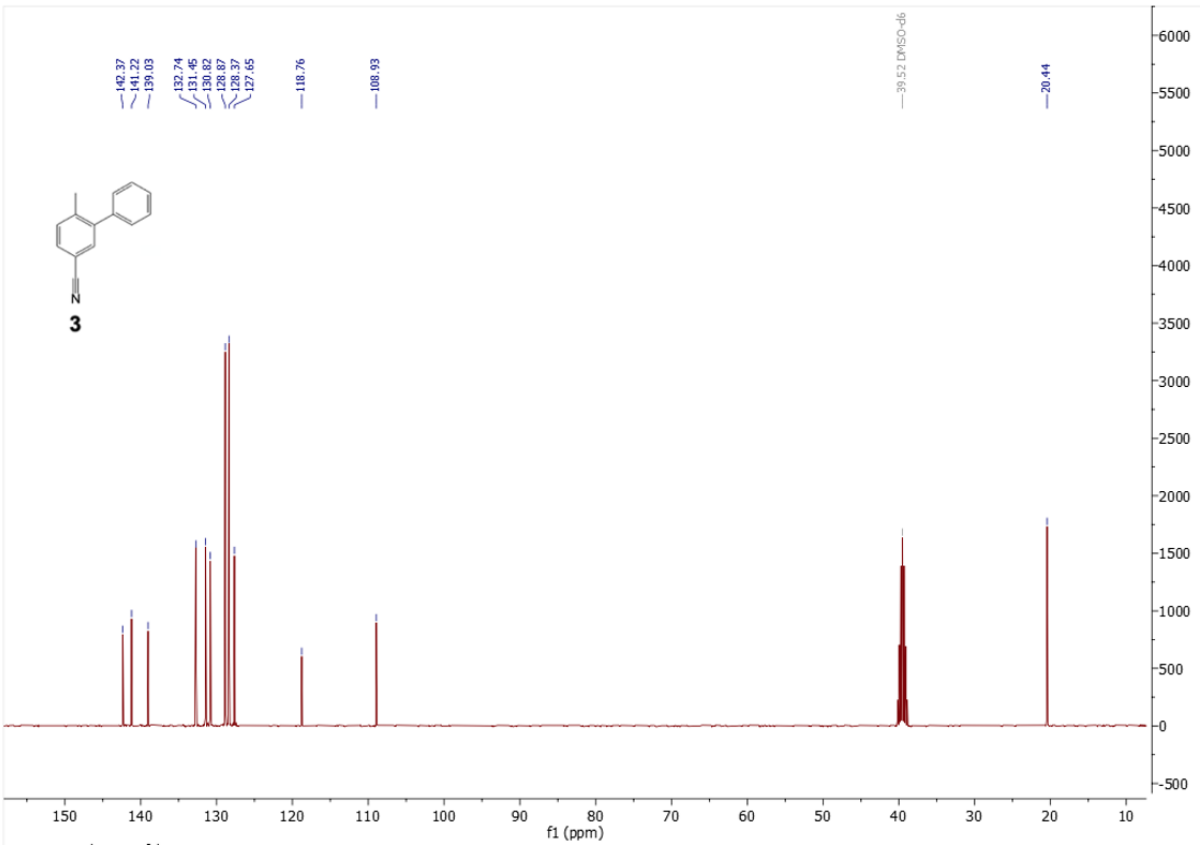


Figure S6: ^{13}C NMR (101 MHz, DMSO) of 6-Methyl-[1,1'-biphenyl]-3-carbonitrile (**3**)

