2021

Chemical Tools for Exploring Metabolite Interactions with Nuclear Receptors and Beyond

Taku Tsukidate

CHEMICAL TOOLS FOR EXPLORING METABOLITE INTERACTIONS WITH NUCLEAR RECEPTORS AND BEYOND

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfilment of the Requirements for
the degree of Doctor of Philosophy

by
Taku Tsukidate
June 2021

CHEMICAL TOOLS FOR EXPLORING METABOLITE INTERACTIONS WITH NUCLEAR RECEPTORS AND BEYOND

Taku Tsukidate, Ph.D.
The Rockefeller University 2021

The identification and functional characterization of specific metabolite–protein interactions remains a major challenge in chemical biology and drug development. The functionalization of native metabolites and synthetic small molecules with bio-orthogonal detection tags (alkyne, azide, and others) has afforded chemical reporters for the biochemical analysis of metabolically labelled proteins, whereas functionalization with photo-crosslinkers enables non-covalent metabolite–protein interactions to be captured. While we and many others have employed this approach to different metabolites and synthetic ligands, the installation of bio-orthogonal detection tags and photo-crosslinkers to more complex and diverse small molecules can be challenging and limiting. In this regard, activity-based protein profiling (ABPP) using active site–directed probes in combination with small molecule competition studies allows the discovery of potential substrates and inhibitors of selective enzyme families. Moreover, the development of proteome-wide labelling of select amino acid (cysteine and lysine) has revealed new sites in proteins for function modulation. However, some key protein classes and specific ligand-binding domains have been inaccessible by these methods. New activity-based probes and targeted chemical reporters of specific protein families are thus needed to characterize their endogenous and exogenous ligands for fundamental biology and drug discovery. In this thesis, we describe two projects aimed at developing chemical probes to assist in small molecule target engagement and target deconvolution studies.

In **Chapter 1**, we provide an overview of the origin and evolution of ABPP and related chemical reporter–based chemical proteomics methods. We also summarise recent developments in label-free methods that work with underivatized molecules.

In **Chapter 2**, we describe the development of a ligand-directed probe (alk-GW9662) for a subset of nuclear receptors (NRs). NRs are a family of ligand-activated transcription factors that regulate diverse physiological processes in animals and are key targets for therapeutic development. Recent studies implicated NRs in host–microbiota interactions and suggested that various microbiota-derived metabolites may serve as ligands for NRs. We demonstrate alk-GW9662's utility to assess target engagement of candidate ligands of peroxisome proliferator–activated receptor (PPAR) γ. We also explore repurposing alk-GW9662. We profile its target proteins using chemical proteomics and demonstrate that the probe can be used in target engagement study of other proteins. Overall, our results suggest that alk-GW9662 may be useful for target engagement analysis of candidate PPAR ligands and offer a starting scaffold in developing probes for other proteins.

In **Chapter 3**, we describe efforts towards the combined use of proteolysis-targeting chimeras (PROTACs) and chemical proteomics in target deconvolution studies. We describe the design and synthesis of PROTACs that recruit putative E3 ligases.

Overall, the projects described herein underline the utility of chemical tools such as ligand-directed probes and PROTACs in target engagement and target deconvolution studies.

ACKNOWLEDGMENTS

I would like to thank my advisor, Howard Hang, for his mentorship for the past five years. You were always there whenever I needed guidance and encouragement. Thank you to the Hang lab members of the past and present, especially: Yen Chih Wang for friendship; Qiang Li for his delightfulness and technical advice on chemical synthesis; Nathan Westcott, Emmanuelle Thinon, and Zhenrun "Jerry" Zhang for technical advice on biological experiments. I would also like to thank the Proteomics Resource Centre and High-throughput Screening Research Core staffs, especially Henrik Molina, Soeren Heissel, Joseph Fernandez, and Carolina Adura Alcaino for their expertise. It was my pleasure to work with all of you.

I greatly appreciate the guidance, support, and insight of my committee members, Sean Brady, Thomas Sakmar, and Daniel Bachovchin. In particular, I would like to thank Daniel Bachovchin for taking me as a rotation student in his lab five years ago and offering continued mentorship ever since. I would also like to thank Alan Saghatelian for graciously agreeing to be on my committee and review my dissertation.

Heaps of thanks and arigatou to my family and friends, who made the completion of graduate school possible. Thank you to my best friend, Lauren, for believing in me even when I couldn't. Thank you to my friends back home, Gaku, Kaneko, Kenji, and Nakayama-kun, for making time to revitalise me whenever I visited Japan. Thank you to my teammates of New York Harriers, especially Leland, David, and Aaron, for running the "PhD marathon" (and countless other runs) with me.

Last, but not least, I thank Takenaka Scholarship Foundation who has offered financial and moral support for the past five years.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ..... **iii**
TABLE OF CONTENTS ..... **iv**
LIST OF FIGURES ..... **v**
LIST OF TABLES ..... **vi**
CHAPTER 1. The origin and evolution of chemoproteomic methods **1**
Active site–directed probes and activity-based protein profiling 1
Non-directed probes and reactivity profiling 3
Label-free methods for target identification and target engagement studies 5
CHAPTER 2. Nuclear receptor chemical reporter enables domain-specific analysis of ligands in mammalian cells **7**
Development of alk-GW9662 for target engagement analysis of PPARγ ligands 8
PPARγ engagement analysis of microbiota-derived metabolites using alk-GW96622 . 14
Repurposing alk-GW9662 for other proteins 18
Discussion 25
CHAPTER 3. Studies towards untargeted protein degradation strategy **26**
Synthesis of PPARγ-recruiting PROTACs based on (*R*)-SR9034 27
Synthesis of GW9662 derivatives 29
Putative ubiquitin E3 ligase activity of PPARγ 32
Targeted degradation of BRD4 32
Target engagement of GW9662-derivatives 32
Discussion 33
CHAPTER 4. Overall discussion and outlook **34**
METHODS. **36**
Chapter 2 36
Chapter 3 51
REFERENCES **86**

LIST OF FIGURES

Figure 1-1. Active site–directed probes and activity-based protein profiling. 2
Figure 1-2. Non-directed probes and reactivity profiling. 3
Figure 1-3. Label-free methods for target identification and target engagement studies. .. 5
Figure 2-1. Mechanisms of action for microbiota-derived metabolites. 7
Figure 2-2. Targeted analysis of PPARγ–ligand interactions using alk-GW9662. 9
Figure 2-3. Chloronitrobenzene is less reactive than iodoacetamide. 9
Figure 2-4. Conserved cysteine residues in NR LBD can be covalently targeted. 10
Figure 2-5. Synthesis of alk-GW9662. 10
Figure 2-6. Alk-GW9662 site-specifically labels PPARγ C283. 11
Figure 2-7. Alk-GW9662 site-specifically labels PPAR LBD. 11
Figure 2-8. The cysteine-to-alanine mutants are functionally equivalent to wild types. .. 12
Figure 2-9. Competitive labelling strategy for target engagement analysis. 13
Figure 2-10. Comparison of in-lysate and in-cell competition IC_{50} *versus in vitro* binding affinity. 13
Figure 2-11. Kinetic analysis of in-lysate PPARγ labelling. 14
Figure 2-12. Kinetic analysis of in-cell PPARγ labelling. 14
Figure 2-13. PPARγ engagement analysis of microbiota-derived metabolites in cells. ... 16
Figure 2-14. PPARγ engagement analysis of tryptophan metabolites in lysate. 18
Figure 2-15. Alk-GW9662 labels endogenous PPARγ in HT-29 cells. 19
Figure 2-16. Proteome-wide analysis of alk-GW9662 targets. 20
Figure 2-17. Visualization of missing patterns. 21
Figure 2-18. Peptide profile plots. 22
Figure 2-19. Druggability profile of target proteins. 22
Figure 2-20. Repurposing alk-GW9662 for TSPO ligand engagement analysis. 24
Figure 3-1. Targeted protein degradation by hetero-bifunctional molecules. 26
Figure 3-2. Design of PPARγ-recruiting PROTACs based on (*R*)-SR9034. 27
Figure 3-3. Putative ubiquitin E3 ligase activity of PPARγ. 32

LIST OF TABLES

Table 2-1. Abbreviated names of microbiota-derived metabolites. 17
Table 2-2. Druggable hits. 23
Table 2-3. Chemical proteomics hits (differential abundance analysis) with molecular weight within 10–20 kDa. 24
Table 3-1. Ubiquitin E3 ligases targeted by alk-GW9662. 29
Table 3-2. PPARγ engagement of GW9662 derivatives in cell lysate. 33

CHAPTER 1. THE ORIGIN AND EVOLUTION OF CHEMOPROTEOMICS METHODS

Mechanism of action studies of small molecules originating from cellular metabolism and the environment (diet, microbiota, and therapeutics) can offer insights into biological pathways and accelerate drug development. In mechanism of action studies, target deconvolution and target engagement analysis often represent key steps. Chemical proteomics enables unbiased mapping of small molecule–protein interactions *in situ* and facilitates target deconvolution and target engagement analysis. This chapter provides an overview of the origin and evolution of chemo-proteomic approaches focusing on activity-based protein profiling and related chemical reporter–based methods. It also summarises recent developments in chemo-proteomics methods that work with underivatized molecules.

Active site–directed probes and activity-based protein profiling

Singer et al. coined the term affinity labelling in 1962 to describe a general method to site-selectively label the active sites of antibodies and enzymes.[1] They demonstrated site-specific labelling of antigen binding site of pure anti-benzenearsonic acid antibody by *p*-(arsonic acid)-benzenediazonium fluoborate. Similarly, Shaw et al. demonstrated site-specific labelling of chymotrypsin active site by a reactive substrate mimic, L-1-tosyl-amido-2-phenylethyl chloromethyl ketone.[2] A key aspect of the affinity labelling strategy is that labelling reagents are *directed* to so-called active sites (i.e., ligand binding sites) by their intrinsic affinity and selective chemical reactivity.

In the late '90s, the approaching dawn of the post-genomic era sparked needs for new strategies for functional analysis of proteins. While it was becoming possible to measure variations in protein level in bottom-up proteomics, an accurate assessment of the functional state of proteins in cellular environment required new approach.[3,4] Activity-based protein profiling (ABPP) emerged as a key technology in the evolution of functional proteomics (**Figure 1-1A**).[5] In the seminal study by the Cravat group in 1999, the authors described the synthesis and utility of an active site–directed probe, FP-biotin, for simultaneously monitoring the activities of multiple serine hydrolases (**Figure 1-1B**). FP-biotin was designed based on fluorophosphate (FP) derivatives that selectively inhibit serine hydrolases over cysteine, aspartyl, and metallohydrolases. Reactivity of FPs with serine hydrolases requires the enzyme to be in catalytically-active state. Accordingly, FP-biotin reacts with active enzymes, but not with inactive enzymes. The authors demonstrated sensitive detection of numerous serine hydrolase in tissue homogenates. Similarly, Bogyo et al. developed active site–directed probes for the proteasome (e.g. ^{125}I-NIP-L_3VS).[6,7] Thus, ABPP is defined as a proteomic method that employs active site–directed probes to collectively characterize many enzymes in cell, tissue, and fluid samples.[8]

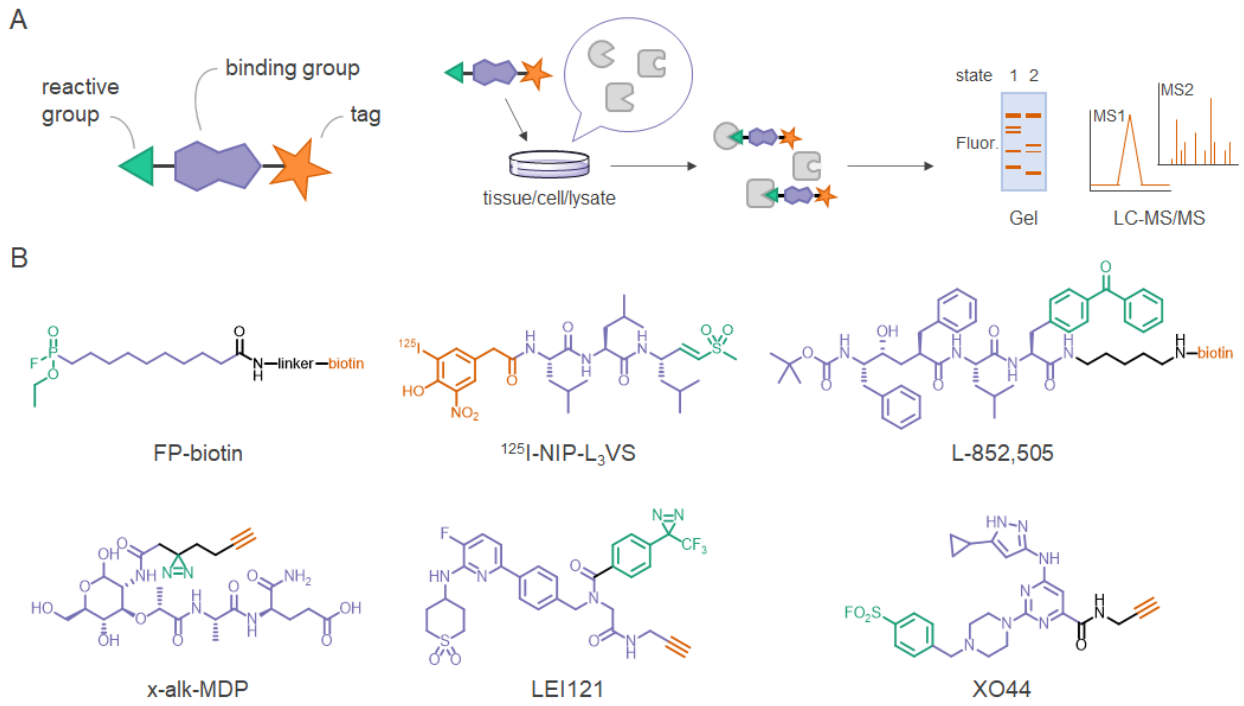


Figure 1-1. Active site–directed probes and activity-based protein profiling (ABPP). (A) General scheme of activity-based protein profiling and related chemical reporter–based chemoproteomic methods. (B) Chemical structures of select active site–directed (ligand-directed) probes. FP-biotin: serine hydrolases; ^{125}I-NIP-L$_3$-VS: proteasome; L-852,505: γ-secretase; x-alk-MDP: NOD2; LEI121: cannabinoid receptor 2; XO44: kinases.

Early ABPP probes carried bulky tags, which limited cellular uptake. Thus, ABPP protocols typically required cells and tissues to be homogenized first before treatment with probes. As a result, proteins were removed from their native cellular environment, and the effects of subcellular localization and endogenous modulators on protein function may be obscured. In addition, bulky tags may influence probe distribution within the cell. To address this limitation, the Cravatt group and the Overkleeft group independently proposed two-step labelling strategy using bio-orthogonal chemistry. The Cravatt group demonstrated that Cu(I)-catalyzed alkyne-azide cycloaddition reaction (CuAAC) could be applied to profile enzyme activity in whole proteome by synthesising a rhodamine-alkyne tag and an azide-derivatized phenyl sulfonate ester probe.[9] The authors later reported detailed characterization of reaction parameters that affect click chemistry–based ABPP.[10] The Overkleeft group applied a bio-orthogonal Staudinger ligation method developed by Bertozzi et al. to detect active proteasomes in living cells.[11,12] Hang et al. demonstrated that the Staudinger ligation enables the use of a single probe for both live-cell imaging and affinity enrichment of labelled proteins.[13]

While many ABPP probes are designed to target conserved nucleophilic residues in active sites, this approach cannot be directly applied to enzymes that lack such residues including aspartic and metalloproteases. As an alternative strategy to label the active site of γ-secretase, Li et al. installed a photo-crosslinker as well as a biotin tag onto a potent inhibitor (**Figure 1-1B**).[14] The authors demonstrated pull-down of γ-secretase from cell membrane by the photo-activatable probe. Similarly, Hagenstein et al. and Saghatelian et al. developed active site–directed photo-activatable

probes respectively for kinases and matrix metalloproteases.[15,16] Although photo-activatable probes had long been used in a variety of biological experiments, these studies introduced photo-activatable probes to ABPP and related chemical proteomics methods.[17]

Many ligand-directed (affinity-based) probes for various proteins of therapeutic interest have been synthesised (**Figure 1-1B**). Recent examples of targeted proteins include G protein–coupled receptors, kinases, nuclear receptors, and bromodomain-containing proteins to name a few.[18–23] The Hang group developed photo-activatable probes for nucleotide-binding oligomerization domain-containing protein 1 and 2 (NOD1 and NOD2).[24] The authors demonstrated pull-down of these receptors from intact cells by the photo-activatable probes. This study also revealed an unpredicted ligand-induced interaction of NOD2 and GTPases. These probes are expected to aid in small molecule–protein target engagement and selectivity profiling studies.

Non-directed probes and reactivity profiling

Active site–directed probes typically incorporate known binding groups or affinity labels. Yet, many proteins still lack known ligands. To extend ABPPs to these proteins, Sorensen et al. introduced non-directed strategy for probe discovery (**Figure 1-2A**).[25,26] The authors synthesised a library of biotinylated sulfonates bearing variable binding groups and evaluated their proteome reactivity. Specific protein targets were identified on the basis of heat-sensitive probe reactivity, which represents a simple and effective primary screen for activity-based protein labelling events, and/or selective reactivity with a subset of the probe library. Several enzyme cofactors were found to influence the sulfonate probe reactivity of specific proteins, which suggested that some of the sulfonate labelling events occurred in the active-site of cofactor-dependent enzymes.

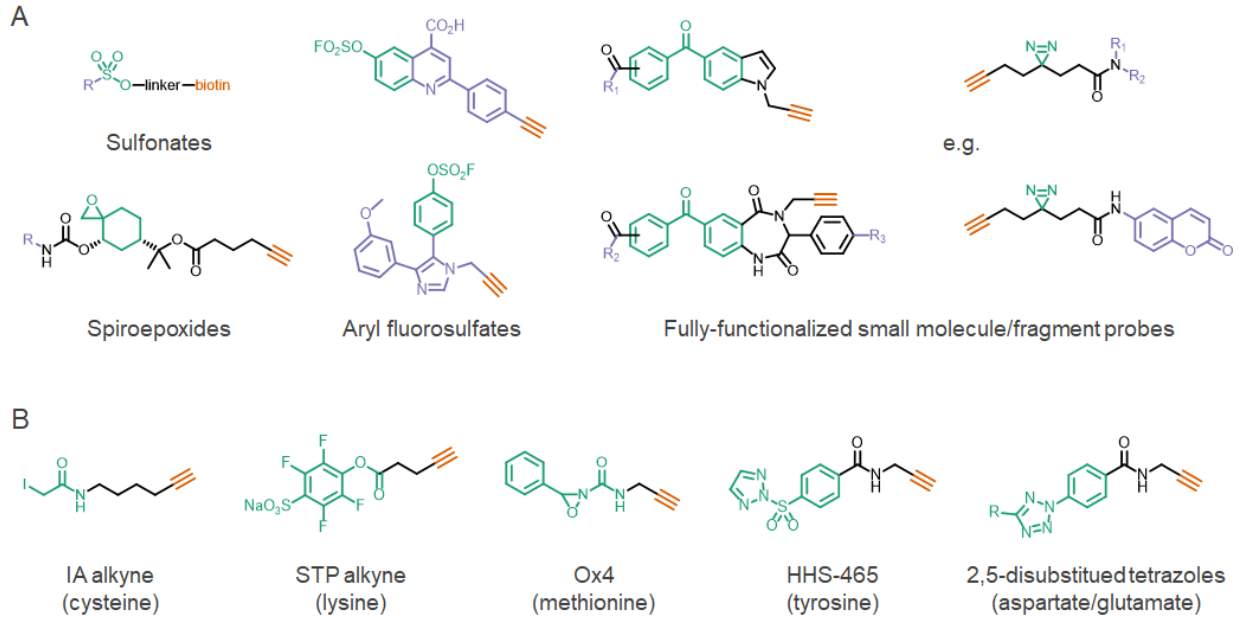


Figure 1-2. Non-directed probes and reactivity profiling. (A) Chemical structures of select non-directed probes. (B) Chemical structures of representative probes for reactivity profiling.

Insights into the proteome reactivity of electrophiles are crucial for designing activity-based probes, particularly for proteins lacking covalent ligands. In this regard, Weerapana et al. profiled proteome reactivity of carbon electrophiles.[27] The authors demonstrated different classes of carbon

electrophiles exhibit markedly distinct amino acid labelling profiles in the proteome. To achieve relative quantification of amino acid side chain reactivity, Weerapana et al. developed a method termed isotopic tandem orthogonal proteolysis–activity-based protein profiling (isoTOP-ABPP).[28] Key to the isoTOP-ABPP workflow is an isotopically labelled TEV protease cleavable linker which is conjugated to iodoacetamide (IA) alkyne probe via CuAAC to enable capture, isotope encoding, and release of probe-modified peptides. The ratio in signal intensities between light and heavy tag–conjugated peptides reflects the relative extent of labelling between two conditions. Alternatively, the cleavable linker can be replaced by isotopically-labelled desthiobiotin azide to simplify the protocol.[29] These studies revealed the enrichment of highly-reactive cysteine residues at functional sites in both mammalian and bacterial proteomes.

An important area of application for non-directed ABPP is target deconvolution studies. Early studies in this area incorporated protein-reactive compounds into chemical genomics screens as a means to discover targets of bioactive small molecules in living systems (**Figure 1-2A**).[30] Evans et al. synthesised a library of covalent chemical probes bearing natural product–inspired spiroepoxide reactive group and an alkyne handle. A member of the library was found to inhibit cell proliferation, and a target protein was identified on the basis of selective reactivity for this probe compound. This approach was later extended to non-covalent small molecules by constructing a library of photo-activatable probes bearing a photo-crosslinker and an alkyne handle. These photo-activatable probes were called fully-functionalised small molecule probes and used for chemical genomics screening and small molecule–protein interaction mapping.[31–36]

While library synthesis can be amenable to incorporating a bio-orthogonal handle, derivatisation of natural products often poses formidable challenges.[37] Furthermore, identifying, let al.one quantifying, the sites of modification of large natural products by mass spectrometry can be challenging. To address these limitations, Wang et al. introduced a competitive version of isoTOP-ABPP that enabled the quantification of the reactivity of electrophilic compounds against >1,000 cysteines in parallel in the human proteome.[38] In this modified isoTOP-ABPP workflow, the IA probe labelling follows treatment of proteome samples with electrophiles. Using this approach, the authors identified a set of cysteines that exhibit selective reactivity with a lipid-derived electrophile or another and constitute 'ligandable hot spots'.

The competitive isoTOP-ABPP workflow has also enabled large-scale proteome-wide covalent ligand discovery. Backus et al. massively extended the scope of 'ligandability profiling' by demonstrating a quantitative analysis of cysteine-reactive small-molecule fragments screened against thousands of proteins in human proteomes and cells.[39] The authors screened an ~ 50-member covalent fragment library and identified ligands for > 700 cysteines found in both druggable proteins and proteins that are challenging to address with non-covalently binding small molecules. Interestingly, several fragment electrophiles collectively covered large proportion of the ligandable cysteines. These fragment electrophiles were later referred to as 'scout fragments' and used to map ligandable cysteines in proteome.[40] Similarly, Vidogradova et al. recently reported reactivity profiling and ligandability profiling of cysteine residues in primary human T cells.[41] It is noteworthy that the authors complemented isoTOP-ABPP with a tandem mass tagging (TMT)–based ABPP method to balance quantitative accuracy and greater multiplexing capability.[42]

While the IA probe is most widely used for reactivity profiling of cysteine residues, studies reported the use of alternative probes can improve coverage or sensitivity.[43,44] Besides cysteine residues, the reactivity of other amino acid residues such as lysine, methionine, tyrosine, aspartic

acid, and glutamic acid has been profiled (**Figure 1-2B**).[2,11,45,46] In addition, the use of latent electrophiles for target-agnostic probe discovery was recently proposed.[47] These studies extend the scope of non-directed ABPP and facilitates target deconvolution of covalently-acting small molecules and covalent ligand discovery.

Label-free methods for target deconvolution and target engagement studies

ABPP and related chemical reporter–based methods require small molecules to be derivatised with reporter tag(s) (and a photo-crosslinker). On one hand, derivatisation enables enrichment of target proteins, which can be beneficial for low-abundance targets. On the other hand, derivatisation can be synthetically challenging and will almost certainly alter overall target profile.[48] While competitive ABPP can partly address this limitation by obviating the need for direct derivatisation, it still biases target profile to a set of proteins and sites that are accessible to the probe compound used in the experiment. Thus, complementary methods that work directly with underivatized molecules are desired.

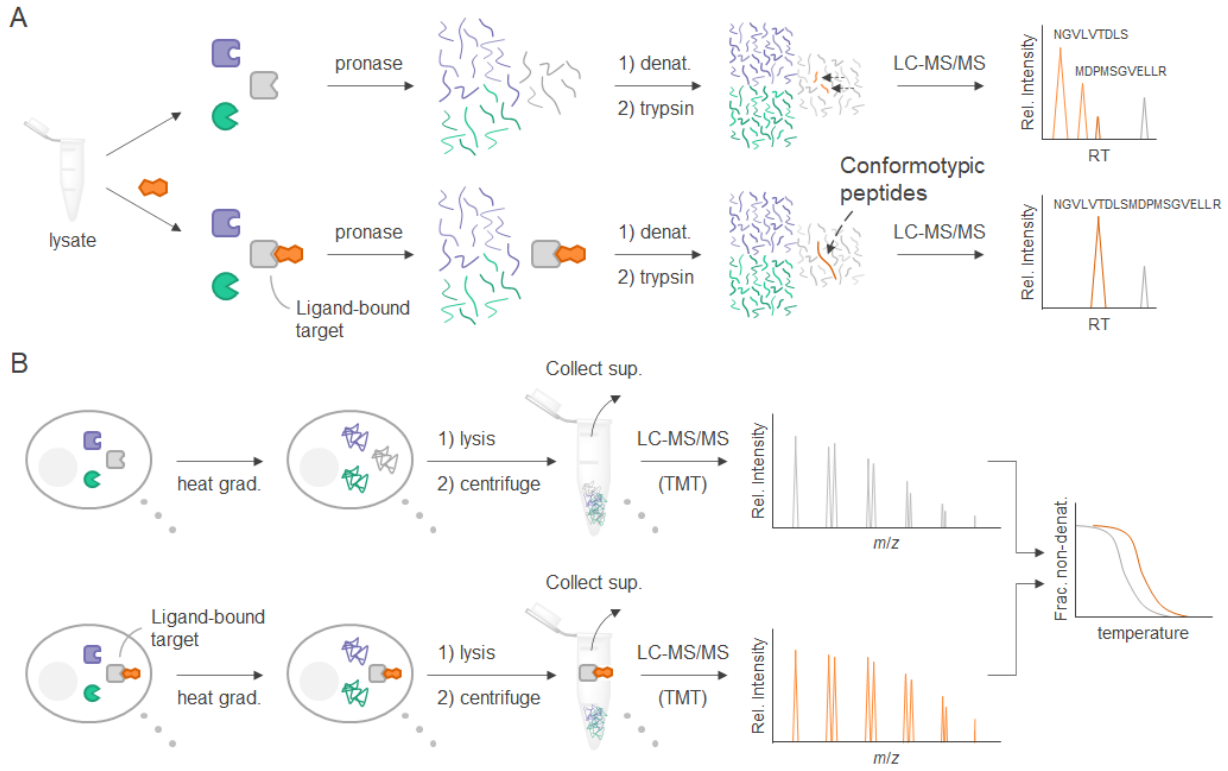


Figure 1-3. Label-free methods for target deconvolution and target engagement studies. (A) Limited proteolysis-small molecule mapping. (B) Thermal proteome profiling.

Label-free methods rely on compounds altering biophysical characteristics of the target proteins—particularly, thermal stability and proteolytic susceptibility. The proteolytic methods take advantage of a reduction in the proteolytic susceptibility (limited proteolysis) of the target proteins upon ligand binding (**Figure 1-3A**). The original protocol based on this principle is Drug Affinity–Responsive Target Stability (DARTS).[49] In DARTS, cell lysate is treated either with vehicle or drug, digested with proteases, and analysed by SDS-PAGE. Proteins that exhibit differential bands on the gel are deemed targets. Similarly, Piazza et al. reported a method that combined limited

proteolysis and mass spectrometry–based proteomics for globally mapping small molecule–protein interactions, which the authors termed limited proteolysis-small molecule mapping (LiP-SMap).[50] In the LiP-SMap workflow, cell lysate is exposed to vehicle or a small molecule of interest and subjected to proteolytic digestion using a broad-spectrum protease to generate structure-specific protein fragments. Fragments are then digested with trypsin to generate peptide mixtures amenable to bottom-up proteomics. Peptides are analysed by LC-MS/MS, and proteomes are compared between the absence and presence of the small molecule using label-free quantification. Proteins with differential proteolytic patterns are considered putative targets. The authors benchmarked LiP-SMap in bacterial cell lysate that are exposed to different metabolites and found the method was capable of mapping known interactions, although higher-than-physiological concentrations of metabolites were needed to detect weak interactions. Using the LiP-SMap workflow, the authors mapped > 1600 putative interactions collectively for 20 metabolites in gel-filtered bacterial cell lysate and inferred binding sites from conformotypic peptides. Recently, the Picotti and the Reiter groups adapted LiP-SMap for complex eukaryotic proteomes by employing a machine learning framework (LiP-Quant).[51] It is noteworthy that these proteolytic methods can map binding sites at peptide-level resolution, which remains challenging with photo-activatable probes.

Alternatively, thermal proteome profiling (TPP) also known as cellular thermal stability assay followed by MS (CETSA MS) takes advantage of altered (increased or reduced) thermal stability of the target proteins upon ligand binding (**Figure 1-3B**). Thus, CETSA is similar to conventional thermal shift assays (TSA) except that CETSA can be performed in complex protein samples and living cells while TSA is performed in purified recombinant proteins. In the original CETSA protocol, multiple aliquots of cells or cell lysate are exposed to vehicle or a small molecule and heated to different temperatures.[52] After cooling, the samples are centrifuged to separate soluble fractions from precipitated proteins. The presence of the target protein in the soluble fraction is quantified by Western blotting. Protein melt curves can shift in either direction upon ligand binding. Isothermal dose–response experiment can be used to estimate relative ligand potencies. While CETSA has been primarily applied to soluble and solubilised proteins, a robust protocol for membrane proteins was recently published.[53,54] TPP combines CETSA and TMT-based quantitative mass spectrometry to study the effect of small molecules on the thermal profile of a proteome.[55] The authors validated the ability of TPP to identify binding targets using broad-specificity kinase inhibitors, which induced shifts in melting temperatures of many kinase targets and other proteins, including regulatory subunits of kinase complexes. TPP has been used for several target identification studies.[56] In addition, a method that combines TPP with the selective labelling and enrichment of sialylated glycoproteins was recently reported.[57] Future method development studies will continue to extend the scope of TPP.

CHAPTER 2. NUCLEAR RECEPTOR CHEMICAL REPORTER ENABLES DOMAIN-SPECIFIC ANALYSIS OF LIGANDS IN MAMMALIAN CELLS

Determining the mechanisms by which intestinal microbiota modulate host metabolism and immunity is important for understanding the host–microbiota interactions and would help develop new therapeutic strategies towards metabolic and inflammatory diseases (**Figure 2-1**). Recent studies highlighted associations between altered composition of intestinal microbiota and multiple metabolic and inflammatory diseases including type 2 diabetes mellitus and inflammatory bowel disease[58,59]. However, molecular mechanisms through which intestinal microbiota influence homeostasis remains largely unknown.

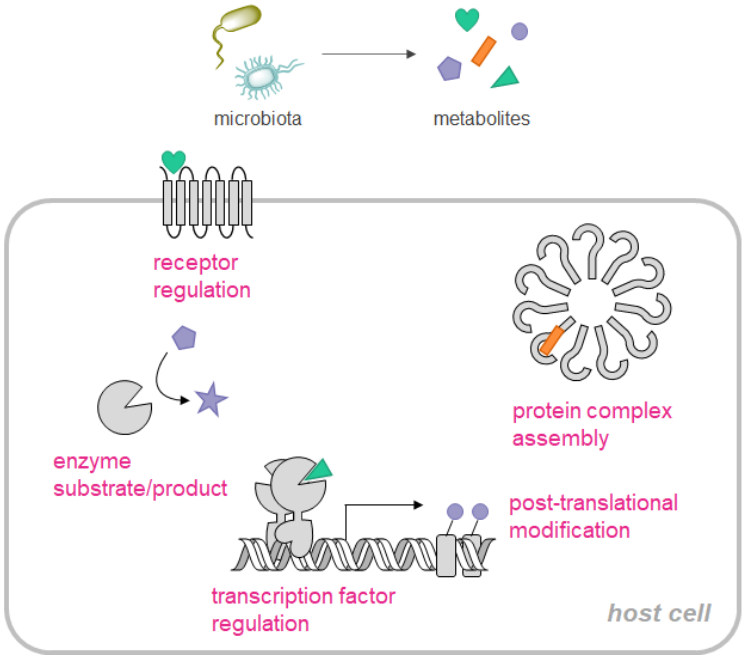


Figure 2-1. Mechanisms of action for microbiota-derived metabolites.

Recent studies identified crucial roles of innate immune recognition of microbial surface molecules and nucleotides in the maintenance of intestinal homeostasis. Deficiency in pattern recognition receptor (PRR) signalling alters composition of intestinal microbiota and disposes hosts to inflammatory pathologies[60,61]. For example, polymorphisms in nucleotide-binding oligomerization domain containing 2 (NOD2) are associated with increased risk of Crohn's disease[62].

In addition to PRR signalling, intestinal microbiota modulates host metabolism and immunity via diet-dependent production of various nutrients and metabolites including amino acids, bile acids, fatty acids, and vitamins[59,63]. Production of these small molecules can influence host physiology via alterations in the availability or the use of energy substrates or via specific receptors that can activate signal-transduction pathways and transcriptional programmes. Fermentation of dietary fibre by intestinal bacteria yields short-chain fatty acids (SCFAs) such as acetate, propionate, and butyrate, which serve as energy source for colonic epithelial cells and augment barrier function[64]. SCFAs also modulate development and function of immune cells by activating specific G-protein-coupled receptor signalling[65,66] or by inhibiting histone deacetylases[67]. Several bacterial metabolites of dietary tryptophan can support mucosal integrity via multiple pathways[68–71].

Nuclear receptors (NRs) are ligand-activated transcription factors that regulate diverse functions such as homeostasis, reproduction, development, and metabolism[72,73]. Humans express 48 NRs, which share a common modular structure consisting of a regulatory domain referred to as activation function-1 (AF-1) domain at the N-terminus, a DNA binding domain in the middle, and a ligand binding domain (LBD) at the C-terminus. LBD–ligand interaction regulates transcriptional activity of most NRs. Ligands of NRs are typically small lipophilic molecules such as steroids, thyroid hormones, retinoids, and fatty acids that readily traverse the plasma membrane. At least three NRs are so far known to mediate activity of microbiota-derived metabolites: farnesoid X receptor (NR1H4, FXR)[74], pregnane X receptor (NR1I2, PXR)[71], and peroxisome proliferator–activated receptor γ (NR1C3, PPARγ)[75,76].

PPARγ was initially characterised as the master regulator of adipogenesis[77]. It controls fatty acid metabolism and inflammation in various types of cells. Recent studies linked intestinal microbiota and PPARγ by demonstrating that intestinal bacteria regulate its activity[78,79] and attenuate inflammation[75,80,81]. PPARγ also appears to be required for the maintenance of symbiosis[76,82,83,84]. Other NRs such as hepatocyte nuclear factor-4α (NR2A1, HNF4α)[85] may also mediate activity of microbiota-derived metabolites. However, it is unclear if these and other NRs directly bind and sense dietary and microbiota-derived metabolites. As PPARγ and NRs are therapeutic targets for metabolic and inflammatory diseases, identification of microbiota-derived ligands of these receptors may help rationally select probiotic bacteria for prevention and treatment of relevant diseases and may provide novel drug candidates.

Development of alk-GW9662 for target engagement analysis of PPARγ ligands

We developed a PPAR chemical reporter to evaluate target engagement of candidate ligands in live cells using bio-orthogonal detection methods (**Figure 2-2**). Herein, target engagement is defined as small molecule–target protein binding event in cells but is occasionally used to refer to such binding events in cell lysate.[86] We designed a chemical reporter alk-GW9662 for PPARs based on a covalent antagonist of PPARγ, GW9662 (**Figure 2-2A**).[87] The chloronitrobenzene moiety is a reactive warhead and undergoes nucleophilic aromatic substitution with sulfhydryl group of the cysteine residue inside the ligand binding pocket.[87] The chloronitrobenzene is less reactive than iodoacetamide and its reactivity is finely tunable, which we considered desirable for our targeted approach (**Figure 2-3**). GW9662 also covalently modifies ligand binding domain of PPARα and PPARβ/δ, presumably by reacting with the conserved cysteine residues (**Figure 2-4**). In addition, a recent study confirmed that GW9662 covalently targets an analogous cysteine residue in RORγt.[88] Previous studies have explored activated aryl-halides as a reactive warhead targeting proteome-wide cysteine and lysine residues, but these reagents failed to target PPARs.[44,89] Herein we demonstrate the chemical reporter can probe PPAR ligand binding pocket.

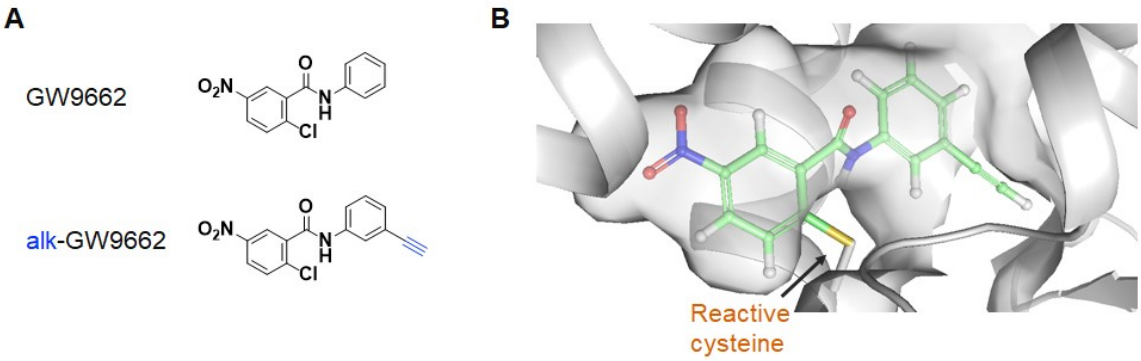


Figure 2-2. Targeted analysis of PPARγ–ligand interactions using alk-GW9662. (A) Chemical structures of GW9662 and alk-GW9662. (B) Covalent docking of alk-GW9662 in ligand binding pocket of PPARγ.

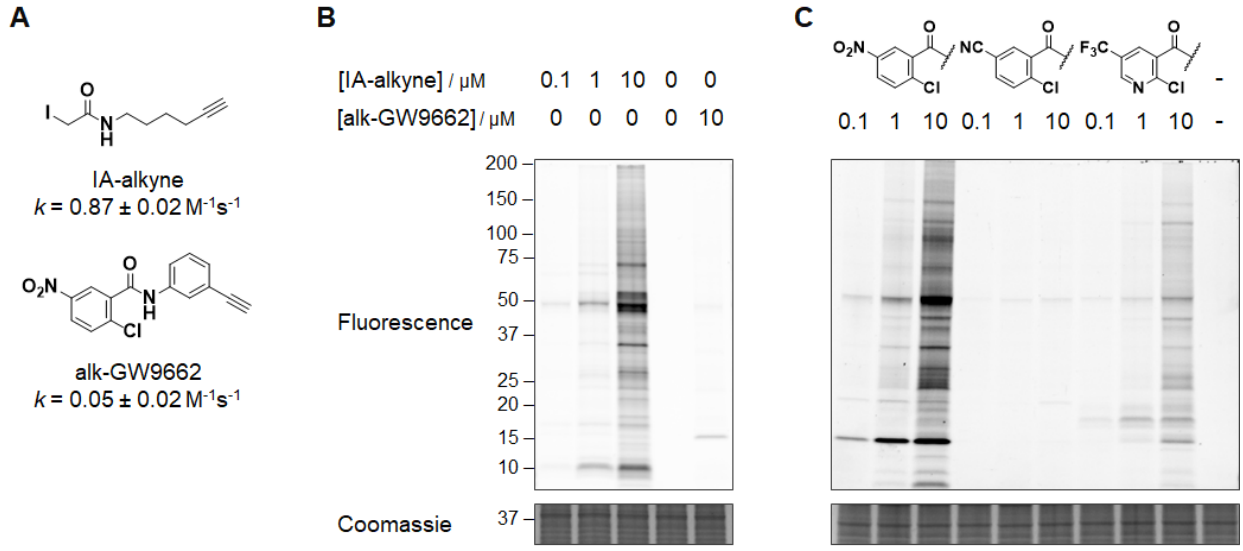


Figure 2-3. Chloronitrobenzene is less reactive than iodoacetamide. (A) Chemical structures and second-order reaction rate constants of IA-alkyne and alk-GW9662. The rate constants were determined by DTNB assay. (B) Proteome labelling by IA-alkyne and alk-GW9662. (C) Reactivity tuning of electrophilic chloroaryls.

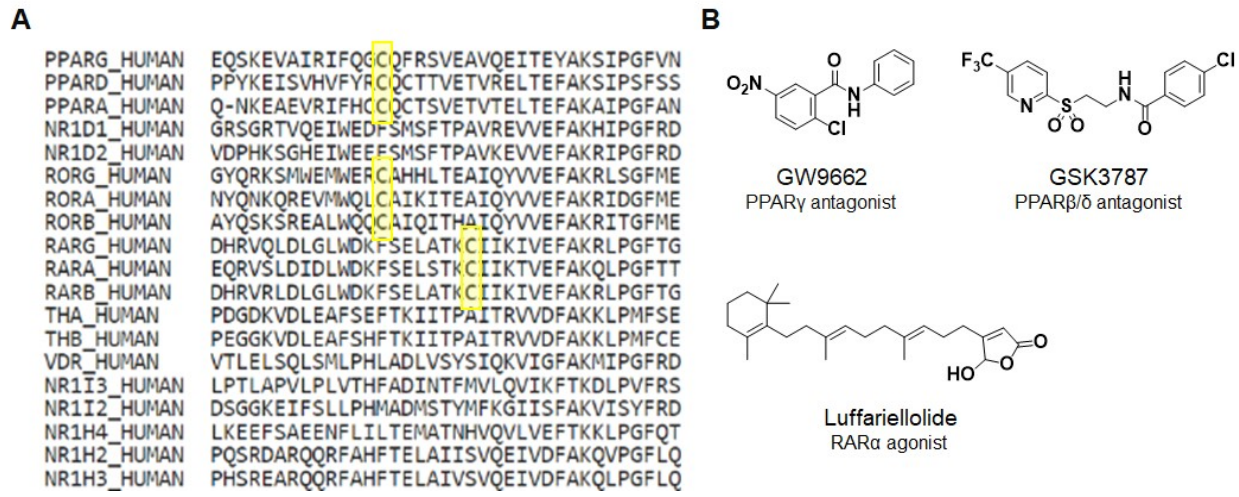


Figure 2-4. Conserved cysteine residues in NR LBD can be covalently targeted. (A) Sequence alignment of NR1 subfamily. (B) Known covalent ligands of NRs.

To achieve bio-orthogonal detection of covalently labelled PPARs in cells, we appended an alkyne on GW9662 as this phenyl ring can tolerate modification (**Figures 2-5, Figure 2-6A**).[90–92] To evaluate site-specific labelling of PPARγ, we transfected HEK293T cells with either wild-type or C283A mutant and treated with varying concentrations of alk-GW9662 (**Figure 2-6B**). After incubation for 1 h proteins were harvested, conjugated with azide-rhodamine B via Cu^{I}-catalyzed azide-alkyne cycloaddition (CuAAC) reaction and separated by SDS-PAGE. In-gel fluorescence profiling revealed a dose-dependent band at 50 kDa in the wild-type PPARγ samples, but not in the mock-transfection and C283A samples. In addition, GW9662 dose-dependently inhibited alk-GW9662 PPARγ labelling (**Figure 2-6C**). We also confirmed alk-GW9662 labelled both PPARα and PPARβ/δ dependently on the conserved cysteine residues using corresponding cysteine-to-serine mutants (**Figure 2-7**), which were functional in cell-based transactivation assays (**Figure 2-8**). These results support that alk-GW9662 site-specifically labels PPAR ligand binding domains in live cells.

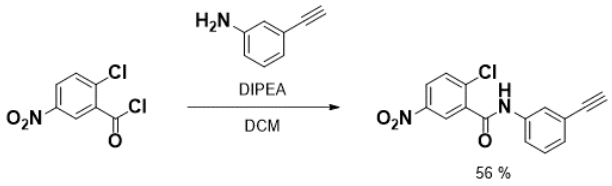


Figure 2-5. Synthesis of alk-GW9662.

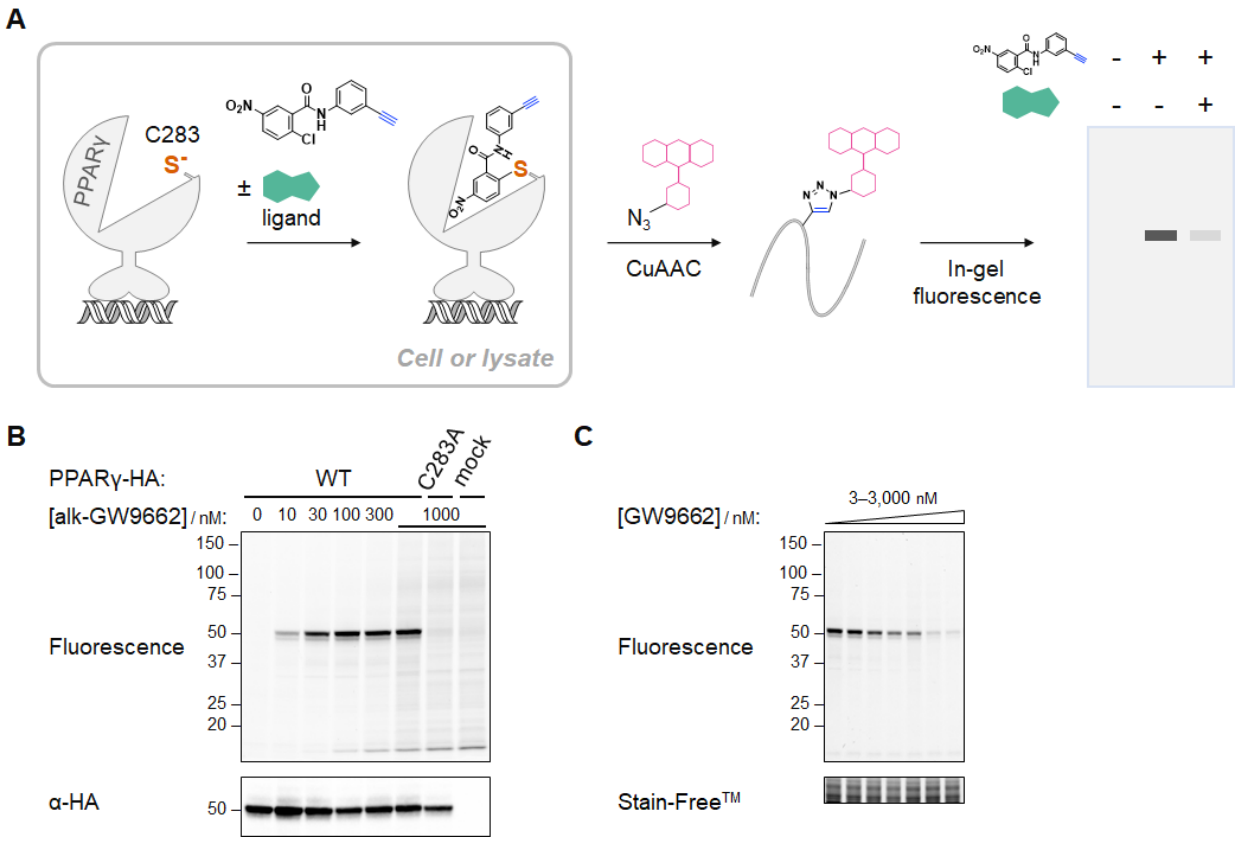


Figure 2-6. Alk-GW9662 site-specifically labels PPARγ C283. (A) Scheme describing the alk-GW9662 PPARγ labelling experiment. (B) In-gel fluorescence analysis of alk-GW9662-labelled proteins in HEK293T cells transfected with PPARγ. (C) Dose-dependent competition of alk-GW9662 labelling by GW9662.