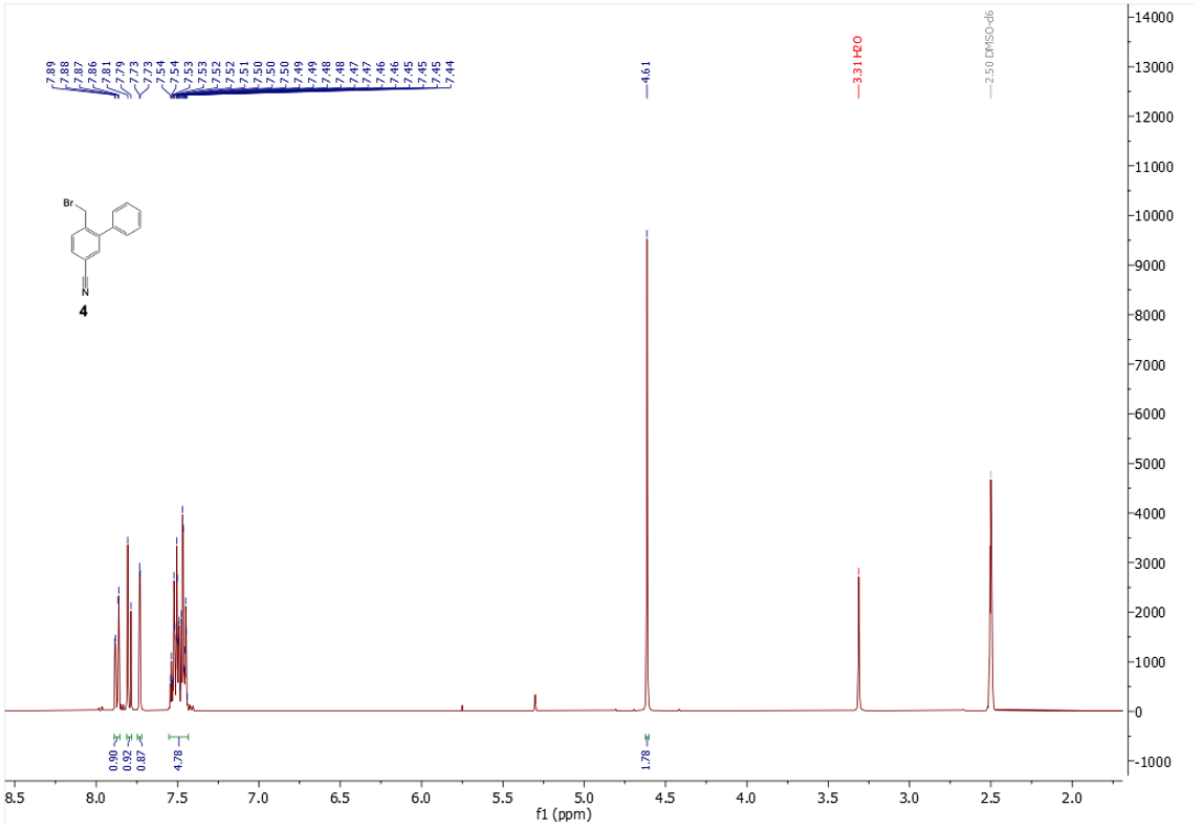


Figure S7: ^{1}H NMR (400 MHz, DMSO) of 6-Bromomethyl-[1,1'-biphenyl]-3-carbonitrile (**4**)

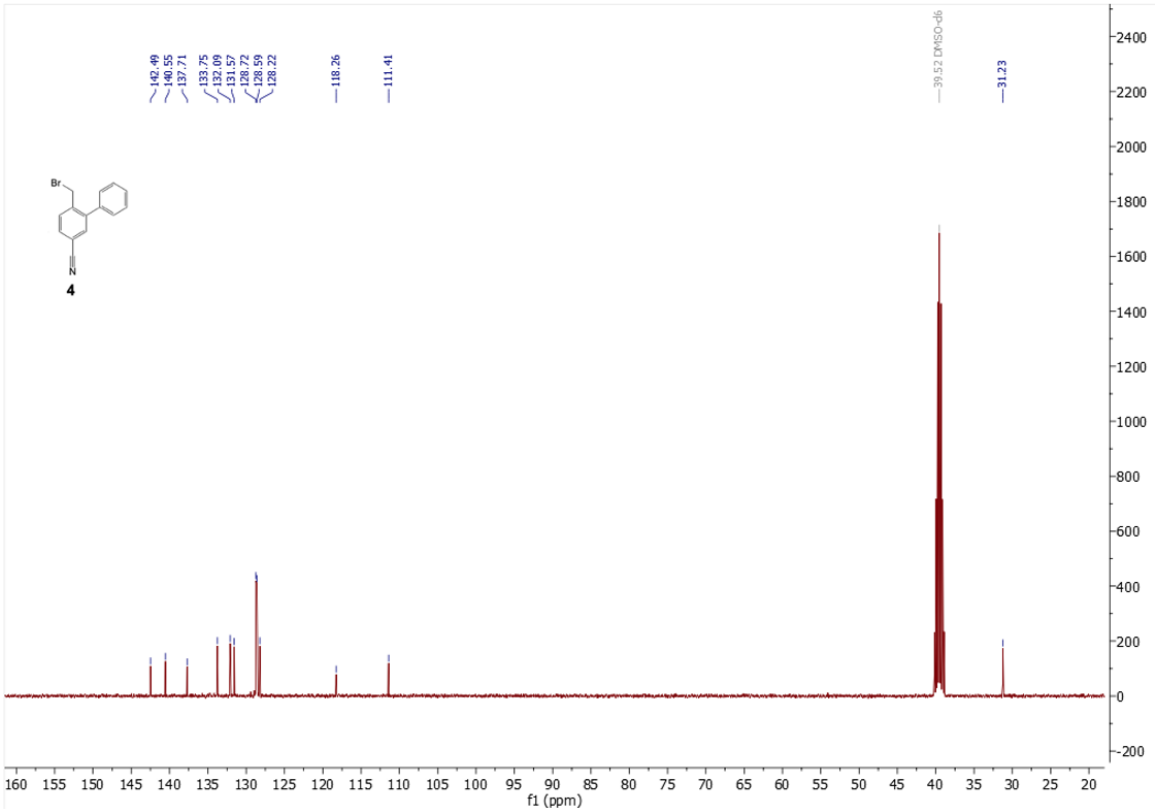


Figure S8: ^{13}C NMR (101 MHz, DMSO) of 6-Bromomethyl-[1,1'-biphenyl]-3-carbonitrile (**4**)