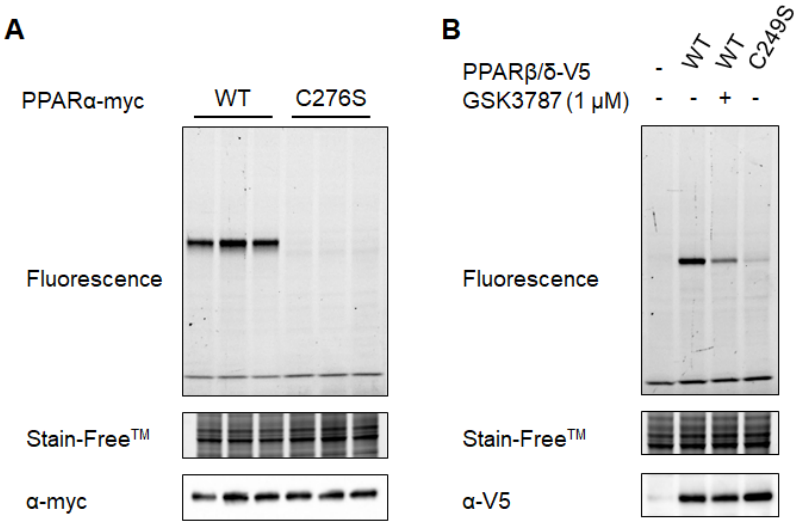


Figure 2-7. Alk-GW9662 site-specifically labels PPAR LBD. (A) PPARα. (B) PPARβ/δ. GSK3787 is a covalent PPARβ/δ ligand (Figure 2-3).

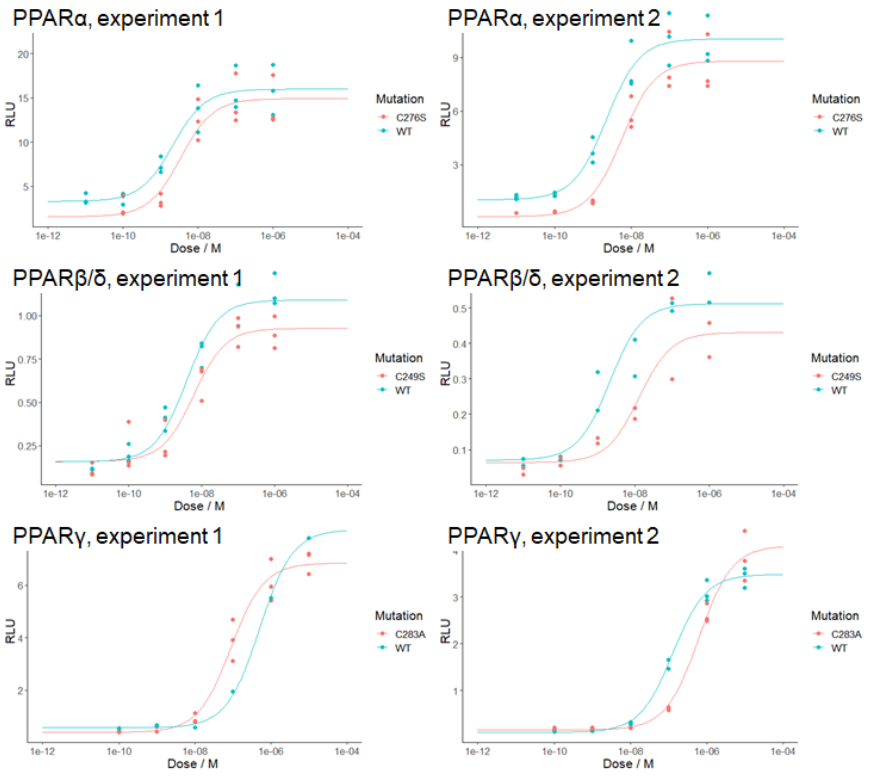


Figure 2-8. The cysteine-to-alanine mutants are functionally equivalent to wild types.

Having demonstrated alk-GW9662 site-specifically labels PPAR ligand binding domains in live cells, we evaluated its utility to determine the occupancy of PPARγ ligand binding pocket. Given that the conserved cysteine residue is functionally dispensable and that there is no known SNP for this cysteine residue, we were primarily interested in non-covalent ligands. To this end, we incubated cells with alk-GW9662 (0.1 μM) and varying concentrations of rosiglitazone, a non-covalent PPARγ ligand that occupies the same binding pocket as GW9662, and quantified alk-GW9662 labelling (**Figure 2-9**). Fluorescence intensity decreased as the dose of rosiglitazone increased, suggesting alk-GW9662 labels a saturable pocket. To establish alk-GW9662 labelling is affinity-dependent, we incubated cell lysate and cells with varying concentrations of drug-like ligands and alk-GW9662 (**Figure 2-10**). We determined pIC_{50} values by fitting sigmoidal curve to fluorescence intensity. The strong correlation between the pIC_{50} value and *in vitro* binding affinity (pK_i value) reported in literature indicate alk-GW9662 labelling indeed provides a quantitative readout of binding despite the competing ligands binding PPARγ non-covalently with varying degree of binding site overlap.[93–95] These results demonstrate alk-GW9662 enables quantitative assessment of PPARγ engagement in cells.

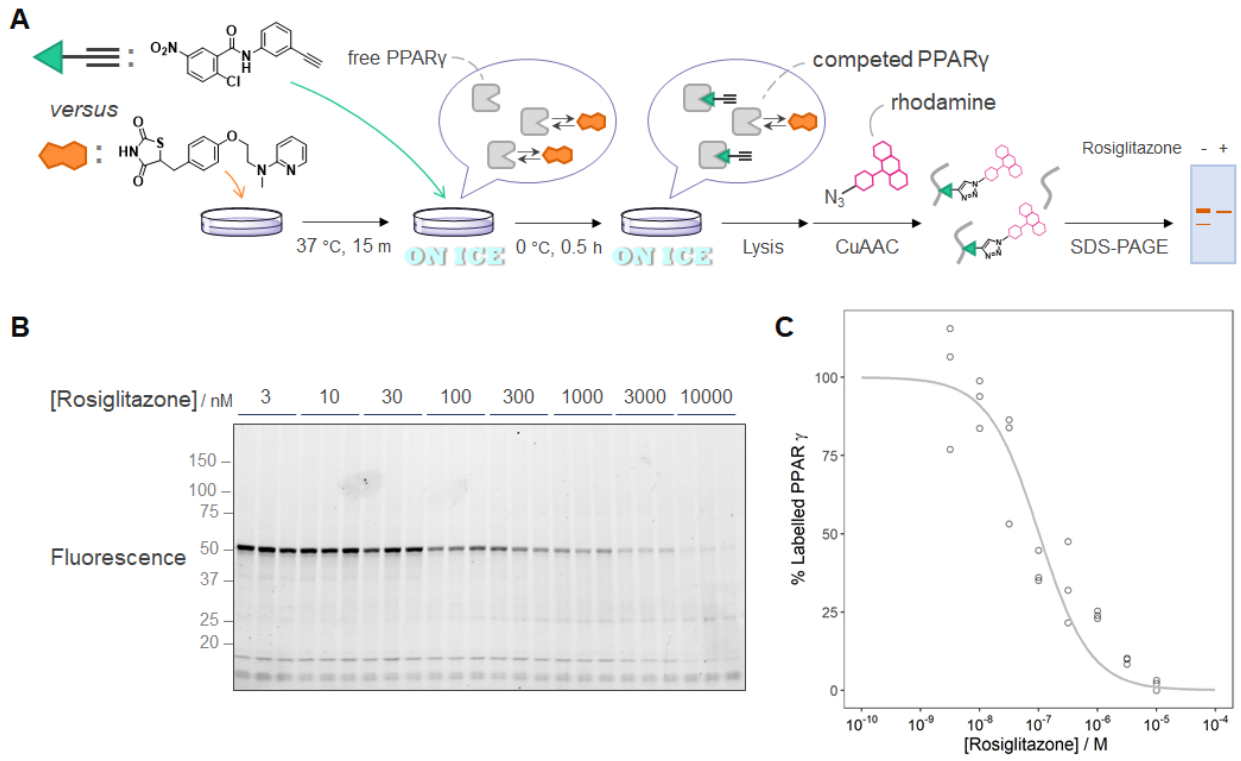


Figure 2-9. Competitive labelling strategy for target engagement analysis. (A) Scheme describing the alk-GW9662 PPARγ competitive labelling experiment. (B) Dose-dependent competition of alk-GW9662 labelling by GW9662. (C) Quantification of the gel shown in B.

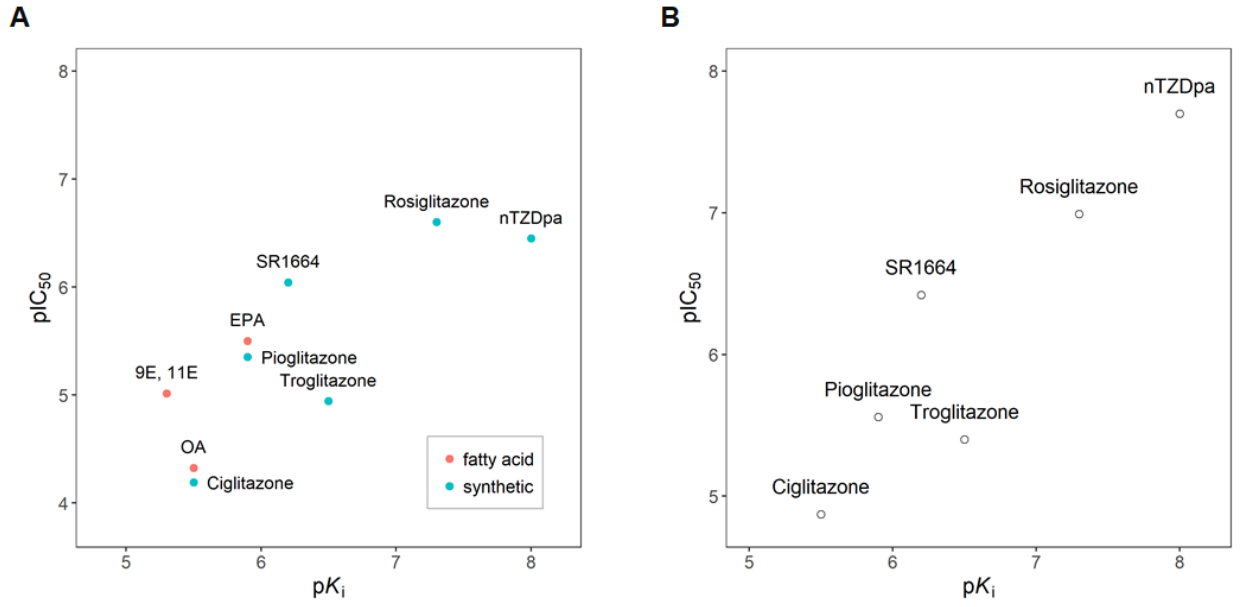


Figure 2-10. Comparison of in-lysate and in-cell competition IC_{50} *versus in vitro* binding affinity. (A) Comparison of in-lysate competition IC_{50} and *in vitro* binding affinity. (B) Comparison of in-cell competition IC_{50} and *in vitro* binding affinity. Pearson's r = 0.91.

Note that labelling conditions were optimized for reproducibility and signal/noise ratio. Thus, in-lysate labelling was performed at 25 °C using a heat block (**Figure 2-11**), and in-cell competitive labelling was performed on ice (**Figure 2-12**).

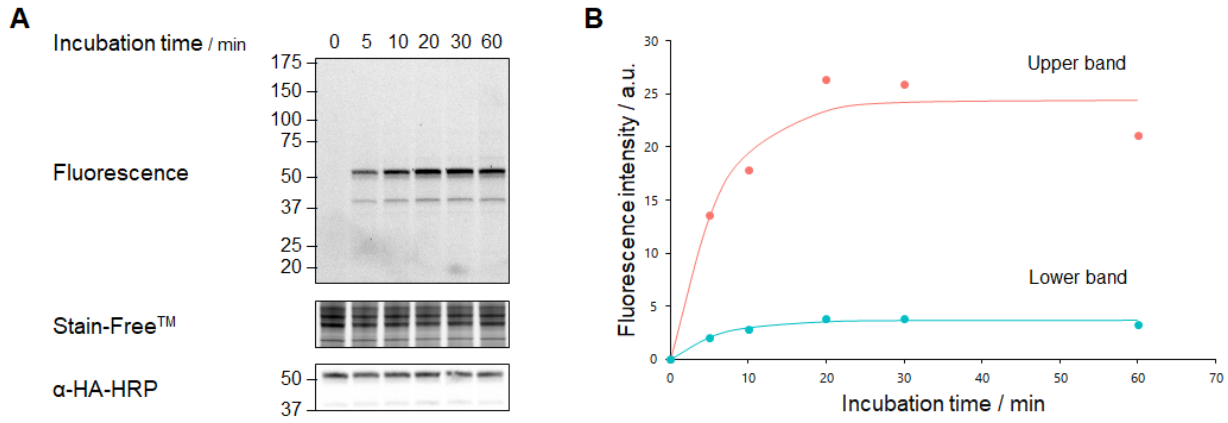


Figure 2-11. Kinetic analysis of in-lysate PPARγ labelling. (A) In-gel fluorescence analysis of labelling kinetics. (B) Quantification of the gel shown in A.

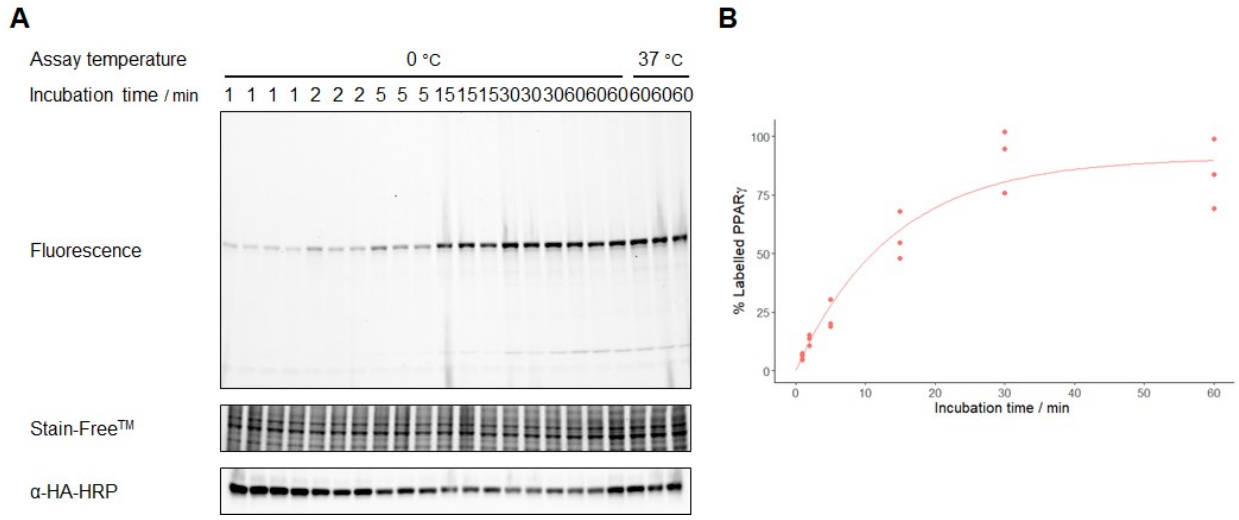


Figure 2-12. Kinetic analysis of in-cell PPARγ labelling. (A) In-gel fluorescence analysis of labelling kinetics. (B) Quantification of the gel shown in A.

PPARγ engagement analysis of microbiota-derived metabolites using alk-GW9662

With in-cell PPARγ engagement assay in hands, we next screened a panel of select microbiota-derived metabolites (**Figure 2-13A**). We hypothesized that PPARγ could mediate anti-inflammatory effects of microbiota-derived metabolites given the immune-regulatory role of PPARγ and recent reports implicating NRs in host–microbiota interactions.[71,74,76,96] Our panel consisted of conjugated linoleic acids (CLAs), cyclopropane fatty acids (CPFAs), tryptophan metabolites (indoles), and short-chain fatty acids (SCFAs). Among these compounds, 9E, 11E (CLA) inhibited alk-GW9662 labelling more than 50 % in cells when tested at 10 µM. In contrast, its geometric isomers failed to reach this threshold despite a previous study reporting similar *in vitro* binding affinities for these compounds, indicating *in vitro* binding affinity fails to translate

to target engagement in cells for this group of compounds.[97] Given literature precedents associating anti-inflammatory effects of CLAs with PPARγ in intestines, we decided to preliminarily characterize 9E, 11E.[75,98] In a confirmatory in-cell target engagement experiment, 9E, 11E exhibited pIC_{50} value of 5.3, which is comparable to the level of target engagement by pioglitazone (**Figure 2-13B**). Previous studies, however, reported marginal receptor activation even at concentrations as high as 100 μM.[99] Thus, we hypothesized that 9E, 11E would be a partial agonist and analyzed dose-responsiveness of transactivation assay (**Figure 2-13C**). Our dose-response analysis indicated pEC_{50} value of 4.7 with maximal receptor activation being 24 % of a full agonist rosiglitazone. We also incubated cells with EC_{80} concentration of rosiglitazone in the presence and absence of 9E, 11E and observed significant reduction of receptor activation when 9E, 11E was present (**Figure 2-13D**). These results suggest 9E, 11E could be a microbiota-derived partial agonist of PPARγ.

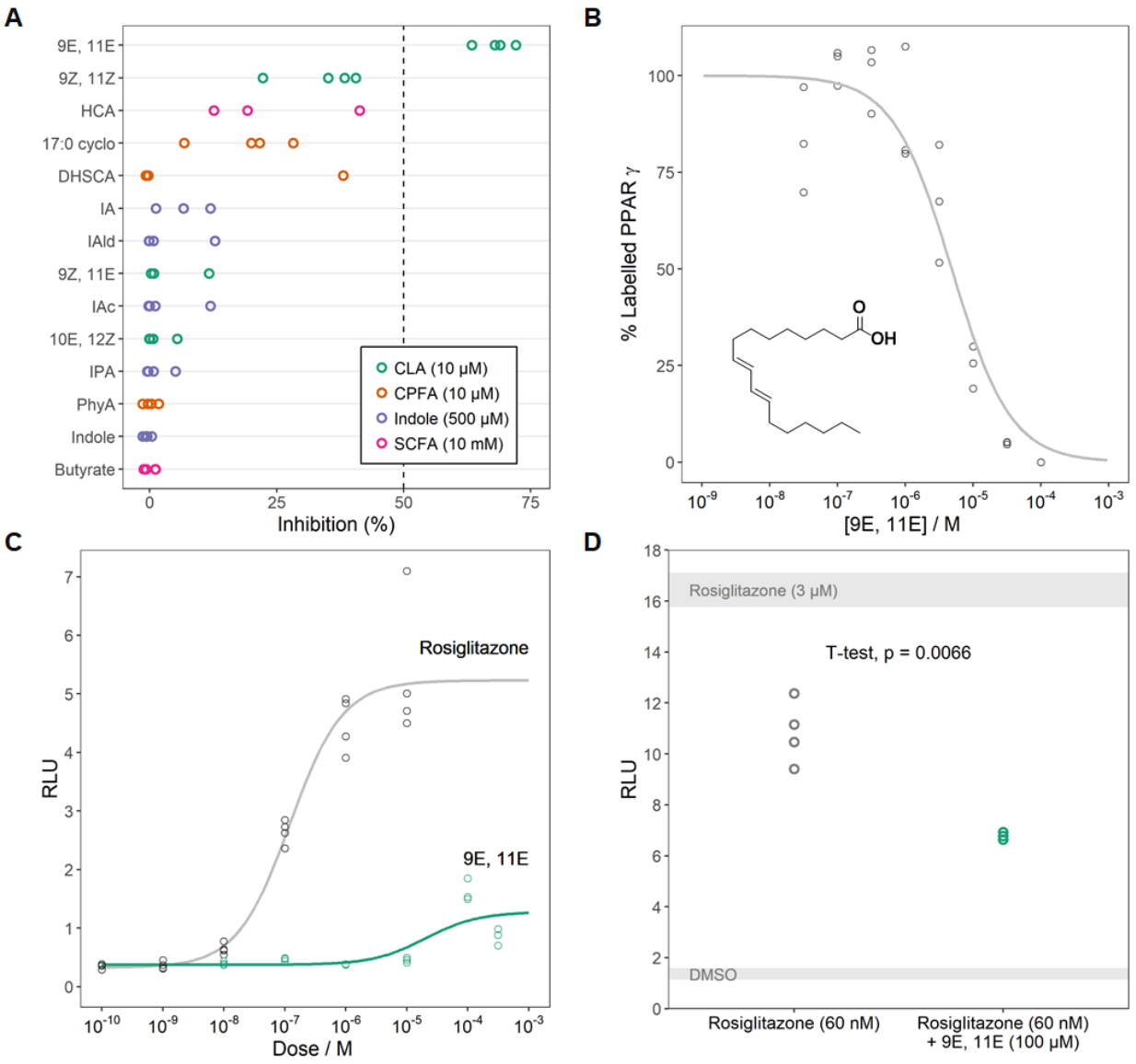


Figure 2-13. PPARγ engagement analysis of microbiota-derived metabolites in cells. (A) Microbiota-derived metabolite screening. (B) Dose-dependent inhibition of alk-GW9662 labelling by 9E, 11E. (C) Cell-based transactivation assay with increasing concentrations of 9E, 11E. (D) Transactivation assay with rosiglitazone (EC_{80} = 60 nM) ± 9E, 11E (100 µM). Grey bands indicate baseline (DMSO) and maximum (rosiglitazone, 3 µM) activations and bandwidth corresponds to standard deviation. Refer to **Table 2-1** for abbreviations used.

Table 2-1. Abbreviated names of microbiota-derived metabolites.

Abbreviation	Formal name	SMILES
9E, 11E	9E,11E-octadecadienoic acid	CCCCCC\C=C\C=C\CCCCCCCC(=O)O
9Z, 11Z	9Z,11Z-octadecadienoic acid	CCCCCC/C=C\C=C/CCCCCCCC(O)=O
HCA	(*E*)-3-(4-hydroxyphenyl)prop-2-enoic acid	C1=CC(=CC=C1/C=C/C(=O)O)O
17:0 cyclo	8-(2-hexylcyclopropyl)octanoic acid	CCCCCCC1CC1CCCCCCCC(=O)O
DHSCA	(1R,2S)-*rel*-2-octyl-cyclopropaneoctanoic acid	OC(CCCCCCC[C@H](C1)[C@H]1CCCCCCCC)=O
IA	2-(1H-indol-3-yl)acetic acid	C1=CC=C2C(=C1)C(=CN2)CC(=O)O
IAld	1*H*-indole-3-carbaldehyde	C1=CC=C2C(=C1)C(=CN2)C=O
9Z, 11E	9Z,11E-octadecadienoic acid	CCCCCC/C=C/C=C\CCCCCCCC(O)=O
IAc	(*E*)-3-(1*H*-indol-3-yl)prop-2-enoic acid	C1=CC=C2C(=C1)C(=CN2)/C=C/C(=O)O
10E, 12Z	10E,12Z-octadecadienoic acid	CCCCC/C=C\C=C\CCCCCCCCC(=O)O
IPA	3-(1*H*-indol-3-yl)propanoic acid	C1=CC=C2C(=C1)C(=CN2)CCC(=O)O
PhyA	2-hexyl-cyclopropanedecanoic acid	CCCCCCC1CC1CCCCCCCCCC(=O)O
Indole	1*H*-indole	C1=CC=C2C(=C1)C=CN2
Butyrate	sodium butanoate	CCCC(=O)[O-].[Na+]
HIA	2-(5-hydroxy-1*H*-indol-3-yl)acetic acid	C1=CC2=C(C=C1O)C(=CN2)CC(=O)O
MIA	2-(5-methoxy-1*H*-indol-3-yl)acetic acid	COC1=CC2=C(C=C1)NC=C2CC(=O)O
Trp	(2*S*)-2-amino-3-(1*H*-indol-3-yl)propanoic acid	C1=CC=C2C(=C1)C(=CN2)C[C@@H](C(=O)O)N
5-HT	3-(2-aminoethyl)-1*H*-indol-5-ol	C1=CC2=C(C=C1O)C(=CN2)CCN

We were also interested in the tryptophan metabolites because these compounds have been associated with microbiota regulation of host immunity[100,101]. None of the tryptophan metabolites inhibited alk-GW9662 labelling in cells even though structurally related metabolites such as 5-methoxyindole-3-acetate (MIA) and 5-hydroxyindole-3-acetate (HIA) were previously reported to bind PPARγ *in vitro*.[102] Given the presence of a carboxylic acid group, which is likely negatively charged under physiological conditions, we hypothesized these compounds failed to pass through cell membranes. To test this hypothesis, we repeated the screening experiment in cell lysate (**Figure 2-14A**). Indoles with a carboxylic acid group inhibited alk-GW9662, suggesting the carboxylic acid groups mediate key interactions. In a confirmatory experiment, we determined the pIC_{50} value of indole propionic acid to be 3.3, which is within the expected range of affinity for typical fragment hits (**Figure 2-14B**). In fact, a previous study identified the structurally related compound 5-methoxyindole-3-propionate in a fragment screening.[103] These results demonstrate alk-GW9662 labelling can be used in PPARγ engagement analysis of low-affinity metabolites in lysate.

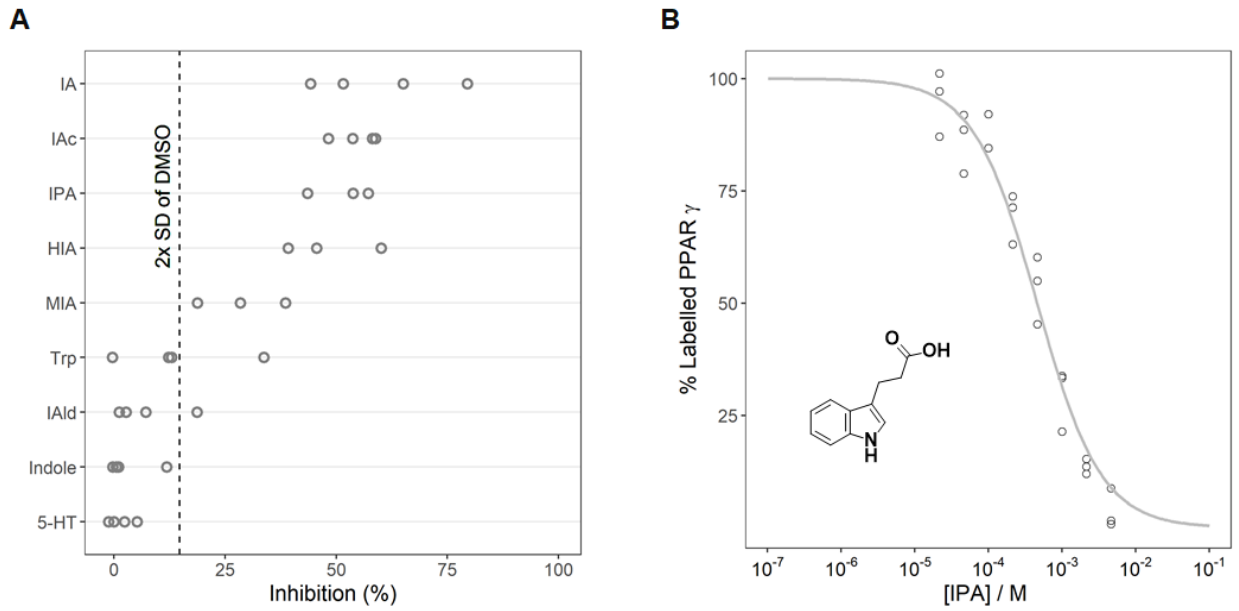


Figure 2-14. PPARγ engagement analysis of tryptophan metabolites in lysate. (A) Tryptophan metabolite screening in lysate. (B) Dose-dependent inhibition of alk-GW9662 labelling by IPA. Refer to **Table 2-1** for abbreviations used.

Repurposing alk-GW9662 for other proteins

The alk-GW9662 PPARγ engagement assay requires transfection and over-expression of PPARγ, which may not recapitulate small molecule interactions of endogenous PPARγ. Besides, chemical tools are most useful when they are used in cells and tissues that are less amenable to genetic techniques. To test the feasibility of using alk-GW9662 to assess endogenous—as opposed to transfected—PPARγ, we incubated HT-29 cells with varying concentrations of alk-GW9662 (**Figure 2-15A**). After incubation for 1 h proteins were harvested, conjugated with azide-rhodamine B and separated by SDS-PAGE. In-gel fluorescence analysis revealed a number of bands, which suggested alk-GW9662 can target several proteins in the absence of over-expressed PPARγ and precluded gel-based competition assay. To confirm alk-GW9662 labels endogenous PPARγ, we performed a pull-down assay (**Figure 2-15B**). Labelled proteins were conjugated with azide-PEG3-biotin, enriched on Neutravidin beads, and analysed by Western blot using anti-PPARγ antibody. The Western blot analysis confirmed that alk-GW9662 indeed labels endogenous PPARγ and enables its enrichment (**Figure 2-15C**). Unfortunately, the absence of promising candidate ligands did not merit further optimization of alk-GW9662 and/or the assay protocol. Instead, we explored repurposing alk-GW9662 for other proteins.

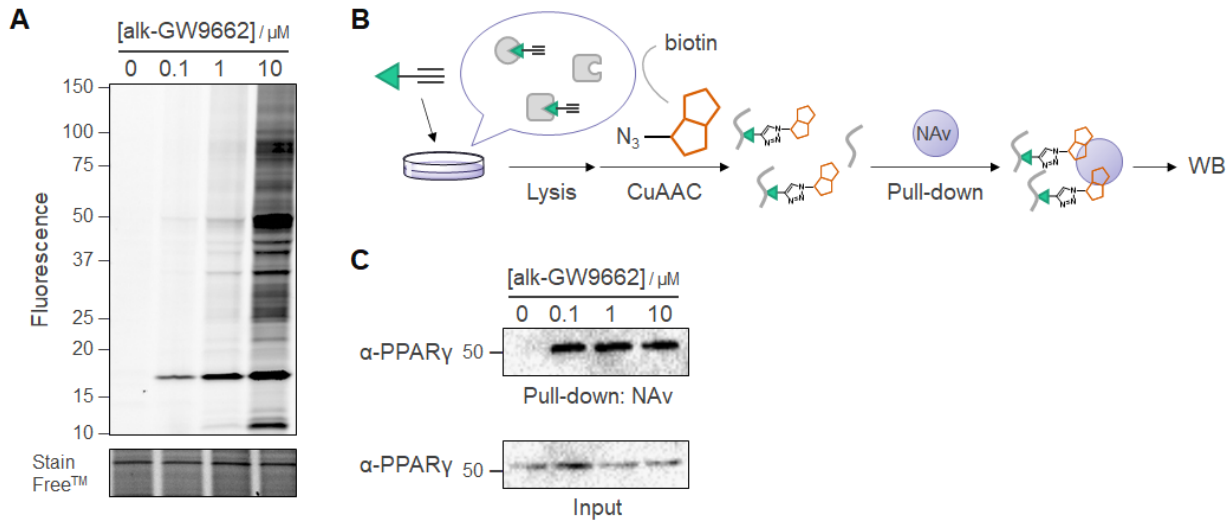


Figure 2-15. Alk-GW9662 labels endogenous PPARγ in HT-29 cells. (A) In-gel fluorescence analysis of alk-GW9662-labelled proteins. (B) Scheme of pull-down assay. (C) Alk-GW9662-mediated pull-down of PPARγ.

To identify alk-GW9662 targets, we performed an LC-MS/MS analysis of labelled proteins (**Figure 2-16A**). As a preliminary experiment, HT-29 cells were treated with vehicle or alk-GW9662 (10 µM). Proteins that were identified with at least two peptides were quantified (label-free quantification) and used for differential abundance analysis comparing vehicle- and probe-treated samples. Overall, 243 proteins satisfied fold-change threshold ($\log_2$FC > 1) and significance threshold (Benjamini Hochberg–adjusted p-value < 0.05). These proteins included PPARγ. In addition, alk-GW9662 targeted 71 more proteins that are considered druggable and/or have been liganded (**Figure 2-16B**).[104,105] The vast majority of targets (171 / 243), however, are neither considered druggable nor liganded. It is noteworthy that only three of the alk-GW9662 targets (RTN3, ALDH2, and RPS5) are known to possess hyper-reactive cysteines.[28] These results suggest potential repurposing of alk-GW9662 in an effort to ligand these challenging proteins.

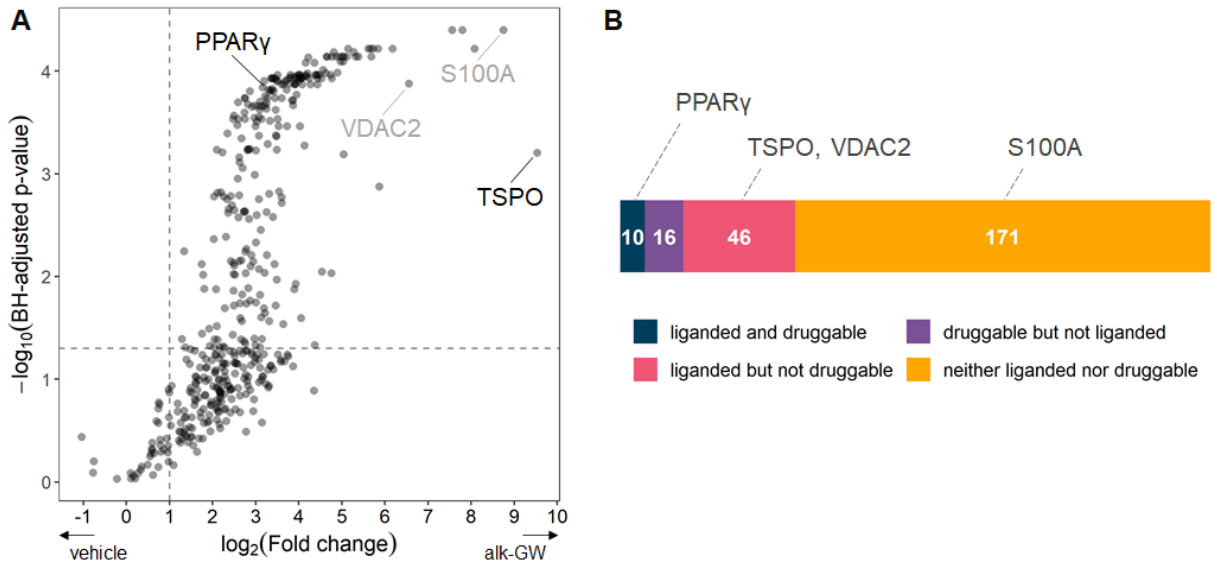


Figure 2-16. Proteome-wide analysis of alk-GW9662 targets. (A) Volcano plot of fold-enrichment using alk-GW9662 (10 μM) compared to vehicle versus significance (empirical Bayes t-test, Benjamini-Hochberg correction, n = 3). (B) Druggability profile of target proteins.

Differential abundance analysis is suitable for experiments where log-fold change is likely zero for most proteins (e.g., probe alone versus probe plus competitor). This null hypothesis, however, is unplausible for current experiment where most peptides/proteins were expected to be only observed in probe-treated samples (**Figure 2-15A**). Indeed, the majority of the peptides were missing from the vehicle samples (**Figure 2-17**). In addition, reliable protein quantification requires at least two peptides from a protein.[106] As a result of this heuristics, proteins that are supported by only a single, even if reliable, are discarded. Thus, the differential abundance analysis is under-powered. Alternatively, 366 proteins that are identified with at least one peptide (Protein FDR < 1 %) and only observed in probe-treated samples can be considered hits. These proteins included at least four nuclear receptors (**Figure 2-18**) and 42 druggable proteins (**Figure 2-19** & **Table 2-2**). These results suggest the possibility of collectively characterizing multiple nuclear receptors by alk-GW9662. Regardless of data analysis approaches taken, a critical design flaw of the experiment is the use of vehicle as a control instead of probe plus competitor, which precluded proper use of differential abundance analysis in the first place and offered limited insight into the specificity of probe–protein interactions.

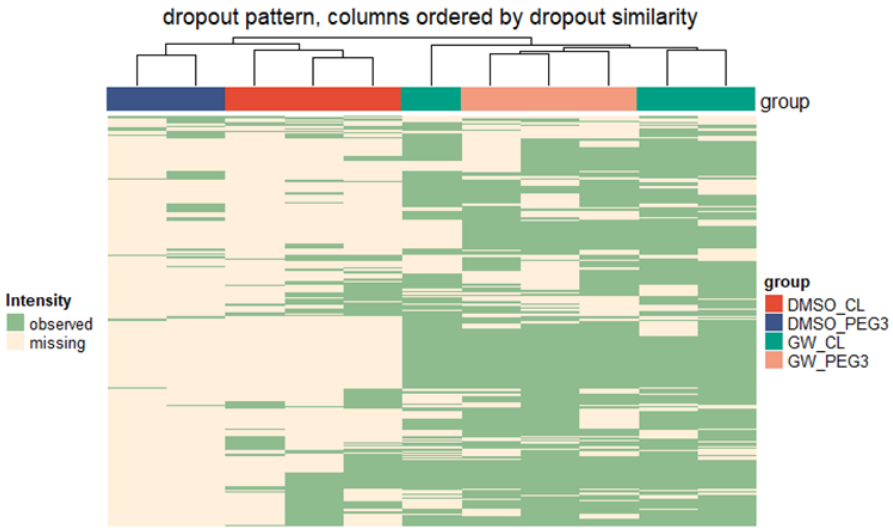


Figure 2-17. Visualization of missing patterns. Each row represents a peptide intensity. Each column represents a sample. DMSO_CL & GW_CL: labelled proteins were enriched using an azobenzene-based cleavable linker (see **Methods** for details); DMSO_PEG3 & GW_PEG3: labelled proteins were enriched using biotin-PEG_3-azide linker.

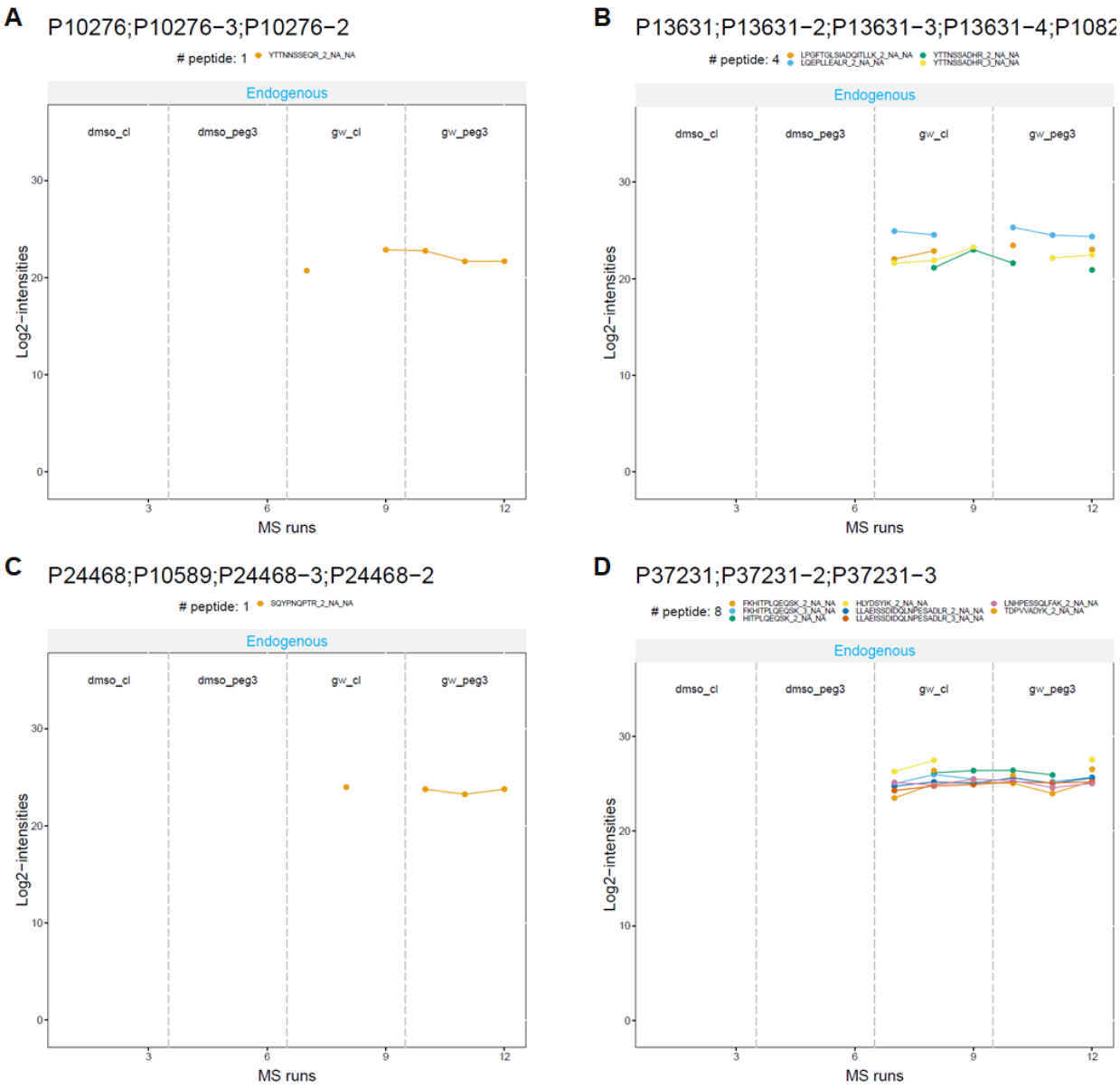


Figure 2-18. Peptide profile plots. (A) RARα (B) RARβ or γ (C) COUP-TF-I or II (D) PPARγ.

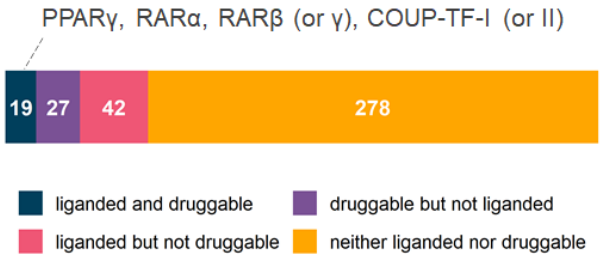


Figure 2-19. Druggability profile of target proteins.

Table 2-2. Druggable hits. Nuclear receptors are highlighted.

UniProtKB entry	HGNC symbol	Druggable domain	InterPro ID	Liganded
Q13370;Q14432	PDE3B;PDE3A	3'5'-cyclic nucleotide phosphodiesterase, catalytic domain	IPR002073	TRUE
Q96CM8	ACSF2	AMP-dependent synthetase/ligase	IPR000873	FALSE
Q9UIG0	BAZ1B	Bromodomain	IPR001487	TRUE
P21359	NF1	CRAL-TRIO lipid-binding domain	IPR001251	FALSE
O95202	LETM1	EF-hand domain	IPR002048	FALSE
P22676;P05937	CALB2;CALB1	EF-hand domain	IPR002048	FALSE
Q0IIM8	TBC1D8B	EF-hand domain	IPR002048	FALSE
Q66K14;Q6ZT07	TBC1D9B;TBC1D9	EF-hand domain	IPR002048	FALSE
Q6KCM7;Q9BV35	SLC25A25;SLC25A23	EF-hand domain	IPR002048	FALSE
Q8IYU8	MICU2	EF-hand domain	IPR002048	FALSE
Q8WXX0	DNAH7	EF-hand domain	IPR002048	FALSE
Q96BS2	TESC	EF-hand domain	IPR002048	FALSE
Q9H223;Q9NZN3	EHD4;EHD3	EF-hand domain	IPR002048	FALSE
P78417	GSTO1	Glutathione S-transferase, C-terminal	IPR004046	TRUE
P78417	GSTO1	Glutathione S-transferase, N-terminal	IPR004045	TRUE
Q14145	KEAP1	Kelch repeat type 1	IPR006652	TRUE
Q9NQX4	MYO5C	Myosin head, motor domain	IPR001609	FALSE
P10276	RARA	Nuclear hormone receptor, ligand-binding domain	IPR000536	TRUE
P13631;P10826	RARG;RARB	Nuclear hormone receptor, ligand-binding domain	IPR000536	TRUE
P24468;P10589	NR2F2;NR2F1	Nuclear hormone receptor, ligand-binding domain	IPR000536	TRUE
P37231	PPARG	Nuclear hormone receptor, ligand-binding domain	IPR000536	TRUE
Q70CQ2	USP34	Peptidase C19, ubiquitin carboxyl-terminal hydrolase	IPR001394	FALSE
P23458	JAK1	Protein kinase domain	IPR000719	TRUE
P51617;P51955	IRAK1;NEK2	Protein kinase domain	IPR000719	TRUE
Q13418	ILK	Protein kinase domain	IPR000719	TRUE
Q15418;P51812;Q15349;Q9UK32	RPS6KA1;RPS6KA3;RPS6KA2;RPS6KA6	Protein kinase domain	IPR000719	TRUE
Q99640	PKMYT1	Protein kinase domain	IPR000719	TRUE
P18031	PTPN1	PTP type protein phosphatase	IPR000242	TRUE
P30740;P50453;O75830	SERPINB1;SERPINB9;SERPINI2	Serpin domain	IPR023796	FALSE
P37059	HSD17B2	Short-chain dehydrogenase/reductase SDR	IPR002347	TRUE
Q99714	HSD17B10	Short-chain dehydrogenase/reductase SDR	IPR002347	TRUE
Q3SXM5	HSDL1	Short-chain dehydrogenase/reductase SDR	IPR002347	FALSE
Q8NBQ5	HSD17B11	Short-chain dehydrogenase/reductase SDR	IPR002347	FALSE
P61586;P08134	RHOA;RHOC	Small GTPase	IPR001806	TRUE
P10301	RRAS	Small GTPase	IPR001806	FALSE
P20340;Q9NRW1;Q14964	RAB6A;RAB6B;RAB39A	Small GTPase	IPR001806	FALSE
Q9H4L7	SMARCAD1	SNF2-related, N-terminal domain	IPR000330	FALSE
Q9NRZ9;Q8TE73	HELLS;DNAH5	SNF2-related, N-terminal domain	IPR000330	FALSE
P61077;P62837;Q9Y2X8;P51668	UBE2D3;UBE2D2;UBE2D4;UBE2D1	Ubiquitin-conjugating enzyme E2	IPR000608	TRUE
P62873;P16520	GNB1;GNB3	WDR	IPR017986	FALSE
Q12788	TBL3	WDR	IPR017986	FALSE
Q13347	EIF3I	WDR	IPR017986	FALSE
Q6P2E9	EDC4	WDR	IPR017986	FALSE
Q8TED0	UTP15	WDR	IPR017986	FALSE
Q9BR76	CORO1B;A9Z1Z3	WDR	IPR017986	FALSE
Q9NRG9	AAAS	WDR	IPR017986	FALSE

As a proof-of-concept, we followed up on TSPO because of its molecular weight (18 kDa) which corresponds to that of the intense band on the gel and the availability of its known ligands (**Table 2-3**). First, we confirmed that alk-GW9662 labels endogenous TSPO expressed in A549 cells by pull-down assay (**Figure 2-20A**). Next, we analyzed target engagement of a TSPO ligand PK-11195 by a gel-based competitive labelling assay (**Figure 2-20B**). Gratifyingly, PK-11195 dose-dependently inhibited alk-GW9662 labelling of endogenous TSPO (**Figure 2-20C**). These results demonstrate alk-GW9662 can be repurposed for target engagement analysis of TSPO ligands.

Table 2-3. Chemical proteomics hits (differential abundance analysis) with molecular weight within 10–20 kDa. uniport_id: Uniprot ID; uniport_ac: Uniprot accession number; mw_kda: molecular weight in kDa; log2FC: log2-transformed fold-change; p(p_val): negative of log10-transformed adjusted p-value; rank: ranking by log2FC.

uniprot_id	uniprot_ac	mw_kda	log2FC	p(p_val)	rank
VKORL_HUMAN	Q8N0U8	19.8	7.6	4.4	5
PTH2_HUMAN	Q9Y3E5	19.2	3.5	3.4	100
TSPO_HUMAN	P30536	18.8	9.5	3.2	1
PPIA_HUMAN	P62937	18.0	4.4	4.1	45
OCAD2_HUMAN	Q56VL3	17.0	2.5	2.6	203
CYB5B_HUMAN	O43169	16.7	5.7	4.1	11
LYSC_HUMAN	P61626	16.5	2.6	3.1	192
RS23_HUMAN	P62266	15.8	3.9	3.9	70
RL30_HUMAN	P62888	12.8	3.3	2.8	118
RL35A_HUMAN	P18077	12.5	3.8	3.9	81
S10A4_HUMAN	P26447	11.7	8.8	4.4	2
SUMO2_HUMAN	P61956	10.9	3.1	1.8	139
SELK_HUMAN	Q9Y6D0	10.6	5.4	4.1	14
BAF_HUMAN	O75531	10.1	2.8	1.9	167

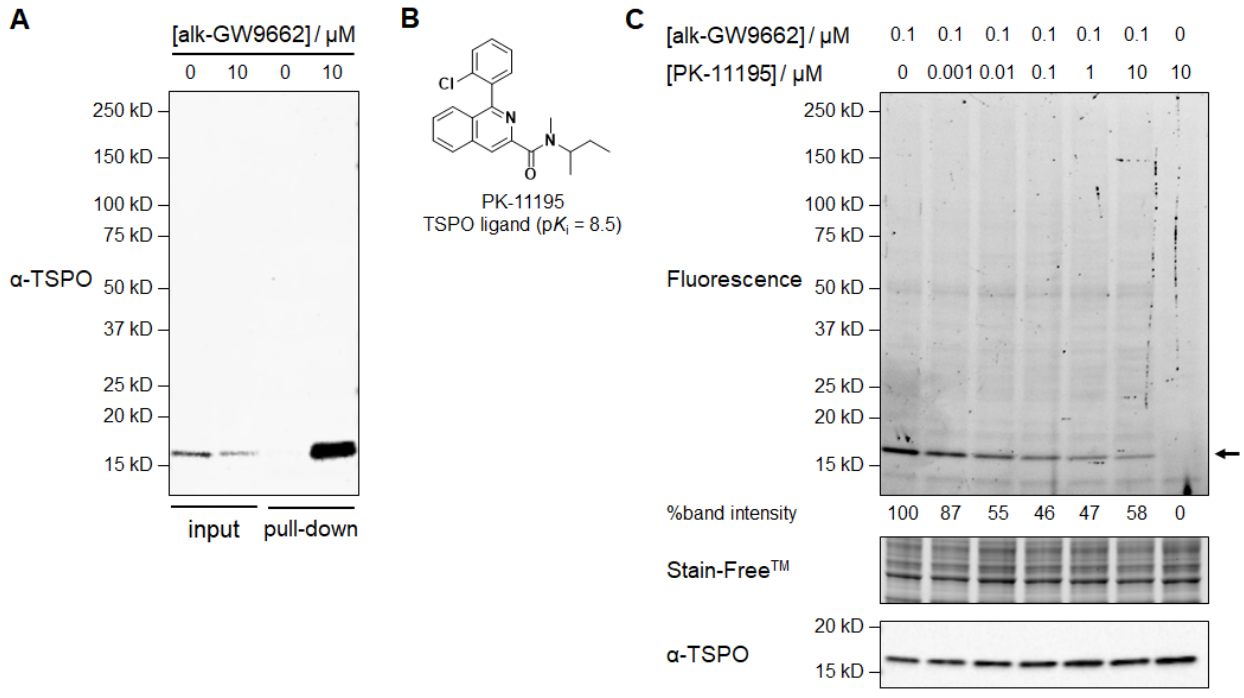


Figure 2-20. Repurposing alk-GW9662 for TSPO ligand engagement analysis. (A) Pull-down assay confirming TSPO alk-GW9662 labeling. (B) Chemical structure of a TSPO ligand. (C) Gel-based target engagement analysis using alk-GW9662.

Discussion

In summary, we developed a PPAR chemical reporter alk-GW9662, which enabled us to quantitatively analyze PPARγ–metabolite interactions in live cells. In many studies of NRs, candidate ligands are postulated to bind the target based on *in vitro* binding assay with recombinant proteins and cell-based transactivation assay.[107,108] While these assays are convenient and suitable for initial candidate screening, one should still confirm and quantify direct target engagement in cells or in tissues.[109] Previous studies have demonstrated target engagement by metabolomic analysis of NR–ligand complex isolated from a tissue sample.[110] This metabolomics approach may be challenging since relevant ligands might be lost during sample preparation. The chemical reporter strategy complements these approaches in quantifying target engagement in cells. With this strategy we characterized a specific conjugate linoleic acid isomer 9E, 11E as a partial agonist of PPARγ. PPARγ partial agonists are actively pursued as a new generation of PPARγ modulators that possess significant insulin-sensitizing and anti-inflammatory effects with fewer side effects of full agonists.[111] Our characterization of the conjugated linoleic acid isomer should inspire future investigations. It is fair to note that studies reported the 9E, 11E isomer is a relatively minor component of fatty acids identified in mouse gut.[112,113] However, more focused analysis of bacterial species and specific strains such as *Bifidobacterium breve* LMG 13194 and NCFB 2258 that produce this isomer may reveal anti-inflammatory microbiota and probiotics as well as help identify synthetic analogues for therapeutic applications.[114–116] In addition, our LC-MS/MS-based proteomics study identified a number of alk-GW9662 targets including several nuclear receptors. Thus, alk-GW9662 may also serve as a starting scaffold towards probe development to enable collective characterization of metabolite interactions with multiple nuclear receptors in cells.

CHAPTER 3. STUDIES TOWARDS UNTARGETED PROTEIN DEGRADATION STRATEGY

ABPP and related chemical reporter–based chemical proteomics provides a list of proteins that likely interact with a small molecule of interest. In typical target deconvolution studies, a team of biologists then narrow down this list to a handful of proteins by considering biological relevance and conduct follow-up studies. While multiple data and knowledge bases are available, selecting candidate(s) is not always straightforward and relies on each biologist's expertise and intuition. For example, chemoproteomic profiling of microbiota-associated metabolites and lipids identified hundreds of proteins that interact with these metabolites (Li et al., unpublished results; Yang et al., unpublished results; Zhao et al., unpublished results). Furthermore, knowledge-based selection tends to favour well-studied proteins and bias against less-studied proteins.[117]

The emergence of new modalities such as proteolysis-targeting chimera (PROTAC) [118] and specific and non-genetic inhibitor of apoptosis protein–dependent protein eraser (SNIPER) [119] represents unique opportunity to identify functionally-relevant protein–metabolite interactions. PROTACs and SNIPERs are chimeric small molecules which bring a protein of interest to a close proximity of a ubiquitin E3 ligase and induce its ubiquitination and proteosomal degradation (**Figure 3-1**).[120] A recent study has demonstrated a chemoproteomics approach to systematically query degradable kinases by combining *in vitro* KINOME*scan* screening and proteome-wide quantification of kinase expression levels.[121] Another study has proposed a workflow to identify novel cereblon modulators by combining focused library synthesis, phenotypic screening, and expression proteomics.[122] These studies suggest that PROTACs could complement chemoproteomic approaches in target deconvolution studies.

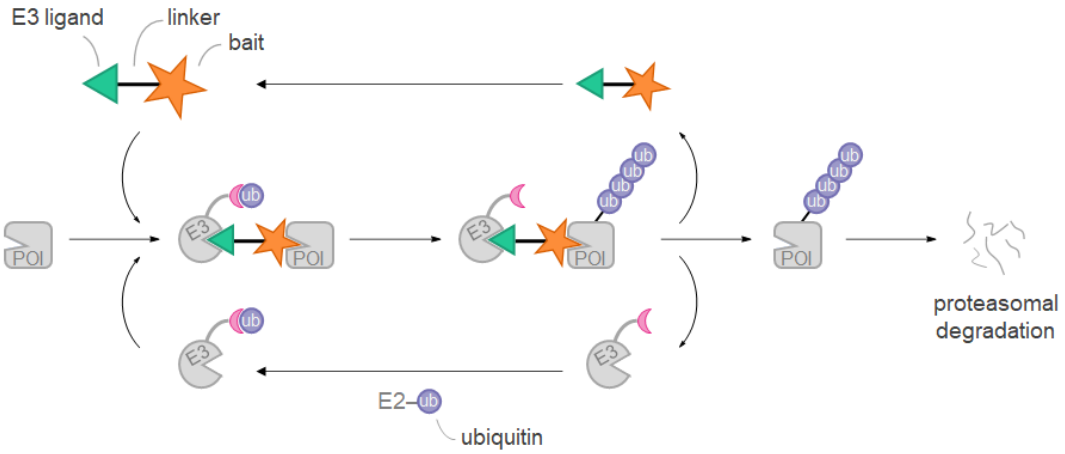


Figure 3-1. Targeted protein degradation by hetero-bifunctional molecules.

While human genome encodes approximately 600 (putative) E3 ligases, only a handful of E3 ligases have been employed in the targeted protein degradation technology.[123,124] Most of the reported compounds using either cereblon or Von Hippel-Lindau because of the availability of drug-like small molecules that recruit them.[125] Expanding the set of E3 ligases that can be recruited by PROTACs will extend the scope of this modality to a wider range of proteins and cell type–selective degradation. In this regard, it is noteworthy that recent studies characterized the E3 ligase activity of PPARγ.[126–128] PPARγ is a druggable protein with numerous synthetic ligands reported in literature.[129] In addition, recent studies reported an enrichment of macrophages that highly

express PPARγ in tumour tissues of clear cell renal cell carcinoma and lung adenocarcinoma patients.[130,131] PPARγ has been associated with immunosuppressive phenotype.[132] These studies suggest that PPARγ-dependent protein degraders could target specific populations of tumour associated macrophages (TAMs) and contribute to the development of emerging anticancer therapies focused on TAMs.[133] Thus, we aimed to develop PROTACs that recruit PPARγ with ultimate goals of applying these tools in target deconvolution of immunomodulatory metabolites in immune cells that express PPARγ.

Synthesis of PPARγ-recruiting PROTACs based on (*R*)-SR9034

We designed and synthesised PPARγ-recruiting PROTACs to target BRD4 for degradation. To achieve this, a BRD4 ligand (+)-JQ1 and a PPARγ ligand (*R*)-SR9034 were connected via ethylene glycol linkers of variable lengths (**Figure 3-2A**). (*R*)-SR9034 is a potent PPARγ partial agonist (IC_{50} = 0.4 nM, EC_{50} = 13 nM) that binds the alternative binding pocket (**Figure 3-2B**).[134,135] The carboxyl group is accessible to solvent and forms salt bridges/hydrogen bonds with Arg288 and Ser342. Replacing it with amides leads to only slight decrease in affinity. In addition, the *S*-isomer is approximately 30-fold less potent, which could be used to design less-active enantiomer controls.

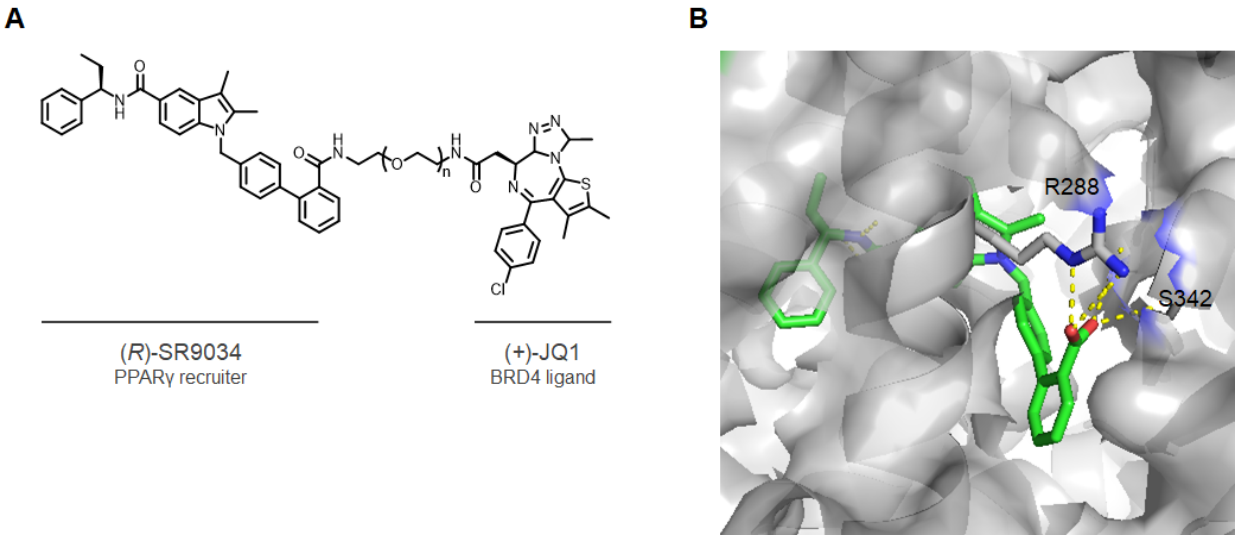
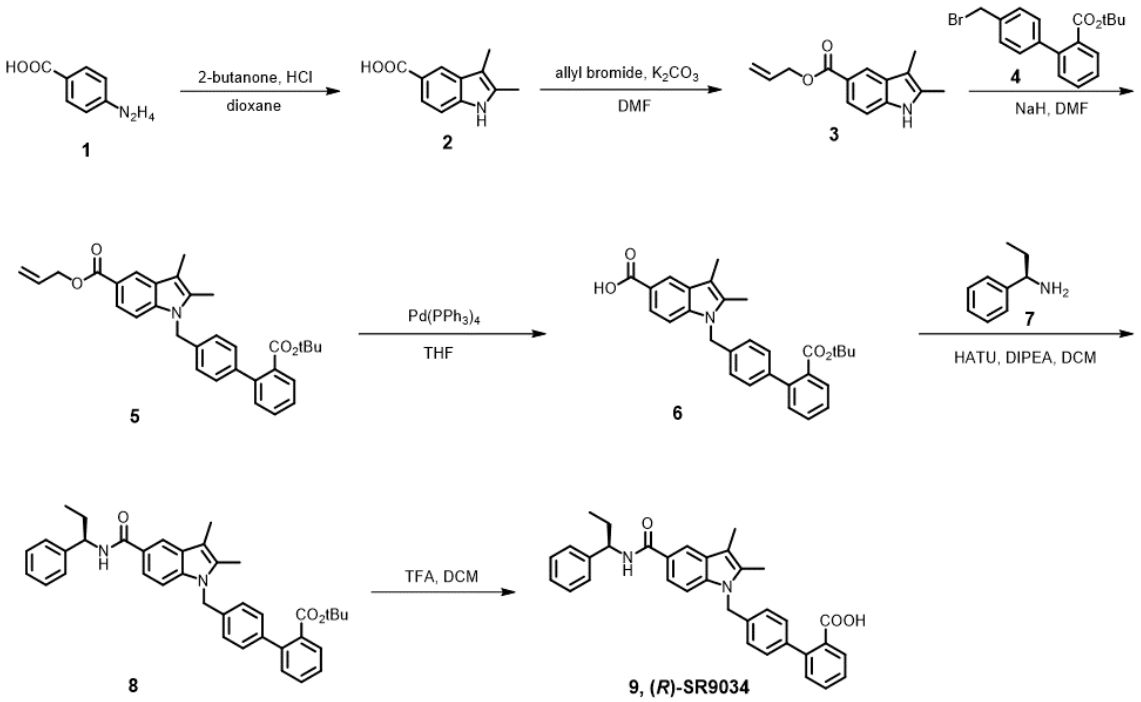
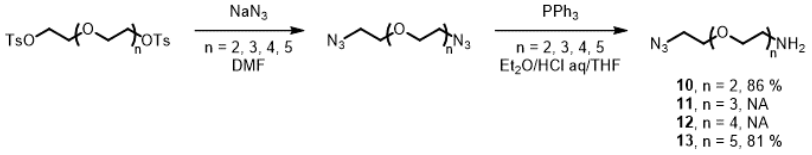


Figure 3-2. Design of PPARγ-recruiting PROTACs based on (*R*)-SR9034. (A) Chemical structure of PROTACs. (B) X-ray co-crystal structure of (*R*)-SR9034 in the PPARγ ligand binding domain (PDB access code: 3KMG).

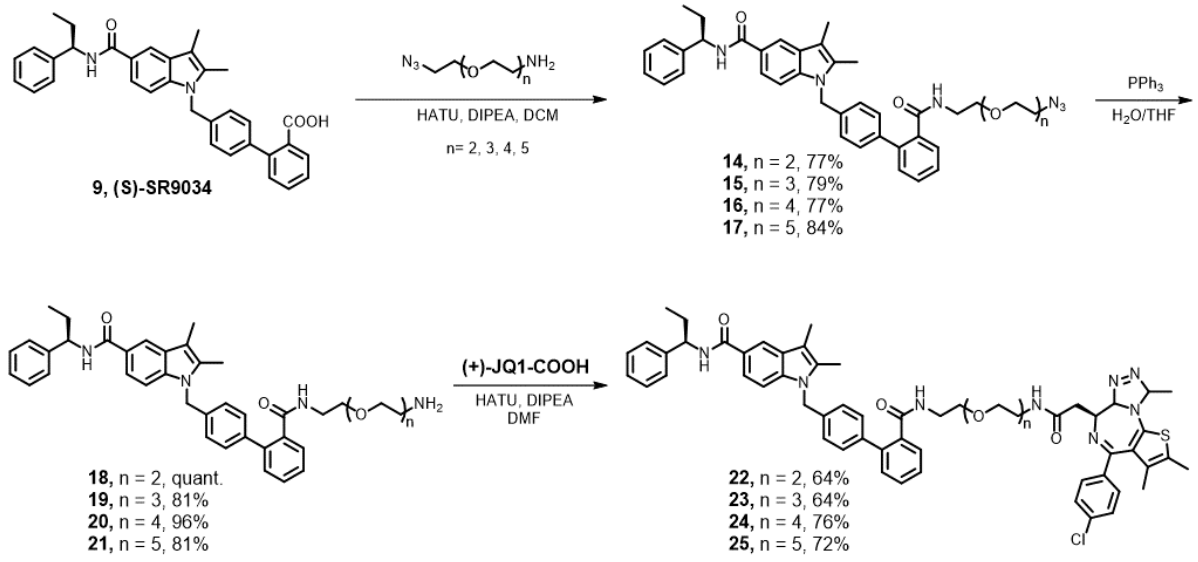
The PPARγ ligand (*R*)-SR9034 was synthesised as previously described (**Scheme 3-1**).[135] Azide-PEG-NH_2 linkers were synthesised from commercially-available ethylene glycol di-*p*-toluenesulfonates (**Scheme 3-2**). (*R*)-SR903 was coupled with the azide-PEG-NH_2 linkers by HATU-mediated condensation reaction (**Scheme 3-3**). The resulting azides were reduced by Staudinger reaction to provide amines which were subsequently coupled with (+)-JQ1-CO_2H. Alternatively, aminobenzoic acid was inserted between the PEG_6 linker and (+)-JQ1-CO_2H for additional distance (**Scheme 3-4**).



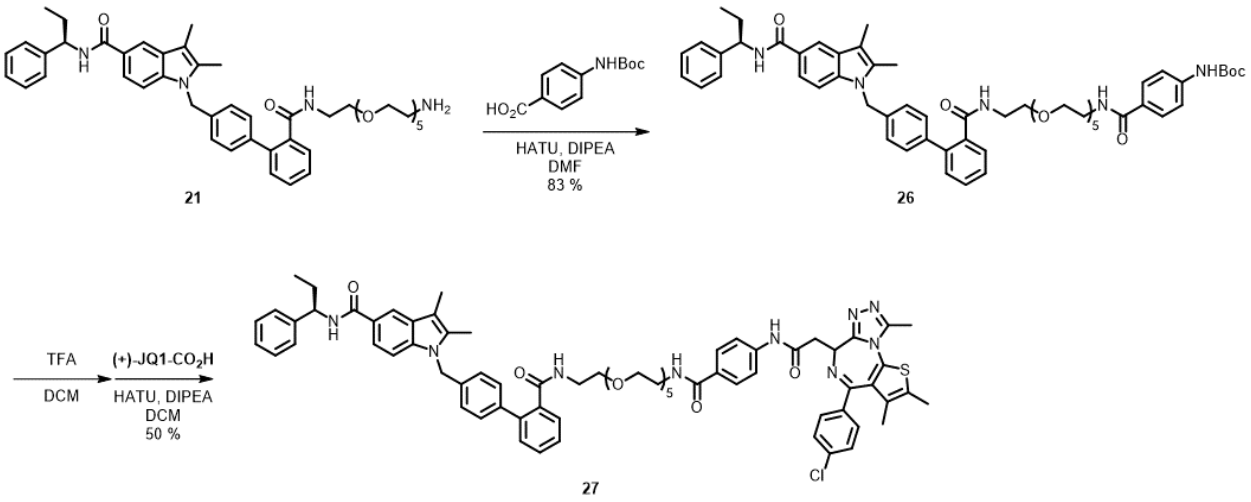
Scheme 3-1. Synthesis of (*R*)-SR9034.



Scheme 3-2. Synthesis of azide-PEG-NH_2 linkers **10**–**13**.



Scheme 3-3. Synthesis of (R)-SR9034–based PROTACs **22**–**25**.



Scheme 3-4. Synthesis of (R)-SR9034–based PROTAC **27**.

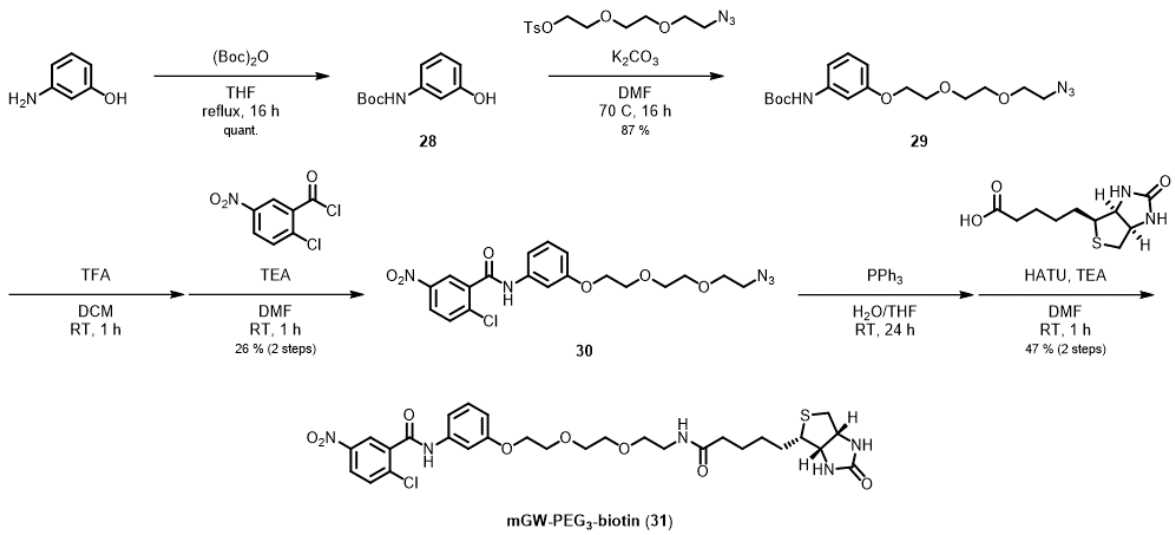
Synthesis of GW9662 derivatives

The chemoproteomic profiling study revealed that alk-GW9662 potentially targets several putative E3 ligases in addition to PPARγ (**Chapter 2, Table 3-1**). In addition, recent studies demonstrated strategies to identify E3 ligases that can be used for targeted degradation application using covalent warheads.[136–138] Thus, we were interested to test whether GW9662 could be used to recruit potentially multiple E3 ligases in the targeted protein degradation strategy.

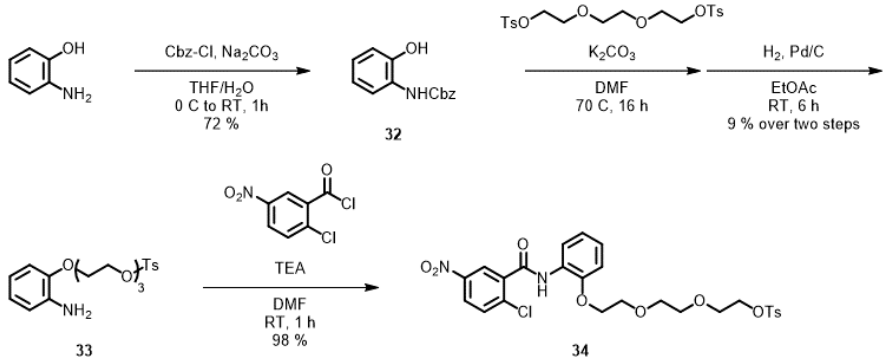
Table 3-1. Ubiquitin E3 ligases targeted by alk-GW9662.

UniProt KB entry	# Peptides	HGNC symbol	Druggable domain	InterPro ID	Liganded
Q14145	5	KEAP1	Kelch repeat type 1	IPR006652	TRUE
Q96SW2;Q96SW2-2	2	CRBN	NA	NA	TRUE
Q12933;Q12933-4;Q12933-3;Q12933-2	6	TRAF2	NA	NA	FALSE
Q14669;Q14669-3;Q14669-4;Q14669-2	1	TRIP12	NA	NA	FALSE
Q7Z6Z7;Q7Z6Z7-2;Q7Z6Z7-3	1	HUWE1	NA	NA	FALSE
Q8WU17	1	RNF139	NA	NA	FALSE
Q969V5	1	MUL1	NA	NA	FALSE
Q9P253	2	VPS18	NA	NA	FALSE
Q9P2G1	9	ANKIB1	NA	NA	FALSE
Q9Y575;Q9Y575-2;Q9Y575-3	1	ASB3	NA	NA	FALSE

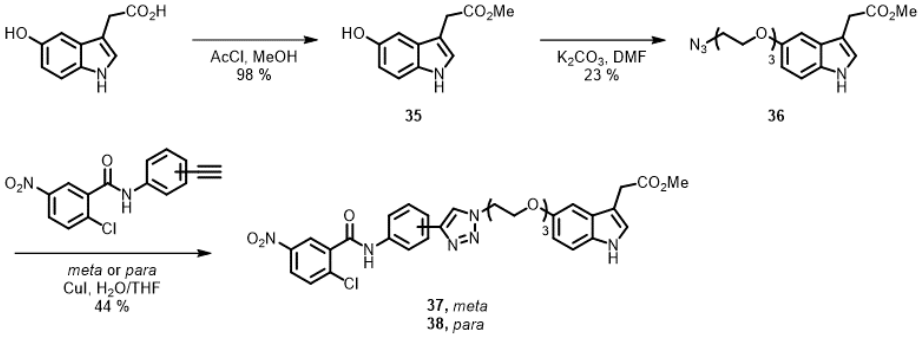
A biotinylated derivative **m-GW-PEG$_3$-biotin** was synthesised from 3-aminophenol in six steps (**Scheme 3-5**). The biotin moiety was introduced in the final step for the ease of purification of intermediates. *Ortho* substitution was explored by synthesising a tosylated derivative (**Scheme 3-6**). As an alternative linking method, copper (I)-catalysed alkyne–azide cycloaddition reaction (CuAAC) was explored for compounds bearing indole-3-acetic acid methyl ester as the bait moiety (**Scheme 3-7**). In addition, a hydrophobic linker was explored by the synthesis of SB1453-derivatives (Schemes **3-8**, **3-9 & 3-10**).[91]



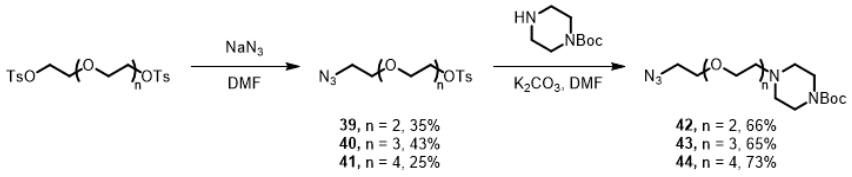
Scheme 3-5. Synthesis of mGW-PEG3-biotin (**31**).



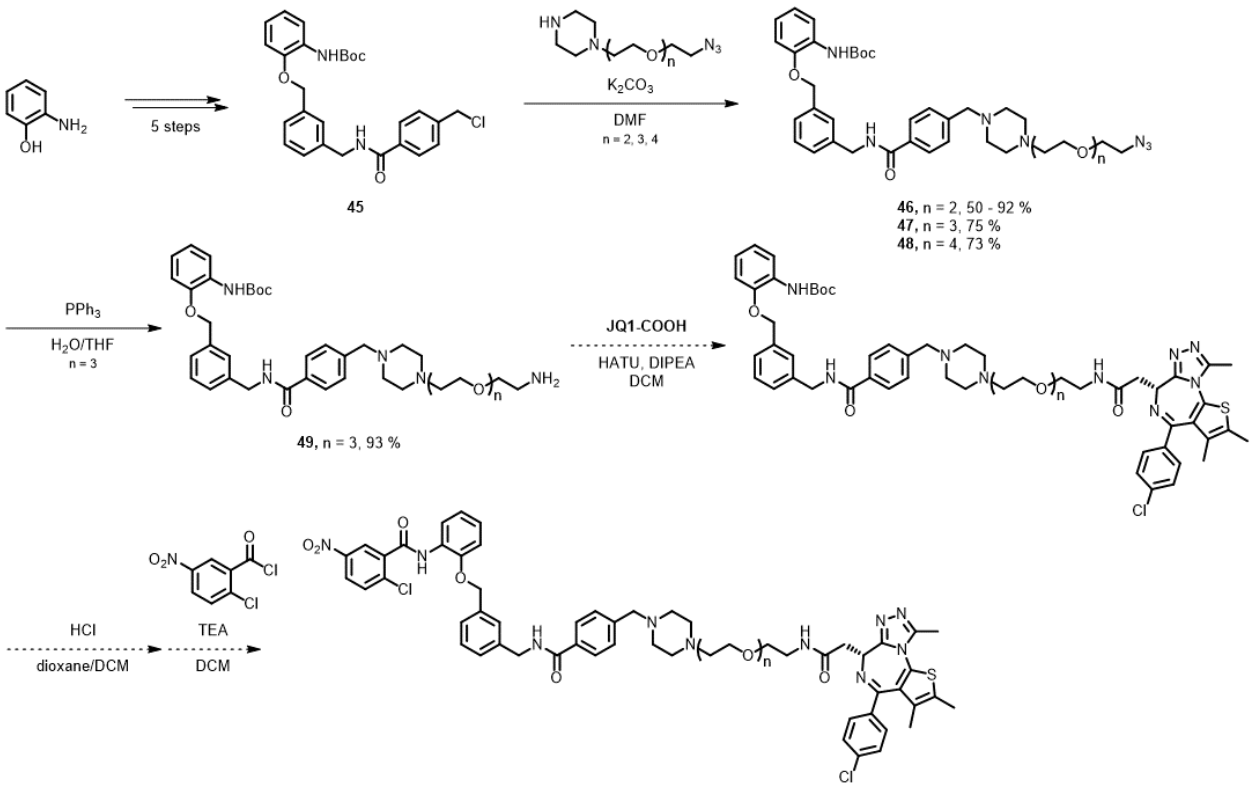
Scheme 3-6. Synthesis of an *ortho*-substituted derivative (**34**).



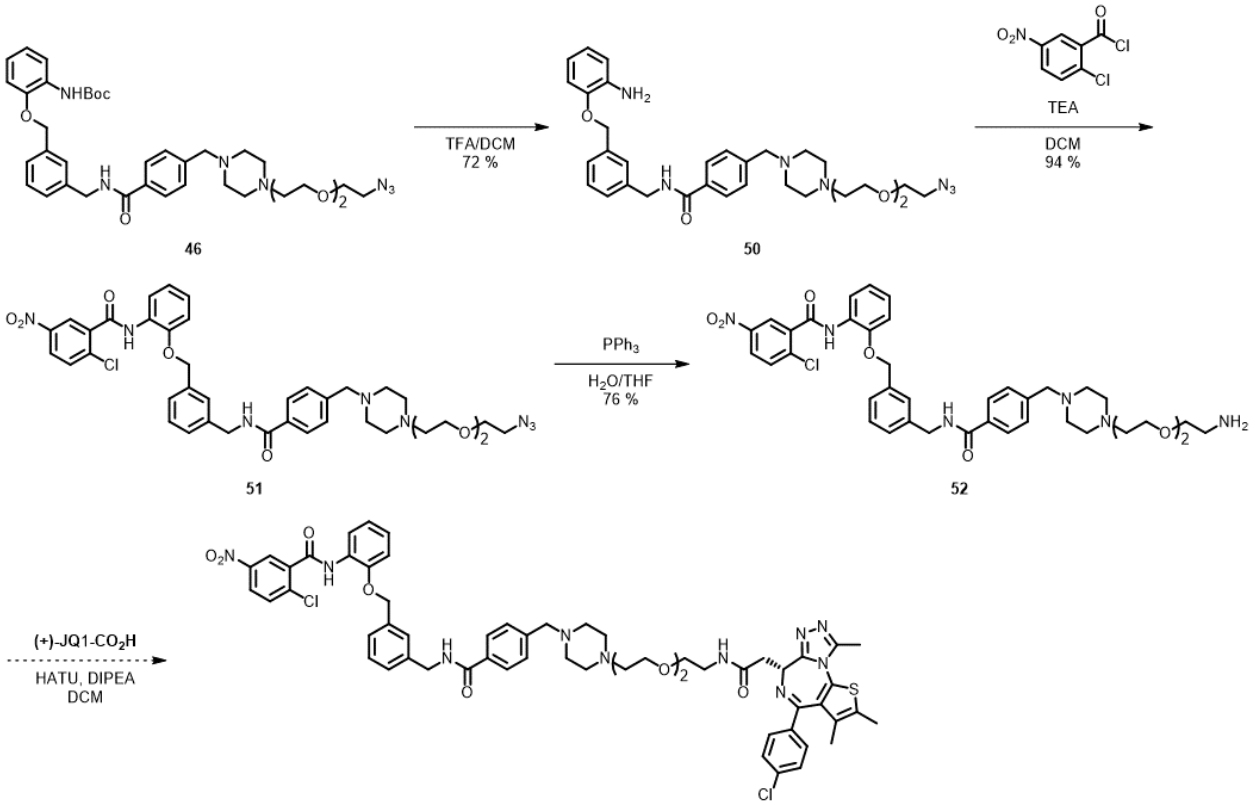
Scheme 3-7. Synthesis of GW9662-derivatives (**37** & **38**) by CuAAC.



Scheme 3-8. Synthesis of azide-PEG-OTs/piperazine linkers (**39**–**44**).



Scheme 3-9. Synthesis of SB1453-derivatives. Dashed arrows indicate prophetic reactions.



Scheme 3-10. Synthesis of SB1453-derivatives. Dashed arrows indicate prophetic reactions.

Putative ubiquitin E3 ligase activity of PPARγ

Hou et al. reported characterisation of the ubiquitin E3 ligase activity of PPARγ.[128] We attempted to replicate their key findings. HEK293T cells were transiently transfected with human PPARγ2 for 16 h and then treated with cycloheximide for up to 8 h (**Figure 3-3A**). The NFκB p65 subunit protein level was assessed by Western blot. In contrast to the published results, the p65 level remained unaffected. We performed similar experiments by transfecting cells with human PPARγ1 (**Figure 3-3B**) or murine PPARγ2 (**Figure 3-3C**) vectors, but the p65 level was still unaffected. Because we could not independently verify the putative E3 ligase activity of PPARγ, we terminated this project.

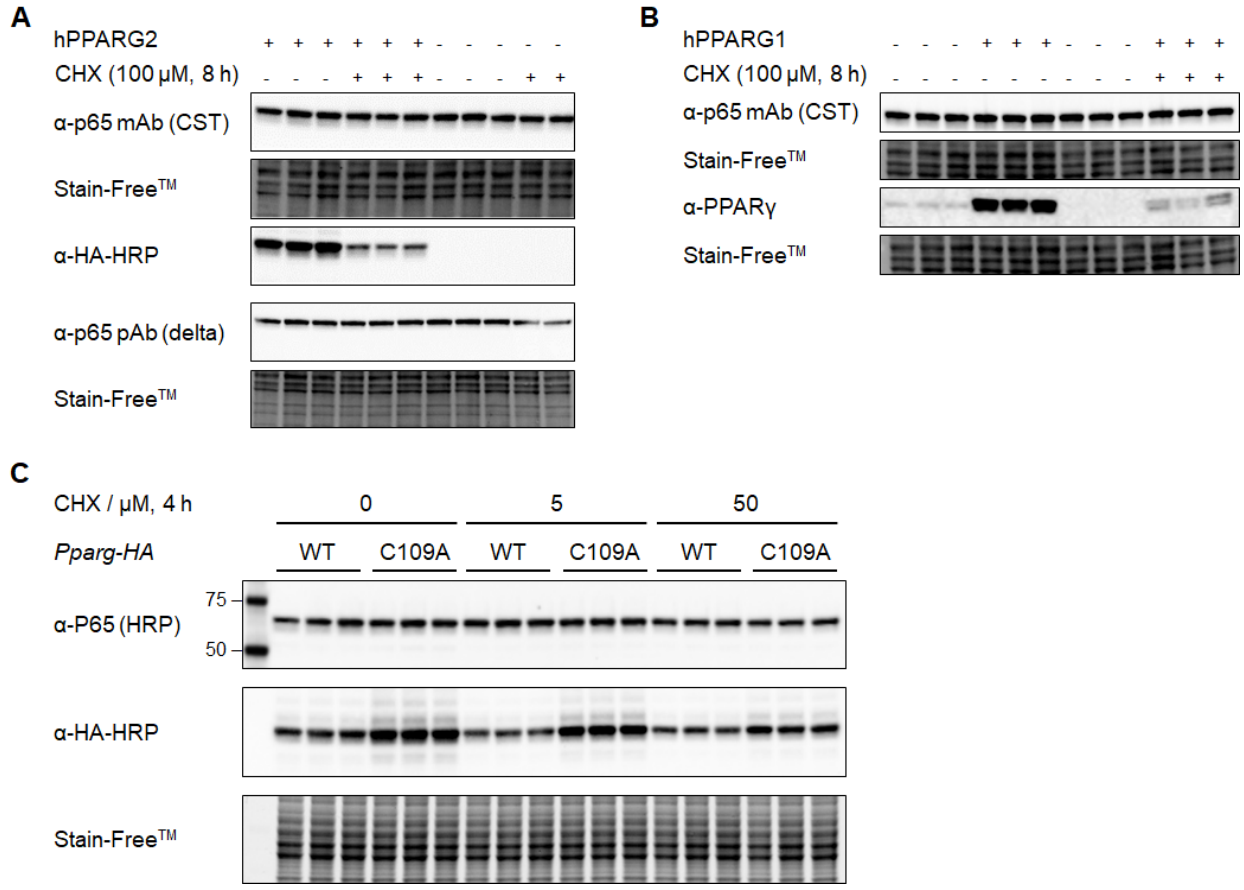


Figure 3-3. Putative ubiquitin E3 ligase activity of PPARγ. (A) Human PPARγ2. (B) Human PPARγ1. (C) Murine PPARγ2. C109A is a putative catalytically-inactive mutant.

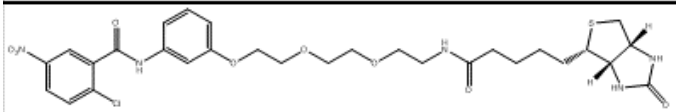
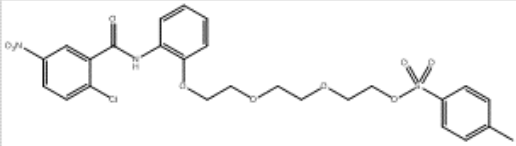
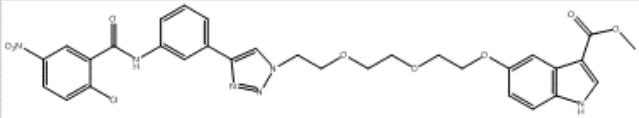
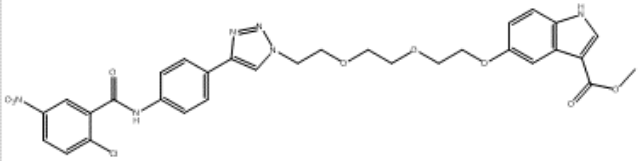
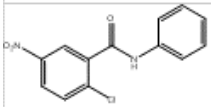
Targeted degradation of BRD4

In accordance with the lack of robust E3 ligase activity, PPARγ-recruiting PROTACs **18**–**21** and **23** failed to induce BRD4 degradation (data not shown).

Target engagement of GW9662-derivatives

We analysed target engagement of GW9662-derivatives in cell lysate using the alk-GW9662 competitive labelling assay (**Table 3-2**). All derivatives poorly engaged PPARγ.

Table 3-2. PPARγ engagement of GW9662 derivatives in cell lysate. [alk-GW9662] = 0.1 μM, [compound] = 1 μM.

Compound	Res. fluor. (mean ± SD)
	88 ± 14 %
	96 ± 5 %
	73 ± 15 %
	99 ± 28 %
	47 ± 21 %

Discussion

Unfortunately, PPARγ failed to exhibit robust E3 ligase activity in our hands. This result makes it untenable to develop PROTACs that solely rely on PPARγ for E3 ligase activity. In contrast, the multi-E3 ligase approach based on GW9662 could be still viable. However, there are several challenges associated with this approach. First, the preliminary nature of the chemoproteomics experiment (i.e., the lack of competition) makes it difficult to judge specificity of identified interactions. It is unclear whether GW9662 and related compounds can achieve sufficient occupancy of the putative E3 ligases. Second, validating E3 ligase activity is not always straightforward because many E3 ligases function as multi-protein complex and their native substrates are often unknown. Despite these challenges, GW9662 still has a role to play thanks to the S_NAr mechanism that may be advantageous in cases where more commonly-used warheads such as chloroacetamide and acrylamide fail to provide tractable hits.[139] Finally, well-validated E3 ligases such as cereblon and Von Hippel-Lindau should be the first choice in target deconvolution studies.

CHAPTER 4. OVERALL DISCUSSION AND OUTLOOK

Activity-based protein profiling (ABPP) and chemical reporter–based chemical proteomics facilitate target deconvolution and target engagement studies.[8,140] Since their inception almost 20 years ago, the scope and utility of these chemical proteomics methods was significantly extended through continuous method and probe development and they have become widely adopted among chemical biology community (**Chapter 1**). Yet, some key protein classes and specific ligand-binding domains still remain inaccessible. In this thesis, we described two projects aimed at developing chemical probes to further empower and complement these methods.

In **Chapter 2**, we described the development of a ligand-directed probe alk-GW9662 for a subset of nuclear receptors (NRs). NRs are a family of ligand-activated transcription factors that regulate diverse physiological processes in animals and are key targets for therapeutic development. Humans express 48 NRs, which share a common modular structure consisting of a regulatory domain referred to as activation function-1 domain at the N-terminus, a DNA binding domain in the middle, and a ligand binding domain (LBD) at the C-terminus. Endogenous small molecules such as steroids, thyroid hormones, retinoids, and fatty acids regulate transcriptional activity of most NRs through interaction with the LBD. Recent studies implicated several NRs in host–microbiota interactions and suggested that various microbiota-derived metabolites may also serve as ligands for NRs.[71,74–76,78–84,96] However, whether these metabolites directly engage NRs remains unclear. Despite having clear importance in fundamental biology and drug development, NRs lack ligand-directed probes and largely remain inaccessible by chemical proteomics approaches.

Alk-GW9662 site-specifically targets conserved cysteine residues in PPAR LBDs. We demonstrated its utility to assess target engagement of candidate ligands of peroxisome proliferator–activated receptor (PPAR) γ in live cells. In many studies of NRs, candidate ligands are postulated to engage the target based on *in vitro* binding assay with recombinant proteins and cell-based transactivation assay.[107,108] While these assays are convenient and suitable for initial candidate screening, one should still confirm and quantify direct target engagement in cells or in tissues.[109] Previous studies have demonstrated target engagement by metabolomic analysis of NR–ligand complex isolated from a tissue sample.[110] This metabolomics approach may be challenging since relevant ligands might be lost during sample preparation. The chemical reporter strategy complements these approaches in quantifying target engagement in cells.

The alk-GW9662 target engagement assay revealed that a specific conjugate linoleic acid isomer 9E, 11E may serve as a partial agonist of PPARγ. PPARγ partial agonists are actively pursued as a new generation of PPARγ modulators that possess significant insulin-sensitizing and anti-inflammatory effects with fewer side effects of full agonists.[111] Our characterization of the conjugated linoleic acid isomer should inspire future investigations. It is fair to note that studies reported the 9E, 11E isomer is a relatively minor component of fatty acids identified in mouse gut.[112,113] However, more focused analysis of bacterial species and specific strains that produce this isomer may reveal anti-inflammatory microbiota and probiotics as well as help identify synthetic analogues for therapeutic applications.[114–116]

We also explored repurposing alk-GW9662. We profiled its target proteins using LC-MS/MS–based proteomics and demonstrated that the probe can be used in target engagement study of other proteins such as TSPO. Taken together, our results suggest that alk-GW9662 may be useful for target engagement analysis of candidate PPAR ligands and offer a starting scaffold in developing probes for other proteins.

In **Chapter 3**, we described efforts towards the combined use of proteolysis-targeting chimeras (PROTACs) and chemical proteomics in target deconvolution studies. ABPP and related chemical reporter–based chemical proteomics provides a list of proteins that likely interact with a small molecule of interest. However, the functional relevance of identified proteins is often unclear, and follow-up studies require careful selection of a tractable numbers of candidates. While multiple data and knowledge bases are available, selecting candidate(s) is not always straightforward and relies on each biologist's expertise and intuition. Furthermore, knowledge-based selection tends to favour well-studied proteins and bias against less-studied proteins.[117] Targeted protein degradation technologies such as PROTAC could be used as complementary methods to modulate and identify functionally-relevant interactions in a similar manner to genome-wide CRISPR screening.[122,141]

While human genome encodes approximately 600 (putative) E3 ligases, only a handful of E3 ligases have been employed in the PROTAC technology.[123,124] Most of the reported compounds using either cereblon or Von Hippel-Lindau because of the availability of drug-like small molecules that recruit them.[125] Expanding the set of E3 ligases that can be recruited by PROTACs will extend the scope of this modality to a wider range of proteins and cell type–selective degradation. Thus, we aimed to develop PROTACs that recruit PPARγ given its putative E3 ligase activity and enrichment in immunosuppressive macrophages.[126–128,130–132] We described the design and synthesis of these reagents. Unfortunately, PPARγ lacked robust E3 activity in our hands. The reagents we synthesised failed to induce protein degradation.