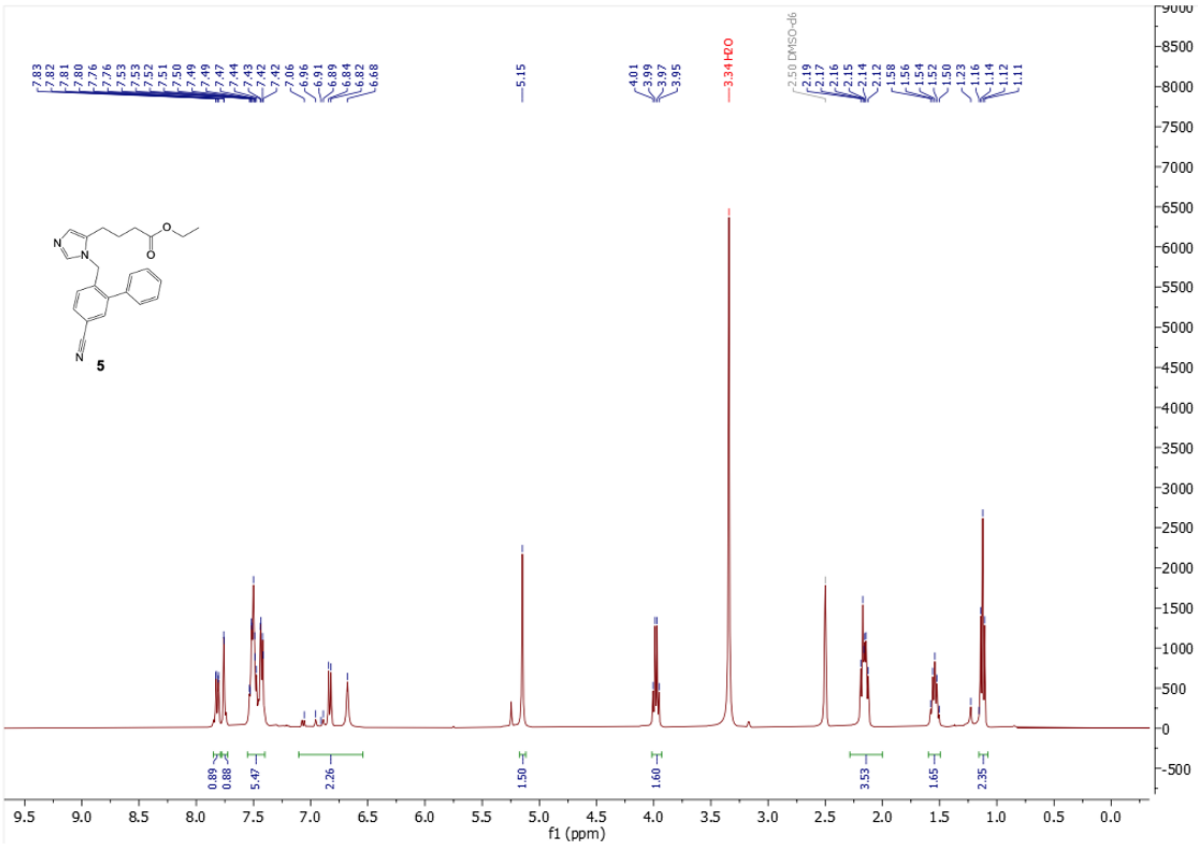


Figure S9: ^{1}H NMR (400 MHz, DMSO) of Ethyl-4-(1-((5-cyano-[1,1'-biphenyl]-2-yl)methyl)-1H-imidazol-5-yl) butanoate (**5**)

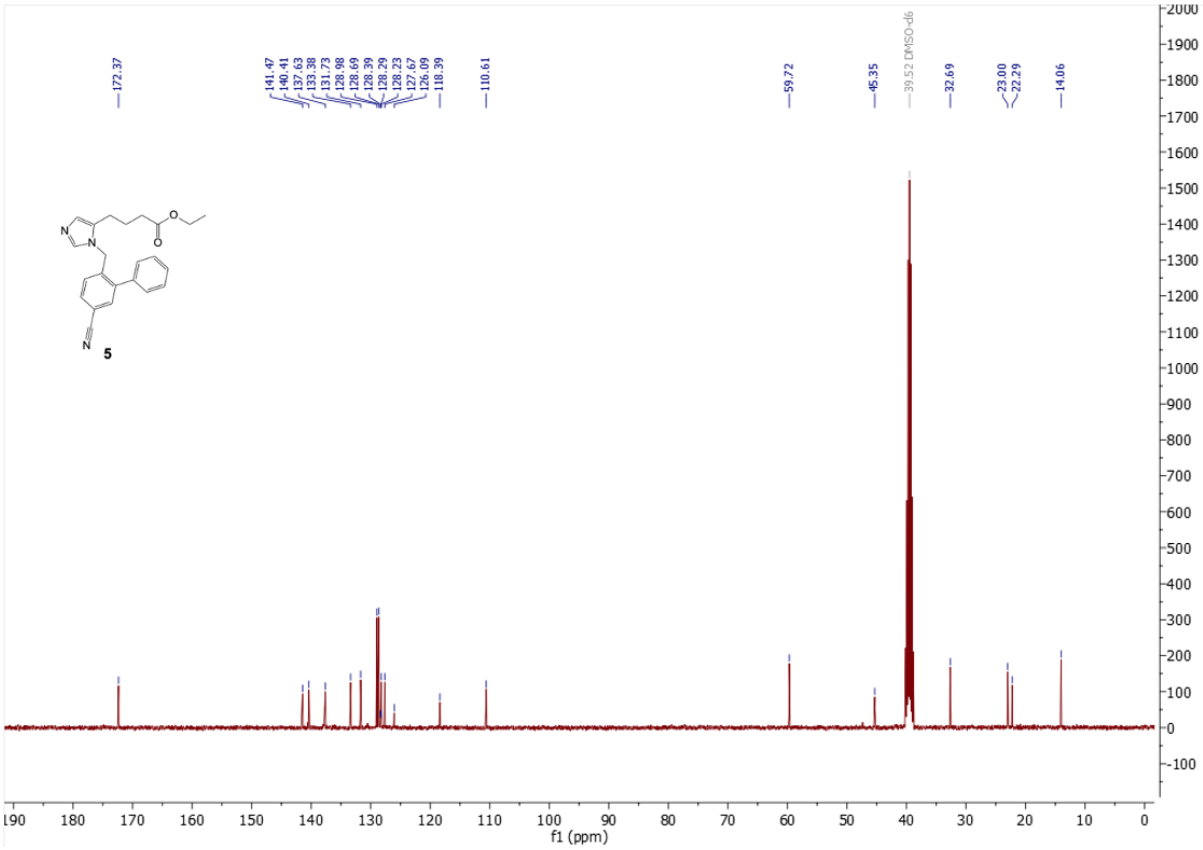


Figure S10: ^{13}C NMR (101 MHz, DMSO) of Ethyl-4-(1-((5-cyano-[1,1'-biphenyl]-2-yl)methyl)-1H-imidazol-5-yl) butanoate (**5**)