Overall, the projects described herein underline the utility of chemical tools such as ligand-directed probes and PROTACs in target engagement and target deconvolution studies. ABPP and related chemical reporter–based chemical proteomics methods enable sensitive detection of small molecule–protein interactions in cellular environment. Active site (ligand)–directed probes are useful for target engagement and selectivity profiling studies, whereas non-directed probes are useful for target deconvolution studies. Both types of probes form covalent bonds with interacting proteins, which facilitates their enrichment and identification of low-abundance targets. On the other hand, probe design and synthesis can be challenging and limiting. This is not only an issue for complex natural products but also for virtually any small molecules. For example, the derivatisation of fragment-sized metabolites with a photo-crosslinker will affect their physicochemical properties and complicate target deconvolution. Proteome-wide competition experiments using chemo-selective probes such as iodoacetamide-alkyne obviates the need for direct derivatisation. This approach also allows relatively straightforward site identification. One should be careful about making the oft-used assumption that compounds containing potential electrophiles act covalently. While these compounds may indeed covalently bind some proteins, their activity may derive from non-covalent interactions with other proteins. Similarly, the success of proteome-wide competition profiling using photoaffinity probes relies on proper design of the photoaffinity probes. In contrast, label-free methods such as limited proteolysis and thermal proteome profiling directly work with underivatized molecules. These methods trade the chemical probe–related complications with lower sensitivity for less abundant proteins. A hybrid method that combines chemical reported–mediated enrichment of cell surface proteins and thermal proteome profiling has been proposed to improve sensitivity.[57] While these chemical proteomics methods can facilitate the identification of small molecule–protein interactions, it can be challenging to judge which one(s) of those interactions are functionally more relevant than others. In this regard, the *non-directed* protein degradation strategy may offer a solution.

METHODS

Chapter 2

Reagents, plasmids, cell lines, and antibodies. 2-ethylnyl aniline, 4-ethynyl aniline, and 2-chloro-5-nitrobenzoyl chloride were purchased from Tokyo Chemical Industry Co., Ltd. The PPAR ligands and the TSPO ligand were purchased from Alfa Aesar, Avanti Polar Lipids, Cayman Chemical, Sigma–Aldrich, or Tokyo Chemical Industry Co., Ltd. Az-Rho was previously synthesised in the lab.[142] Biotin-PEG3-azide and TBTA were purchased from Click Chemistry Tools, $CuSO_4$-$5H_2O$ was purchased from Sigma–Aldrich, and TCEP-HCl was purchased from Thermo Fisher Scientific. Wild-type PPARα (HG12080-CM) and PPARγ (MG53363-CY) plasmids were purchased from Sino Biological. PPARβ/δ plasmid was sub-cloned from pAdTrack-PPARdelta (Bert Vogelstein, Addgene plasmid # 16529) into pLX302 (David Root, Addgene plasmid # 25896) by Gateway™ Cloning Technology (Thermo Fisher Scientific). Mutants were obtained with QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent). PPRE X3-TK-luc was a gift from Bruce Spiegelman (Addgene plasmid # 1015). pRL-TK was purchased from Promega (AF025846). Human embryonic kidney 293T (HEK293T) and HT-29 cell lines were obtained from American Type Culture Collection and was cultured in Gibco™ DMEM, high glucose, pyruvate (Thermo Fisher Scientific, # 11995065) supplemented with 10 % foetal bovine serum (FBS) at 37 °C under 5 % CO_2 atmosphere. Anti-HA-HRP was purchased from Sigma–Aldrich (rat, clone 3F10, # 12013819001), anti-myc-HRP (mouse, clone 9B11, # 2040S) and anti-PPARγ (rabbit, clone D69, #2430S) were purchased from Cell Signalling Technology and anti-V5 antibody (mouse, clone E10, ab53418) and anti-mouse IgG-HRP (goat, ab97023) were purchased from abcam.

Transactivation assay. HEK293T cells were seeded on a 96-well plate and cultured overnight. Cells were transfected with a plasmid encoding PPAR (100 ng/well), PPRE X3-TK-luc (2.5 ng/well) and pRL-TK (1 ng/well) for 6 h using 0.2 µL/well Lipofectamine® 2000 (Thermo Fisher Scientific) reagent for 6 h and then were incubated with compound for additional 18 h. Dual-Luciferase® Reporter Assay (Promega) was performed according to the manufacturer's protocol.

Competition assay in lysate. HEK293T cells were seeded in a 10-cm dish overnight before transfection with a plasmid encoding PPARγ using Lipofectamine® 2000 (Thermo Fisher Scientific) for 24 h. Cells were washed with ice-cold PBS, pelleted, snap-frozen and stored at -80 °C until use. The pellet was suspended in a hypotonic buffer (20 mM HEPES, 10 mM KCl, 1.5 mM $MgCl_2$, pH 7.4) and lysed by three freeze-thaw cycles. Insoluble material was removed by centrifugation (4 °C, 12,000 rcf, 10 min). Protein concentration was adjusted to 2 mg/mL by diluting with the hypotonic buffer. Lysate (20 µL) was aliquoted in 1.7 mL tubes and equilibrated at 25 °C for 3 min. Then the lysate was incubated with compound (100x stock in DMSO, 0.2 µL) and alk-GW9662 (100x stock in DMSO, 0.2 µL) at 25 °C for 20 min (Figure S9). After this time labelling reaction was quenched by the addition of 12% SDS (4 µL) and diluted with the hypotonic buffer (20 µL). The lysate was reacted with 5 µL click reaction cocktail (final concentrations: 100 µM Az-Rho, 1 mM TCEP, 100 µM TBTA, 1 mM $CuSO_4$) at room temperature for 1 h. Proteins were precipitated by adding methanol (200 µL) at room temperature and incubating at -80 °C for at least 1 h, followed by centrifugation (4 °C, 18,000 rcf, 10 min). The protein pellet was briefly washed with methanol (200 µL), air-dried for 5 min, dissolved in 25 µL 1x Laemmli sample buffer (Bio Rad, #161-0747) and boiled at 95 °C for 5 min.

Competition assay in live cells. HEK293T cells were seeded in a 12-well plate overnight before transfection with a plasmid encoding for a PPAR using Lipofectamine® 2000 for 18 h. Culture media were exchanged with fresh media devoid of FBS (1 mL). To each well compound (1000x stock in DMSO, 1 µL) was added and incubated at 37 °C under 5 % CO_2 atmosphere for 15 min. Then the plate was cooled on ice for 15 min before adding alk-GW9662 (1000x stock in DMSO, 1 µL). After 30 min on ice, media were exchanged with ice-cold PBS containing GW9662 (10 µM) to terminate labelling. Performing in-cell labelling on ice was necessary to slow down labelling reaction and obtain reproducible results (Figure S10). Cells were collected, centrifuged (4 °C, 800 rcf, 3 min), and lysed with 0.1 % SDS in the hypotonic buffer (16 µL) containing 250 U/mL Benzonase® endonuclease (EMD Millipore, #71206) followed by the addition of 12% SDS (4 µL) and dilution with the hypotonic buffer (final concentrations: 50 µg protein/45 µL, *ca.* 0.2% SDS). The lysate was reacted with the click reaction cocktail (5 µL) at room temperature for 1 h, and proteins were precipitated as above. The pellet was dissolved in 30 µL 1x Laemmli sample buffer (Bio Rad, #161-0747) and boiled at 95 °C for 5 min.

Alk-GW9662 labelling in HT-29 cells. HT-29 cells were grown to confluence and incubated with alk-GW9662 in fresh media (DMEM, no FBS) in 5 % CO_2 incubator at 37 °C for 1 h. Cells were washed with ice-cold PBS three times, collected by scraping and centrifugation, lysed with 0.1 % SDS in the hypotonic buffer containing 250 U/mL Benzonase® endonuclease followed by the addition of 12% SDS, and diluted with the hypotonic buffer to the final concentrations: 50 µg protein/45 µL; *ca.* 0.2% SDS. The lysate was reacted with the click reaction cocktail containing either Az-Rho or biotin-PEG3-azide at room temperature for 1 h, and proteins were precipitated as above. For in-gel fluorescence analysis, 50 µg pellet was dissolved in 30 µL 1x Laemmli sample buffer and boiled at 95 °C for 5 min. For pull-down assay, 2 mg pellet was dissolved in the hypotonic buffer containing 0.1 % SDS (4 mL) and incubated with Neutravidin agarose resin (100 µL, 50 % slurry, Pierce) at room temperature for 1 h followed by wash with following buffers: 0.1% SDS / PBS, PBS x 2, 4 M urea / PBS, 50 mM $NaHCO_3$ / PBS, and PBS. Proteins were eluted by boiling in 1x Laemmli sample buffer at 95 °C for 5 min.

SDS-PAGE and Western blotting. Samples were loaded on the Criterion™ TGX Stain-Free™ precast gel (Bio Rad) and resolved at 200 V for *ca.* 45 min. SDS-PAGE gels were imaged on ChemiDoc XRS+ system (Bio Rad) for rhodamine and Stain-Free™. For Western blotting, the imaged gel was transferred to a 0.45 µm pore-size nitrocellulose membrane (Bio Rad). The membrane was blocked with 5% skim milk/TBS-T (anti-HA-HRP) or 3 % BSA/TBS-T (anti-myc-HRP and anti-V5) at room temperature for 1 h, incubated with anti-HA-HRP (1:10,000) at room temperature for 1 h, anti-myc (1:1,000), anti-V5 (1:1,000) or anti-PPARγ (1:1000) at 4 °C overnight followed by washing three times with TBS-T. For anti-V5, the blot was treated with anti-mouse IgG-HRP (1:10,000) at room temperature for 1 h and washed three times with TBS-T. The blots were visualised with Clarity ECL Western Substrate (Bio Rad) according to the manufacturer's protocol, imaged with ChemiDoc XRS+ system (Bio Rad) and processed with Image Lab™ (Bio Rad).

***In silico* modelling of PPARγ–alk-GW9662 covalent complex.** The crystal structure of PPARγ–GW9662 complex (3B0R) was used to generate the model of PPARγ–alk-GW9662 complex. The Covalent Docking Panel[143] from Maestro was used in default settings.

Proteomics sample preparation. HT-29 cells were incubated with alk-GW9662 (10 µM) harvested as described above. Proteome (2 mg) was conjugated with biotin-PEG3-azide or the

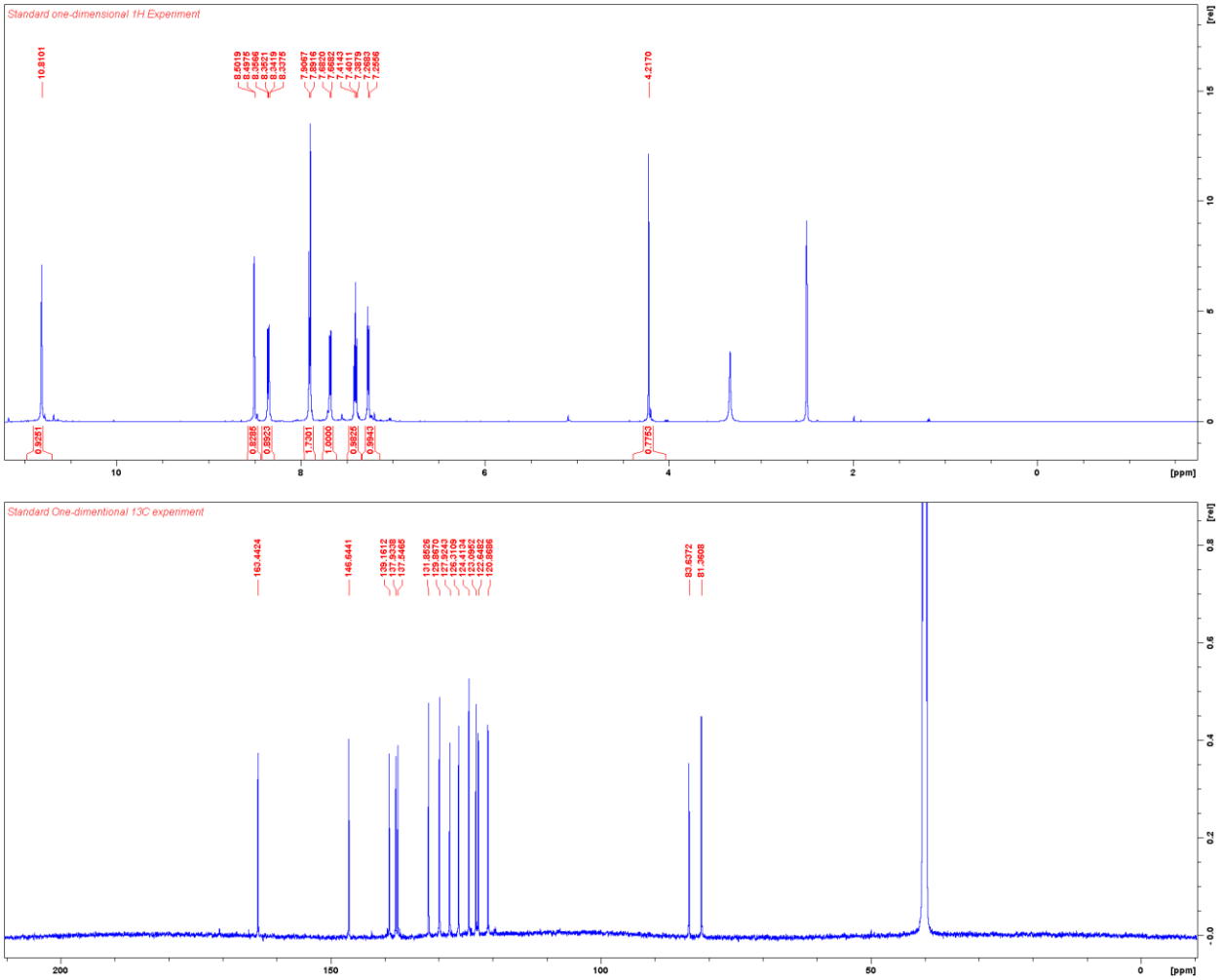
azobenzene-based cleavable linker and precipitated as described above. Pellet was resuspended in 6 M urea / 0.2 % SDS / PBS (500 µL) and dissolved by sonication. Disulfides were reduced with 200 mM TCEP (50 µL, pre-neutralized with potassium carbonate) for 30 min at 37 °C. Reduced thiols were then alkylated with 400 mM iodoacetamide / PBS (70 µL) for 30 min at ambient temperature in dark. Each solution was diluted with SDS / PBS (final SDS concentration: 0.2 %; final total volume: 4 mL) and incubated with the neutravidin agarose resin (150 µL, 50 % slurry) for 1.5 h on a rotator. The neutravidin beads were collected by centrifugation (3,000 g, 2 min) and sequentially washed with 0.1% SDS / PBS (4 mL), PBS (4 mL x 2), 4 M urea / PBS (4 mL), 50 mM NH_4HCO_3 (4 mL) and PBS (4 mL). The resin was transferred to a Protein LoBind tube (Eppendorf) and bound proteins were digested on-bead overnight at 37 °C in 2 M urea / 1 mM CaCl2 / PBS (200 µL) containing sequencing grade trypsin + LysC (2 µg, Promega). The proteolyzed supernatant was transferred to a fresh Protein LoBind tube, acidified with formic acid (5% final). The peptides were extracted with C18 column (6 layers, EMD Millipore) and eluted with 5% formic acid in 30% water / 70% acetonitrile (150 µL) and 5 % formic acid in 10% water / 90% iso-propanol (150 µL). The samples were dried with Speed-vacuum and submitted for analysis. Note that the cleavable linker was used without chemical cleavage.

Proteomics data analysis. Acquired tandem MS spectra were extracted using MaxQuant software and queried against the Uniprot complete human database.[144] Outputs from MaxQuant were processed in R using MSstats and limma packages.[145,146] Proteins identified with a minimum of two unique peptides were considered for downstream analysis. Peptide intensities were log2-transformed and median-centred, and missing values were replaced with random draws from normal distribution centred at estimated censoring thresholds (10^{-4} quantile) with variance tuned to the median of the peptide-wise estimated variances.[147] Peptide intensities were then summarised using median polish algorithm implemented in the MSstats package. Resulting data was subjected to empirical Bayes t-test to assess statistical significance implemented in the limma package. The nominal p-values were corrected for multiple testing by controlling for the false discovery rate (FDR), as proposed by Benjamini and Hochberg.[148] The druggability analysis was performed using a druggable genome as proposed by Wang et al. and liganded proteins curated in BindingDB.[104,105]

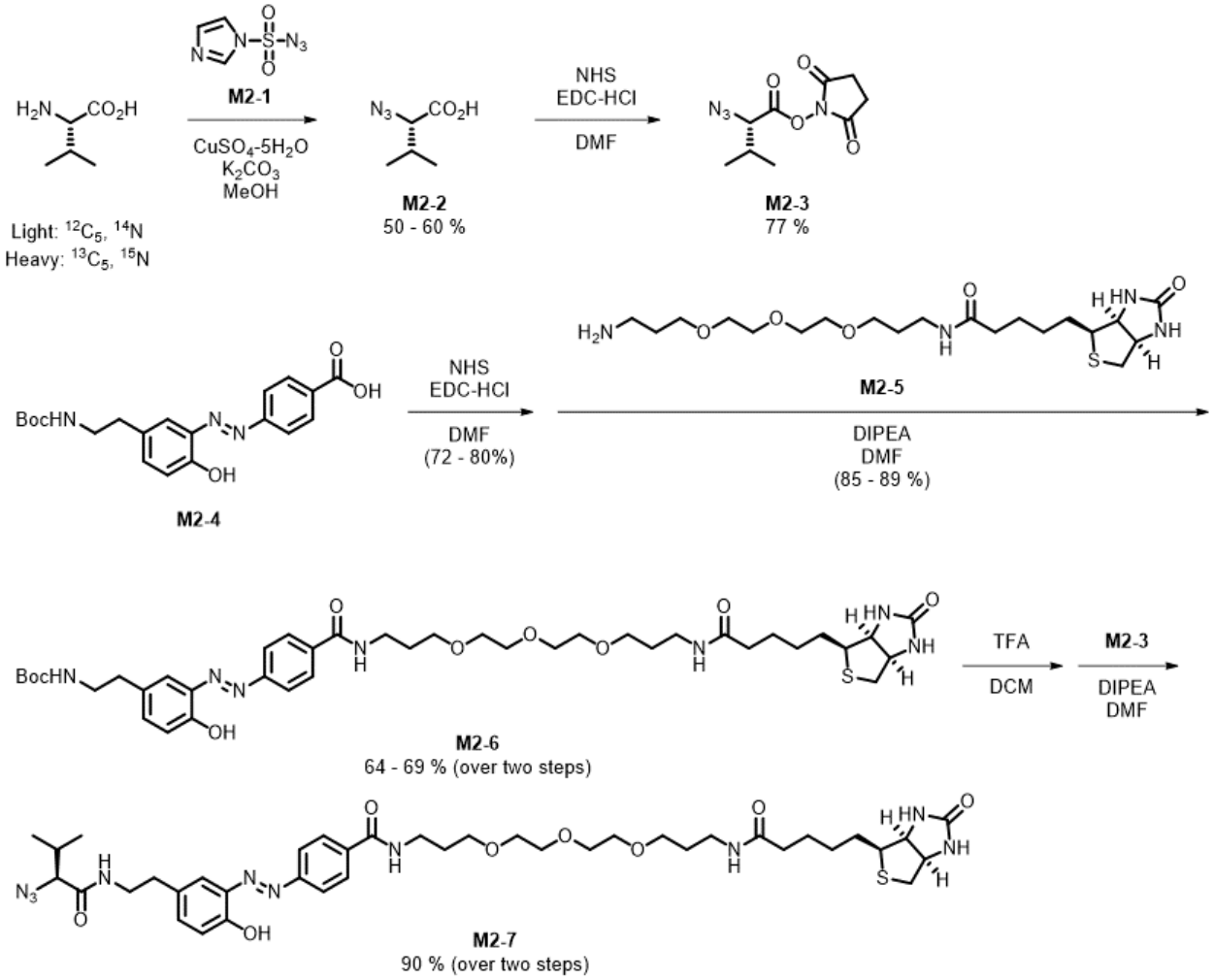
Alternative proteomics data analysis. Proteins identified with a minimum of one unique peptide were considered for downstream analysis.

Data analysis and visualization. Data were analysed and visualised in R. Dose-response analyses were performed using drc package.[149]

Syntheses of alk-GW9662. To a solution of 3-ethynyl aniline (28 µL, 250 µmol, 1.1 equiv.) and diisopropylethylamine (44 µL, 250 µmol, 1.1 equiv.) in dichloromethane (DCM, 200 µL) was added a solution of 2-chloro-5-nitrobenzoyl chloride (50 mg, 227 µmol, 1.0 equiv.) in DCM (300 µL) dropwise on ice. After 15 min, the reaction was allowed to warm to room temperature for 45 min. The reaction mixture was diluted with ethyl acetate (EtOAc, 10 mL), washed with 1 N HCl (10 mL x 3), water (10 mL), sat. $NaHCO_3$ (10 mL) and brine (10 mL). The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude material was crystalized from EtOAc to yield the title compound as white solid (38 mg, 126 µmol, 56 %). ^{1}H NMR (600 MHz, DMSO): δ 10.81 (s, 1H), 8.50 (s, 1 H), 8.34 (d, J = 8.8 Hz, 1 H), 7.89 (d, J = 9.3 Hz, 1 H), 7.89 (s, 1 H), 7.66 (d, J = 8.0 Hz, 1 H), 7.40 (dd, J = 8.5 Hz, 1 H), 7.25 (d, J = 7.5 Hz, 1 H), 4.22 (s, 1 H). ^{13}C NMR (150 MHz, DMSO): δ 163.4 (s), 146.6 (s), 139.2 (s), 137.9 (s), 137.5 (s), 131.9 (s), 129.9 (s), 127.9 (s), 126.3 (s), 124.4 (s), 123.1 (s), 122.6 (s), 120.9 (s), 83.6 (s), 81.4 (s). HR-MS (ESI): $[C_{15}H_9ClN_2O_3 + H]^+$ expected, 301.03745; observed, 301.03775.



Synthesis of the azobenzene-based cleavable linker.



Synthesis of **M2-1**

The diazo transfer reagent **M2-1** was synthesised in three steps from imidazole and thionyl chloride and used in situ as previously described.[150]

Step 1: 1,1'-sulfonylbis(1H-imidazole)

To a solution of imidazole (20 g, 294 mmol) in DCM (160 mL) was added thionyl chloride (5 mL, 62 mmol) dropwise on ice. The reaction mixture was allowed to warm to room temperature overnight and filtered to remove precipitates. The filtrate was concentrated and crystalised from hot isopropanol (40 mL). Yield: 9.82 g (50 mmol, 80 %)

Step 2: 1-((1H-imidazol-1-yl)sulfonyl)-3-methyl-1H-imidazol-3-ium trifluoromethanesulfonate

To a solution of 1,1'-sulfonylbis(1H-imidazole) (292 mg, 1.48 mmol) in DCM (3 mL) was added methyl trifluoromethanesulfonate (146 µL, 1.33 mmol) was added dropwise on ice. After being stirred on ice for 2 h, the reaction mixture was filtered to collect the precipitates. The precipitates were dried in vacuo to thoroughly remove residual DCM.

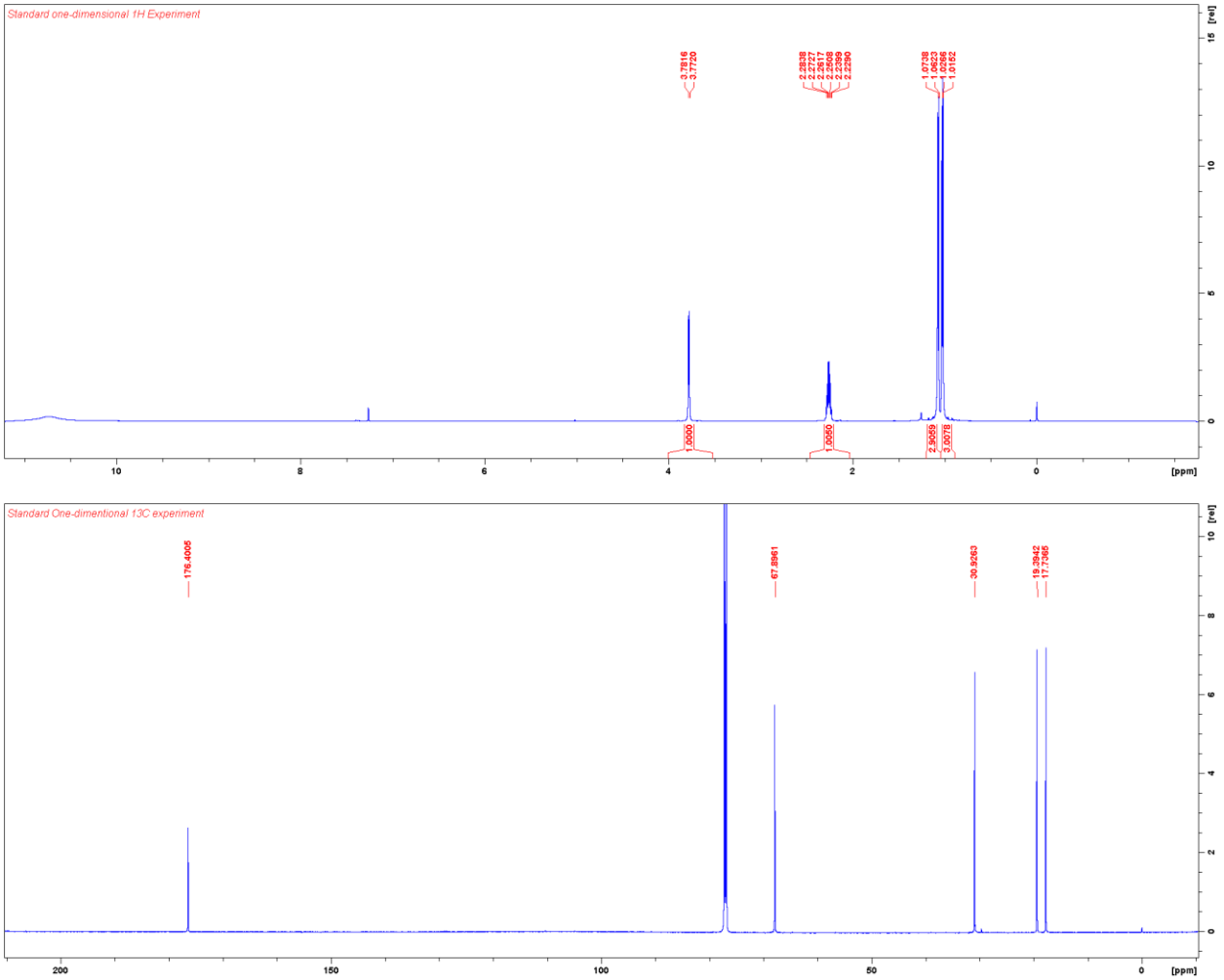
Step 3: 1H-imidazole-1-sulfonyl azide (**M2-1**)

To a solution of the imidazolium salt in water (1.6 mL) and EtOAc (1.6 mL) was added sodium azide (109 mg, 1.67 mmol) portion-wise on ice. The reaction mixture was stirred on ice for 1 h. The organic layer containing the diazo transfer reagent was separated from the aqueous layer and used without further purification.

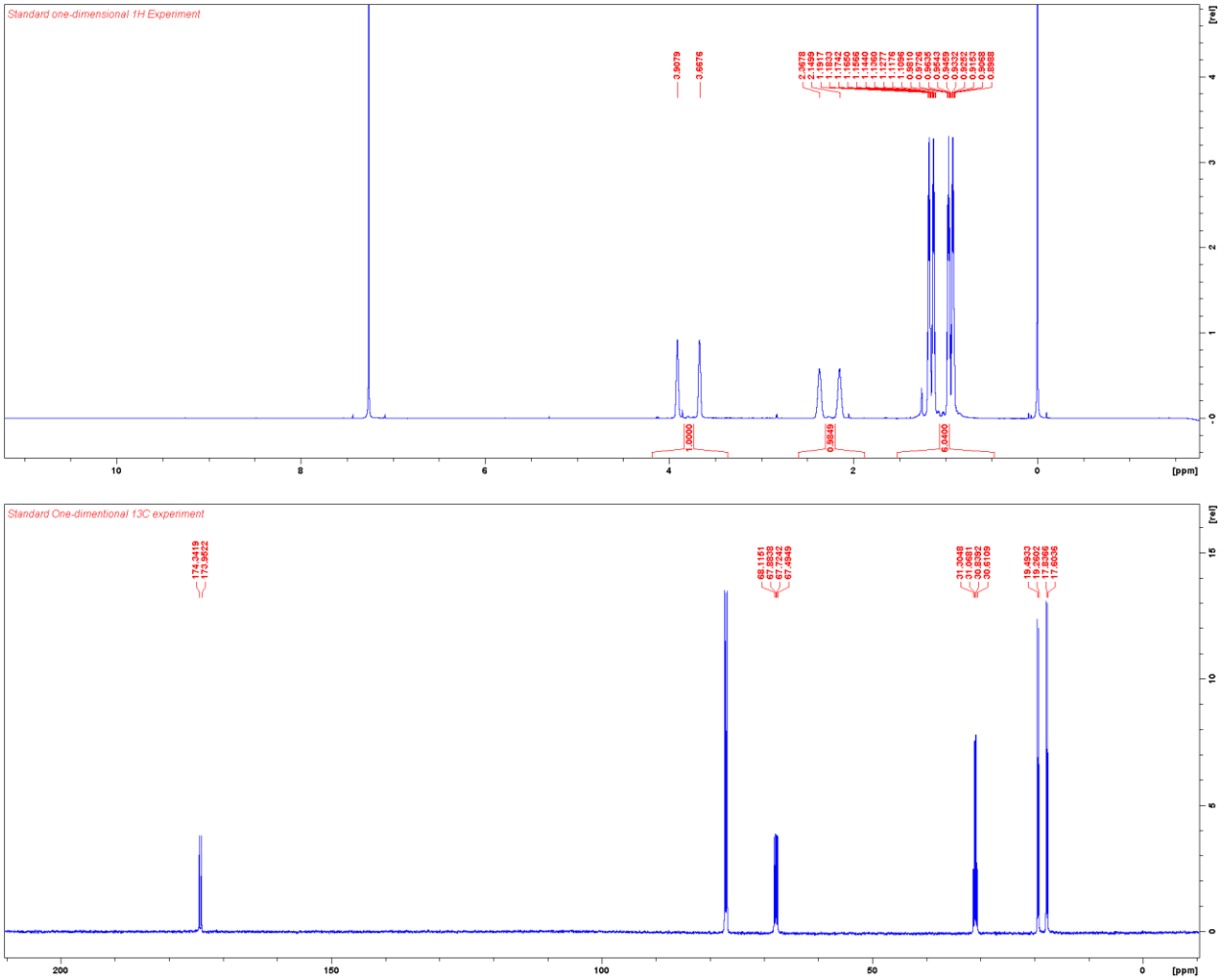
Synthesis of **M2-2**

To a solution of valine (94 mg, 1.0 mmol), K_2CO_3 (166 mg, 1.5 mmol), and $CuSO_4$-$5H_2O$ (2 mg) in MeOH was added the diazo transfer reagent in EtOAc (~ 1.2 mmol) dropwise on ice. The reaction mixture was allowed to warm to room temperature over 2 h, concentrated, diluted with 1 N HCl, and extracted thrice with EtOAc. The organic layer was dried over $MgSO_4$, filtered, and concentrated. Column chromatography using 20 % EtOAc/hexanes + 1 % AcOH yielded the title compound as oil (85 mg, 0.59 mmol, 59%).

M2-2 (light) ^{1}H NMR (600 MHz, $CDCl_3$): δ 3.78 (d, J = 5.8 Hz, 1H), 2.26 (m, 1H), 1.07 (d, J = 6.9 Hz, 3H), 1.02 (d, J = 6.9 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 176.4, 67.9, 30.9, 19.4, 17.7.



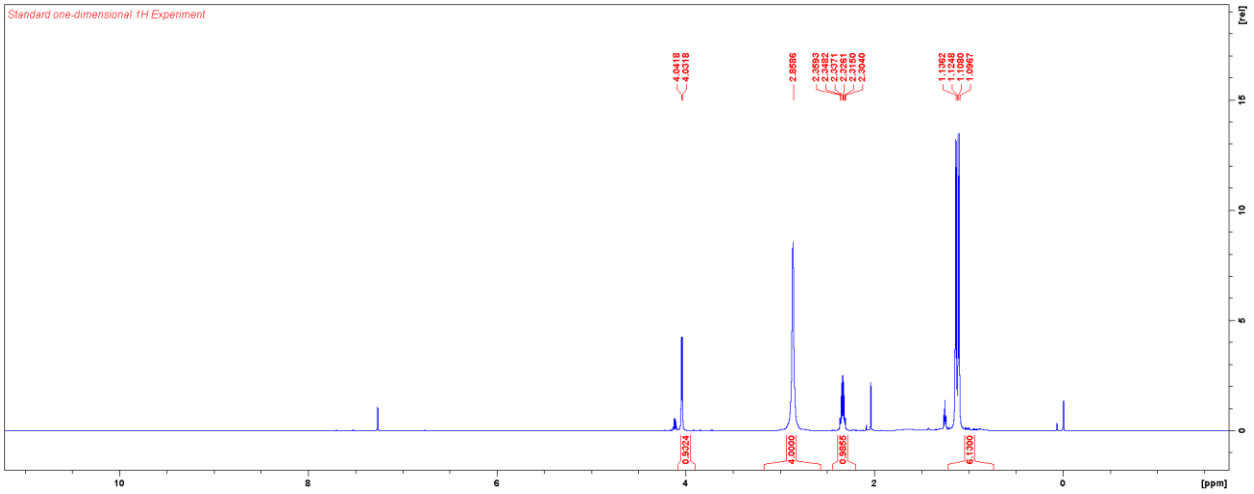
M2-2 (heavy) ^{1}H NMR (600 MHz, $CDCl_3$): δ 3.78 (d, J = 144 Hz, 1H), 2.26 (dd, J = 131 Hz, 3.1 Hz, 1H), 1.07 (m, 3H), 1.02 (m, 3H). ^{13}C NMR (150 MHz, $CDCl_3$): δ 174.18 (d, J = 58.5 Hz), 67.82 (dd, J = 58.5 Hz, 34.7 Hz), 30.9 (ddd, J = 35.0 Hz, 35.0 Hz, 34.7 Hz), 19.40 (d, J = 35.0 Hz), 17.74 (d, J = 35.0 Hz).



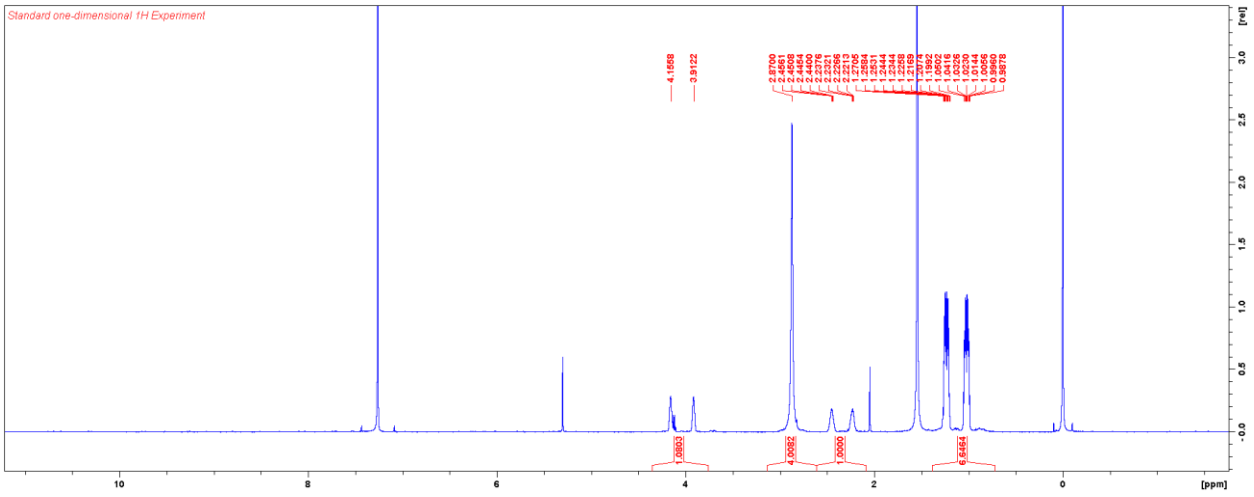
Synthesis of **M2-3**

To a solution of **M2-3** (25 mg, 0.18 μmol) and *N*-hydroxysuccinimide (24 mg, 0.21 mmol) in DMF (2 mL) was added EDC-HCl (40 mg, 0.21 mmol). After being stirred at room temperature overnight, the reaction mixture was diluted with EtOAc, washed with brine and 1 N HCl, dried over $MgSO_4$, filtered, and concentrated. This material was used without further purification. Yield: 0.13 mmol, 77 %.

M2-3 (light) ^{1}H NMR (600 MHz, $CDCl_3$): δ 4.04 (d, J = 6.0 Hz, 1H), 2.87 (br, 4H), 2.33 (m, 1H), 1.13 (d, J = 6.8 Hz, 3H), 1.10 (d, J = 6.8 Hz, 3H).



M2-3 (heavy) ^{1}H NMR (600 MHz, $CDCl_3$): δ 4.03 (d, J = 146 Hz, 1H), 2.89 (br, 4H), 2.34 (dm, J = 131 Hz, 1H), 1.14 (dm, J = 127 Hz, 3H), 1.11 (dm, J = 127 Hz, 3H).

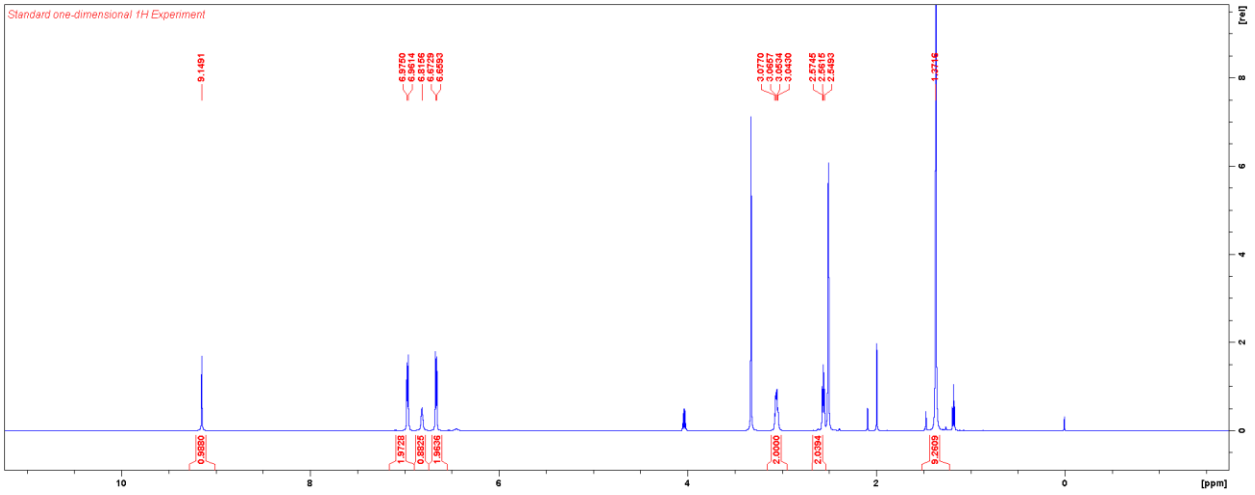


Synthesis of **M2-4**

The azobenzene **M2-4** was synthesised in two steps from tyramine as previously described.[150]

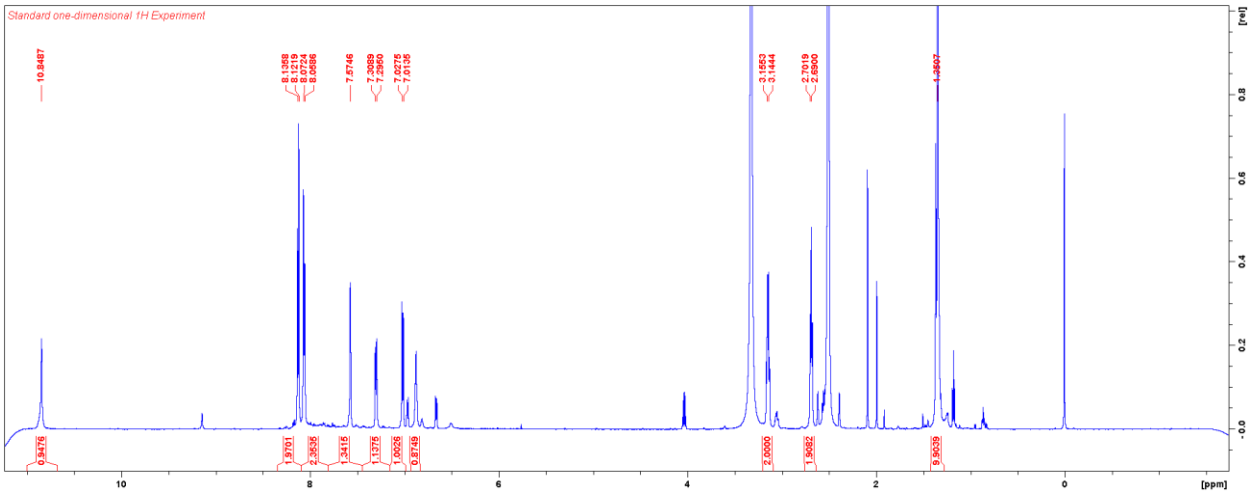
Step 1: *N*-Boc tyramine

A solution of tyramine (2 g, 14.6 mmol) and Boc_2O (3.5 g, 16.0 mmol) in MeOH (20 mL) was stirred at room temperature overnight. The reaction mixture was concentrated, diluted with EtOAc, washed with water, dried over $MgSO_4$, filtered, and concentrated. Column chromatography using 25 % EtOAc/hexanes yielded *N*-Boc tyramine quantitatively. ^{1}H NMR (600 MHz, $(CD_3)_2SO$): δ 9.15 (s, 1H), 6.97 (d, J = 8.2 Hz, 2H), 6.82 (br, 1H), 6.67 (d, J = 8.2 Hz), 3.06 (m, 2H), 2.56 (t, J = 7.6 Hz), 1.37 (s, 9H).



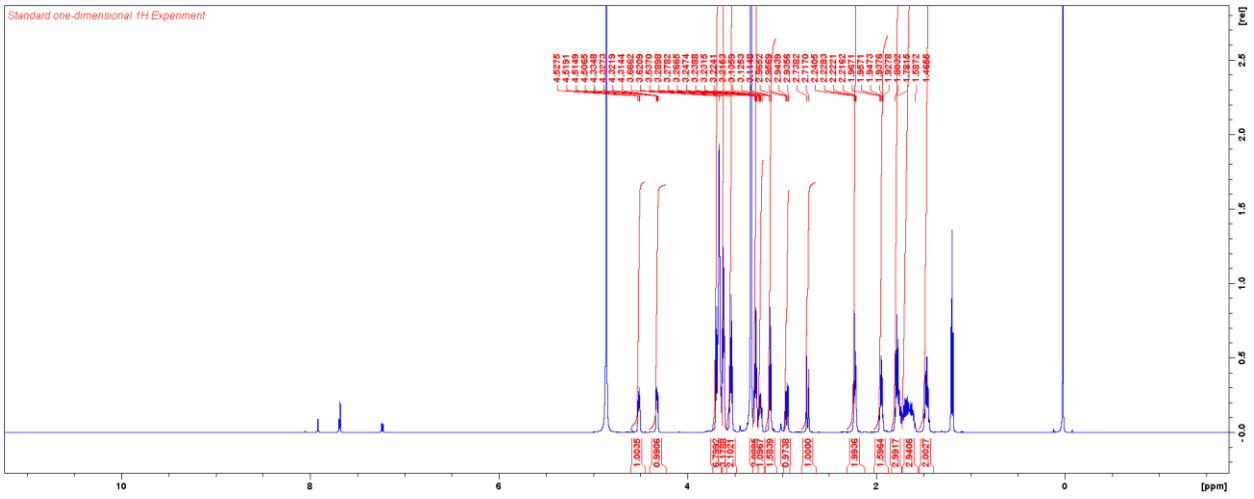
Step 2: (E)-4-((5-(2-((tert-butoxycarbonyl)amino)ethyl)-2-hydroxyphenyl)diazenyl)benzoic acid (**M2-4**)

To a solution of aminobenzoic acid (2.41 g, 17.7 mmol) in 6 N HCl (50 mL) was added $NaNO_2$ (3.09 g, 44.8 mmol) on ice. The reaction mixture was stirred on ice for 20 min. To this *N*-Boc tyramine (4.20 g, 17.7 mmol), sat. $NaHCO_3$ (150 mL), $NaHCO_3$ (20 g), and acetone (20 mL) was added. After being stirred overnight, the reaction mixture was quenched with 1 N HCl (500 mL). The precipitates were collected by filtration and washed with water and acetone to yield **M2-4** as a brown solid. Yield: 3.2 g (8.3 mmol, 47%). ^{1}H NMR (600 MHz, $(CD_3)_2SO$): δ 10.85 (s, 1H), 8.13 (d, J = 8.3 Hz, 2H), 8.07 (d, J = 8.3 Hz, 2H), 7.57 (s, 1H), 7.30 (d, J = 8.3 Hz, 1H), 7.02 (d, J = 8.3 Hz), 6.88 (br, 1H), 3.15 (m, 2H), 2.69 (t, J = 7.1 Hz, 2H), 1.36 (s, 9H).



Synthesis of **M2-5**

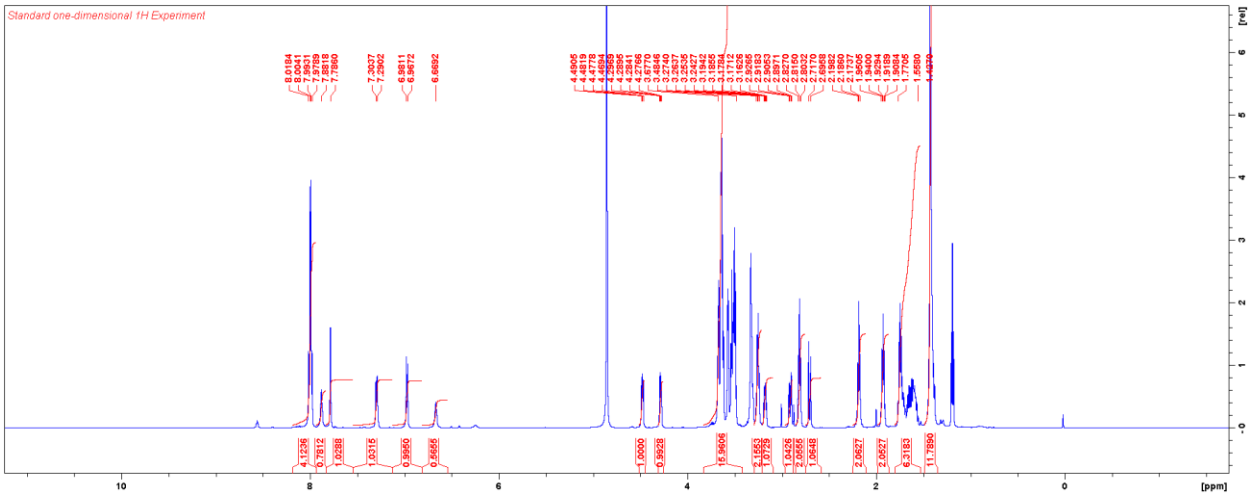
Biotin (1.00 g, 4.09 mmol) was dissolved in DMF (20 mL) at 70 °C and cooled to room temperature. To this were added triethylamine (1.00 mL, 7.18 mmol) and pentafluorophenyl trifluoroacetate (1.00 mL, 5.82 mmol). After being stirred at room temperature for 30 min, the reaction mixture was concentrated. The residue was triturated with diethyl ether and dissolved in DMF (40 mL). This was added dropwise over 15 min to a mixture of 4,7,10-trioxa-1,13-tridecanediamine (3.76 g, 17.1 mmol) and TEA (0.8 mL, 5.74 mmol) on ice. After being stirred on ice for 1 h, the reaction mixture was evaporated. The residue was triturated with diethyl ether to yield the title compound as a white solid (700 mg, 50 %). ^{1}H NMR (600 MHz, CD_3OD): δ 4.51 (m, 1H), 4.33 (m, 1H), 3.70–3.61 (m, 10H), 3.54 (m, 2H), 3.28 (t, J = 7.0 Hz, 2H), 3.23 (m, 1H), 3.13 (t, J = 6.4 Hz, 2H), 2.95 (dd, J = 12.9 Hz, 5.0 Hz, 1H), 2.73 (d, J = 12.7 Hz, 1H), 2.23 (m, 2H), 1.95 (m, 2H), 1.80–1.59 (m, 6H), 1.47 (m, 2H).

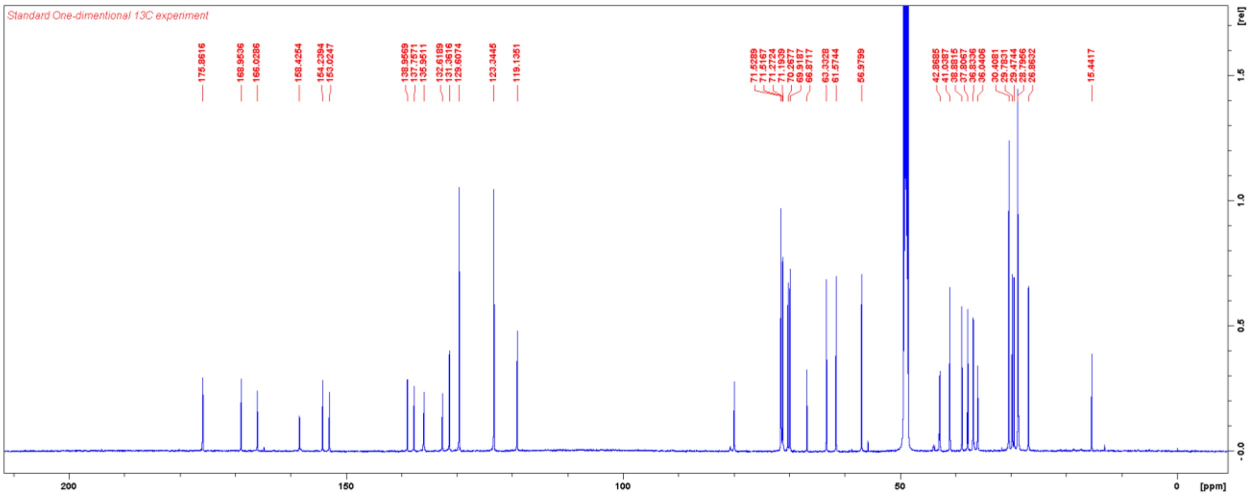


Synthesis of **M2-6**

To a solution of **M2-4** (150 mg, 0.39 mmol) and *N*-hydroxysuccinimide (54 mg, 0.47 mmol) in DMF (5 mL) was added EDC-HCl (90 mg, 0.47 mmol). After being stirred at room temperature overnight, the reaction mixture was diluted with EtOAc, washed with brine, dried over $MgSO_4$, filtered, concentrated, and flash chromatographed using 33 % EtOAc/hexanes. Yield: 135 mg (0.28 mmol, 72 %).

To a solution of the NHS-ester (135 mg, 0.28 mmol) and DIPEA (73 µL, 0.42 mmol) in DMF (5 mL) was added **M2-5** (188 mg, 0.42 mmol). After being stirred at room temperature for 1 h, the reaction mixture was diluted with chloroform, washed with 1 N HCl and brine, dried over $MgSO_4$, filtered, concentrated, and flash chromatographed using 5 to 10 % MeOH/DCM. Yield: 202 mg (0.25 mmol, 89 %). ^{1}H NMR (600 MHz, CD_3OD): δ 7.99 (m, 4 H), 7.88 (br, 1H), 7.79 (s, 1H), 7.30 (d, J = 8.1 Hz, 1H), 6.98 (d, J = 8.3 Hz, 1H), 6.67 (br, 1H), 4.48 (m, 1H), 4.28 (m, 1H), 3.68–3.48 (m, 16H), 3.26 (m, 2H), 3.17 (m, 1H), 2.91 (dd, J = 12.9 Hz, 5.0 Hz, 1H), 2.82 (t, J = 7.2 Hz, 2H), 2.71 (d, J = 12.7 Hz, 1H), 2.19 (m, 2H), 1.93 (m, 2H), 1.77–1.56 (m, 6H), 1.43 (m, 11H). ^{13}C NMR (150 MHz, CD_3OD): δ 175.9,169.0, 166.0, 158.4, 154.2, 153.0, 139.0, 137.8, 136.0, 132.6, 131.4, 129.6, 123.3, 119.1, 71.5, 71.5, 71.3, 71.2, 70.3, 69.9, 66.9, 63.3, 61.6, 57.0, 42.9, 41.0, 38.9, 37.8, 36.8, 36.0, 30.4, 29.8, 29.5, 28.8, 26.9, 15.4.

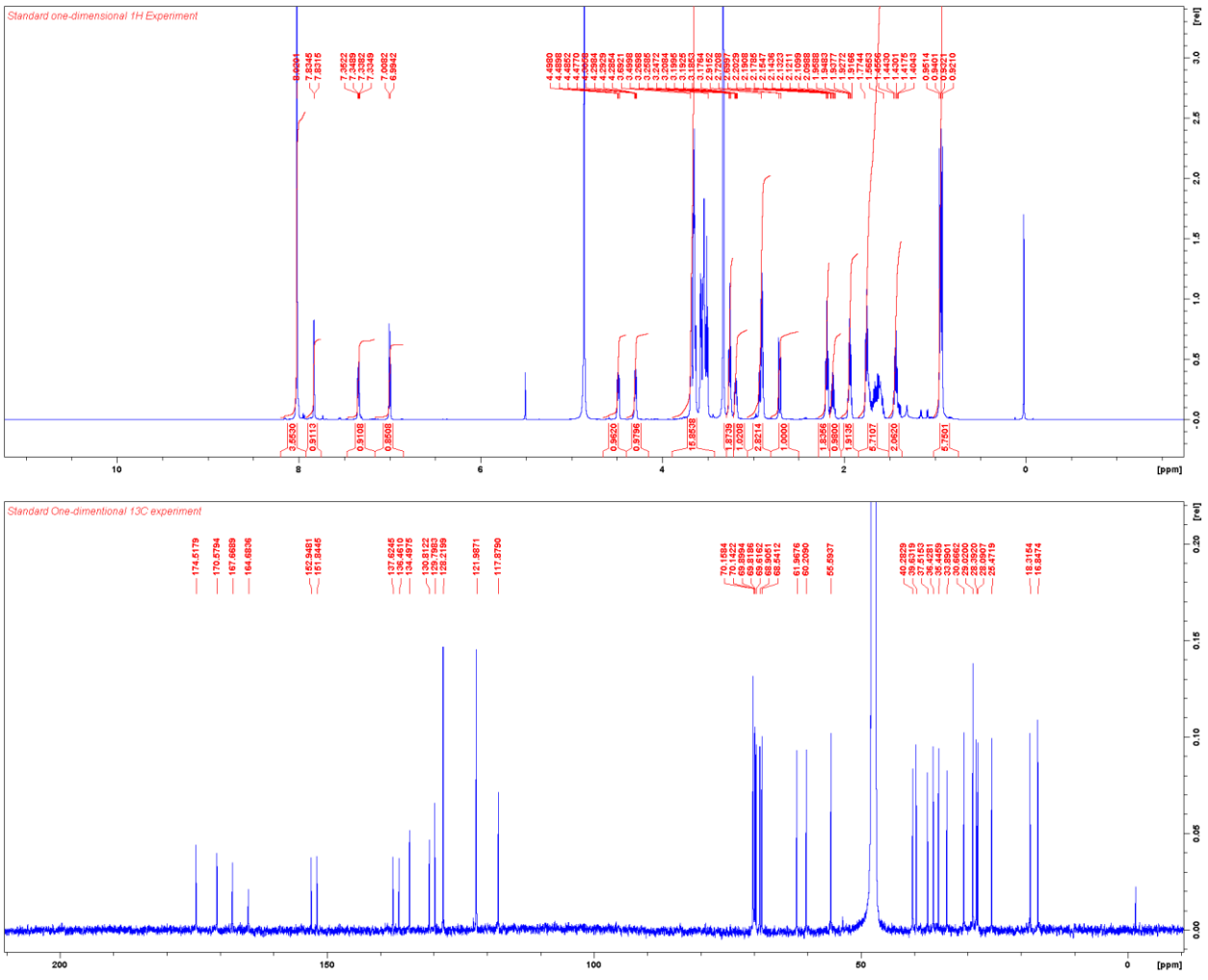



Standard One-dimentional 13C experiment
175.8616
168.9536
166.0286
158.4254
154.2394
153.0247
138.9569
137.7571
135.9511
132.6189
131.3616
129.6074
123.3445
119.1351
71.5289
71.5167
71.2724
71.1939
70.2677
69.9187
66.8717
63.3328
61.5744
56.9799
42.8685
41.0387
38.8815
37.8067
36.8336
36.0406
30.4081
29.7831
29.4744
28.7956
26.8632
15.4417
[rel]
1.5
1.0
0.5
-0.0
200
150
100
50
0
[ppm]

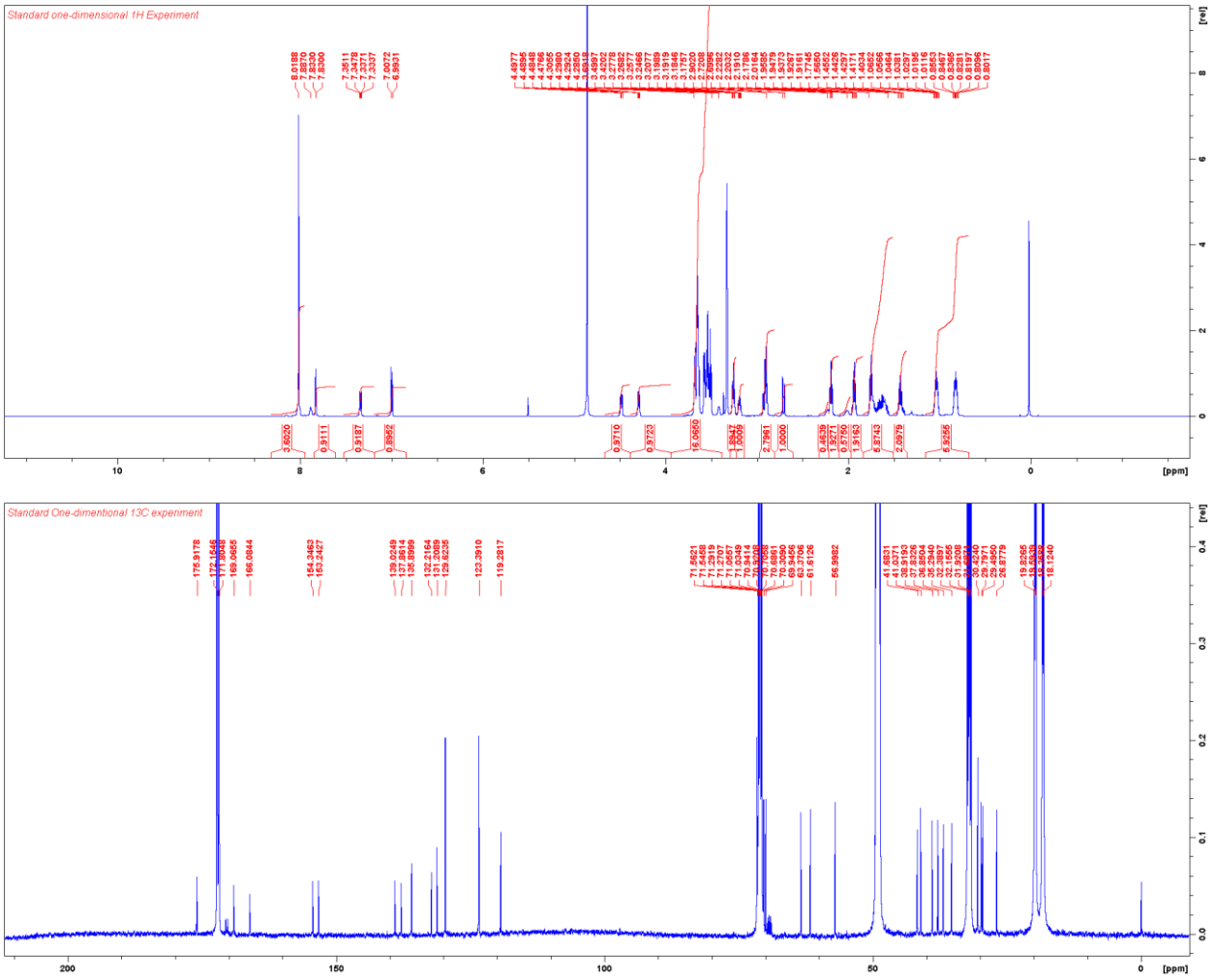
Synthesis of **M2-7**

To a solution of **M2-6** (40 mg, 50 µmol) in DCM (1 mL) was added TFA (0.5 mL). After being stirred for 1 h, the reaction mixture was evaporated and azeotroped with toluene. The residue was dissolved in DMF (0.5 mL). To this were added DIPEA (88 µL, 500 µmol) and **M2-3** (10 mg, 40 µmol). After being stirred at room temperature for 1 h, the reaction mixture was diluted with DCM, washed with 1 N HCl and brine, dried over $MgSO_4$, filtered, concentrated, and flash chromatographed using 5 to 10 % MeOH/DCM. Yield: 38 mg (45 µmol, 90 %).

M2-7 (light) $[C_{40}H_{58}N_{10}O_8S+H]^+$ expected, 839.4233; observed, 839.4225. ^{1}H NMR (600 MHz, CD_3OD): δ 8.02 (m, 4 H), 7.83 (d, J = 1.8 Hz, 1H), 7.34 (dd, J = 6.4 Hz, 2.0 Hz, 1H), 7.00 (d, J = 6.4 Hz, 1H), 4.48 (m, 1H), 4.29 (m, 1H), 3.69–3.50 (m, 16H), 3.26 (t, J = 6.8 Hz, 2H), 3.19 (m, 1H), 2.92 (m, 3H), 2.71 (d, J = 12.7 Hz, 1H), 2.19 (m, 2H), 2.13 (m, 1H), 1.94 (m, 2H), 1.77–1.56 (m, 6H), 1.43 (m, 2H), 0.95 (d, J = 6.7 Hz, 3H), 0.93 (d, J = 6.7 Hz, 3H). ^{13}C NMR (150 MHz, CD_3OD): δ 175.9,172.0, 169.1, 166.1, 154.4, 153.2, 139.0, 137.9, 135.9, 132.2, 131.2, 129.6, 123.4, 119.3, 71.6, 71.5, 71.3, 71.2, 71.0, 70.3, 69.9, 63.4, 61.6, 57.0, 41.7, 41.0, 38.9, 37.8, 36.8, 35.3, 32.1, 30.4, 29.8, 29.5, 26.9, 19.7, 18.3.



M2-7 (heavy) $[C_{35}{}^{13}C_5H_{58}N_9{}^{15}NO_8S+H]^+$ expected, 845.4360; observed, 845.4358. ^{1}H NMR (600 MHz, CD_3OD): δ 8.02 (m, 4 H), 7.83 (d, J = 1.8 Hz, 1H), 7.34 (dd, J = 6.4 Hz, 2.0 Hz, 1H), 7.00 (d, J = 6.4 Hz, 1H), 4.48 (m, 1H), 4.29 (m, 1H), 3.69–3.42 (m, 16H), 3.26 (m, 2H), 3.19 (m, 1H), 2.92 (m, 3H), 2.71 (d, J = 12.7 Hz, 1H), 2.19 (m, 2H), 2.12 (d, J = 127 Hz, 1H), 1.94 (m, 2H), 1.77–1.56 (m, 6H), 1.43 (m, 2H), 0.93 (dm, J = 126 Hz, 6H). ^{13}C NMR (150 MHz, CD_3OD): δ 175.92, 172.00 (d, J = 52.5 Hz), 169.07, 166.08, 154.35, 153.24, 139.02, 137.86, 135.90, 132.22, 131.21, 129.62, 123.39, 119.28, 71.56, 71.55, 70.99 (ddd, J = 52.5 Hz, 35.3 Hz, 3.1 Hz), 70.31, 69.95, 63.37, 61.61, 57.00, 41.68, 41.04, 38.92, 37.83, 36.85, 35.29, 32.04 (ddd, J = 35.2 Hz, 35.1 Hz, 35.1 Hz), 30.42, 29.80, 29.50, 26.88, 19.71 (d, J = 34.9 Hz), 18.24 (d, J = 35.2 Hz).



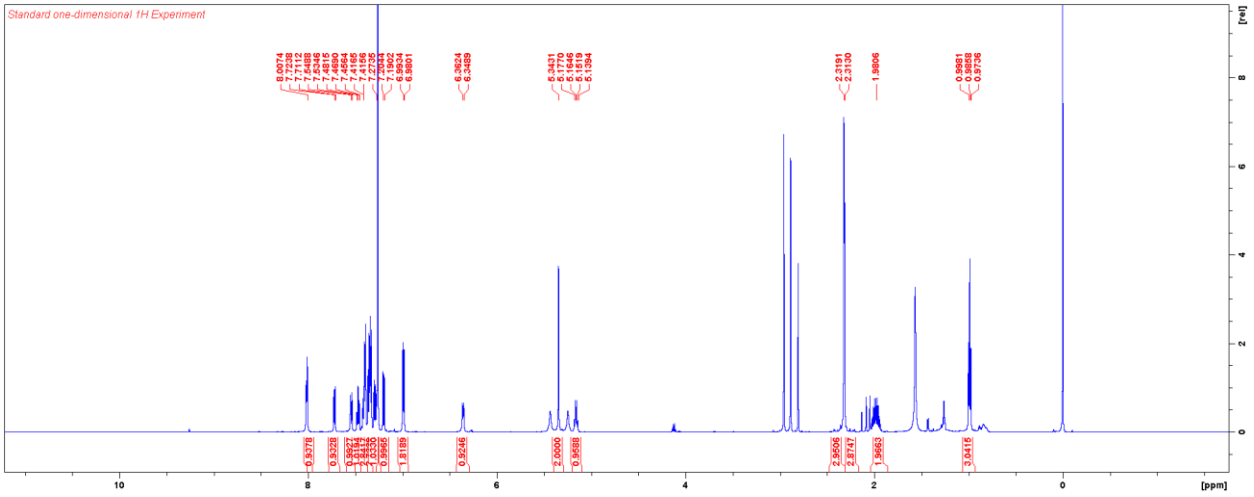
Chapter 3

List of E3 ligases was adopted from UbiHub.[124]

Synthesis of (*R*)-SR9034 (compound **9**)

(*R*)-SR9034 was synthesised as previously described.[135]

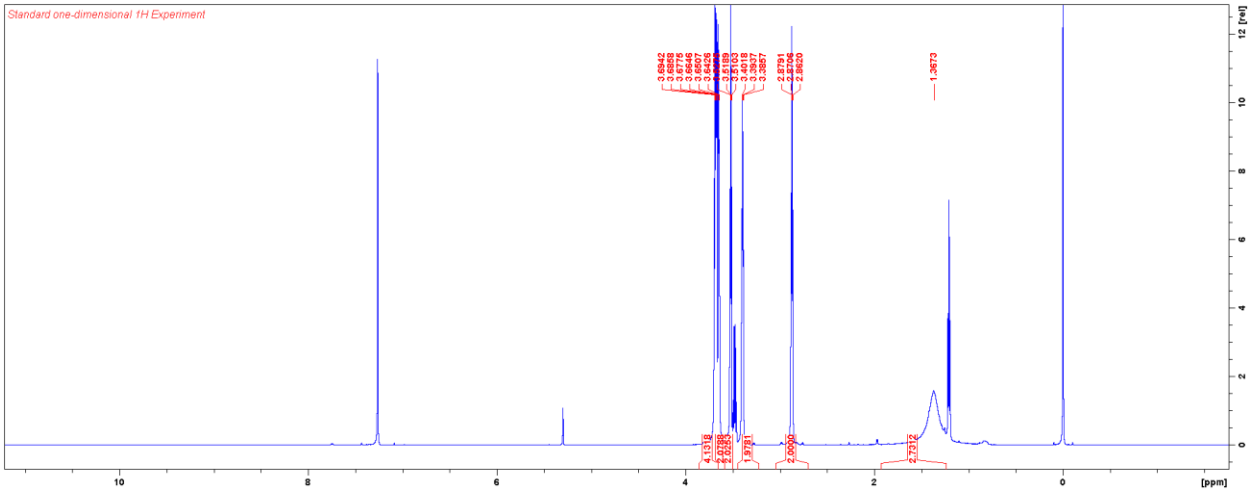
^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.01 (s, 1H), 7.72 (d, J = 7.6 Hz, 1H), 7.54 (d, J = 8.5 Hz, 1H), 7.47 (t, J = 7.6 Hz, 1H), 7.42–7.27 (m, 9H), 7.20 (d, J = 8.5 Hz, 1H), 6.99 (d, J = 8.0 Hz, 2H), 6.36 (d, J = 8.1 Hz, 1H), 5.34 (s, 2H), 5.16 (q, J = 7.5 Hz, 1H), 2.32 (s, 3H), 2.31 (s, 3H), 1.98 (m, 2H), 0.99 (t, J = 7.5 Hz, 3H).



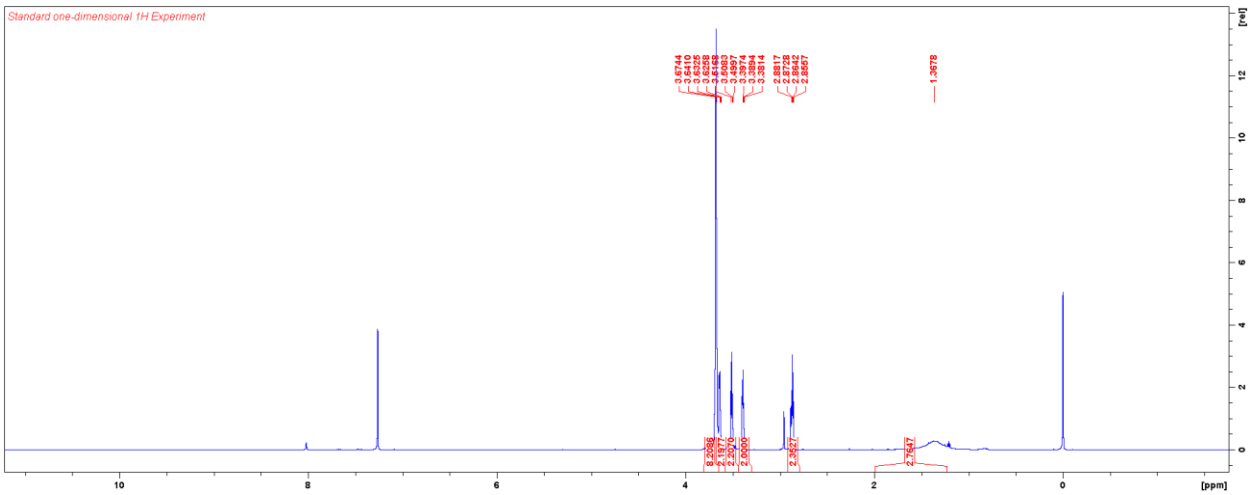
General method for the syntheses of azide-PEG-NH_2 linkers (compounds **10**–**13**)

A commercially-available PEG bis(tosylate) (1 equiv.) was dissolved in DMF (1 M). To this sodium azide (4 equiv.) was added. The reaction mixture was heated at 70 °C overnight. DMF was removed by evaporation, and the residue was dissolved in Et_2O. Insoluble material was removed by filtration. The Et_2O was removed by evaporation to yield a diazide intermediate. The diazide intermediate was dissolved in Et_2O/1 N HCl/THF (5:5:1) (~ 0.3 M). To this triphenylphosphine in Et_2O (1 equiv.) was added dropwise. The reaction mixture was stirred at room temperature overnight. Insoluble material was removed by filtration. The organic layer was extracted three times with 4 N HCl. Combined aqueous layer was washed with Et_2O, basified with NaOH, and extracted with DCM. Combined DCM layer was dried over Na_2SO_4, filtered, and concentrated to yield an azide-PEG-NH_2 linker (81–86 %).

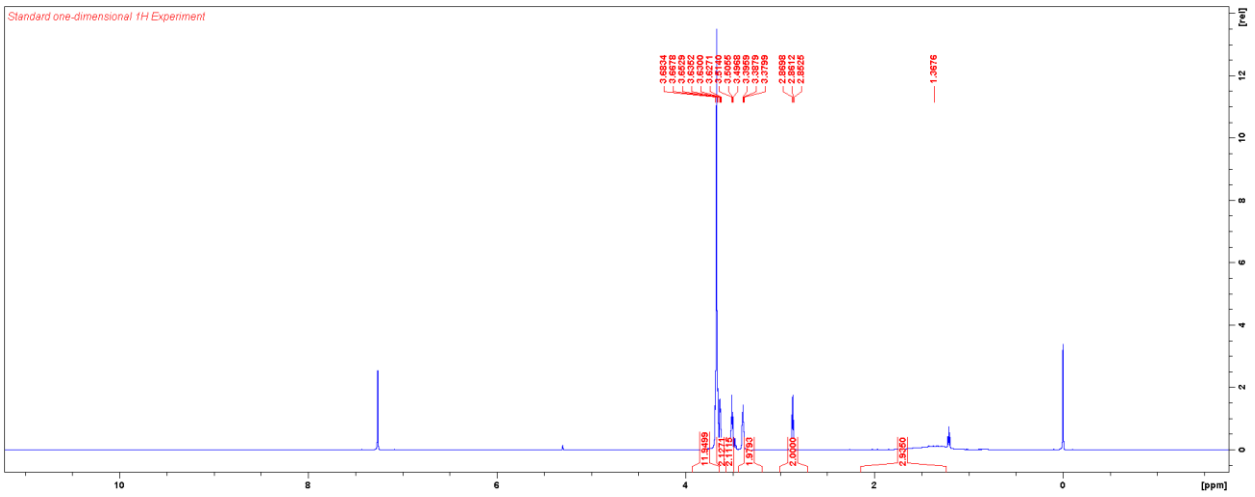
Azide-PEG_2-NH_2 linker (**10**); Yield: 86 % from triethylene glycol bis(*p*-toluenesulfonate); ^1H NMR ($CDCl_3$, 600 MHz): δ 3.68 (m, 4H), 3.64 (m, 2H), 3.52 (t, J = 5.2 Hz, 2H), 3.39 (t, J = 4.8 Hz, 2 H), 2.87 (t, J = 5.2 Hz, 2H), 1.37 (br, 2H).



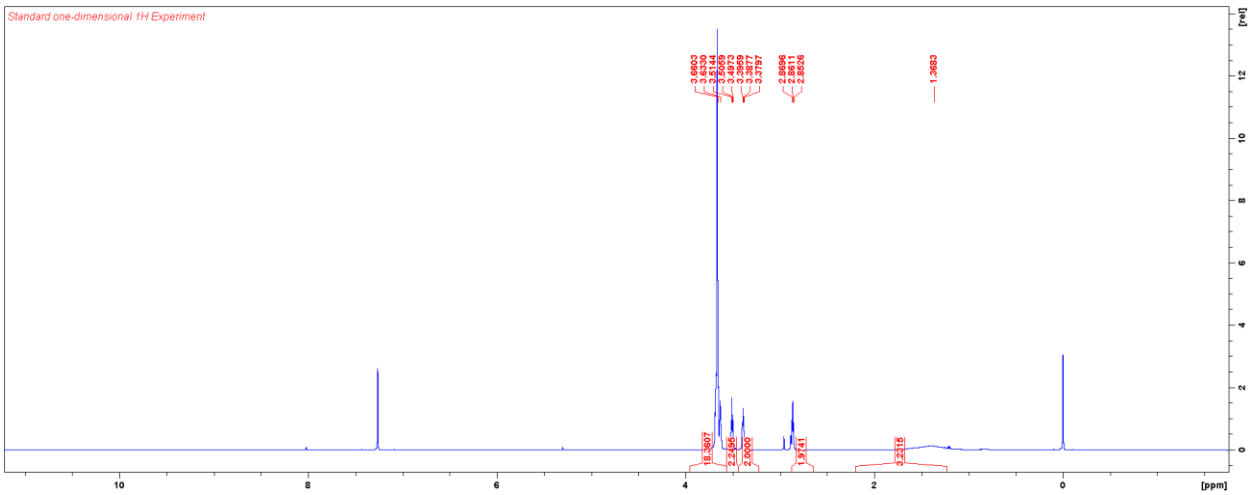
Azide-PEG$_3$-NH$_2$ linker (**11**); Yield: n.d. from tetraethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 3.67 (m, 8H), 3.63 (m, 2H), 3.51 (t, J = 5.2, 2H), 3.39 (t, J = 4.8, 2H), 2.86 (t, J = 5.2, 2H), 1.37 (br, 2H).



Azide-PEG$_4$-NH$_2$ linker (**12**); Yield: n.d. from pentaethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 3.67 (m, 12H), 3.63 (m, 2H), 3.51 (t, J = 5.2, 2H), 3.39 (t, J = 4.8, 2H), 2.86 (t, J = 5.2, 2H), 1.37 (br, 2H).



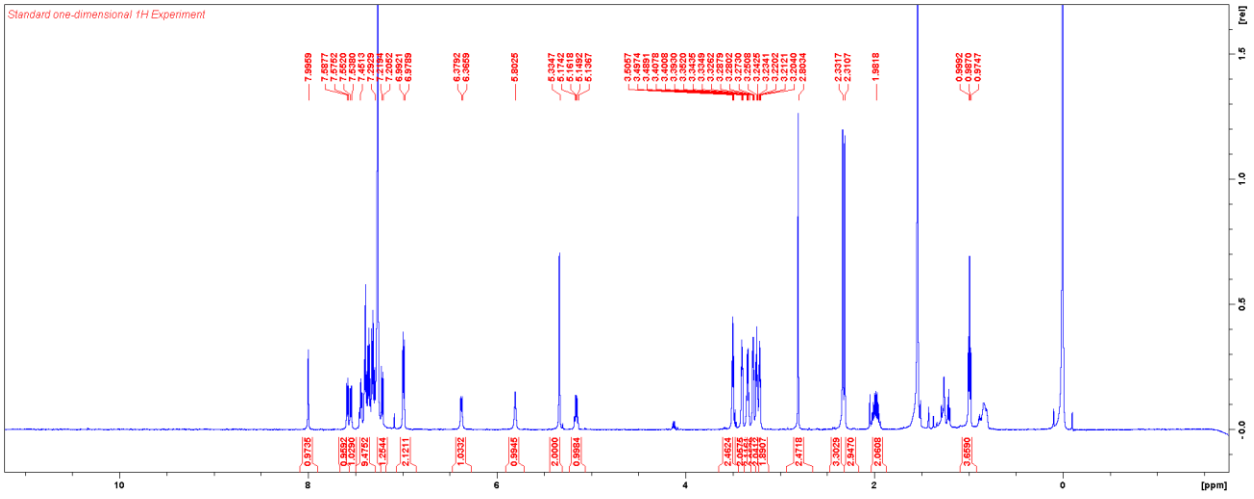
Azide-PEG$_5$-NH$_2$ linker (**13**); Yield: 81 % from hexaethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 3.66 (m, 16H), 3.63 (m, 2H), 3.51 (t, J = 5.2, 2H), 3.39 (t, J = 4.8, 2H), 2.86 (t, J = 5.2, 2H), 1.37 (br, 2H).



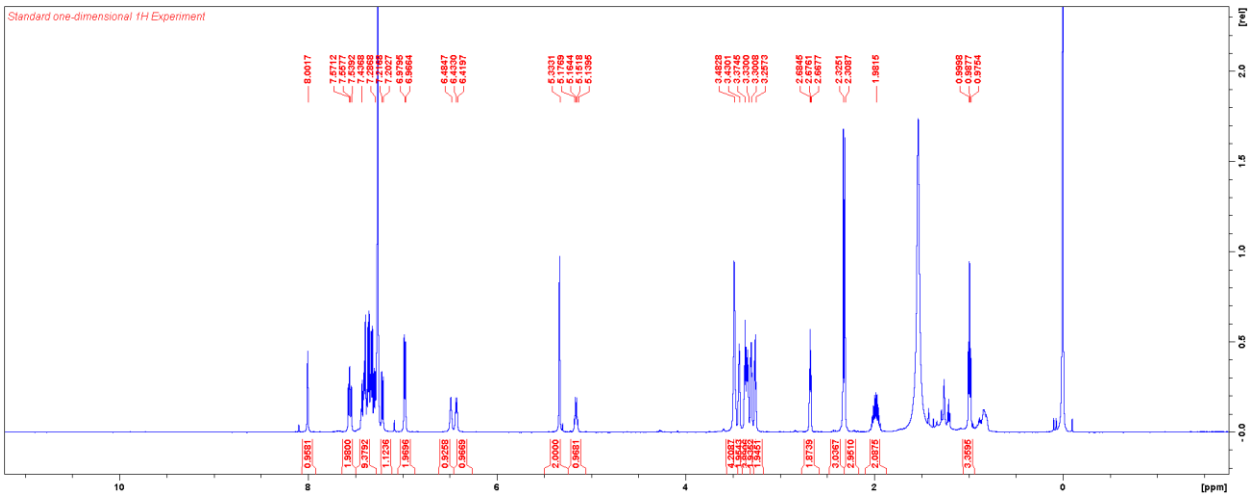
General method for the synthesis of (*R*)-SR9034-PEG-N_3 (compounds **14**–**17**)

To a solution of (*R*)-SR9034 (1 equiv.) and HATU (1 equiv.) in DCM (0.2 M) was added DIPEA (5 equiv.). The reaction mixture was stirred on ice for 5 min. To this an azide-PEG-NH_2 linker in DCM (1 equiv., 0.2 M) was added. The reaction mixture was stirred at room temperature for 1 h. The reaction mixture was diluted with EtOAc, washed with sat. $NaHCO_3$, sat NH_4Cl, and sat NaCl. The organic layer was dried over Na_2SO_4, filtered, and concentrated. Purification on CombiFlash Rf using 0–100 % EtOAc/DCM or 0–10% MeOH/DCM yielded a (*R*)-SR9034-PEG-N_3.

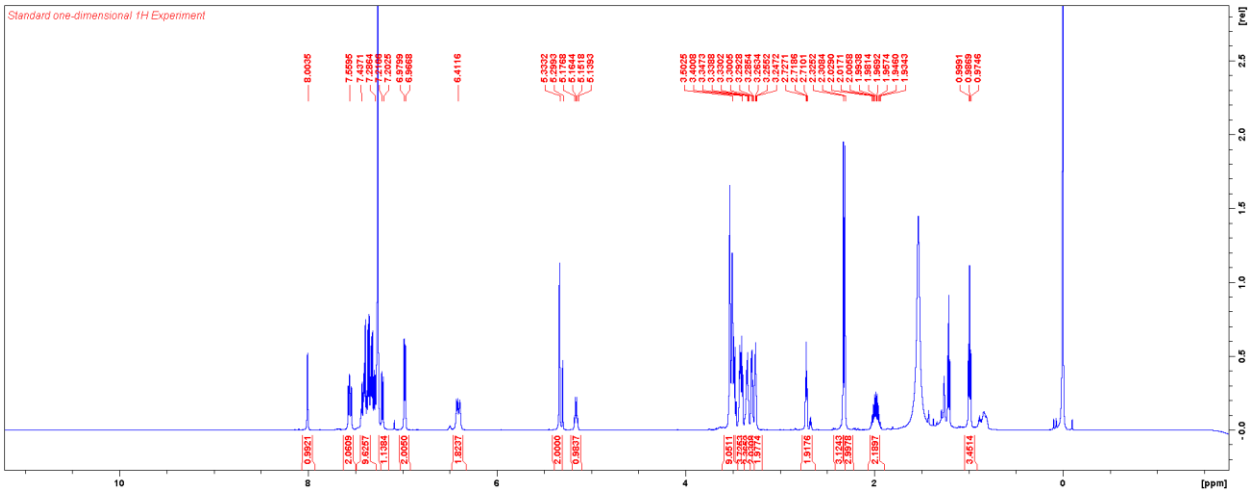
(*R*)-SR9034-PEG_2-N_3 (compound **14**); 25 mg (37 µmol, 77 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.00 (s, 1H), 7.58 (d, J = 7.2 Hz, 1H), 7.54 (d, J = 8.4 Hz, 1H), 7.45–7.29 (m, 9H), 7.21 (d, J = 8.4 Hz, 1H), 6.98 (d, J = 7.8, 2H), 6.37 (d, J = 7.8 Hz, 1H), 5.80 (br, 1H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.50 (t, J = 4.8, 2 H), 3.40 (t, J = 4.2 Hz, 2H), 3.34 (m, 2H), 3.28 (t, J = 4.2 Hz, 2H), 3.24 (t, J = 4.8 Hz 2H), 2.80 (br, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.97 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



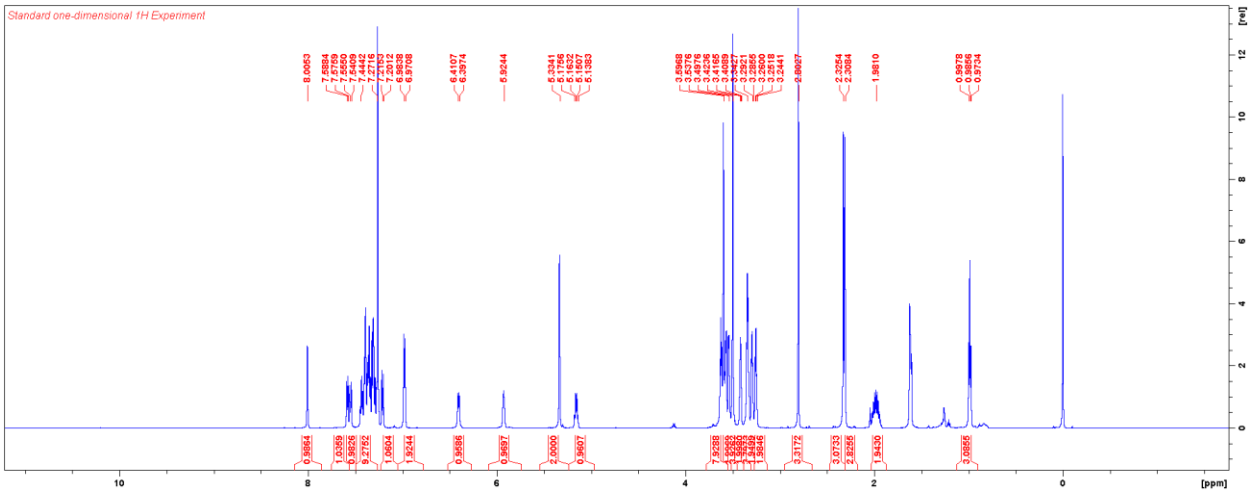
(*R*)-SR9034-PEG$_3$-N$_3$ (compound **15**); 55 mg (77 µmol, 79 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.00 (s, 1H), 7.56 (m, 2H), 7.44–7.25 (m, 9H), 7.20 (d, J = 8.5 Hz, 1H), 6.98 (d, J = 7.8, 2H), 6.48 (br, 1H), 6.42 (d, J = 7.8 Hz, 1 H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.48 (m, 4H), 3.43 (t, J = 4.2 Hz, 2H), 3.35 (m, 4H), 3.30 (t, J = 4.2 Hz, 2H), 3.26 (t, J = 4.8 Hz, 2H), 2.68 (t, J = 5.0, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.98 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



(*R*)-SR9034-PEG$_4$-N$_3$ (compound **16**); 53 mg (70 µmol, 77 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.00 (s, 1H), 7.56 (m, 2H), 7.44–7.25 (m, 9H), 7.20 (d, J = 8.5 Hz, 1H), 6.98 (d, J = 7.8, 2H), 6.41 (m, 2H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.50 (m, 8H), 3.40 (m, 4H), 3.34 (t, J = 5.2 Hz, 2H), 3.29 (t, J = 4.4 Hz, 2H), 3.26 (t, J = 5.0 Hz, 2H), 2.72 (t, J = 5.1, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.98 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



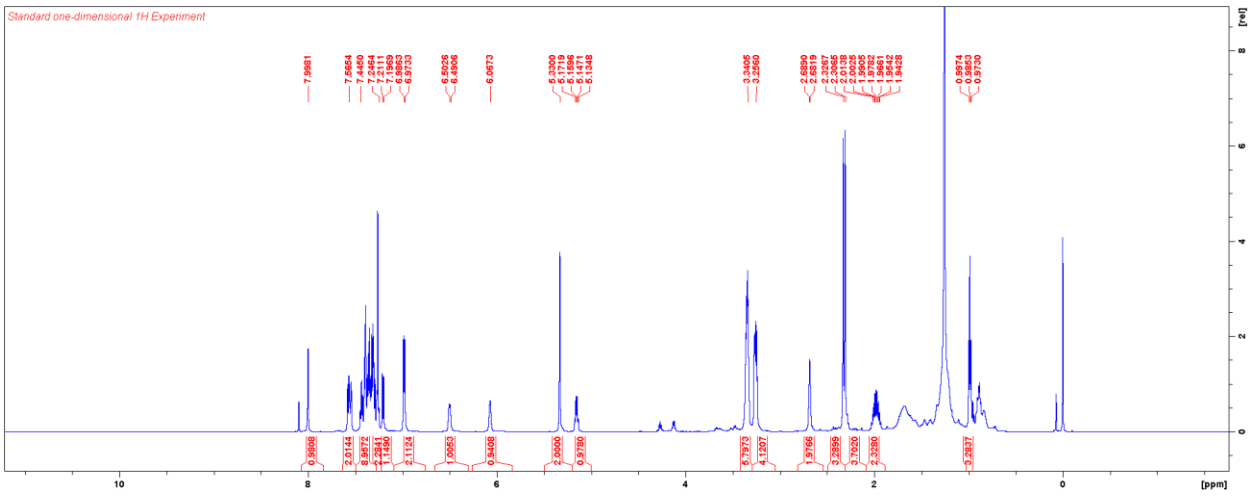
(*R*)-SR9034-PEG$_5$-N$_3$ (compound **17**); 79 mg (98 µmol, 84 %); ^{1}H NMR (CDCl$_3$, 600 MHz): δ 8.00 (s, 1H), 7.58 (d, J = 7.8 Hz, 1H), 7.55 (d, J = 8.4 Hz, 1 H), 7.44–7.27 (m, 9H), 7.20 (d, J = 8.5 Hz, 1H), 6.98 (d, J = 7.8, 2H), 6.40 (d, J = 6.8 Hz, 1H), 5.92 (br, 1H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.60 (m, 6H), 3.54 (m, 2H), 3.50 (m, 4H), 3.42 (t, J = 4.4, 2H), 3.34 (m, 4 H), 3.29 (t, J = 4.0 Hz, 2H), 3.25 (t, J = 5.0 Hz, 2H), 2.80 (s, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.98 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



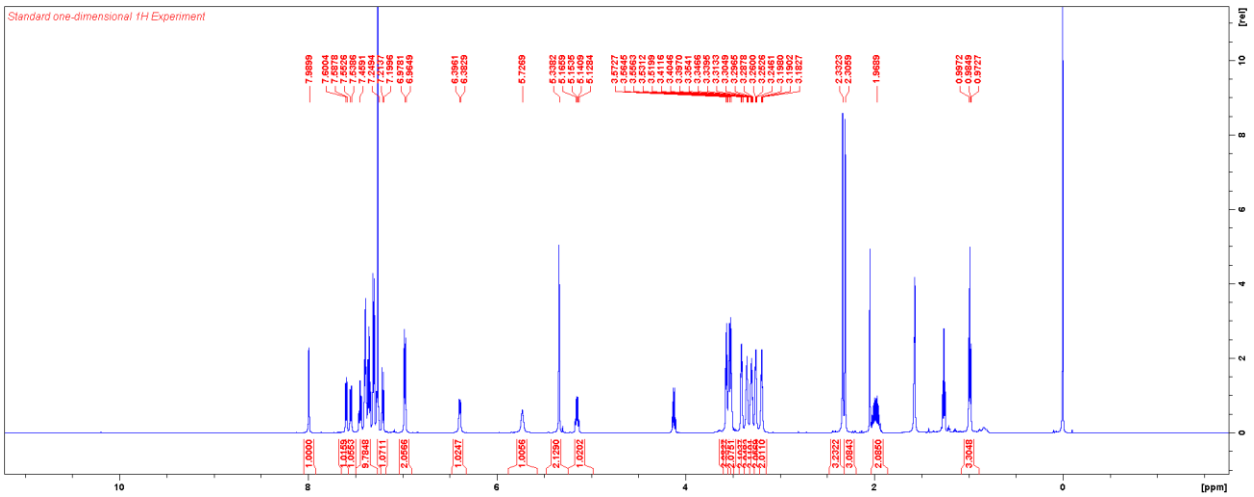
General method for the synthesis of (*R*)-SR9034-PEG-NH_2 (compounds **18**–**21**)

To a solution of a (*R*)-SR9034-PEG-N_3 (1 equiv.) in 10 % H_2O/THF (0.1 M) was added triphenylphosphine (1.5 equiv.). The reaction mixture was stirred at room temperature overnight. The reaction mixture was concentrated. Purification using gravity column chromatography (3 or 5 % MeOH/DCM + 1 % NH_4OH) yielded a (*R*)-SR9034-PEG-NH_2.