Diketopiperazine formation

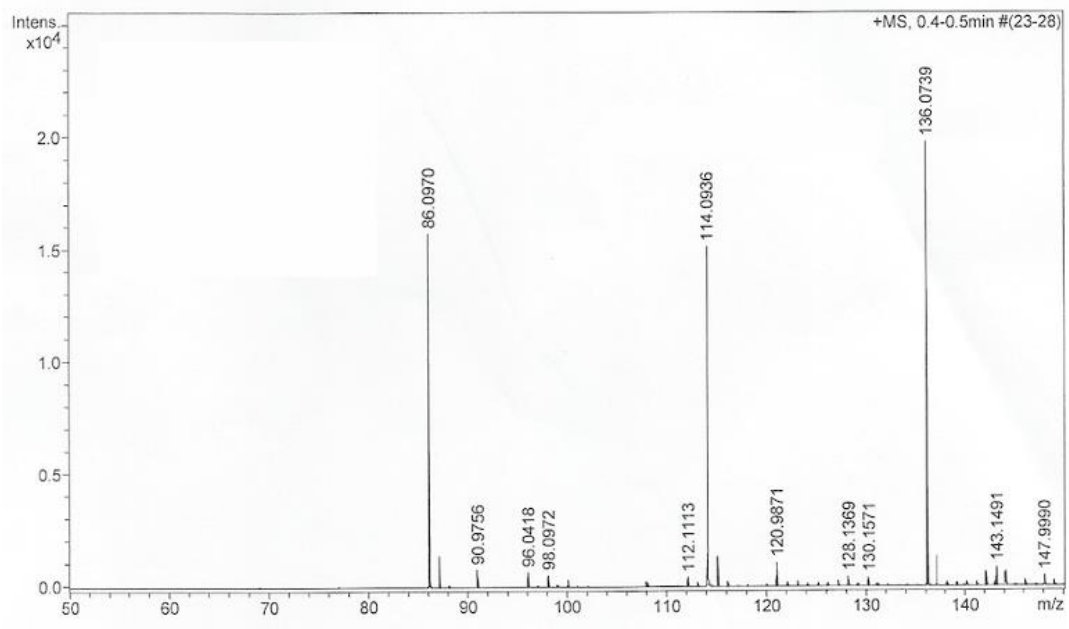


Figure S11: ESI+ HRMS of diketopiperazine (cyclic diglycine dipeptide). m/z expected $[M+H]^+$ = 114, found 114. m/z expected $[M+Na]^+$ = 136, found 136.

NEDD8 Synthesis: Test Cleave 1

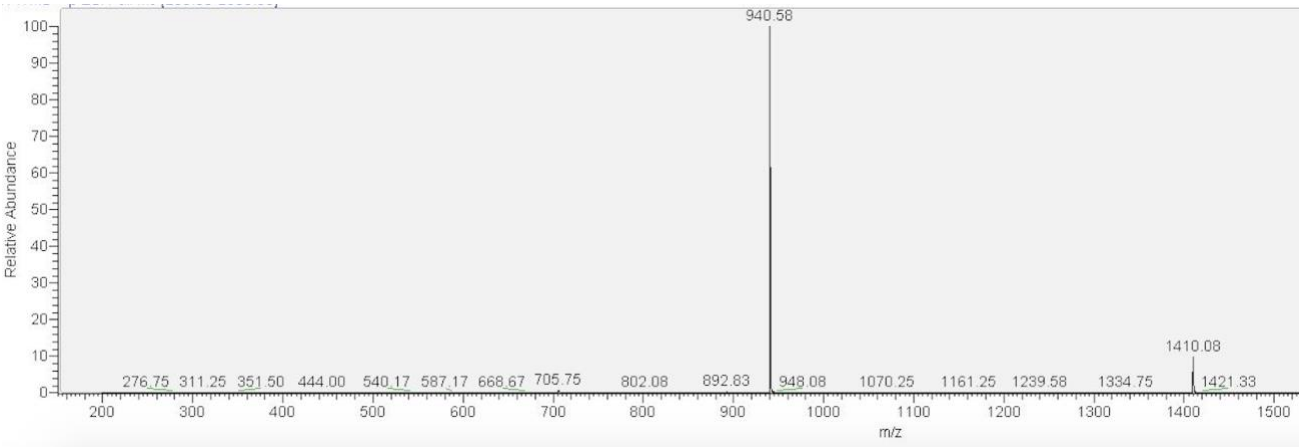


Figure S12: LC-MS of NEDD8 residues 52-76 + Fmoc. m/z expected [M/3] = 940, found 940. m/z expected [M/2] = 1410, found 1410.