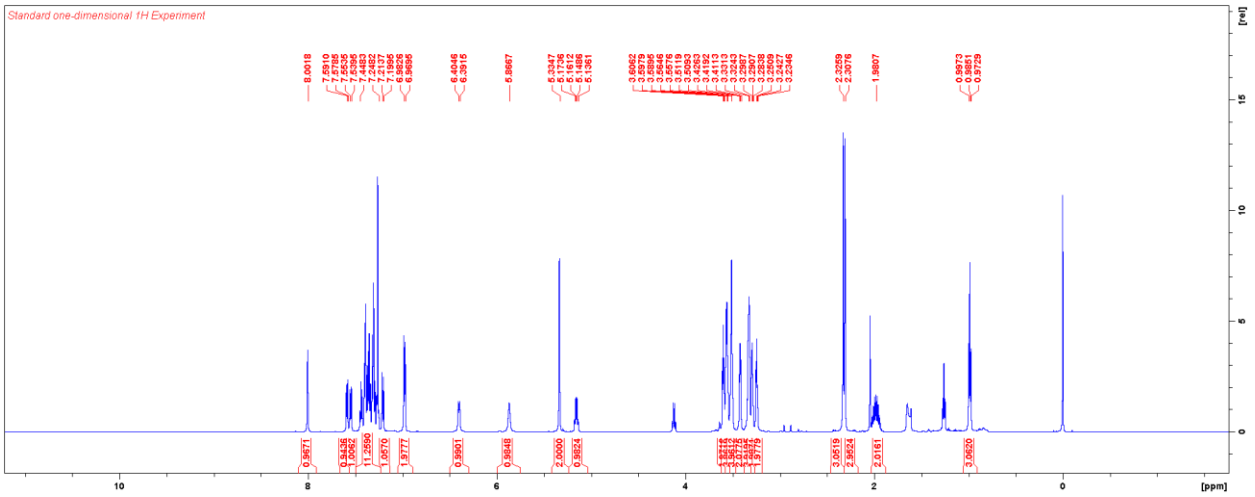
(*R*)-SR9034-PEG_2-NH_2 (compound **18**); 73 mg (110 µmol, quant.); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.00 (s, 1H), 7.57 (m, 2H), 7.45–7.25 (m, 9H), 7.20 (d, J = 8.5 Hz, 1H), 6.98 (d, J = 7.8, 2H), 6.49 (d, J = 7.2, 1H), 6.07 (br, 1H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.34 (m, 6H), 3.26 (m, 4H), 2.68 (br, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.97 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



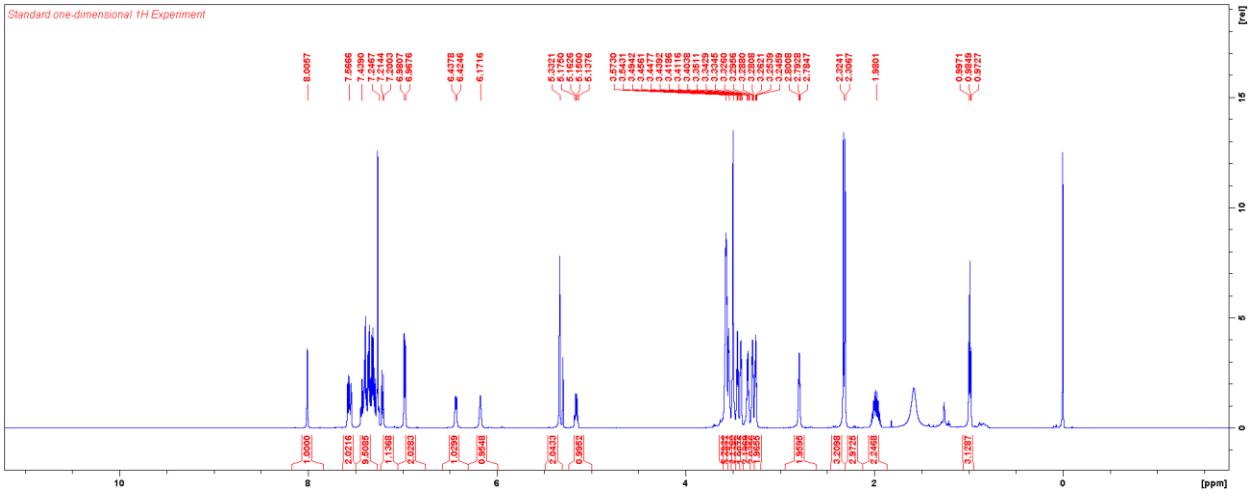
(*R*)-SR9034-PEG_3-NH_2 (compound **19**); 36 mg (62 µmol, 81 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.99 (s, 1H), 7.59 (d, J = 7.4 Hz, 1H), 7.54 (d, J = 8.4 Hz, 1H), 7.46–7.25 (m, 9H), 7.20 (d, J = 8.4 Hz, 1H), 6.97 (d, J = 7.8, 2H), 6.38 (d, J = 7.8 Hz, 1H), 5.73 (br, 1H), 5.34 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.56 (t, J = 4.8, 2 H), 3.53 (br, 2H), 3.52 (br, 2H), 3.40 (t, J = 4.2 Hz, 2H), 3.35 (t, J = 4.2 Hz, 2H), 3.30 (m, 2H), 3.25 (t, J = 4.4 Hz, 2 H), 3.19 (t, J = 4.6 Hz, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.97 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



(*R*)-SR9034-PEG$_4$-NH$_2$ (compound **20**); 49 mg (67 µmol, 96 %); ^{1}H NMR (CDCl$_3$, 600 MHz): δ 8.00 (s, 1H), 7.58 (d, J = 7.4 Hz, 1H), 7.54 (d, J = 8.4 Hz, 1H), 7.44–7.25 (m, 9H), 7.20 (d, J = 8.4 Hz, 1H), 6.97 (d, J = 7.8, 2H), 6.40 (d, J = 7.8 Hz, 1H), 5.87 (br, 1H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.59 (t, J = 5.0, 2 H), 3.56 (m, 4H), 3.51 (m, 4H), 3.42 (t, J = 4.8 Hz, 2H), 3.33 (m, 2H), 3.29 (t, J = 4.2 Hz, 2H), 3.25 (t, J = 4.8 Hz, 2H), 2.33 (s, 3H), 2.31 (s, 3H), 1.98 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



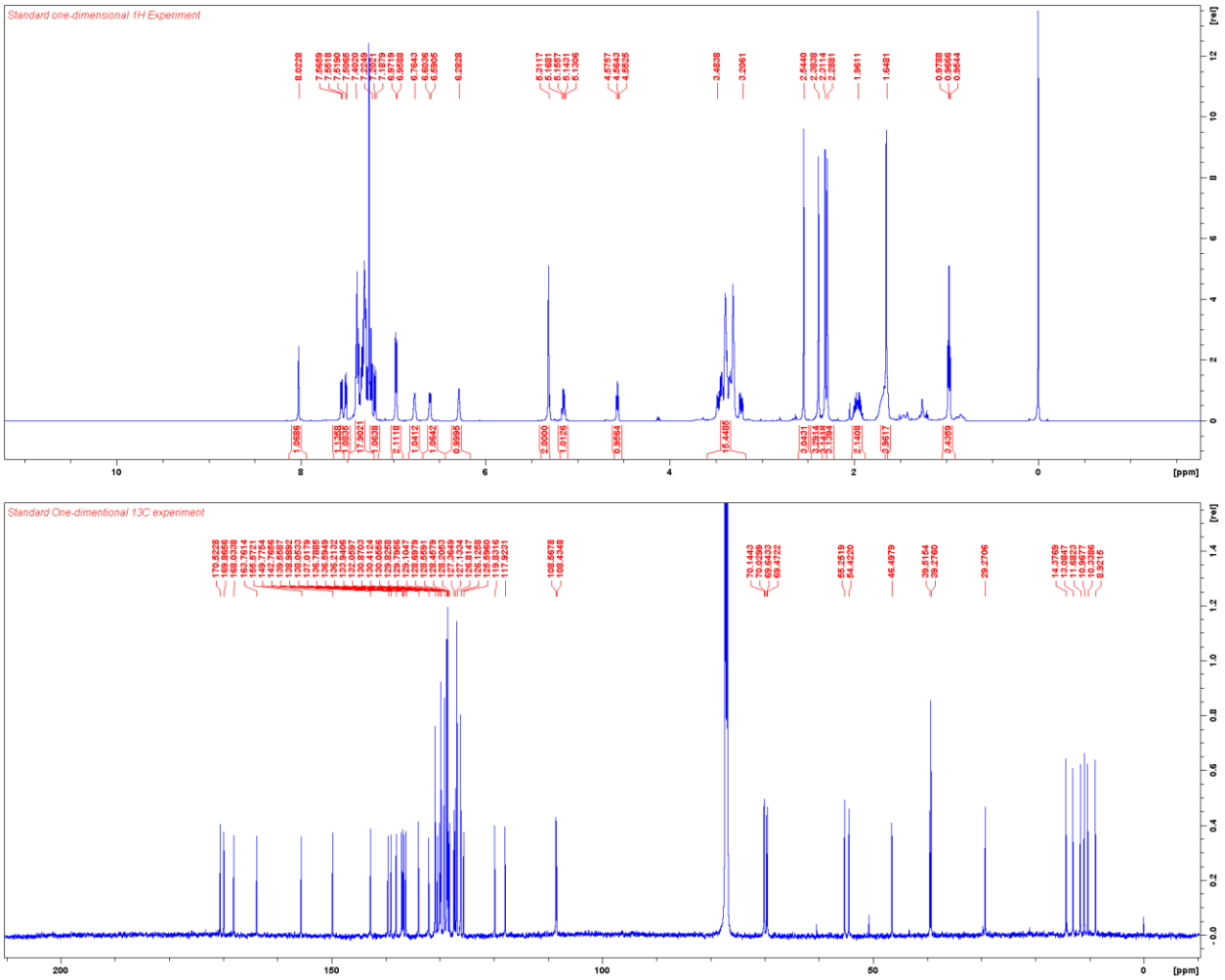
(*R*)-SR9034-PEG$_5$-NH$_2$ (compound **21**); 59 mg (76 µmol, 81 %); ^{1}H NMR (CDCl$_3$, 600 MHz): δ 8.01 (s, 1H), 7.57 (m, 2H), 7.44–7.25 (m, 9H), 7.20 (d, J = 8.4 Hz, 1H), 6.97 (d, J = 7.8, 2H), 6.43 (d, J = 7.8 Hz, 1H), 6.17 (br, 1H), 5.33 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 3.57 (m, 6H), 3.54 (br, 2H), 3.49 (m, 4H), 3.45 (t, J = 5.1 Hz, 2H), 3.41 (t, J = 4.2 Hz, 2H), 3.33 (m, 2H), 3.29 (t, J = 4.4 Hz, 2H), 3.25 (t, J = 4.8 Hz, 2H), 2.79 (t, J = 6.4 Hz, 2H), 2.32 (s, 3H), 2.31 (s, 3H), 1.98 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).



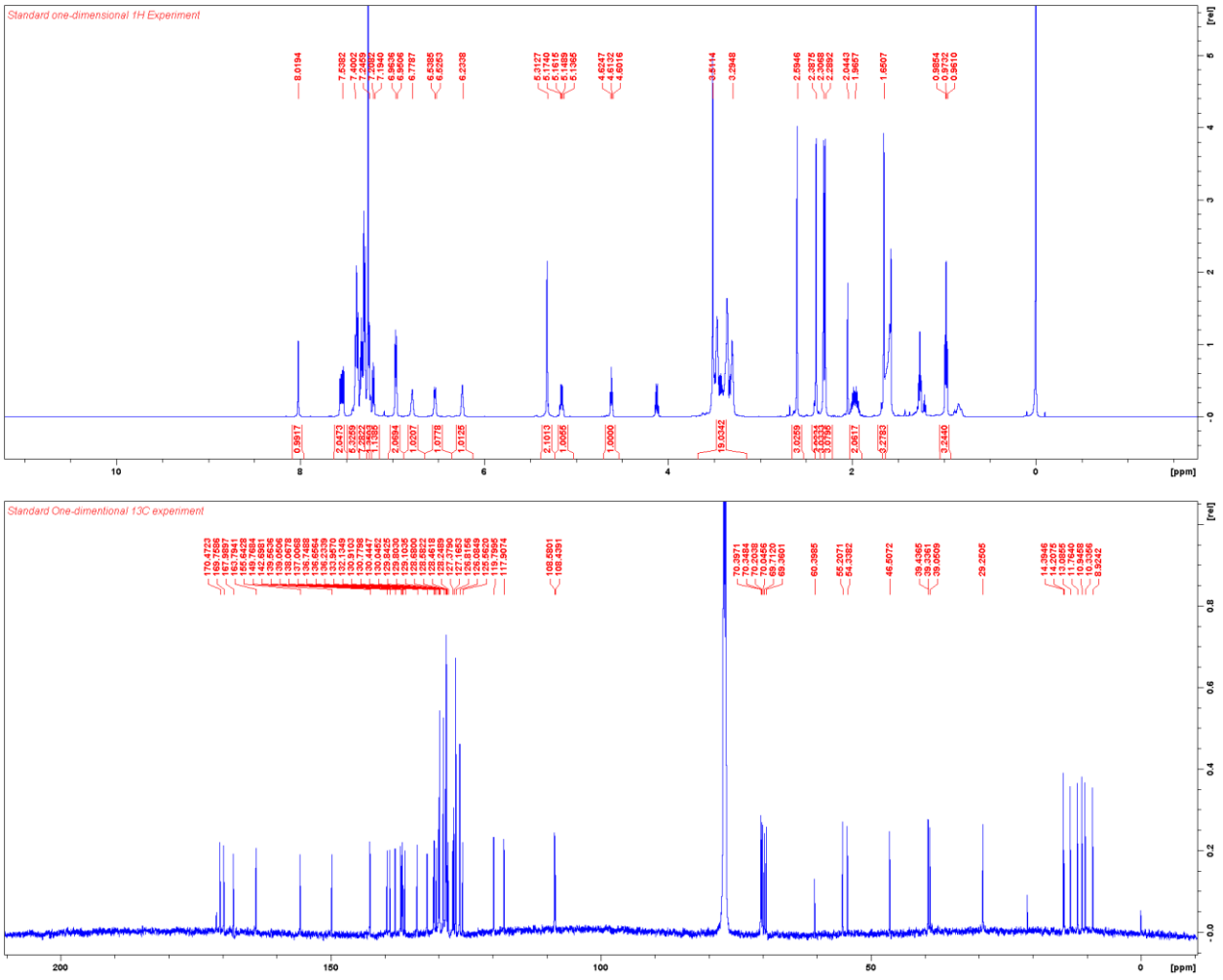
General method for the synthesis of (*R*)-SR9034-PEG-(+)-JQ1 (compounds **22**–**25**)

The commercially-available (+)-JQ1 (MedChemExpress) was converted to free acid (+)-JQ1-acid by treatment with TFA/DCM (1:2) on ice. The reaction was allowed to warm to room temperature over 5 h. To a solution of (+)-JQ1-acid (0.9 equiv.) and HATU (0.9 equiv.) in DCM (0.1 M) was added DIPEA (10 equive). The reaction mixture was stirred on ice for 5 min. To this a solution of (*R*)-SR9034-PEG-NH_2 (1 equiv.) in DCM (0.1 M) was added. The reaction was stirred at room temperature for 2 h. The reaction mixture was diluted with EtOAc, washed with sat. $NaHCO_3$, sat NH_4Cl, and sat NaCl. The organic layer was dried over Na_2SO_4, filtered, and concentrated. Purification on CombiFlash Rf using 0–15 % MeOH/DCM yielded a (*R*)-SR9034-PEG-(+)-JQ1.

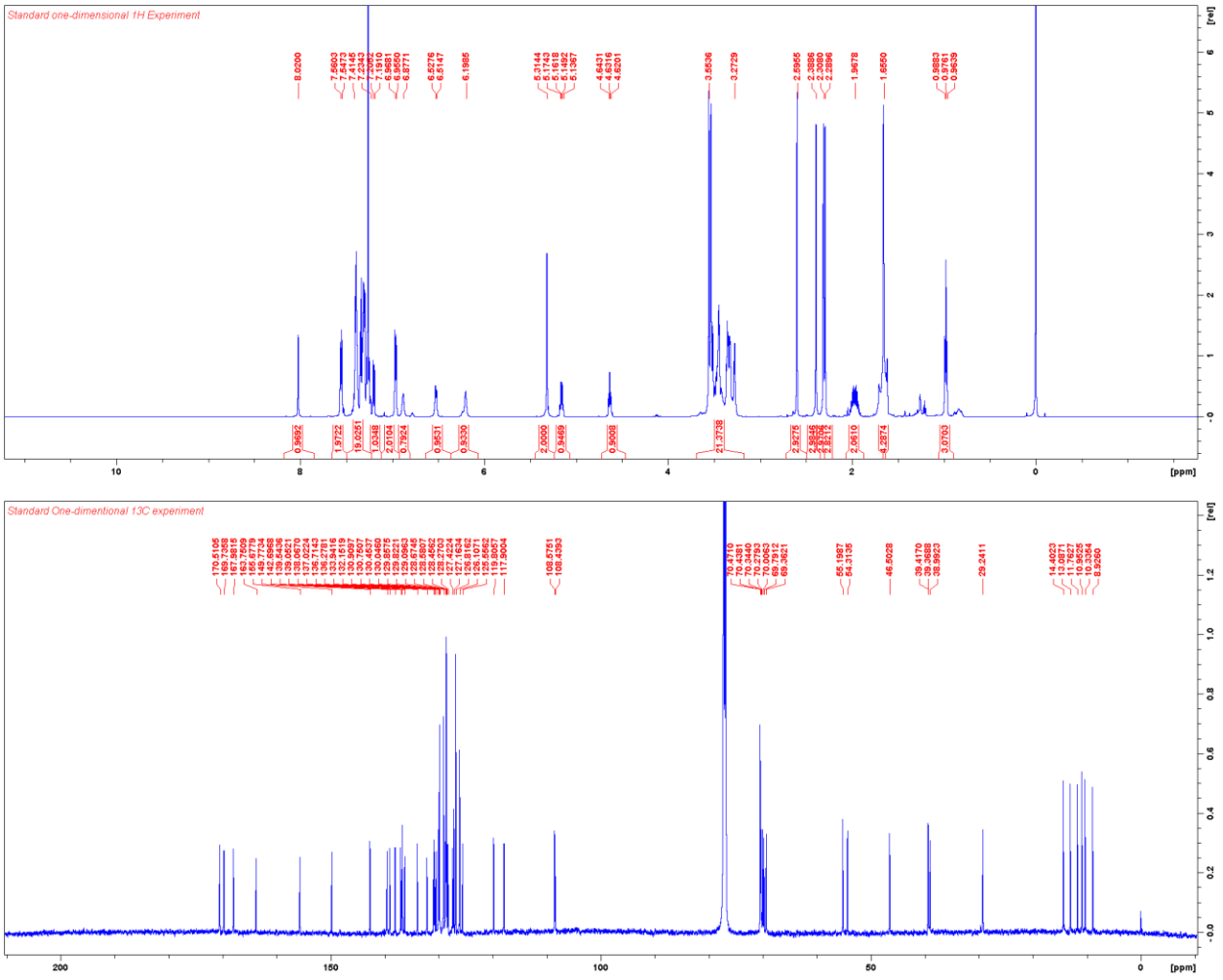
(*R*)-SR9034-PEG_2-(+)-JQ1 (compound **22**); 32 mg (31 µmol, 64 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.02 (s, 1H), 7.56 (d, J = 8.4, 1H), 7.51 (d, J = 7.8 Hz, 1H), 7.40–7.22 (m, 13H), 7.19 (d, J = 8.4 Hz, 1H), 6.96 (d, J = 7.8, 2H), 6.76 (br, 1H), 6.59 (d, J = 7.8 Hz, 1H), 6.28 (br, 1H), 5.31 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 4.56 (t, J = 7.0 Hz, 1H), 3.51–3.29 (m, 14H), 2.54 (s, 3H), 2.38 (s, 3H), 2.31 (s, 3H), 2.29 (s, 3H), 1.96 (m, 2H), 1.65 (s, 3H), 0.99 (t, J = 7.3 Hz, 3H). ^{13}C NMR ($CDCl_3$, 150 MHz): δ 170.5, 169.9, 168.0, 163.8, 155.6, 149.8, 142.8, 139.6, 139.0, 138.1, 137.0, 136.8, 136.6, 136.2, 133.9, 132.1, 130.9, 130.4, 130.1, 129.8, 129.8, 129.1, 128.7, 128.6, 128.5, 128.2, 127.4, 127.1, 126.8, 126.1, 125.6, 119.8, 117.9, 108.6, 108.4, 70.1, 70.0, 69.6, 69.5, 55.3, 54.4, 46.5, 39.5, 39.3, 29.3, 14.4, 13.1, 11.7, 11.0, 10.3, 8.9.



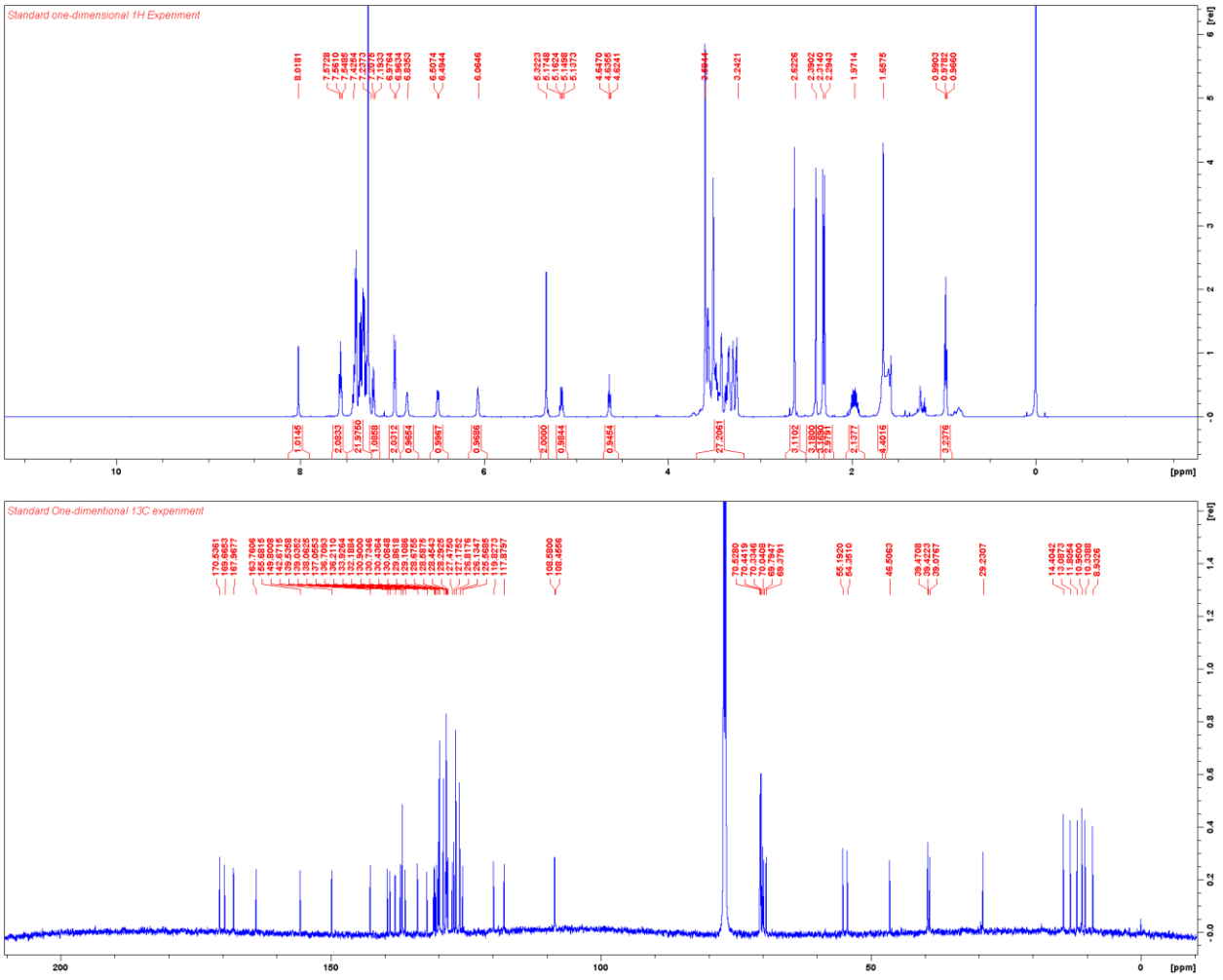
(*R*)-SR9034-PEG$_3$-(+)-JQ1 (compound **23**); 20 mg (19 µmol, 64 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.02 (s, 1H), 7.53 (m, 2H), 7.46–7.25 (m, 13H), 7.20 (d, J = 8.4 Hz, 1H), 6.96 (d, J = 7.8, 2H), 6.78 (br, 1H), 6.54 (d, J = 7.8 Hz, 1H), 6.23 (br, 1H), 5.31 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 4.61 (t, J = 7.0 Hz, 1H), 3.51–3.29 (m, 18H), 2.59 (s, 3H), 2.39 (s, 3H), 2.31 (s, 3H), 2.29 (s, 3H), 1.97 (m, 2H), 1.65 (s, 3H), 0.99 (t, J = 7.3 Hz, 3H). ^{13}C NMR ($CDCl_3$, 150 MHz): δ 170.5, 169.8, 168.0, 163.8, 155.6, 149.8, 142.7, 139.6, 139.1, 138.1, 137.0, 136.7, 136.7, 136.2, 134.0, 132.1, 130.9, 130.8, 130.4, 130.0, 129.8, 129.8, 129.1, 128.7, 128.6, 128.5, 128.2, 127.4, 127.2, 126.8, 126.1, 125.6, 119.8, 117.9, 108.6, 108.4, 70.4, 70.3, 70.2, 70.0, 69.7, 69.4, 55.2, 54.3, 46.5, 39.4, 39.3, 39.1, 29.3, 14.4, 13.1, 11.8, 10.9, 10.3, 8.9.



(*R*)-SR9034-PEG$_4$-(+)-JQ1 (compound **24**); 29 mg (26 µmol, 76 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.02 (s, 1H), 7.55 (m, 2H), 7.41–7.23 (m, 13H), 7.20 (d, J = 8.4 Hz, 1H), 6.96 (d, J = 7.8, 2H), 6.87 (br, 1H), 6.52 (d, J = 7.8 Hz, 1H), 6.20 (br, 1H), 5.31 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 4.63 (t, J = 7.0 Hz, 1H), 3.55–3.27 (m, 22H), 2.60 (s, 3H), 2.39 (s, 3H), 2.31 (s, 3H), 2.29 (s, 3H), 1.97 (m, 2H), 1.66 (s, 3H), 0.99 (t, J = 7.3 Hz, 3H). ^{13}C NMR ($CDCl_3$, 150 MHz): δ 170.5, 169.7, 168.0, 163.8, 155.7, 149.8, 142.7, 139.5, 139.1, 138.1, 137.0, 136.7, 136.3, 133.9, 132.2, 130.9, 130.8, 130.5, 130.0, 129.9, 129.8, 129.1, 128.7, 128.6, 128.5, 128.3, 127.4, 127.2, 126.8, 126.1, 125.6, 119.8, 117.9, 108.6, 108.4, 70.5, 70.4, 70.3, 70.3, 70.0, 69.8, 69.4, 55.2, 54.3, 46.5, 39.4, 39.4, 39.0, 29.2, 14.4, 13.1, 11.8, 11.0, 10.3.

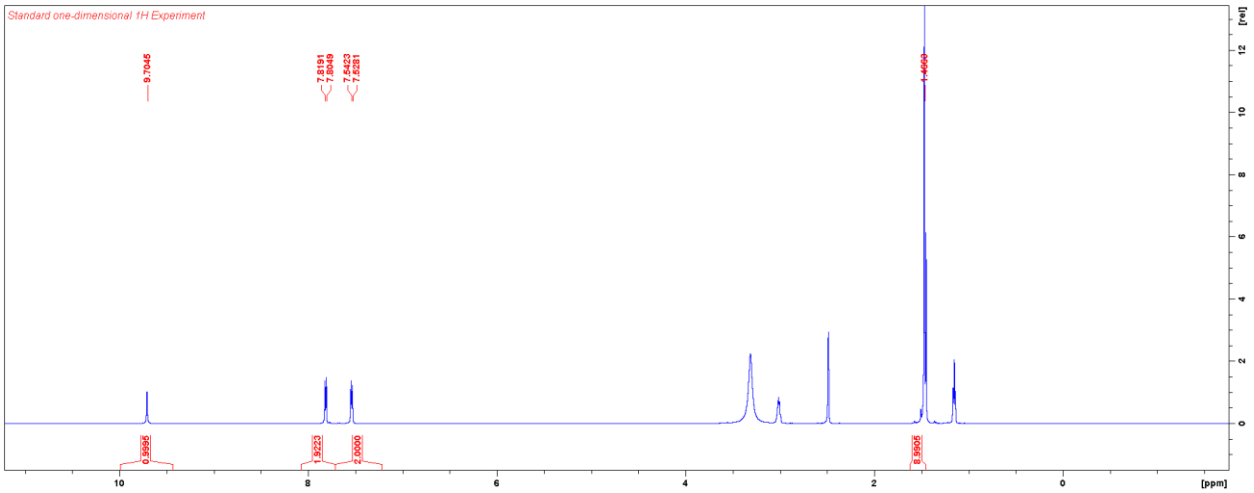


(*R*)-SR9034-PEG$_5$-(+)-JQ1 (compound **25**); 32 mg (28 µmol, 72 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.02 (s, 1H), 7.56 (m, 2H), 7.43–7.23 (m, 13H), 7.20 (d, J = 8.4 Hz, 1H), 6.96 (d, J = 7.8, 2H), 6.84 (br, 1H), 6.49 (d, J = 7.8 Hz, 1H), 6.06 (br, 1H), 5.32 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 4.64 (t, J = 7.0 Hz, 1H), 3.59–3.24 (m, 26H), 2.62 (s, 3H), 2.39 (s, 3H), 2.31 (s, 3H), 2.29 (s, 3H), 1.97 (m, 2H), 1.66 (s, 3H), 0.99 (t, J = 7.3 Hz, 3H). ^{13}C NMR ($CDCl_3$, 150 MHz): δ 170.5, 169.7, 168.0, 163.8, 155.7, 149.8, 142.7, 139.5, 139.0, 138.1, 137.1, 136.7, 136.2, 133.9, 132.2, 130.9, 130.7, 130.4, 130.1, 129.9, 129.1, 128.7, 128.6, 128.5, 128.3, 127.5, 127.2, 126.8, 126.1, 125.6, 119.8, 117.9, 108.6, 108.5, 70.5, 70.4, 70.3, 70.0, 69.8, 69.4, 55.2, 54.4, 46.5, 39.5, 39.4, 39.1, 29.2, 14.4, 13.1, 11.8, 11.0, 10.3, 8.9.



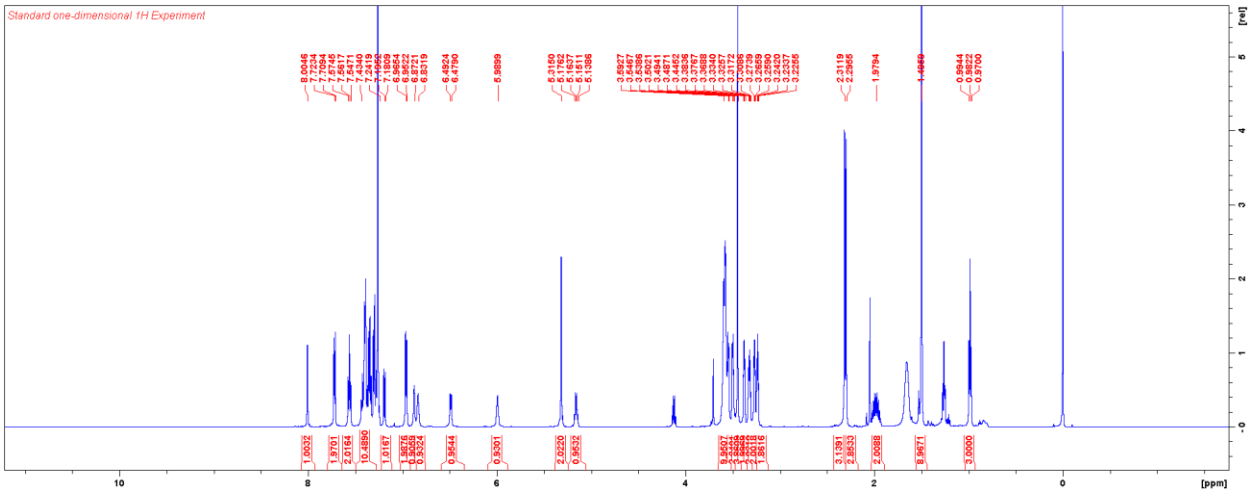
Synthesis of *N*-Boc-aminobenzoic acid

To 4-aminobenzoic acid (0.5 g, 3.65 mmol) in dioxane (10 mL) and water (5 mL) were added TEA (1.02 mL, 7.30 mmol) followed by Boc_2O (1.59 g, 1.67 mL, 7.30 mmol). The reaction mixture was stirred at room temperature for 24 h. Solvent was removed under reduced pressure, and 3 N aqueous hydrochloric acid (5 mL) was added dropwise to the residue. A precipitate was obtained, collected, washed with water, and dried to provide the title compound as a white solid. This material was used in the next step without further purification. ^{1}H NMR (DMSO-d6, 600 MHz): δ 9.70 (s, 1H), 7.81 (d, J = 8.4 Hz, 2H), 7.53 (d, J = 8.4 Hz, 2H), 1.47 (s, 9H).



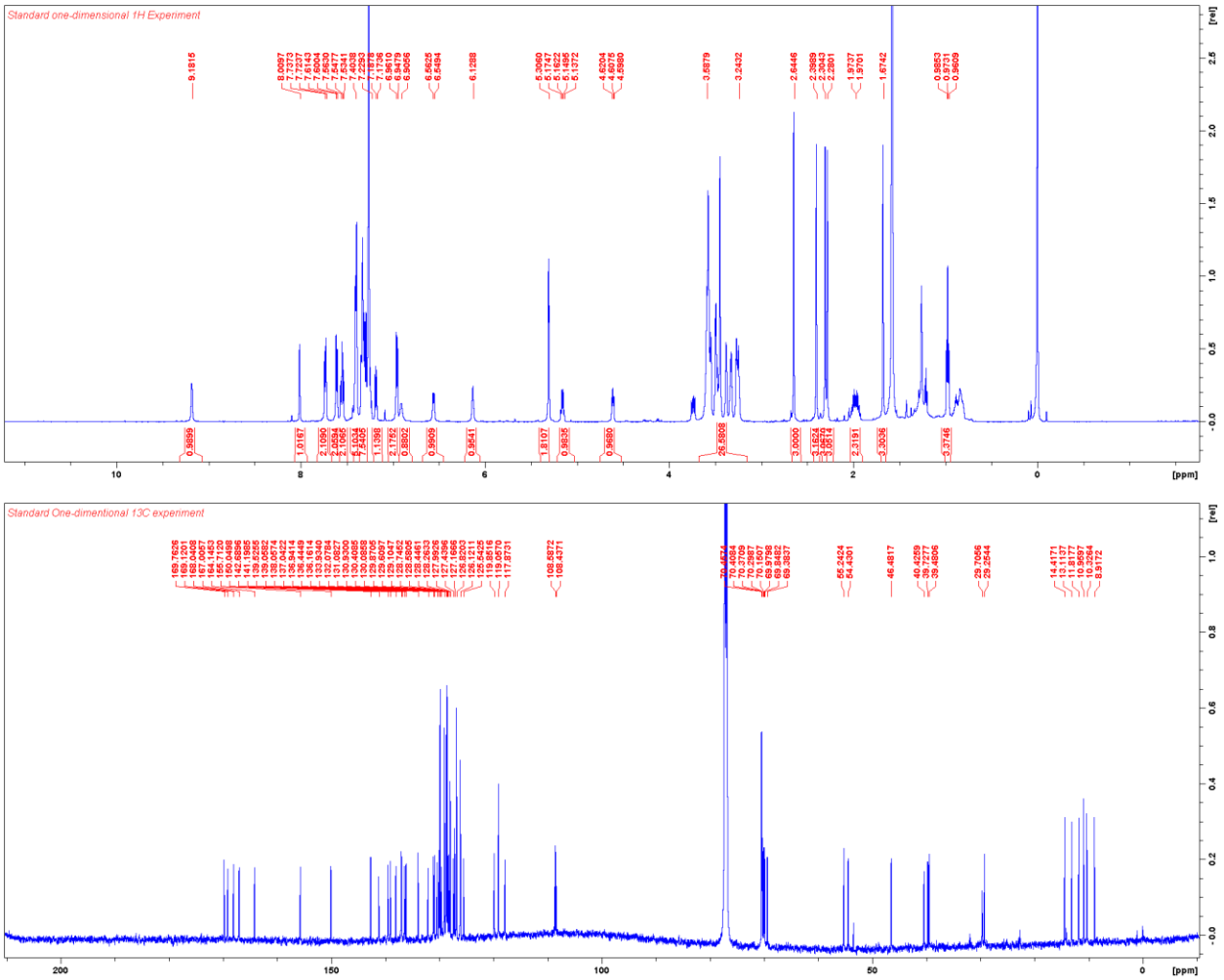
Synthesis of (*R*)-SR9034-PEG$_5$-ab-Boc (compound **26**)

To a solution of *N*-Boc-aminobenzoic acid (10 mg, 42 µmol) and HATU (16 mg, 42 µmol) in DMF (0.4 mL) was added DIPEA (34 µmol, 190 µmol). The reaction mixture was stirred on ice for 5 min. To this (*R*)-SR9034-PEG$_5$-NH$_2$ (30 mg, 39 µmol) was added. The reaction was stirred at room temperature for 1 h. The reaction mixture was diluted with EtOAc, washed with sat. $NaHCO_3$, sat NH_4Cl, and sat NaCl. The organic layer was dried over Na_2SO_4, filtered, and concentrated. Purification on CombiFlash Rf using 0–10 % MeOH/DCM yielded a (*R*)-SR9034-PEG$_5$-ab-Boc. Yield: 35 mg (35 µmol, 83 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.00 (s, 1H), 7.71 (d, J = 8.4 Hz, 2H), 7.56 (m, 2H), 7.43–7.24 (m, 9H), 7.19 (d, J = 8.4 Hz, 1H), 6.97 (d, J = 8.4 Hz, 2H), 6.87 (br, 1H), 6.83 (br, 1H), 6.48 (d, J = 7.8 Hz, 1H), 5.99 (br, 1H), 5.32 (s, 2H), 5.16 (q, J = 7.4 Hz, 1H), 3.59–3.54 (m, 10H), 3.49 (t, J = 4.4 Hz, 2H), 3.44 (m, 4H), 3.38 (t, J = 4.4 Hz, 2H), 3.33 (m, 2H), 3.27 (t, J = 4.4 Hz, 2H), 3.23 (t, J = 4.4 Hz, 2H), 2.31 (s, 3H), 2.30 (s, 3H), 1.98 (m, 2H), 1.50 (s, 9H), 0.98 (t, J = 7.3 Hz, 3H).



Synthesis of (*R*)-SR9034-PEG$_5$-ab-(+)-JQ1 (compound **27**)

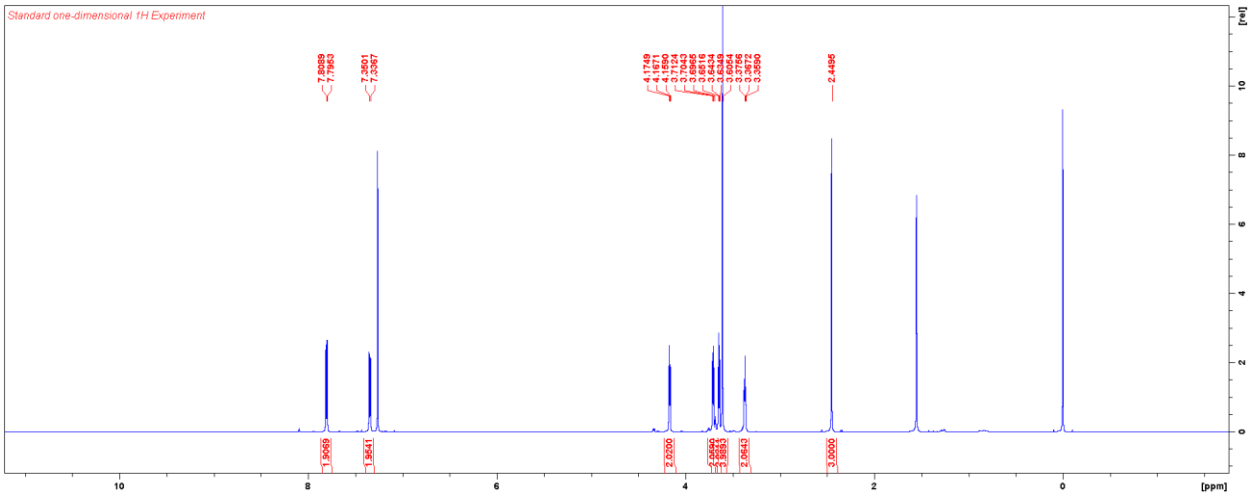
(*R*)-SR9034-PEG$_5$-ab-Boc (35 mg, 35 µmol) was converted to free amine by treatment with TFA/DCM (1:2) for 2 h. To this a solution of (+)-JQ1-acid (17 mg, 42 µmol), HATU (16 mg, 0.042 µmol), and DIPEA (31 µL, 180 µmol) in DCM (400 µL) was added. The reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with EtOAc, washed with sat. $NaHCO_3$, sat NH_4Cl, and sat NaCl. The organic layer was dried over Na_2SO_4, filtered, and concentrated. Purification on CombiFlash Rf using 0–10 % MeOH/DCM and subsequently on gravity column using 4 % MeOH/DCM yielded (*R*)-SR9034-PEG$_5$-ab-(+)-JQ1 as a white solid (21 µmol, 50 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 9.18 (br, 1H), 8.01 (s, 1H), 7.73 (d, J = 8.4 Hz, 2H), 7.61 (d, J = 8.4 Hz, 2H), 7.55 (m, 2H), 7.40–7.22 (m, 13H), 7.18 (d, J = 8.4 Hz, 1H), 6.95 (d, J = 8.4 Hz, 2H), 6.91 (br, 1H), 6.55 (d, J = 7.8 Hz, 1H), 6.13 (br, 1H), 5.31 (s, 2H), 5.15 (q, J = 7.4 Hz, 1H), 4.61 (t, J = 7.0 Hz, 1H), 3.59–3.24 (m, 26H), 2.64 (s, 3H), 2.40 (s, 3H), 2.30 (s, 3H), 2.28 (s, 3H), 1.97 (m, 2H), 1.67 (s, 3H), 0.98 (t, J = 7.3 Hz, 3H); ^{13}C NMR ($CDCl_3$, 150 MHz): δ 169.8, 169.1, 168.0, 167.0, 164.1, 155.7, 150.0, 142.7, 141.2, 139.5, 139.1, 138.1, 137.0, 136.9, 136.4, 136.2, 133.9, 132.1, 131.1, 130.9, 130.4, 130.1, 129.9, 129.6, 129.1, 128.7, 128.6, 128.4, 128.3, 128.0, 127.4, 127.2, 126.8, 126.1, 125.5, 119.9, 119.1, 117.9, 108.6, 108.4, 70.5, 70.4, 70.4, 70.3, 70.2, 70.0, 69.8, 69.4, 55.2, 54.4, 46.5, 40.4, 39.7, 39.5, 29.7, 29.3, 14.4, 13.1, 11.8, 11.0, 10.3, 8.9.



General method for the synthesis of azide-PEG-OTs linkers (compounds **39**–**41**)

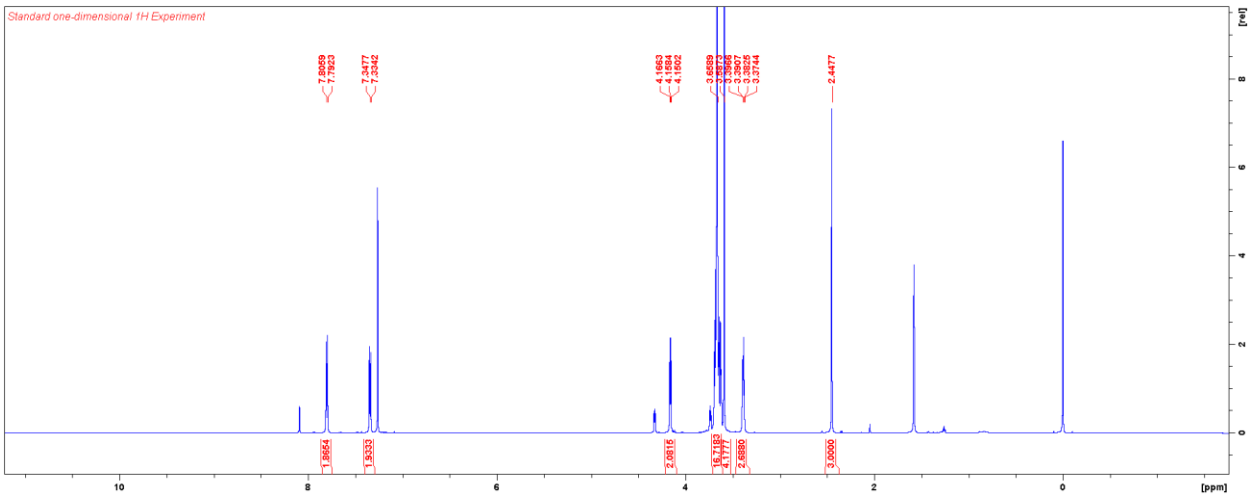
A commercially-available PEG bis(tosylate) (1 equiv.) was dissolved in DMF (1 M). To this sodium azide (1 equiv.) was added. The reaction mixture was heated at 70 °C for 4–16 h. DMF was removed by evaporation, and the residue was dissolved in EtOAc and water. The organic layer was washed with sat. NaCl, dried over Na_2SO_4, filtered, and concentrated. Purification on CombiFlash using 0–100 % EtOAc/hexanes yielded an azide-PEG-OTs linker as colourless oil (25–35 %).

Azide-PEG_2-OTs linker (**39**); Yield: 35 % from triethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.80 (d, J = 4.8 Hz, 2H), 7.34 (d, J = 4.8 Hz, 2H), 4.17 (t, J = 4.8 Hz, 2H), 3.70 (t, J = 4.8 Hz, 2H), 3.64 (t, J = 4.8 Hz, 2H), 3.61 (m, 4H), 3.37 (d, J = 4.8 Hz, 2H), 2.45 (s, 3H).



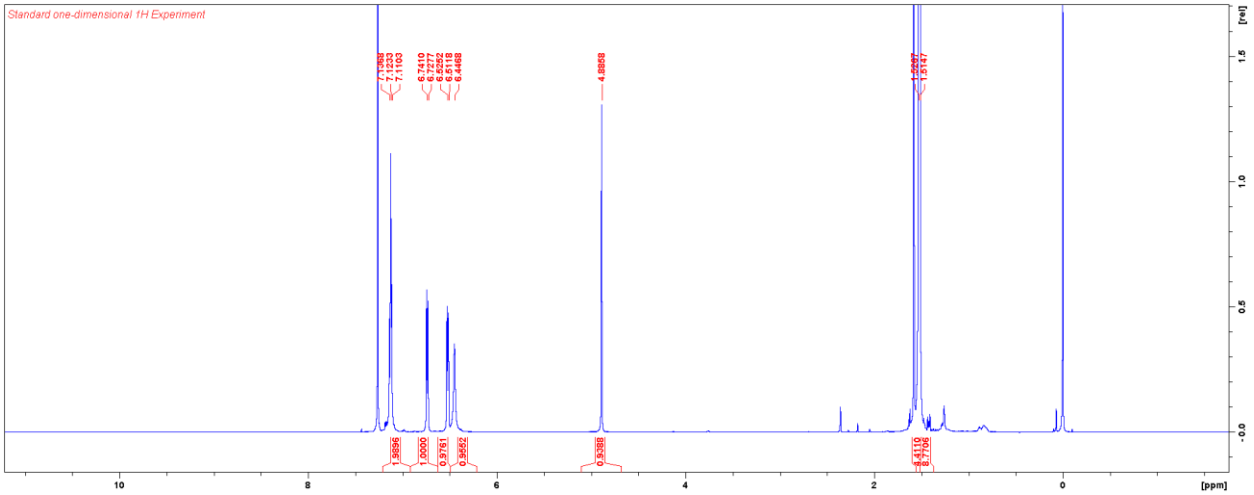
Azide-PEG$_3$-OTs linker (**40**); Yield: n.d. from tetraethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ NA.