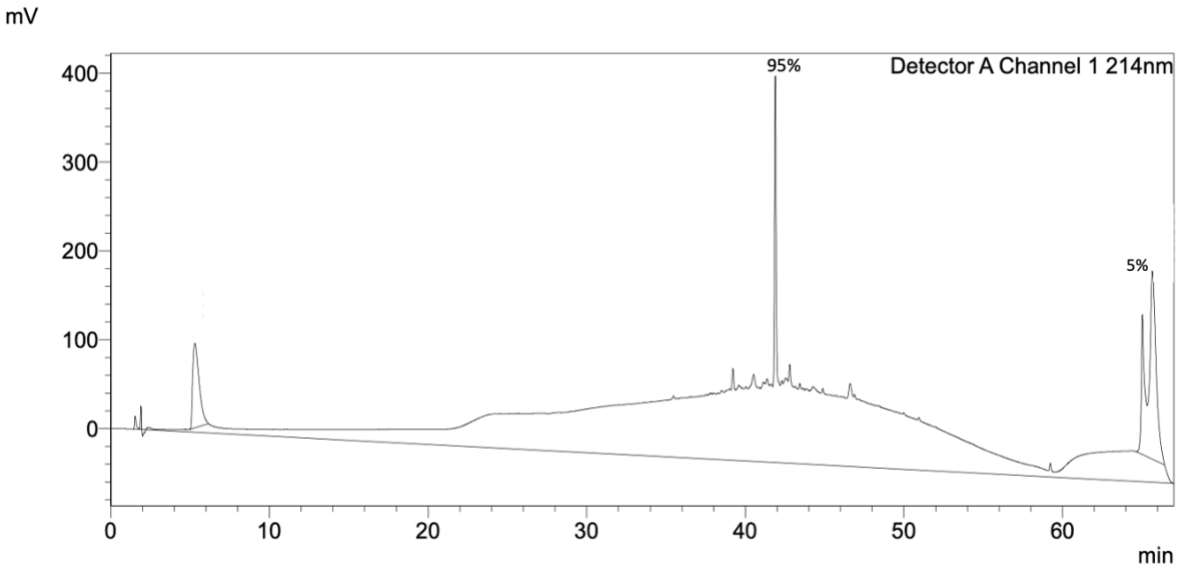


Figure S13: Analytical HPLC of crude NEDD8 residues 52-76 + Fmoc. Rt of product at 42 min.

NEDD8 Synthesis: Test cleave 2

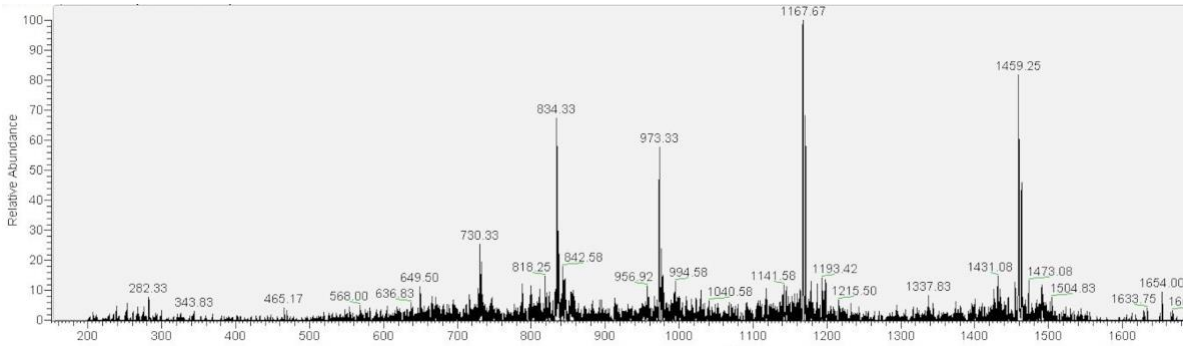


Figure S14: LC-MS of NEDD8 residues 25-76. m/z expected [M/6] = 973, found 973. m/z expected [M/5] = 1167, found 1168. m/z expected [M/4] = 1459, found 1459.

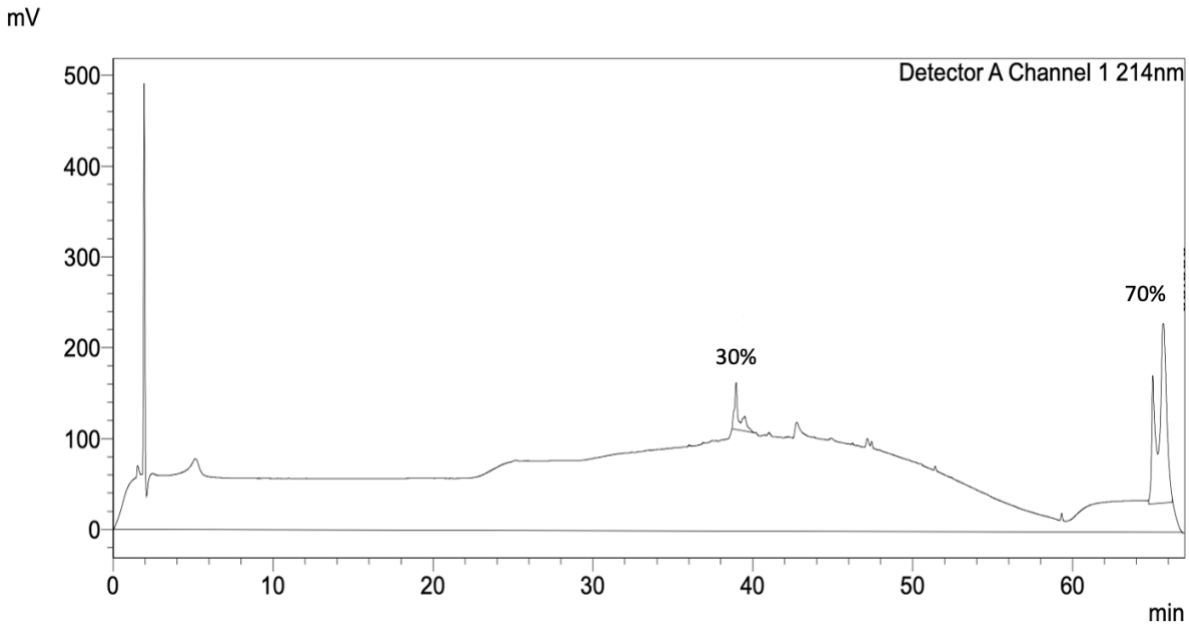


Figure S15: Analytical HPLC of crude NEDD8 residues 25-76 + Fmoc. Rt of product at 39.5 min.