Azide-PEG$_4$-OTs linker (**41**); Yield: 25 % from pentaethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.80 (d, J = 4.8 Hz, 2H), 7.34 (d, J = 4.8 Hz, 2H), 4.16 (t, J = 4.8 Hz, 2H), 3.66 (m, 12H), 3.59 (m, 4H), 3.38 (t, J = 4.8 Hz, 2H), 2.45 (s, 3H).



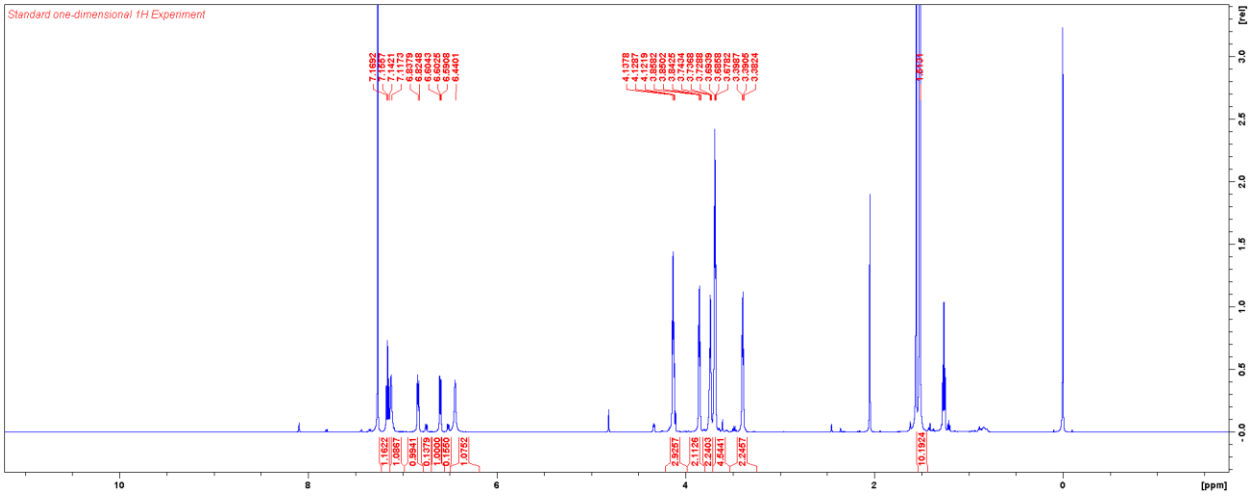
Synthesis of *N*-Boc-3-aminophenol (compound **28**)

To a solution of 3-aminophenol (1 g, 9.0 mmol) in THF (20 ml) was added $(Boc)_2O$ (2.1 g, 9.9 mmol). After being stirred at reflux for 24 h, the reaction mixture was dissolved in EtOAc, washed with 1 N HCl, sat. $NaHCO_3$, and sat NaCl, dried over Na_2SO_4, filtered, concentrated, and azeotroped twice with toluene. This yielded *N*-Boc-3-aminophenol (1.9 g, quant.) as an off-white solid. ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.12 (t, J = 7.8 Hz, 1H), 6.73 (d, J = 7.8 Hz, 1H), 6.51 (d, J = 7.8 Hz, 1H), 6.45 (s, 1H), 4.89 (s, 1H), 1.51 (s, 9H).



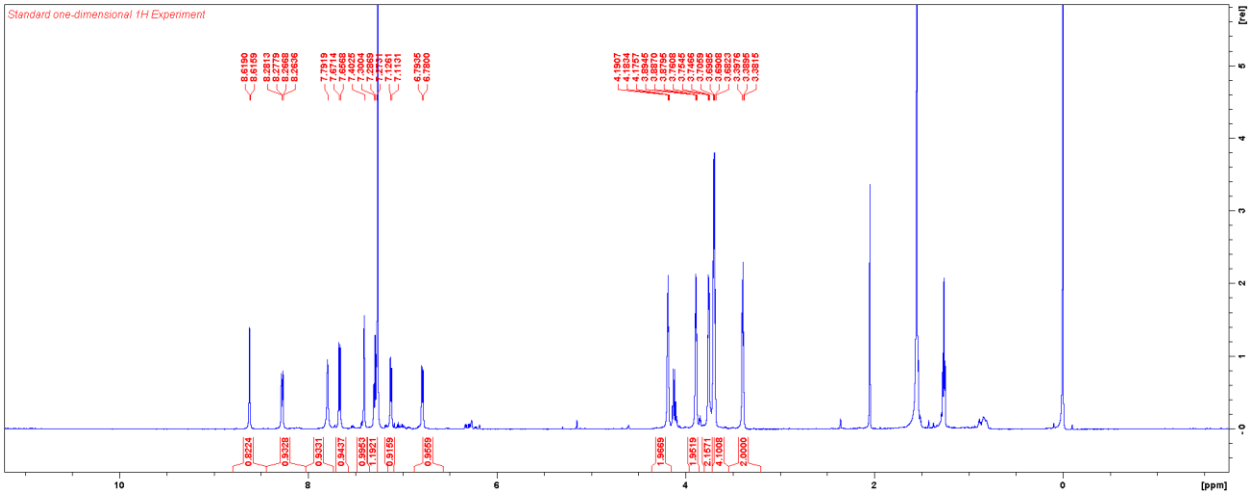
Synthesis of *N*-Boc-3-aminophenol-PEG$_2$-azide (compound **29**)

To a solution of *N*-Boc-3-aminophenol (42 mg, 0.2 mmol) and azide-PEG$_2$-OTs (66 mg, 0.2 mmol) in DMF (0.4 mL) was added K_2CO_3 (55 mg, 0.4 mmol). After being stirred at 70 °C for 24 h, the reaction mixture was diluted with EtOAc, washed with sat NaCl, dried over Na_2SO_4, filtered and concentrated. Purification on CombiFlash Rf using 0–100 % EtOAc/hexanes yielded the desired product in 9:1 mixture with *N*-Boc-3-aminophenol (61 mg, 0.17 mmol, 87 %). This mixture was used in the next step without further purification. ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.16 (t, J = 7.8 Hz, 1H), 7.11 (br, 1H), 6.83 (d, J = 7.8 Hz, 1H), 6.60 (d, J = 7.8 Hz, 1H), 6.44 (br, 1H), 4.13 (t, J = 4.2 Hz, 2H), 3.85 (t, J = 4.8 Hz, 2H), 3.74 (t, J = 4.8 Hz, 2 H), 3.69 (m, 4H), 3.39 (t, J = 4.8 Hz, 2H), 1.51 (s, 9H).



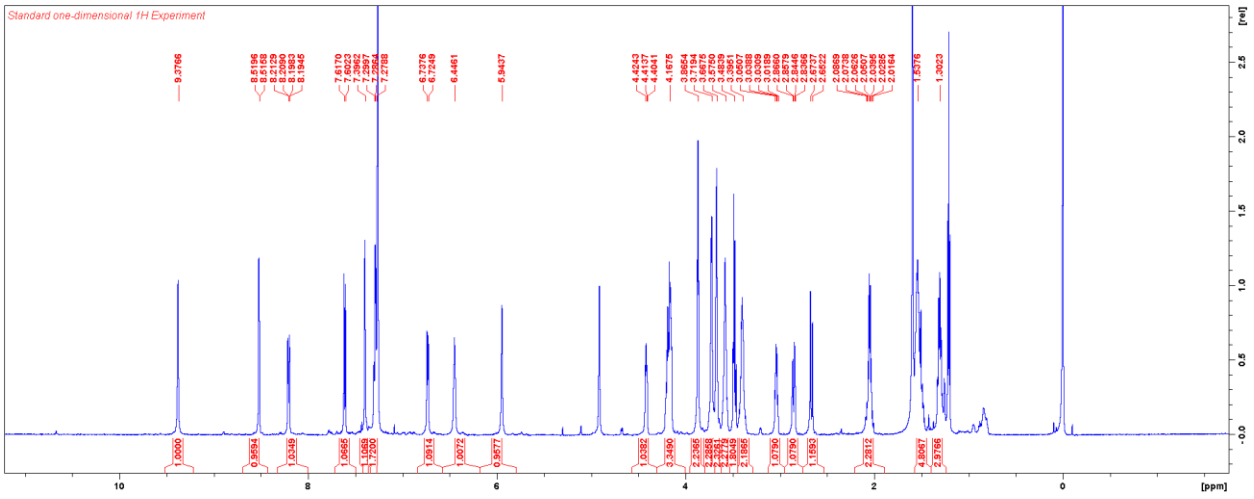
Synthesis of mGW9662-PEG$_2$-azide (compound **30**)

To a solution of 2-chloro-5-nitrobenzoyl chloride (37 mg, 0.17 mmol) and triethylamine (95 µL, 0.68 mmol) in DCM (1.5 mL) was added *N*-Boc-3-aminophenol-PEG$_2$-azide (0.17 mmol) in DCM (0.5 mL). After being stirred on ice for 1 h, the reaction mixture was diluted with EtOAc, washed with sat. $NaHCO_3$, sat. NH_4Cl, and sat. NaCl, dried over Na_2SO_4, filtered and concentrated. Purification on CombiFlash using 0–100 % EtOAc/hexanes yielded the title compound (20 mg, 0.044 mmol, 26 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.62 (d, J = 1.9 Hz, 1H), 8.27 (dd, J = 8.7 Hz, 1.9 Hz, 1H), 7.79 (br, 1H), 7.66 (d, J = 8.7 Hz, 1H), 7.40 (s, 1H), 7.29 (t, J = 7.8 Hz, 1H), 7.12 (d, J = 7.8 Hz, 1H), 6.78 (d, J = 7.8 Hz, 1H), 4.18 (t, J = 4.4 Hz, 2H), 3.89 (t, J = 4.4 Hz, 2H), 3.75 (t, J = 4.8 Hz, 2 H), 3.69 (m, 4H), 3.39 (t, J = 4.8 Hz, 2H).



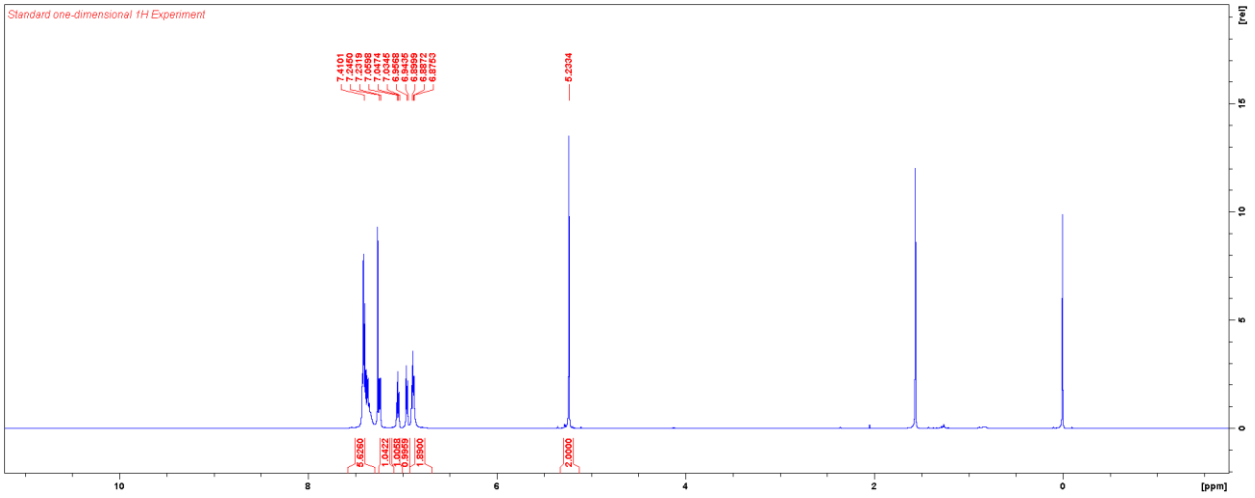
Synthesis of mGW9662-PEG_2-biotin (compound **31**)

To a solution of mGW9662-PEG_2-azide (15 mg, 33 µmol) in 3 % H_2O/THF (1 mL) was added triphenylphosphine (9 mg, 33 µmol). After being stirred at room temperature for 24 h, the reaction mixture was acidified with 1 N HCl, washed with Et_2O, basified with NaOH, and extracted thrice with EtOAc. Combined organic layer was dried over Na_2SO_4, filtered and concentrated to yield amine intermediate. A solution of biotin (9 mg, 35 µmol), HATU (13 mg, 35 µmol), and TEA (10 µL, 71 µmol) in DMF (350 µL) was stirred on ice for 5 min. To this a solution of the amine intermediate in DMF (350 µL) was added. After being stirred at room temperature for 1 h, the reaction mixture was diluted with EtOAc, washed with sat. $NaHCO_3$, sat. NH_4Cl, and sat. NaCl, dried through a Na_2SO_4 column and concentrated. Purification on CombiFlash using 0–15 % MeOH/DCM yielded the title compound (10 mg, 15 µmol, 47 % over two steps). ^1H NMR ($CDCl_3$, 600 MHz): δ 9.38 (s, 1H), 8.52 (d, J = 1.9 Hz, 1H), 8.20 (dd, J = 8.7 Hz, 1.9 Hz, 1H), 7.61 (d, J = 8.7 Hz, 1H), 7.40 (s, 1H), 7.29 (t, J = 7.8 Hz, 1H), 6.73 (d, J = 7.8 Hz, 1H), 6.44 (br, 1H), 5.94 (br, 1H), 4.41 (t, J = 6.6 Hz, 1H), 4.17 (m, 3H), 3.87 (t, J = 4.4 Hz, 2H), 3.72 (br, 2H), 3.67 (br, 2H), 3.58 (br, 2H), 3.48 (br, 2H), 3.40 (br, 2H), 3.03 (dd, J = 12.6 Hz, 7.2 Hz, 1H), 2.85 (dd, J = 12.6 Hz, 4.8 Hz, 1H), 2.66 (d, J = 12.6 Hz, 1H), 2.04 (m, 2H), 1.54 (m, 4H), 1.30 (m, 2H).



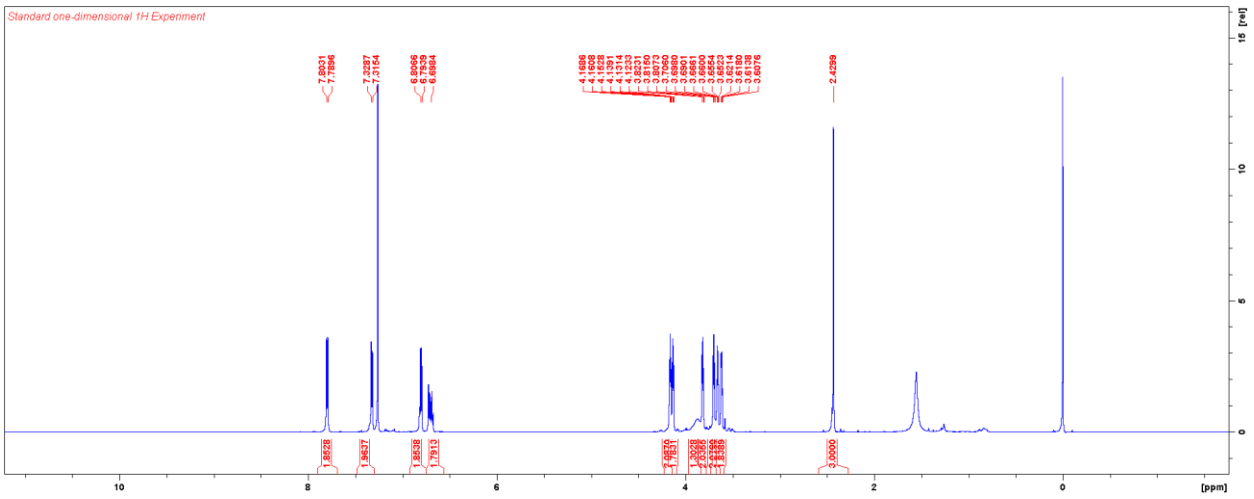
Synthesis of *N*-Cbz-2-aminophenol (compound **32**)

N-Cbz-2-aminophenol was synthesised as previously described. Yield: 72 %. ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.41 (m, 5H), 7.24 (d, J = 8.4 Hz, 1H), 7.05 (t, J = 7.8 Hz, 1H), 6.95 (d, J = 7.8 Hz, 1H), 6.89 (m, 2H), 5.23 (s, 2H).



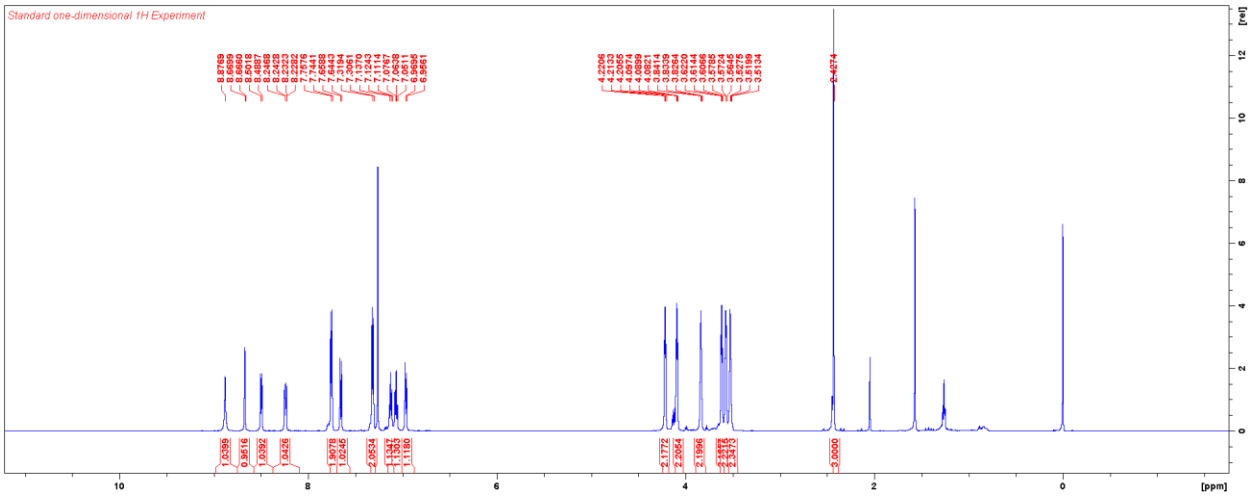
Synthesis of 2-aminophenol-PEG_2-OTs (compound **33**)

To a solution of *N*-Cbz-2-aminophenol (50 mg, 0.21 mmol) and triethylene glycol bis(*p*-toluenesulfonate) (188 mg, 0.41 mmol) in DMF (0.4 mL) was added K_2CO_3 (57 mg, 0.41 mmol). After being stirred at 70 °C for 16 h, the reaction mixture was diluted with EtOAc, washed with sat. NH_4Cl and sat. NaCl, dried over $MgSO_4$, filtered and concentrated to yield crude *N*-Cbz-2-aminophenol-PEG_2-OTs. The crude material was dissolved in EtOAc (4 mL) and subjected to hydrogenolysis condition using Pd/C (11 mg, 0.01 mmol) as catalysis. After being stirred at room temperature for 6 h, the reaction mixture was filtered and concentrated. Purification on CombiFlash using 0–100 % EtOAc/hexanes yielded the title compound (7 mg, 18 µmol, 9 % over two steps). ^1H NMR ($CDCl_3$, 600 MHz): δ 7.79 (d, J = 7.8 Hz, 2H), 7.32 (d, J = 7.8 Hz, 2H), 6.80 (m, 2H), 6.70 (m, 2H), 4.16 (t, J = 4.8 Hz, 2H), 4.13 (d, J = 4.8 Hz, 2H), 3.82 (t, J = 4.8 Hz, 2H), 3.70 (t, J = 4.8 Hz, 2H), 3.66 (m, 2H), 3.62 (m, 2H), 2.43 (s, 3H).



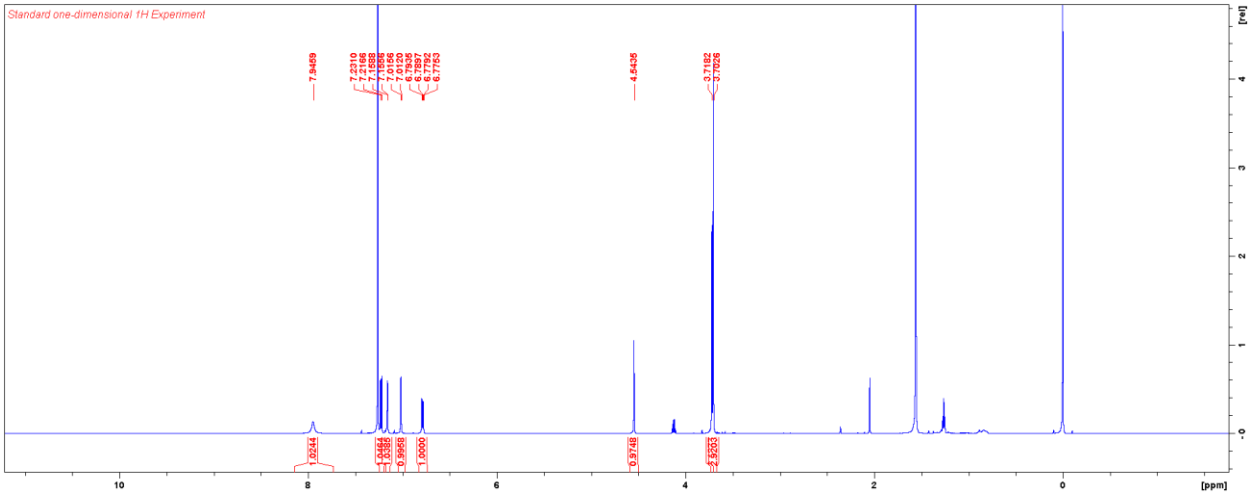
Synthesis of oGW-PEG$_2$-OTs (compound **34**)

To a solution of 2-chloro-5-nitrobenzoic acid (4 mg, 18 µmol) and triethylamine (5 µL, 36 µmol) in DCM (200 µL) was added a solution of 2-aminophenol-PEG$_2$-OTs (7 mg, 18 µmol) in DCM (200 µL). After being stirred at room temperature for 0.5 h, the reaction mixture was diluted with EtOAc, washed with sat. NH_4Cl, dried over $MgSO_4$, filtered and concentrated. Purification on CombiFlash using 0–100 % EtOAc/hexanes yielded the title compound (10 mg, 17 µmol, 98 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.88 (s, 1H), 8.67 (d, J = 1.8 Hz, 1H), 8.50 (d, J = 8.4 Hz, 1H), 8.24 (dd, J = 8.4 Hz, 1.8 Hz, 1H), 7.75 (d, J = 7.8 Hz, 2H), 7.65 (d, J = 8.4 Hz, 1H), 7.31 (d, J = 7.8 Hz, 2H), 7.12 (t, J = 7.8 Hz, 1H), 7.06 (t, J = 7.8 Hz, 1H), 6.97 (d, J = 7.8 Hz, 1H), 4.21 (t, J = 4.5 Hz, 2H), 4.09 (t, J = 4.5 Hz, 2H), 3.83 (t, J = 4.5 Hz, 2H), 3.61 (t, J = 4.7 Hz, 2H), 3.57 (t, J = 4.2 Hz, 2H), 3.52 (t, J = 4.2 Hz, 2H), 2.43 (s, 3H).



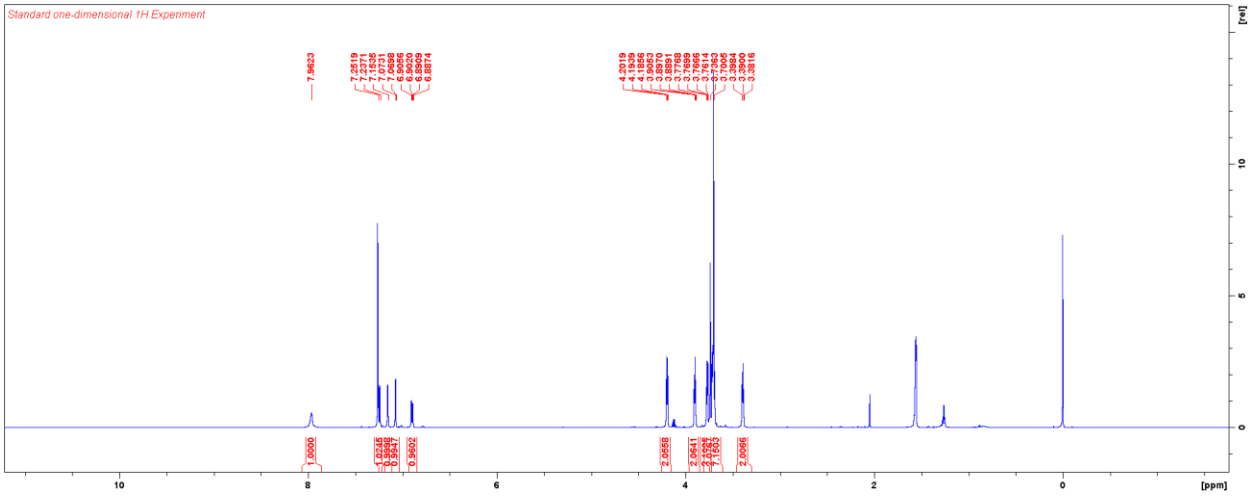
Synthesis of 5-Hydroxy-3-indoleacetic acid methyl ester (IAOMe) (compound **35**)

5-Hydroxy-3-indoleacetic acid (1.00 g, 5.23 mmol) was dissolved in methanol (25 ml). To the solution, acetyl chloride (1.0 ml) was slowly added dropwise, and the mixture was stirred at room temperature for 2 hours. After the reaction was confirmed by TLC to be complete, the reaction was terminated by the addition of sat. $NaHCO_3$, and the solvent was removed under reduced pressure. Then, water (20 ml) was added to the residue, followed by extraction with ethyl acetate (50 ml) three times. The organic layer was washed twice with sat. NaCl and dried over Na_2SO_4. The solvent was distilled off under reduced pressure, and then, the residue was purified using silica gel column chromatography (40 % EtOAc/hexanes) to obtain 5-hydroxy-3-indoleacetic acid methyl ester (0.53 g, 2.58 mmol, 49 %); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.95 (br, 1H), 7.23 (d, J = 8.6 Hz, 1H), 7.15 (d, J = 1.9 Hz, 1H), 7.01 (d, J = 2.2 Hz, 1H), 6.79 (dd, J = 8.6 Hz, 2.3 Hz, 1H), 4.54 (s, 1H), 3.72 (s, 2H), 3.70 (s, 3H).



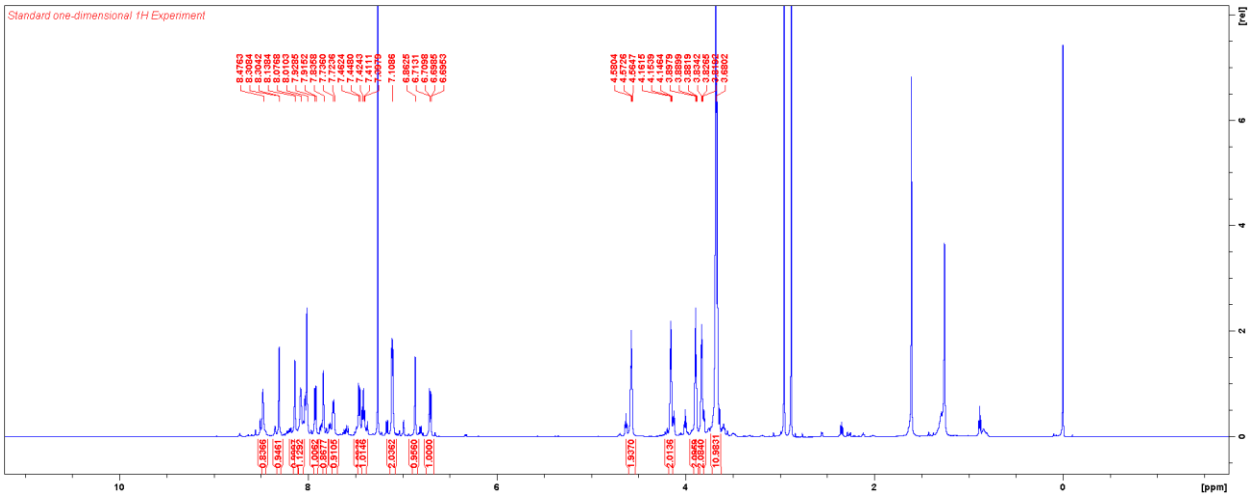
Synthesis of IAOMe-PEG$_2$-azide (compound **36**)

To a solution of 5-Hydroxy-3-indoleacetic acid methyl ester (88 mg, 0.43 mmol) and azide-PEG$_2$-OTs (480 mg, 1.3 mmol) in DMF (2 mL) was added K_2CO_3 (178 mg, 1.3 mmol). After being stirred at 70 °C for 4 days, the reaction mixture was diluted with EtOAc, washed with water, dried over $MgSO_4$, filtered and concentrated. Purification on CombiFlash using 0–100 % EtOAc/hexanes yielded the title compound (36 mg, 99 µmol, 23 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.96 (br, 1H), 7.24 (d, J = 8.9 Hz, 1H), 7.15 (s, 1H), 7.07 (d, J = 2.0 Hz, 1H), 6.90 (dd, J = 8.9 Hz, 2.0 Hz, 1H), 4.19 (t, J = 5.0 Hz, 2H), 3.90 (t, J = 5.0 Hz, 2H), 3.77 (t, J = 4.4 Hz, 2H), 3.74 (m, 2H), 3.70 (m, 7H), 3.39 (t, J = 5.0 Hz, 2H).



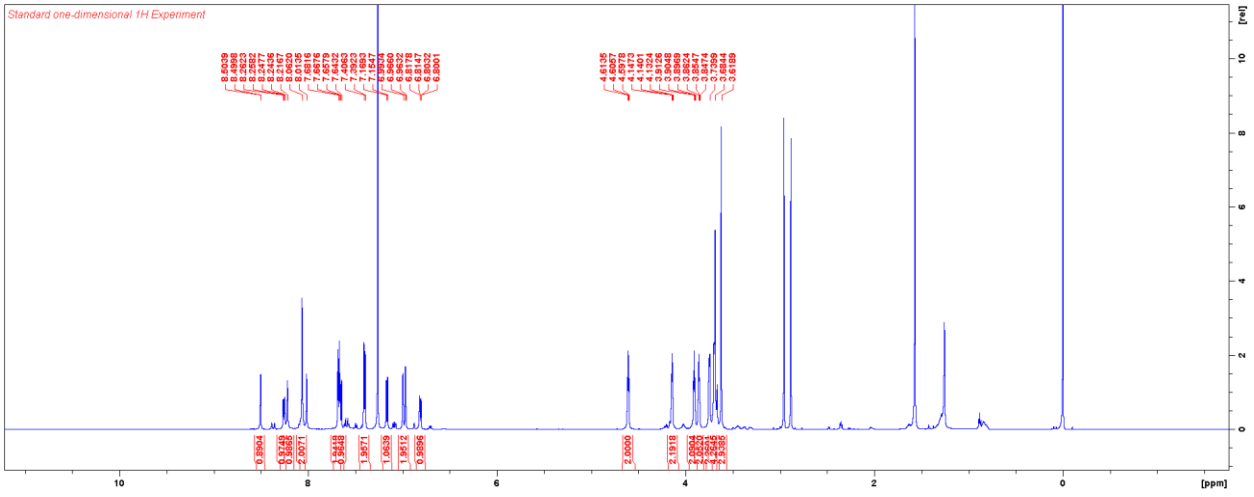
Synthesis of *m*-triazole-GW9662-PEG$_2$-IAOMe (compound **37**)

To a solution of *m*-alk-GW9662 (6 mg, 19 µmol) and IAOMe-PEG$_2$-azide (7 mg, 19 µmol) in DMF (200 µL) was added CuI (~ 1 mg). After being stirred at room temperature overnight, the reaction mixture was directly loaded on CombiFlash (0–10 % MeOH/DCM). Yield: 6 mg (8 µmol, 44 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.48 (s, 1H), 8.30 (d, J = 2.5 Hz, 1H), 8.14 (s, 1H), 8.08 (s, 1H), 7.92 (d, J = 7.8 Hz, 1H), 7.84 (s, 1H), 7.73 (d, J = 7.8 Hz, 1H),7.45 (d, J = 8.4 Hz, 1H), 7.41 (t, J = 8.4 Hz, 1H), 7.11 (m, 2H), 6.86 (s, 1H), 6.71 (dd, J = 8.9 Hz, 2.1 Hz, 1H), 4.57 (t, J = 4.8 Hz, 2H), 4.15 (t, J = 4.5 Hz, 2H), 3.89 (t, J = 4.8 Hz, 2H), 3.83 (t, J = 4.5 Hz, 2H), 3.68 (m, 9H).



Synthesis of *p*-triazole-GW9662-PEG$_2$-IAOMe (compound **38**)

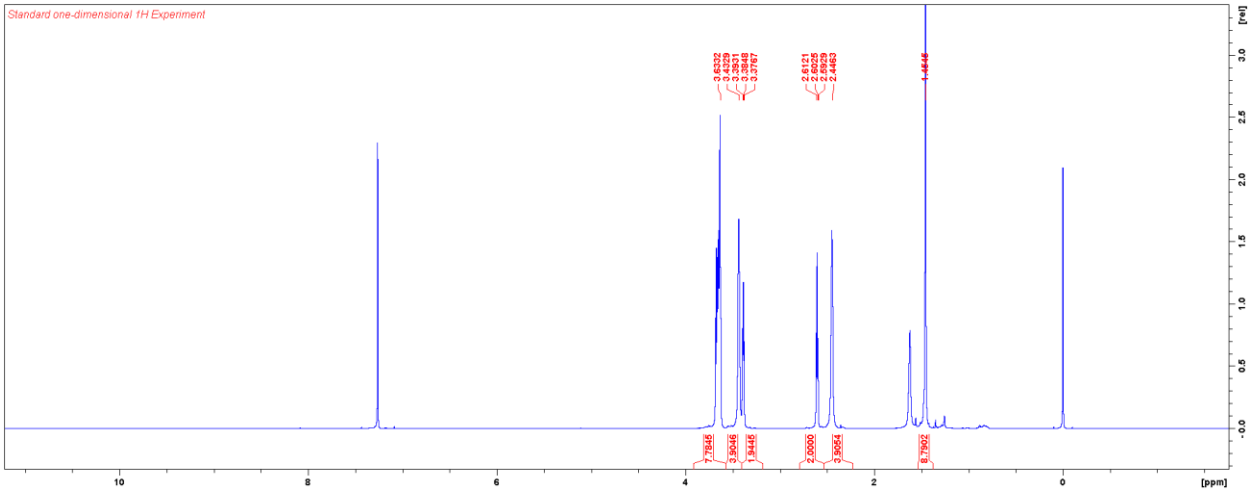
p-Triazole-GW9662-PEG$_2$-azide was synthesised in the same manner as *m*-triazole-GW9662-PEG$_2$-azide. ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.50 (d, J = 2.5 Hz, 1H), 8.26 (dd, J = 8.7 Hz, 2.5 Hz, 1H), 8.22 (s, 1H), 8.08 (m, 2H), 7.68 (d, J = 8.4 Hz, 2H), 7.65 (d, J = 8.7 Hz, 1H), 7.40 (d, J = 8.4 Hz, 1H), 7.16 (d, J = 8.8 Hz, 1H), 6.99 (s, 1H), 6.97 (d, J = 1.9 Hz, 1H), 6.81 (dd, J = 8.7 Hz, 1.9 Hz, 1H), 4.61 (t, J = 4.8 Hz, 2H), 4.14 (t, J = 4.5 Hz, 2H), 3.90 (t, J = 4.8 Hz, 2H), 3.85 (t, J = 4.5 Hz, 2H), 3.74 (br, 2H), 3.69 (m, 4H), 3.62 (s, 3H).



General method for the synthesis of Boc-piperazine-PEG-azide linkers (compounds **42**–**44**)

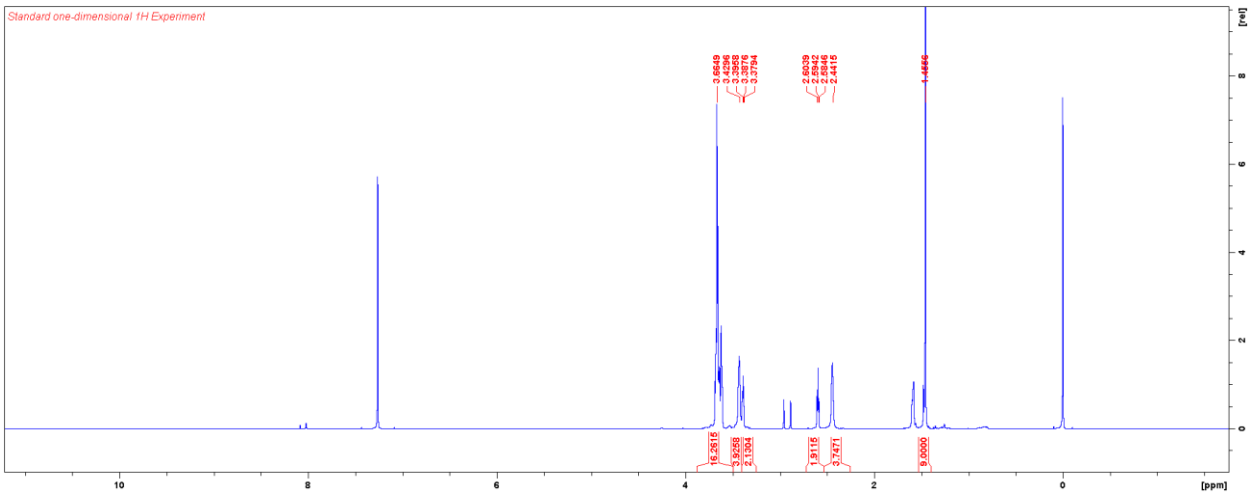
To a solution of 1-Boc-piperazine (1 equiv.) and an azide-PEG-OTs (1 equiv.) in DMF (0.5 M) was added K_2CO_3 (2 equiv.). After being stirred at 60 °C for 24 h, the reaction mixture was diluted with Et_2O, washed with 1 N NaOH, dried over Na_2SO_4, filtered and concentrated. Purification on CombiFlash using 0–10 % MeOH/DCM yielded a Boc-protected linker.

Boc-piperazine-PEG_2-azide (**42**); Yield: 57–87 % from triethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 3.63 (m, 8H), 3.43 (br, 4H), 3.38 (t, J = 4.9 Hz, 2H), 2.60 (t, J = 5.8 Hz, 2H), 2.45 (br, 4H), 1.45 (s, 9H).



Boc-piperazine-PEG$_3$-azide (**43**); Yield: 65 % from tetraethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ NA.

Boc-piperazine-PEG$_4$-azide (**44**); Yield: 73 % from pentaethylene glycol bis(*p*-toluenesulfonate); ^{1}H NMR ($CDCl_3$, 600 MHz): δ 3.66 (m, 16H), 3.43 (br, 4H), 3.39 (t, J = 4.9 Hz, 2H), 2.59 (t, J = 5.8 Hz, 2H), 2.44 (br, 4H), 1.46 (s, 9H).



General method for the synthesis of piperazine-PEG-azide linkers

A Boc-protected linker was treated with 1 N HCl in dioxane. After being stirred at room temperature overnight, the reaction mixture was concentrated in vacuo and azeotroped twice with toluene. The residue was suspended in DCM, and 1 N NaOH was added. The organic layer was extracted thrice with DCM, dried over Na_2SO_4, filtered, and concentrated to yield a piperazine-PEG-azide linker. This material was used without further purification.

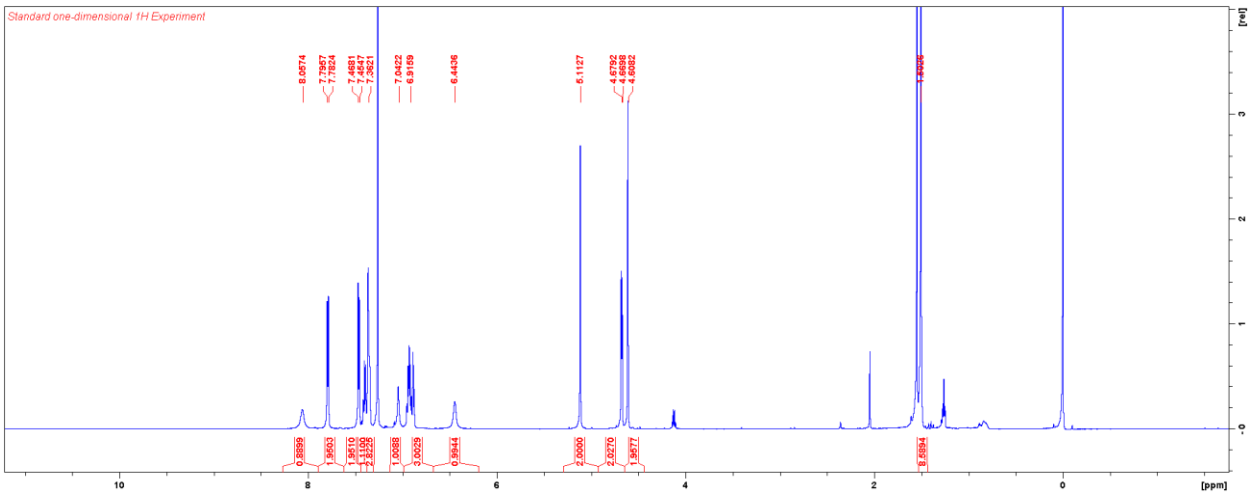
Piperazine-PEG$_2$-azide linker; Yield: 89 %.

Piperazine-PEG$_3$-azide linker; Yield: 48 %.

Piperazine-PEG$_4$-azide linker; Yield: quant.

Synthesis of compound **45**

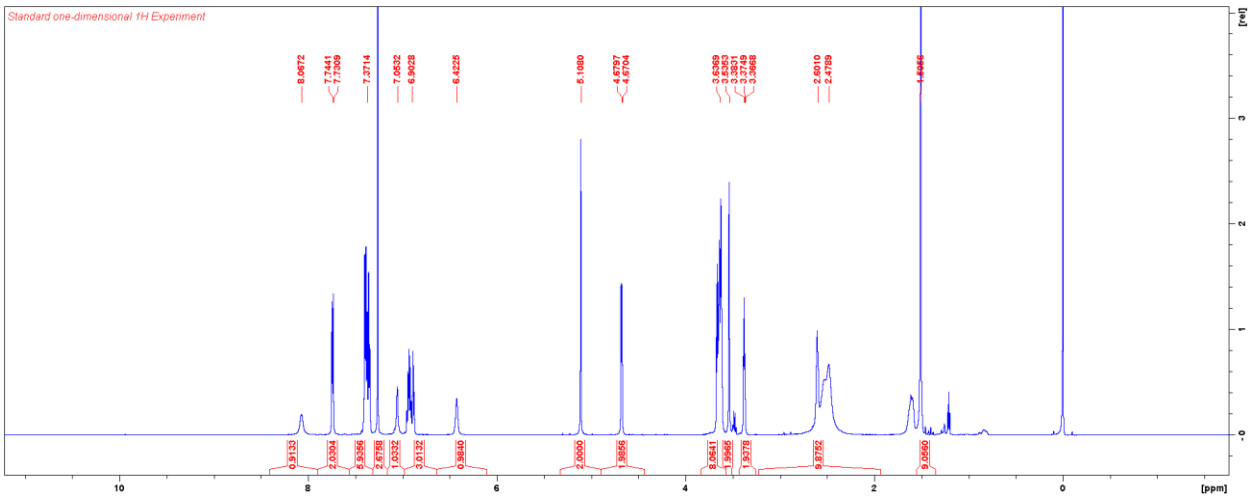
The title compound was synthesised from 2-aminophenol in five steps as previously described.[91] ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.06 (br, 1H), 7.79 (d, J = 8.0 Hz, 2H), 7.46 (d, J = 8.0 Hz, 2H), 7.38 (m, 4H), 7.04 (br, 1H), 6.91 (m, 3H), 6.44 (br, 1H), 5.11 (s, 2H), 4.68 (d, J = 5.6 Hz, 2H), 4.61 (s, 2H), 1.50 (s, 9H).



General method for the synthesis of compounds **46**–**48**

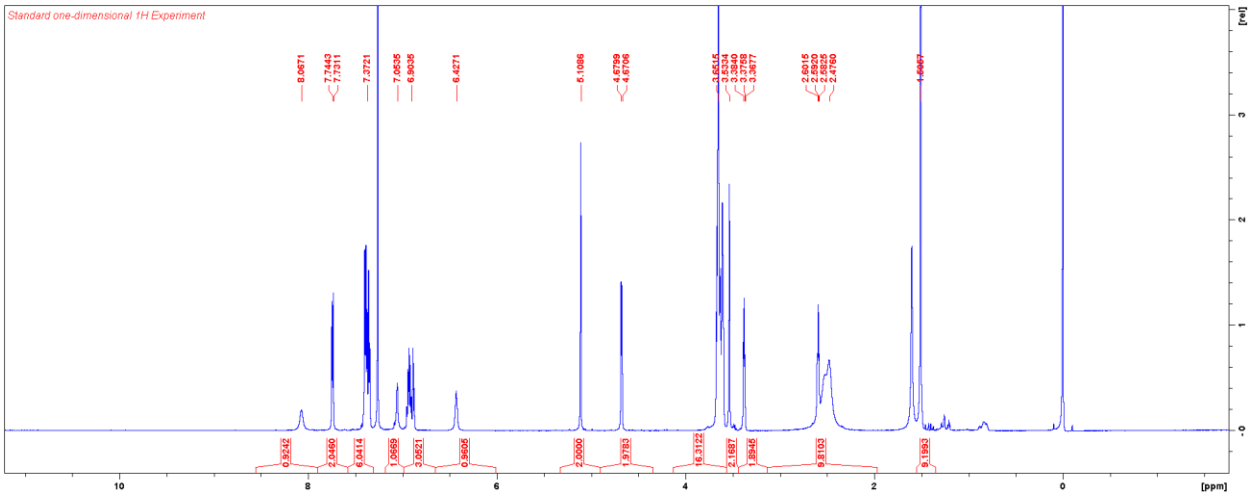
To a solution of the first intermediate 1 (1 equiv.) and a piperazine-PEG-azide linker (1 euiv.) in DMF (0.5 M) was added K_2CO_3 (3 equiv.). After being stirred at 60 °C for 24 h, the reaction mixture was diluted with Et_2O and 1 N NaOH, and extracted thrice with Et_2O. Combined organic layer was dried over $MgSO_4$, filtered and concentrated. Purification on CombiFlash using 0–10 % MeOH/DCM yielded a second intermediate.

Compound **46**; 50–92 % from piperazine-PEG_2-azide linker; ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.07 (br, 1H), 7.74 (d, J = 8.0 Hz, 2H), 7.37 (m, 6H), 7.05 (br, 1H), 6.90 (m, 3H), 6.42 (br, 1H), 5.11 (s, 2H), 4.68 (d, J = 5.6 Hz, 2H), 3.64 (m, 8H), 3.54 (s, 2H), 3.37 (t, J = 4.9 Hz, 2H), 2.60 (br, 2H), 2.48 (8H), 1.51 (s, 9H).