NEDD8 Synthesis: Test cleave 3

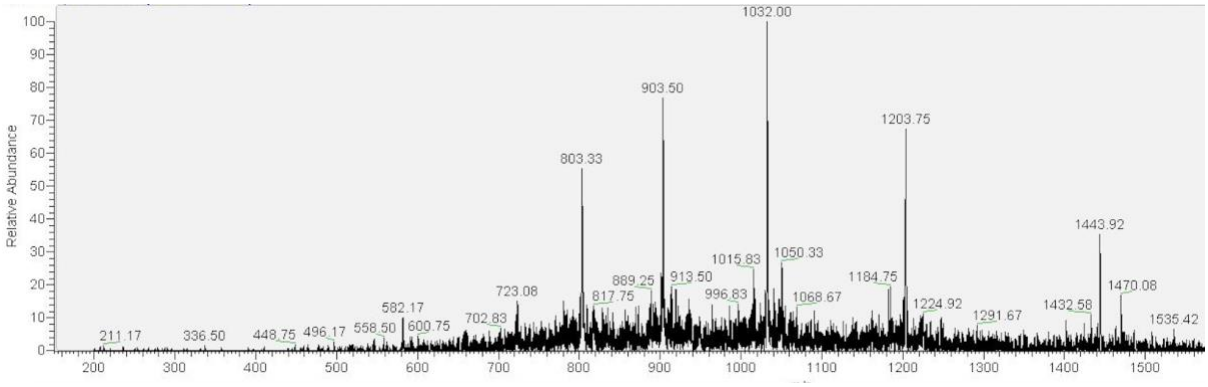


Figure S16: LC-MS of NEDD8 residues 1-76 + Fmoc. m/z expected [M] = 8782, not found. m/z found = 7224 (by ESI calculation).

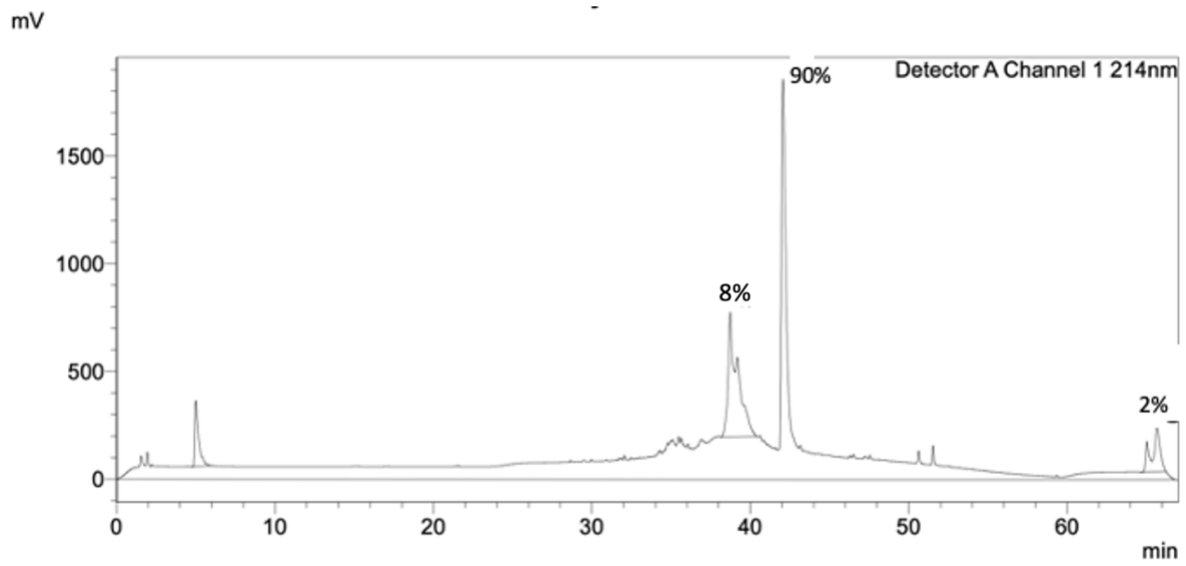
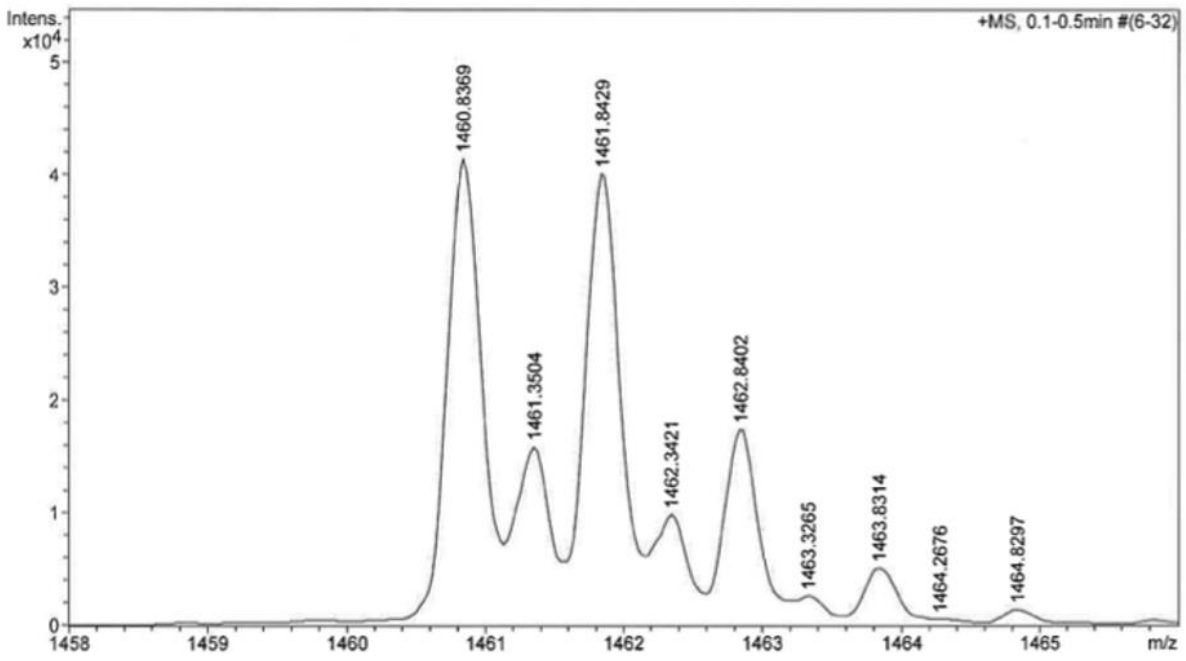


Figure S17: Analytical HPLC of crude NEDD8 residues 1-76 + Fmoc. Rt of observed mass = 39 min.

Biotinylated 12-mer NEDD8 *C*-terminal Peptide



Formula	z	m/z	Meas. m/z	err [ppm]	err [mDa]
	1+		1460.8369		
C 66 H 114 N 19 O 16 S		1460.8406		2.6	3.8

Figure S18: ESI+ HRMS of SVLHLVLALRGG + biotin. m/z expected $[M+H]^+$ = 1461, found 1461.

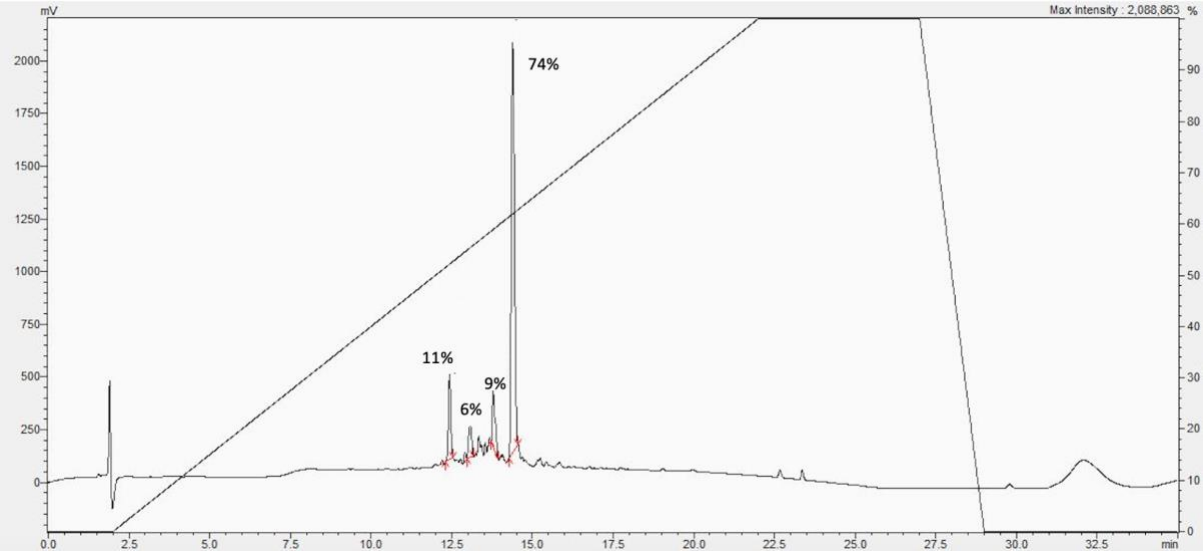


Figure S19: Analytical HPLC of crude SVLHLVLALRGG + biotin. Rt of product = 14.5 min.

Conjugation of 12-mer Peptide to Hydroxylamine

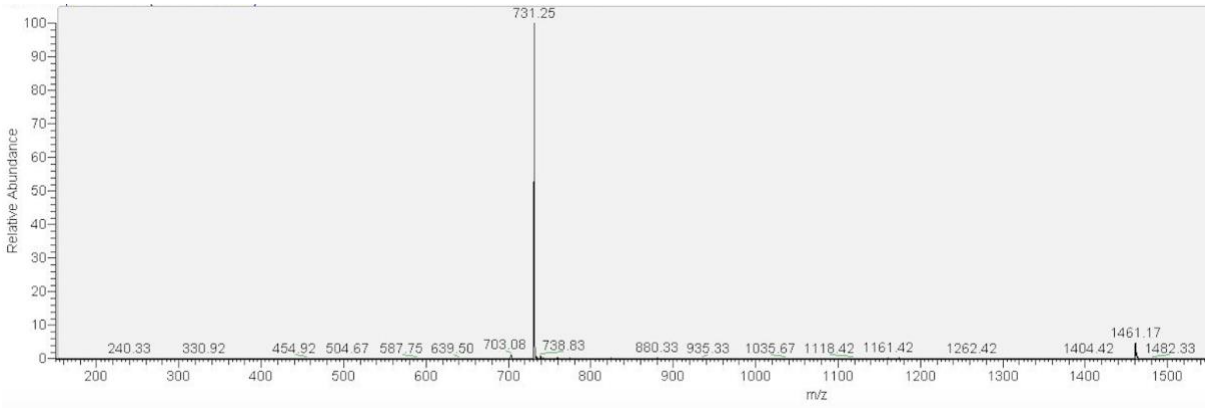


Figure S20: LC-MS of biotin-SVLHLVLALRGG + hydroxylamine. m/z expected [M] = 1476, not found.

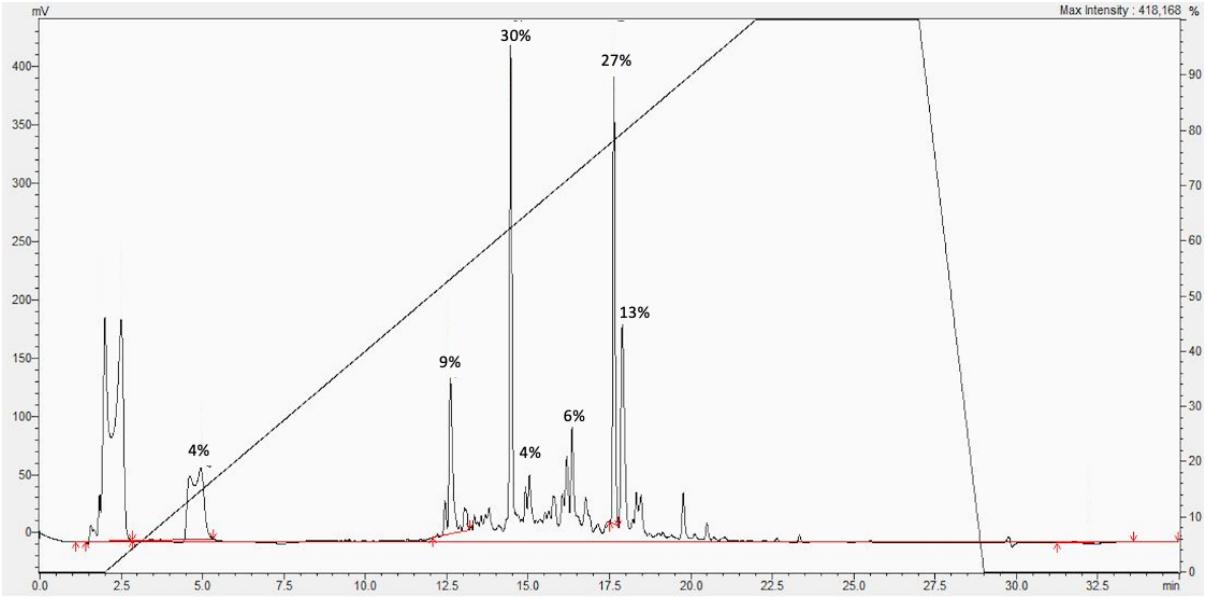


Figure S21: Analytical HPLC of crude biotin-SVLHLVLALRGG + hydroxylamine. Rt of observed mass = 14.5 min

Conjugation of 12-mer Peptide to Imidazole

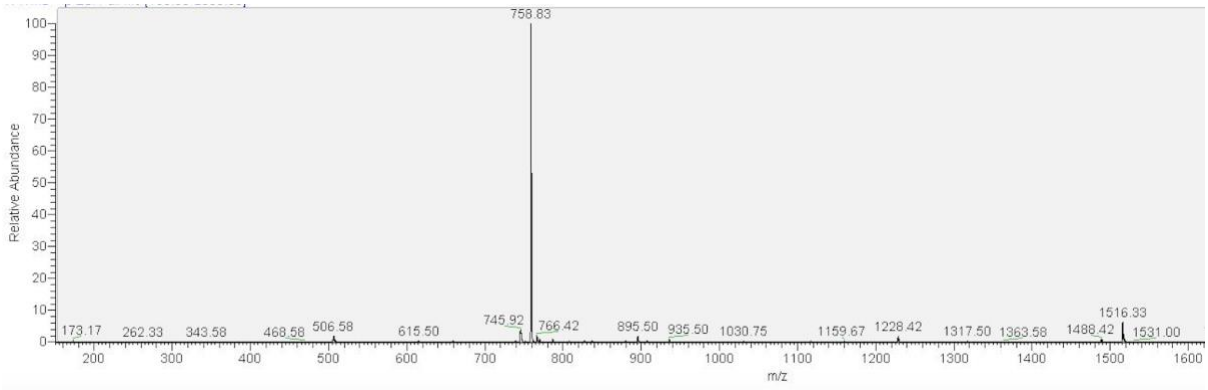


Figure S22: LC-MS of biotin-SVLHLVLALRGG + imidazole. m/z expected [M] = 1512, not found.

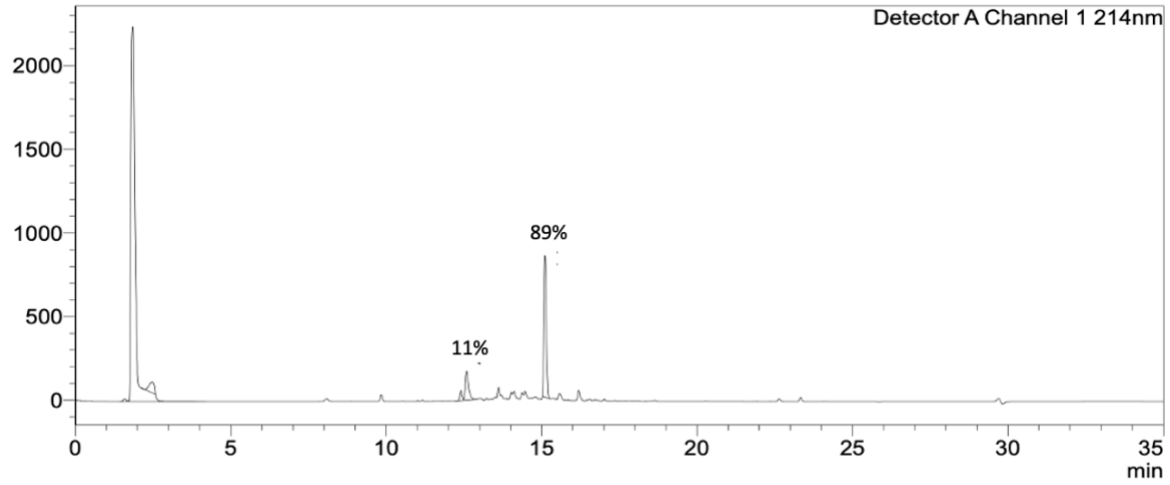


Figure S23: Analytical HPLC of crude biotin-SVLHLVLALRGG + imidazole. Rt of observed mass = 15 min.