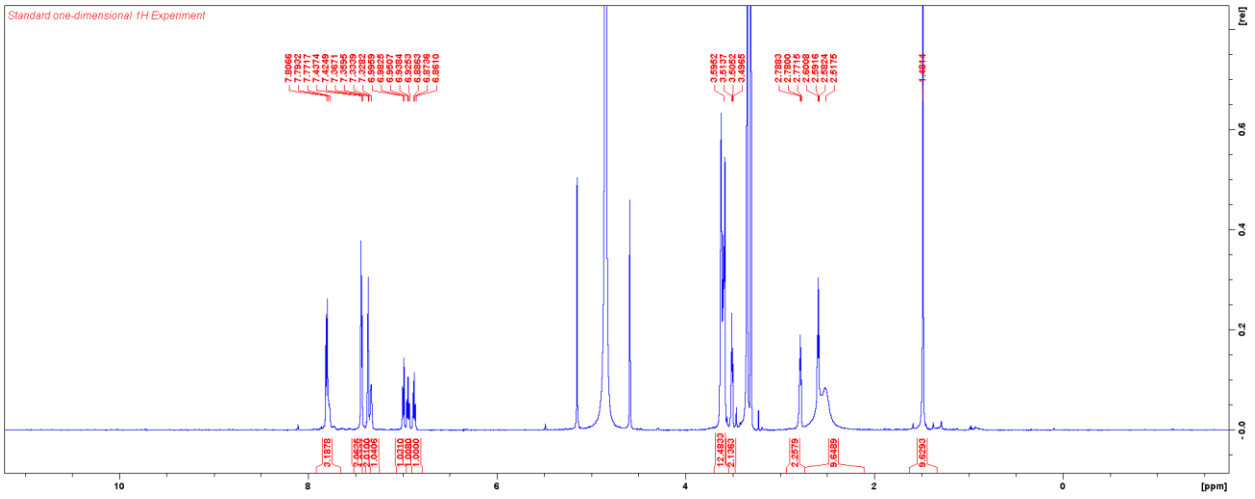
Compound **47**; 75 % from piperazine-PEG_3-azide linker; ^{1}H NMR ($CDCl_3$, 600 MHz): δ NA.

Compound **48**; 73 % from piperazine-PEG_4-azide linker; ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.07 (br, 1H), 7.74 (d, J = 8.0 Hz, 2H), 7.37 (m, 6H), 7.05 (br, 1H), 6.90 (m, 3H), 6.42 (br, 1H), 5.11 (s, 2H), 4.68 (d, J = 5.6 Hz, 2H), 3.65 (m, 16H), 3.53 (s, 2H), 3.38 (t, J = 4.9 Hz, 2H), 2.59 (t, J = 5.7 Hz, 2H), 2.48 (8H), 1.51 (s, 9H).



Synthesis of compound **49**

To a solution of the second intermediate with piperazine-PEG_3-azide linker (138 mg, 0.19 mmol) in 10 % H_2O/THF (2 mL) was added triphenylphosphine (74 mg, 0.28 mmol). After being stirred at room temperature overnight, the reaction mixture was concentrated. Purification on gravity column using 8 % MeOH/DCM + 1 % NH_4OH yielded the free amine (130 mg, 93 %). ^{1}H NMR (MeOD, 600 MHz): δ 7.80 (d, J = 8.0 Hz, 2H), 7.77 (br, 1H), 7.43 (m, 3H), 7.37 (m, 2H), 7.33 (d, J = 3.4 Hz, 1H), 6.99 (d, J = 8.0 Hz, 1H), 6.94 (m, 1H), 6.87 (m, 1H), 3.60 (m, 12H), 3.51 (t, J = 5.2 Hz, 2H), 2.78 (t, J = 5.0 Hz, 2H), 2.59 (t, J = 5.5 Hz, 2H), 2.52 (8H), 1.48 (s, 9H).



Attempted synthesis of **49**-(+)-JQ1

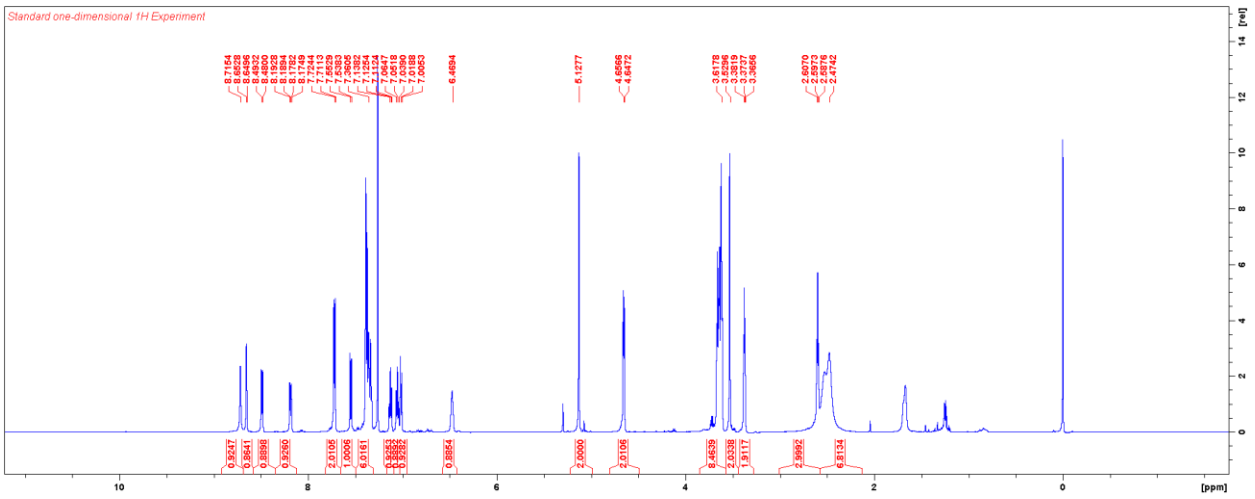
To a solution of (+)-JQ1-acid (19 mg, 48 µmol) and HATU (18 mg, 48 µmol) in DCM (0.5 mL) was added DIPEA (47 µL, 268 µmol). The reaction mixture was stirred on ice for 5 min. To this the free amine intermediate was added. The reaction mixture was allowed to warm to room temperature over 1 h. This resulted in multiple products that proved to be inseparable using normal phase chromatography techniques.

Synthesis of compound **50**

Compound **46** (146 mg, 0.21 mmol) was treated with TFA (0.5 mL) in DCM (1 mL) on ice for 1 h. The solvent was removed under reduced pressure. The residue was azeotroped twice with toluene and purified on CombiFlash using 0–25 % MeOH/$CHCl_3$ to yield an aniline intermediate (100 mg, 0.15 mmol, 72 %).

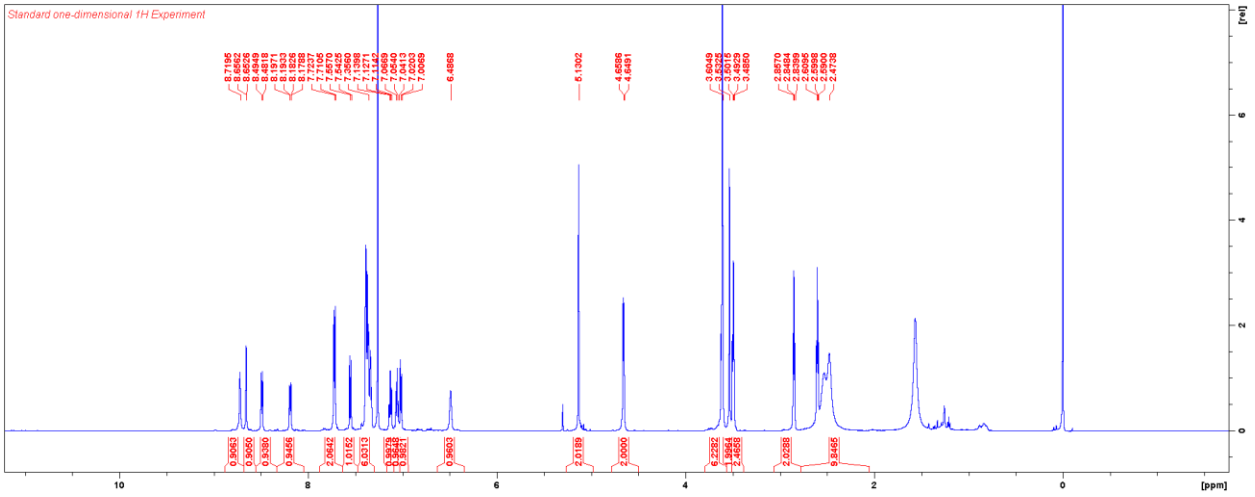
Synthesis of SB1453-PEG_2-azide (compound **51**)

To a solution of 2-chloro-5nitrobenzoic acid (33 mg, 0.15 mmol) and TEA (21 µL, 0.15 mmol) in DCM (0.75 mL) was added the aniline intermediate (100 mg, 0.15 mmol) in DCM (0.75 mL) on ice. After being stirred for 1 h, the reaction mixture was diluted with DCM, washed with NaOH, dried over Na_2SO_4, filtered and concentrated. Purification on CombiFlash using 0–25 % MeOH/DCM yielded SB1453-PEG_2-azide as a white solid (109 mg, 0.14 mmol, 94 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.72 (s, 1H), 8.65 (d, J = 1.9 Hz, 1H), 8.49 (d, J = 8.0 Hz, 1H), 8.19 (dd, J = 8.0 Hz, 2.0 Hz, 1H), 7.72 (d, J = 7.9 Hz, 2H), 7.55 (d, J = 8.8 Hz, 1H), 7.36 (m, 6H), 7.13 (t, J = 7.8 Hz, 1H), 7.05 (t, J = 7.7 Hz, 1H), 7.01 (d, J = 8.1 Hz, 1H), 6.46 (br, 1H), 5.13 (s, 2H), 4.65 (d, J = 5.6 Hz, 2H), 3.62 (m, 8H), 3.53 (s, 2H), 3.37 (t, J = 4.9 Hz, 2H), 2.60 (t, J = 5.8 Hz, 2H), 2.47 (br, 8 H).



Synthesis of SB1453-PEG$_2$-NH$_2$ (compound **52**)

To a solution of SB1453-PEG$_2$-azide (109 mg, 0.14 mmol) in 10 % H$_2$O/THF (1 mL) was added triphenylphosphine (56 mg, 0.21 mmol). After being stirred at room temperature overnight, the reaction mixture was concentrated. Purification on gravity column using 7.5 % MeOH/DCM + 1 % NH_4OH yielded the free amine as yellow oil (80 mg, 0.11 mmol, 76 %). ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.72 (s, 1H), 8.65 (d, J = 1.9 Hz, 1H), 8.49 (d, J = 8.0 Hz, 1H), 8.19 (dd, J = 8.0 Hz, 2.0 Hz, 1H), 7.72 (d, J = 7.9 Hz, 2H), 7.55 (d, J = 8.8 Hz, 1H), 7.36 (m, 6H), 7.13 (t, J = 7.8 Hz, 1H), 7.05 (t, J = 7.7 Hz, 1H), 7.01 (d, J = 8.1 Hz, 1H), 6.49 (br, 1H), 5.13 (s, 2H), 4.65 (d, J = 5.6 Hz, 2H), 3.60 (m, 6H), 3.53 (s, 2H), 3.49 (t, J = 4.9 Hz, 2H), 2.85 (t, J = 5.2 Hz, 2H), 2.60 (t, J = 5.8 Hz, 2H), 2.47 (br, 8 H).



Attempted synthesis of SB1453-PEG$_2$-(+)-JQ1

To a solution of (+)-JQ1-acid (22 mg, 54 µmol) and HATU (20 mg, 54 µmol) in DCM (0.5 mL) was added TEA (37 µL, 268 µmol). The reaction mixture was stirred on ice for 5 min. To this the SB1453-PEG$_2$-NH$_2$ (40 mg, 54 µmol) in DCM (0.5 mL) was added. The reaction mixture was allowed to warm to room temperature over 1 h. This resulted in multiple products that proved to be inseparable using normal phase chromatography techniques.

REFERENCES

(1) Wofsy, L.; Metzger, H.; Singer, S. J. Affinity Labeling—a General Method for Labeling the Active Sites of Antibody and Enzyme Molecules. *Biochemistry* **1962**, *1* (6), 1031–1039. https://doi.org/10.1021/bi00912a013.

(2) Schoellmann, G.; Shaw, E. Direct Evidence for the Presence of Histidine in the Active Center of Chymotrypsin. *Biochemistry* **1963**, *2* (2), 252–255. https://doi.org/10.1021/bi00902a008.

(3) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. *Nat. Biotechnol.* **1999**, *17* (10), 994–999. https://doi.org/10.1038/13690.

(4) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. Direct Analysis of Protein Complexes Using Mass Spectrometry. *Nat. Biotechnol.* **1999**, *17* (7), 676–682. https://doi.org/10.1038/10890.

(5) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. Activity-Based Protein Profiling: The Serine Hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (26), 14694–14699. https://doi.org/10.1073/pnas.96.26.14694.

(6) Bogyo, M.; Shin, S.; McMaster, J. S.; Ploegh, H. L. Substrate Binding and Sequence Preference of the Proteasome Revealed by Active-Site-Directed Affinity Probes. *Chem. Biol.* **1998**, *5* (6), 307–320. https://doi.org/10.1016/S1074-5521(98)90169-7.

(7) Bogyo, M.; McMaster, J. S.; Gaczynska, M.; Tortorella, D.; Goldberg, A. L.; Ploegh, H. Covalent Modification of the Active Site Threonine of Proteasomal β Subunits and the Escherichia Coli Homolog HslV by a New Class of Inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (13), 6629–6634. https://doi.org/10.1073/pnas.94.13.6629.

(8) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **2008**, *77* (1), 383–414. https://doi.org/10.1146/annurev.biochem.75.101304.124125.

(9) Speers, A. E.; Adam, G. C.; Cravatt, B. F. Activity-Based Protein Profiling in Vivo Using a Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J. Am. Chem. Soc.* **2003**, *125* (16), 4686–4687. https://doi.org/10.1021/ja034490h.

(10) Speers, A. E.; Cravatt, B. F. Profiling Enzyme Activities In Vivo Using Click Chemistry Methods. *Chem. Biol.* **2004**, *11* (3), 535–546. https://doi.org/10.1016/j.

(11) Saxon, E.; Bertozzi, C. R. Cell Surface Engineering by a Modified Staudinger Reaction. *Science (80-.).* **2000**, *287* (5460), 2007–2010. https://doi.org/10.1126/science.287.5460.2007.

(12) Ovaa, H.; Van Swieten, P. F.; Kessler, B. M.; Leeuwenburgh, M. A.; Fiebiger, E.; Van den Nieuwendijk, A. M. C. H.; Galardy, P. J.; Van der Marel, G. A.; Ploegh, H. L.; Overkleeft, H. S. Chemistry in Living Cells: Detection of Active Proteasomes by a Two-Step Labeling Strategy. *Angew. Chemie - Int. Ed.* **2003**, *42* (31), 3626–3629. https://doi.org/10.1002/anie.200351314.

(13) Hang, H. C.; Loureiro, J.; Spooner, E.; van der Velden, A. W. M.; Kim, Y. M.; Pollington, A. M.; Maehr, R.; Starnbach, M. N.; Ploegh, H. L. Mechanism-Based Probe for the Analysis of Cathepsin Cysteine Proteases in Living Cells. *ACS Chem. Biol.* **2006**, *1* (11), 713–723. https://doi.org/10.1021/cb600431a.

(14) Li, Y. M.; Xu, M.; Lai, M. T.; Huang, Q.; Castro, J. L.; DiMuzlo-Mower, J.; Harrison, T.; Lellis, C.; Nadin, A.; Neduvelli, J. G.; et al. Photoactivated γ-Secretase Inhibitors Directed to the Active Site Covalently Label Presenilin 1. *Nature* **2000**, *405* (6787), 689–694. https://doi.org/10.1038/35015085.

(15) Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F. Activity-Based Probes for the Proteomic Profiling of Metalloproteases. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (27), 10000–10005. https://doi.org/10.1073/pnas.0402784101.

(16) Hagenstein, M. C.; Mussgnug, J. H.; Lotte, K.; Plessow, R.; Brockhinke, A.; Kruse, O.; Sewald, N. Affinity-Based Tagging of Protein Families with Reversible Inhibitors: A Concept for Functional Proteomics. *Angew. Chemie - Int. Ed.* **2003**, *42* (45), 5635–5638. https://doi.org/10.1002/anie.200352084.

(17) Brunner, J. New Photolabeling and Crosslinking Methods. *Annu. Rev. Biochem.* **1993**, *62*, 483–514. https://doi.org/10.1146/annurev.bi.62.070193.002411.

(18) Soethoudt, M.; Stolze, S. C.; Westphal, M. V.; Van Stralen, L.; Martella, A.; Van Rooden, E. J.; Guba, W.; Varga, Z. V.; Deng, H.; Van Kasteren, S. I.; et al. Selective Photoaffinity Probe That Enables Assessment of Cannabinoid CB 2 Receptor Expression and Ligand Engagement in Human Cells. *J. Am. Chem. Soc.* **2018**, *140* (19), 6067–6075. https://doi.org/10.1021/jacs.7b11281.

(19) Yang, X.; Michiels, T. J. M.; De Jong, C.; Soethoudt, M.; Dekker, N.; Gordon, E.; Van Der Stelt, M.; Heitman, L. H.; Van Der Es, D.; Ijzerman, A. P. An Affinity-Based Probe for the Human Adenosine A2A Receptor. *J. Med. Chem.* **2018**, *61* (17), 7892–7901. https://doi.org/10.1021/acs.jmedchem.8b00860.

(20) D'Ascenzio, M.; Pugh, K. M.; Konietzny, R.; Berridge, G.; Tallant, C.; Hashem, S.; Monteiro, O.; Thomas, J. R.; Schirle, M.; Knapp, S.; et al. An Activity-Based Probe Targeting Non-Catalytic, Highly Conserved Amino Acid Residues within Bromodomains. *Angew. Chemie - Int. Ed.* **2019**, *58* (4), 1007–1012. https://doi.org/10.1002/anie.201807825.

(21) Zhao, Q.; Ouyang, X.; Wan, X.; Gajiwala, K. S.; Kath, J. C.; Jones, L. H.; Burlingame, A. L.; Taunton, J. Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J. Am. Chem. Soc.* **2017**, *139* (2), 680–685. https://doi.org/10.1021/jacs.6b08536.

(22) Li, X.; Wu, Y.; Tian, G.; Jiang, Y.; Liu, Z.; Meng, X.; Bao, X.; Feng, L.; Sun, H.; Deng, H.; et al. Chemical Proteomic Profiling of Bromodomains Enables the Wide-Spectrum Evaluation of Bromodomain Inhibitors in Living Cells. *J. Am. Chem. Soc.* **2019**, *141* (29), 11497–11505. https://doi.org/10.1021/jacs.9b02738.

(23) Tsukidate, T.; Li, Q.; Hang, H. C. Nuclear Receptor Chemical Reporter Enables Domain-Specific Analysis of Ligands in Mammalian Cells. *ACS Chem. Biol.* **2020**, *15* (9), 2324–

2330. https://doi.org/10.1021/acschembio.0c00432.

(24) Wang, Y. C.; Westcott, N. P.; Griffin, M. E.; Hang, H. C. Peptidoglycan Metabolite Photoaffinity Reporters Reveal Direct Binding to Intracellular Pattern Recognition Receptors and Arf GTPases. *ACS Chem. Biol.* **2019**, *14* (3), 405–414. https://doi.org/10.1021/acschembio.8b01038.

(25) Adam, G. C.; Cravatt, B. F.; Sorensen, E. J. Profiling the Specific Reactivity of the Proteome with Non-Directed Activity-Based Probes. *Chem. Biol.* **2001**, *8* (1), 81–95. https://doi.org/10.1016/S1074-5521(00)90060-7.

(26) Adam, G. C.; Sorensen, E. J.; Cravatt, B. F. Proteomic Profiling of Mechanistically Distinct Enzyme Classes Using a Common Chemotype. *Nat. Biotechnol.* **2002**, *20* (8), 805–809. https://doi.org/10.1038/nbt714.

(27) Weerapana, E.; Simon, G. M.; Cravatt, B. F. Disparate Proteome Reactivity Profiles of Carbon Electrophiles. *Nat. Chem. Biol.* **2008**, *4* (7), 405–407. https://doi.org/10.1038/nchembio.91.

(28) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* **2010**, *468* (7325), 790–797. https://doi.org/10.1038/nature09472.

(29) Zanon, P. R. A.; Lewald, L.; Hacker, S. M. Isotopically Labeled Desthiobiotin Azide (IsoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome. *Angew. Chemie - Int. Ed.* **2020**, *59* (7), 2829–2836. https://doi.org/10.1002/anie.201912075.

(30) Evans, M. J.; Saghatelian, A.; Sorensen, E. J.; Cravatt, B. F. Target Discovery in Small-Molecule Cell-Based Screens by in Situ Proteome Reactivity Profiling. *Nat. Biotechnol.* **2005**, *23* (10), 1303–1307. https://doi.org/10.1038/nbt1149.

(31) Kambe, T.; Correia, B. E.; Niphakis, M. J.; Cravatt, B. F. Mapping the Protein Interaction Landscape for Fully Functionalized Small-Molecule Probes in Human Cells. *J. Am. Chem. Soc.* **2014**, *136* (30), 10777–10782. https://doi.org/10.1021/ja505517t.

(32) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; et al. Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **2017**, *168* (3), 527-541.e29. https://doi.org/10.1016/j.cell.2016.12.029.

(33) Cisar, J. S.; Cravatt, B. F. Fully Functionalized Small-Molecule Probes for Integrated Phenotypic Screening and Target Identification. *J. Am. Chem. Soc.* **2012**, *134* (25), 10385–10388. https://doi.org/10.1021/ja304213w.

(34) Niphakis, M. J.; Lum, K. M.; Cognetta, A. B.; Correia, B. E.; Ichu, T. A.; Olucha, J.; Brown, S. J.; Kundu, S.; Piscitelli, F.; Rosen, H.; et al. A Global Map of Lipid-Binding Proteins and Their Ligandability in Cells. *Cell* **2015**, *161* (7), 1668–1680. https://doi.org/10.1016/j.cell.2015.05.045.

(35) Hulce, J. J.; Cognetta, A. B.; Niphakis, M. J.; Tully, S. E.; Cravatt, B. F. Proteome-Wide

Mapping of Cholesterol-Interacting Proteins in Mammalian Cells. *Nat. Methods* **2013**, *10* (3), 259–264. https://doi.org/10.1038/nmeth.2368.

(36) Wang, Y.; Dix, M. M.; Bianco, G.; Remsberg, J. R.; Lee, H.-Y.; Kalocsay, M.; Gygi, S. P.; Forli, S.; Vite, G.; Lawrence, R. M.; et al. Expedited Mapping of the Ligandable Proteome Using Fully Functionalized Enantiomeric Probe Pairs. *Nat. Chem.* **2019**, *11* (12), 1113–1123. https://doi.org/10.1038/s41557-019-0351-5.

(37) Nomura, D. K.; Maimone, T. J. Target Identification of Bioactive Covalently Acting Natural Products. *Curr. Top. Microbiol. Immunol.* **2019**, *420* (1), 351–374. https://doi.org/10.1007/82_2018_121.

(38) Wang, C.; Weerapana, E.; Blewett, M. M.; Cravatt, B. F. A Chemoproteomic Platform to Quantitatively Map Targets of Lipid-Derived Electrophiles. *Nat. Methods* **2013**, *11* (1), 79–85. https://doi.org/10.1038/nmeth.2759.

(39) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-Páez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; et al. Proteome-Wide Covalent Ligand Discovery in Native Biological Systems. *Nature* **2016**, *534* (7608), 570–574. https://doi.org/10.1038/nature18002.

(40) Bar-Peled, L.; Kemper, E. K.; Suciu, R. M.; Vinogradova, E. V.; Backus, K. M.; Horning, B. D.; Paul, T. A.; Ichu, T. A.; Svensson, R. U.; Olucha, J.; et al. Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. *Cell* **2017**, *171* (3), 696-709.e23. https://doi.org/10.1016/j.cell.2017.08.051.

(41) Vinogradova, E. V; Zhang, X.; Remillard, D.; Lazar, D. C.; Suciu, R. M.; Wang, Y.; Bianco, G.; Yamashita, Y.; Crowley, V. M.; Schafroth, M. A.; et al. An Activity-Guided Map of Electrophile-Cysteine Interactions in Primary Human T Cells. *Cell* **2020**, *182* (4), 1009-1026.e29. https://doi.org/10.1016/j.cell.2020.07.001.

(42) Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; et al. Reimagining High-Throughput Profiling of Reactive Cysteines for Cell-Based Screening of Large Electrophile Libraries. *Nat. Biotechnol.* **2021**. https://doi.org/10.1038/s41587-020-00778-3.

(43) Qin, W.; Qin, K.; Zhang, Y.; Jia, W.; Chen, Y.; Cheng, B.; Peng, L.; Chen, N.; Liu, Y.; Zhou, W.; et al. S-Glycosylation-Based Cysteine Profiling Reveals Regulation of Glycolysis by Itaconate. *Nat. Chem. Biol.* **2019**, *15* (10), 983–991. https://doi.org/10.1038/s41589-019-0323-5.

(44) Motiwala, H. F.; Kuo, Y.-H.; Stinger, B. L.; Palfey, B. A.; Martin, B. R. Tunable Heteroaromatic Sulfones Enhance In-Cell Cysteine Profiling. *J. Am. Chem. Soc.* **2020**, *142* (4), 1801–1810. https://doi.org/10.1021/jacs.9b08831.

(45) Yamashita, Y.; Vinogradova, E. V.; Zhang, X.; Suciu, R. M.; Cravatt, B. F. A Chemical Proteomic Probe for the Mitochondrial Pyruvate Carrier Complex. *Angew. Chemie - Int. Ed.* **2020**, *59* (10), 3896–3899. https://doi.org/10.1002/anie.201914391.

(46) Hahm, H. S.; Toroitich, E. K.; Borne, A. L.; Brulet, J. W.; Libby, A. H.; Yuan, K.; Ware,

T. B.; McCloud, R. L.; Ciancone, A. M.; Hsu, K. Global Targeting of Functional Tyrosines Using Sulfur-Triazole Exchange Chemistry. *Nat. Chem. Biol.* **2020**, *16* (2), 150–159. https://doi.org/10.1038/s41589-019-0404-5.

(47) Mortenson, D. E.; Brighty, G. J.; Plate, L.; Bare, G.; Chen, W.; Li, S.; Wang, H.; Cravatt, B. F.; Forli, S.; Powers, E. T.; et al. "Inverse Drug Discovery" Strategy to Identify Proteins That Are Targeted by Latent Electrophiles As Exemplified by Aryl Fluorosulfates. *J. Am. Chem. Soc.* **2018**, *140* (1), 200–210. https://doi.org/10.1021/jacs.7b08366.

(48) Lee, J.; Schapira, M. The Promise and Peril of Chemical Probe Negative Controls. *ACS Chem. Biol.* **2021**, acschembio.1c00036. https://doi.org/10.1021/acschembio.1c00036.

(49) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A.; et al. Target Identification Using Drug Affinity Responsive Target Stability (DARTS). *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (51), 21984–21989. https://doi.org/10.1073/pnas.0910040106.

(50) Piazza, I.; Kochanowski, K.; Cappelletti, V.; Fuhrer, T.; Noor, E.; Sauer, U.; Picotti, P. A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell* **2018**, *172* (1–2), 358-372.e23. https://doi.org/10.1016/j.cell.2017.12.006.

(51) Piazza, I.; Beaton, N.; Bruderer, R.; Knobloch, T.; Barbisan, C.; Chandat, L.; Sudau, A.; Siepe, I.; Rinner, O.; de Souza, N.; et al. A Machine Learning-Based Chemoproteomic Approach to Identify Drug Targets and Binding Sites in Complex Proteomes. *Nat. Commun.* **2020**, *11* (1), 1–13. https://doi.org/10.1038/s41467-020-18071-x.

(52) Molina, D. M.; Jafari, R.; Ignatushchenko, M.; Seki, T.; Larsson, E. A.; Dan, C.; Sreekumar, L.; Cao, Y.; Nordlund, P. Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science (80-.).* **2013**, *341* (6141), 84–87. https://doi.org/10.1126/science.1233606.

(53) Reinhard, F. B. M.; Eberhard, D.; Werner, T.; Franken, H.; Childs, D.; Doce, C.; Savitski, M. F.; Huber, W.; Bantscheff, M.; Savitski, M. M.; et al. Thermal Proteome Profiling Monitors Ligand Interactions with Cellular Membrane Proteins. *Nat. Methods* **2015**, *12* (12), 1129–1131. https://doi.org/10.1038/nmeth.3652.

(54) Kawatkar, A.; Schefter, M.; Hermansson, N. O.; Snijder, A.; Dekker, N.; Brown, D. G.; Lundbäck, T.; Zhang, A. X.; Castaldi, M. P. CETSA beyond Soluble Targets: A Broad Application to Multipass Transmembrane Proteins. *ACS Chem. Biol.* **2019**, *14* (9), 1913–1920. https://doi.org/10.1021/acschembio.9b00399.

(55) Savitski, M. M.; Reinhard, F. B. M.; Franken, H.; Werner, T.; Savitski, M. F.; Eberhard, D.; Molina, D. M.; Jafari, R.; Dovega, R. B.; Klaeger, S.; et al. Tracking Cancer Drugs in Living Cells by Thermal Profiling of the Proteome. *Science (80-.).* **2014**, *346* (6205). https://doi.org/10.1126/science.1255784.

(56) Dai, L.; Prabhu, N.; Yu, L. Y.; Bacanu, S.; Ramos, A. D.; Nordlund, P. Horizontal Cell Biology: Monitoring Global Changes of Protein Interaction States with the Proteome-Wide Cellular Thermal Shift Assay (CETSA). *Annu. Rev. Biochem.* **2019**, *88*, 383–408. https://doi.org/10.1146/annurev-biochem-062917-012837.

(57) Kalxdorf, M.; Günthner, I.; Becher, I.; Kurzawa, N.; Knecht, S.; Savitski, M. M.; Eberl, H. C.; Bantscheff, M. Cell Surface Thermal Proteome Profiling Tracks Perturbations and Drug Targets on the Plasma Membrane. *Nat. Methods* **2021**, *18* (1), 84–91. https://doi.org/10.1038/s41592-020-01022-1.

(58) Tremaroli, V.; Bäckhed, F. Functional Interactions between the Gut Microbiota and Host Metabolism. *Nature* **2012**, *489* (7415), 242–249. https://doi.org/10.1038/nature11552.

(59) Postler, T. S.; Ghosh, S. Understanding the Holobiont: How Microbial Metabolites Affect Human Health and Shape the Immune System. *Cell Metab.* **2017**, *26* (1), 110–130. https://doi.org/10.1016/j.cmet.2017.05.008.

(60) Rakoff-Nahoum, S.; Paglino, J.; Eslami-varzaneh, F.; Edberg, S.; Medzhitov, R. Recognition of Commensal Microflora by Toll- Like Receptors Is Required for Intestinal Homeostasis. *Cell* **2004**, *118* (July 23), 229–241. https://doi.org/10.1016/j.cell.2004.07.002.

(61) Iliev, I. D.; Funari, V. A.; Taylor, K. D.; Nguyen, Q.; Reyes, C. N.; Strom, S. P.; Brown, J.; Becker, C. A.; Fleshner, P. R.; Dubinsky, M.; et al. Interactions between Commensal Fungi and the C-Type Lectin Receptor Dectin-1 Influence Colitis. *Science (80-.).* **2012**, *336* (6086), 1314–1317. https://doi.org/10.1126/science.1221789.

(62) Caruso, R.; Warner, N.; Inohara, N.; Núñez, G. NOD1 and NOD2: Signaling, Host Defense, and Inflammatory Disease. *Immunity* **2014**, *41* (6), 898–908. https://doi.org/10.1016/j.immuni.2014.12.010.

(63) Brestoff, J. R.; Artis, D. Commensal Bacteria at the Interface of Host Metabolism and the Immune System. *Nature Immunology*. 2013, pp 676–684. https://doi.org/10.1038/ni.2640.

(64) Kelly, C. J.; Zheng, L.; Campbell, E. L.; Saeedi, B.; Scholz, C. C.; Bayless, A. J.; Wilson, K. E.; Glover, L. E.; Kominsky, D. J.; Magnuson, A.; et al. Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function. *Cell Host Microbe* **2015**, *17* (5), 662–671. https://doi.org/10.1016/j.chom.2015.03.005.

(65) Smith, P. M.; Howitt, M. R.; Panikov, N.; Michaud, M.; Gallini, C. A.; Bohlooly-Y, M.; Glickman, J. N.; Garrett, W. S. The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science (80-.).* **2013**, *341* (6145), 569–573. https://doi.org/10.1126/science.1241165.

(66) Maslowski, K. M.; Vieira, A. T.; Ng, A.; Kranich, J.; Sierro, F.; Di Yu; Schilter, H. C.; Rolph, M. S.; MacKay, F.; Artis, D.; et al. Regulation of Inflammatory Responses by Gut Microbiota and Chemoattractant Receptor GPR43. *Nature* **2009**, *461* (7268), 1282–1286. https://doi.org/10.1038/nature08530.

(67) Arpaia, N.; Campbell, C.; Fan, X.; Dikiy, S.; Van Der Veeken, J.; Deroos, P.; Liu, H.; Cross, J. R.; Pfeffer, K.; Coffer, P. J.; et al. Metabolites Produced by Commensal Bacteria Promote Peripheral Regulatory T-Cell Generation. *Nature* **2013**, *504* (7480), 451–455. https://doi.org/10.1038/nature12726.

(68) Zelante, T.; Iannitti, R. G.; Cunha, C.; Luca, A. De; Giovannini, G.; Pieraccini, G.;

Zecchi, R.; Angelo, C. D.; Massi-benedetti, C.; Fallarino, F.; et al. Tryptophan Catabolites from Microbiota Engage Aryl Hydrocarbon Receptor and Balance. *Immunity* **2013**, *39* (2), 372–385. https://doi.org/10.1016/j.immuni.2013.08.003.

(69) Bansal, T.; Alaniz, R. C.; Wood, T. K.; Jayaraman, A. The Bacterial Signal Indole Increases Epithelial-Cell Tight-Junction Resistance and Attenuates Indicators of Inflammation. *Proc. Natl. Acad. Sci.* **2010**, *107* (1), 228–233. https://doi.org/10.1073/pnas.0906112107.

(70) Wlodarska, M.; Luo, C.; Kolde, R.; Vlamakis, H.; Porter, J. A.; Xavier, R. J.; Wlodarska, M.; Luo, C.; Kolde, R.; Hennezel, E.; et al. Indoleacrylic Acid Produced by Commensal Peptostreptococcus Species Suppresses Inflammation. *Cell Host Microbe* **2017**, *22* (1), 25-37.e6. https://doi.org/10.1016/j.chom.2017.06.007.

(71) Venkatesh, M.; Mukherjee, S.; Wang, H.; Li, H.; Sun, K.; Benechet, A. P.; Qiu, Z.; Maher, L.; Redinbo, M. R.; Phillips, R. S.; et al. Symbiotic Bacterial Metabolites Regulate Gastrointestinal Barrier Function via the Xenobiotic Sensor PXR and Toll-like Receptor 4. *Immunity* **2014**, *41* (2), 296–310. https://doi.org/10.1016/j.immuni.2014.06.014.

(72) Aranda, A. N. A.; Pascual, A. Nuclear Hormone Receptors and Gene Expression. *Physiol. Rev.* **2001**, *81* (3), 1269–1304.

(73) Evans, R. M.; Mangelsdorf, D. J. Nuclear Receptors, RXR, and the Big Bang. *Cell* **2014**, *157* (1), 255–266. https://doi.org/10.1016/j.cell.2014.03.012.

(74) Sayin, S. I.; Wahlström, A.; Felin, J.; Jäntti, S.; Marschall, H. U.; Bamberg, K.; Angelin, B.; Hyötyläinen, T.; Orešič, M.; Bäckhed, F. Gut Microbiota Regulates Bile Acid Metabolism by Reducing the Levels of Tauro-Beta-Muricholic Acid, a Naturally Occurring FXR Antagonist. *Cell Metab.* **2013**, *17* (2), 225–235. https://doi.org/10.1016/j.cmet.2013.01.003.

(75) Bassaganya-Riera, J.; Viladomiu, M.; Pedragosa, M.; de Simone, C.; Carbo, A.; Shaykhutdinov, R.; Jobin, C.; Arthur, J. C.; Corl, B. A.; Vogel, H.; et al. Probiotic Bacteria Produce Conjugated Linoleic Acid Locally in the Gut That Targets Macrophage PPAR γ to Suppress Colitis. *PLoS One* **2012**, *7* (2), 1–8. https://doi.org/10.1371/journal.pone.0031238.

(76) Byndloss, M. X.; Olsan, E. E.; Rivera-Chávez, F.; Tiffany, C. R.; Cevallos, S. A.; Lokken, K. L.; Torres, T. P.; Byndloss, A. J.; Faber, F.; Gao, Y.; et al. Microbiota-Activated PPAR-γ Signaling Inhibits Dysbiotic Enterobacteriaceae Expansion. *Science (80-.).* **2017**, *357* (6351), 570–575. https://doi.org/10.1126/science.aam9949.

(77) Tontonoz, P.; Spiegelman, B. M. Fat and Beyond: The Diverse Biology of PPARγ. *Annu. Rev. Biochem.* **2008**, *77* (1), 289–312. https://doi.org/10.1146/annurev.biochem.77.061307.091829.

(78) Nepelska, M.; de Wouters, T.; Jacouton, E.; Béguet-Crespel, F.; Lapaque, N.; Doré, J.; Arulampalam, V.; Blottière, H. M. Commensal Gut Bacteria Modulate Phosphorylation-Dependent PPARγ Transcriptional Activity in Human Intestinal Epithelial Cells. *Sci. Rep.* **2017**, *7* (1), 43199. https://doi.org/10.1038/srep43199.

(79) Couvigny, B.; de Wouters, T.; Kaci, G.; Jacouton, E.; Delorme, C.; Doré, J.; Renault, P.; Blottière, H. M.; Guédon, E.; Lapaque, N. Commensal Streptococcus Salivarius Modulates PPARγ Transcriptional Activity in Human Intestinal Epithelial Cells. *PLoS One* **2015**, *10* (5), e0125371. https://doi.org/10.1371/journal.pone.0125371.

(80) Kelly, D.; Campbell, J. I.; King, T. P.; Grant, G.; Jansson, E. A.; Coutts, A. G. P.; Pettersson, S.; Conway, S. Commensal Anaerobic Gut Bacteria Attenuate Inflammation by Regulating Nuclear-Cytoplasmic Shuttling of PPAR-γ and RelA. *Nat. Immunol.* **2004**, *5* (1), 104–112. https://doi.org/10.1038/ni1018.

(81) Are, A.; Aronsson, L.; Wang, S.; Greicius, G.; Lee, Y. K.; Gustafsson, J.-A.; Pettersson, S.; Arulampalam, V. Enterococcus Faecalis from Newborn Babies Regulate Endogenous PPAR Activity and IL-10 Levels in Colonic Epithelial Cells. *Proc. Natl. Acad. Sci.* **2008**, *105* (6), 1943–1948. https://doi.org/10.1073/pnas.0711734105.

(82) Peyrin-Biroulet, L.; Beisner, J.; Wang, G.; Nuding, S.; Oommen, S. T.; Kelly, D.; Parmentier-Decrucq, E.; Dessein, R.; Merour, E.; Chavatte, P.; et al. Peroxisome Proliferator-Activated Receptor Gamma Activation Is Required for Maintenance of Innate Antimicrobial Immunity in the Colon. *Proc. Natl. Acad. Sci.* **2010**, *107* (19), 8772–8777. https://doi.org/10.1073/pnas.0905745107.

(83) Sun, Y.; Zhang, M.; Chen, C. C.; Gillilland, M.; Sun, X.; El-Zaatari, M.; Huffnagle, G. B.; Young, V. B.; Zhang, J.; Hong, S. C.; et al. Stress-Induced Corticotropin-Releasing Hormone-Mediated NLRP6 Inflammasome Inhibition and Transmissible Enteritis in Mice. *Gastroenterology* **2013**, *144* (7), 1478-1487.e8. https://doi.org/10.1053/j.gastro.2013.02.038.

(84) Kempster, S. L.; Belteki, G.; Forhead, A. J.; Fowden, A. L.; Catalano, R. D.; Lam, B. Y.; McFarlane, I.; Charnock-Jones, D. S.; Smith, G. C. S. Developmental Control of the Nlrp6 Inflammasome and a Substrate, IL-18, in Mammalian Intestine. *AJP Gastrointest. Liver Physiol.* **2011**, *300* (2), G253–G263. https://doi.org/10.1152/ajpgi.00397.2010.

(85) Davison, J. M.; Lickwar, C. R.; Song, L.; Breton, G.; Crawford, G. E.; Rawls, J. F. Microbiota Regulate Intestinal Epithelial Gene Expression by Suppressing the Transcription Factor Hepatocyte Nuclear Factor 4 Alpha. *Genome Res.* **2017**, *27* (7), 1195–1206. https://doi.org/10.1101/gr.220111.116.

(86) Robers, M. B.; Friedman-Ohana, R.; Huber, K. V. M.; Huber, K. V. M.; Kilpatrick, L.; Kilpatrick, L.; Vasta, J. D.; Berger, B. T.; Chaudhry, C.; Hill, S.; et al. Quantifying Target Occupancy of Small Molecules within Living Cells. *Annu. Rev. Biochem.* **2020**, *89*, 557–581. https://doi.org/10.1146/annurev-biochem-011420-092302.

(87) Leesnitzer, L. M.; Parks, D. J.; Bledsoe, R. K.; Cobb, J. E.; Collins, J. L.; Consler, T. G.; Davis, R. G.; Hull-Ryde, E. A.; Lenhard, J. M.; Patel, L.; et al. Functional Consequences of Cysteine Modification in the Ligand Binding Sites of Peroxisome Proliferator Activated Receptors by GW9662. *Biochemistry* **2002**, *41* (21), 6640–6650. https://doi.org/10.1021/bi0159581.

(88) Meijer, F. A.; van den Oetelaar, M. C. M.; Doveston, R. G.; Sampers, E. N. R.; Brunsveld, L. Covalent Occlusion of the RORγt Ligand Binding Pocket Allows Unambiguous

Targeting of an Allosteric Site. *ACS Med. Chem. Lett.* **2021**, acsmedchemlett.1c00029. https://doi.org/10.1021/acsmedchemlett.1c00029.

(89) Shannon, D. A.; Banerjee, R.; Webster, E. R.; Bak, D. W.; Wang, C.; Weerapana, E. Investigating the Proteome Reactivity and Selectivity of Aryl Halides. *J. Am. Chem. Soc.* **2014**, *136* (9), 3330–3333. https://doi.org/10.1021/ja4116204.

(90) Ohtera, A.; Miyamae, Y.; Yoshida, K.; Maejima, K.; Akita, T.; Kakizuka, A.; Irie, K.; Masuda, S.; Kambe, T.; Nagao, M. Identification of a New Type of Covalent PPARγ Agonist Using a Ligand-Linking Strategy. *ACS Chem. Biol.* **2015**, *10* (12), 2794–2804. https://doi.org/10.1021/acschembio.5b00628.

(91) Bae, H.; Jang, J. Y.; Choi, S.-S.; Lee, J.-J.; Kim, H.; Jo, A.; Lee, K.-J.; Choi, J. H.; Suh, S. W.; Park, S. B. Mechanistic Elucidation Guided by Covalent Inhibitors for the Development of Anti-Diabetic PPARγ Ligands. *Chem. Sci.* **2016**, *7* (8), 5523–5529. https://doi.org/10.1039/C6SC01279E.

(92) Brust, R.; Lin, H.; Fuhrmann, J.; Asteian, A.; Kamenecka, T. M.; Kojetin, D. J. Modification of the Orthosteric PPARγ Covalent Antagonist Scaffold Yields an Improved Dual-Site Allosteric Inhibitor. *ACS Chem. Biol.* **2017**, *12* (4), 969–978. https://doi.org/10.1021/acschembio.6b01015.

(93) Henke, B. R.; Blanchard, S. G.; Brackeen, M. F.; Brown, K. K.; Cobb, J. E.; Collins, J. L.; Harrington, W. W.; Hashim, M. A.; Hull-Ryde, E. A.; Kaldor, I.; et al. N-(2-Benzoylphenyl)-L-Tyrosine PPARγ Agonists. 1. Discovery of a Novel Series of Potent Antihyperglycemic and Antihyperlipidemic Agents. *J. Med. Chem.* **1998**, *41* (25), 5020–5036. https://doi.org/10.1021/jm9804127.

(94) Berger, J. P.; Petro, A. E.; Macnaul, K. L.; Kelly, L. J.; Zhang, B. B.; Richards, K.; Elbrecht, A.; Johnson, B. A.; Zhou, G.; Doebber, T. W.; et al. Distinct Properties and Advantages of a Novel Peroxisome Proliferator-Activated Protein γ Selective Modulator. *Mol. Endocrinol.* **2003**, *17* (4), 662–676. https://doi.org/10.1210/me.2002-0217.

(95) Choi, S.-S.; Kim, E. S.; Koh, M.; Lee, S.-J.; Lim, D.; Yang, Y. R.; Jang, H.-J.; Seo, K.; Min, S.-H.; Lee, I. H.; et al. A Novel Non-Agonist Peroxisome Proliferator-Activated Receptor γ (PPARγ) Ligand UHC1 Blocks PPARγ Phosphorylation by Cyclin-Dependent Kinase 5 (CDK5) and Improves Insulin Sensitivity. *J. Biol. Chem.* **2014**, *289* (38), 26618–26629. https://doi.org/10.1074/jbc.M114.566794.

(96) Hang, S.; Paik, D.; Yao, L.; Kim, E.; Jamma, T.; Lu, J.; Ha, S.; Nelson, B. N.; Kelly, S. P.; Wu, L.; et al. Bile Acid Metabolites Control TH17 and Treg Cell Differentiation. *Nature* **2019**, *576* (7785), 143–148. https://doi.org/10.1038/s41586-019-1785-z.

(97) Belury, M. A.; Moya-Camarena, S. Y.; Lu, M.; Shi, L.; Leesnitzer, L. M.; Blanchard, S. G. Conjugated Linoleic Acid Is an Activator and Ligand for Peroxisome Proliferator-Activated Receptor-Gamma (PPARγ). *Nutr. Res.* **2002**, *22* (7), 817–824. https://doi.org/10.1016/S0271-5317(02)00393-7.

(98) Bassaganya-Riera, J.; Reynolds, K.; Martino-Catt, S.; Cui, Y.; Hennighausen, L.; Gonzalez, F.; Rohrer, J.; Benninghoff, A. U.; Hontecillas, R. Activation of PPAR γ and δ by Conjugated Linoleic Acid Mediates Protection from Experimental Inflammatory

Bowel Disease. *Gastroenterology* **2004**, *127* (3), 777–791. https://doi.org/10.1053/j.gastro.2004.06.049.

(99) Viladomiu, M.; Hontecillas, R.; Bassaganya-Riera, J. Modulation of Inflammation and Immunity by Dietary Conjugated Linoleic Acid. *Eur. J. Pharmacol.* **2016**, *785*, 87–95. https://doi.org/10.1016/j.ejphar.2015.03.095.

(100) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discov.* **2016**, *15* (9), 605–619. https://doi.org/10.1038/nrd.2016.109.

(101) Rooks, M. G.; Garrett, W. S. Gut Microbiota, Metabolites and Host Immunity. *Nature Reviews Immunology*. Nature Publishing Group 2016, pp 341–352. https://doi.org/10.1038/nri.2016.42.

(102) Waku, T.; Shiraki, T.; Oyama, T.; Maebara, K.; Nakamori, R.; Morikawa, K. The Nuclear Receptor PPARγ Individually Responds to Serotonin- and Fatty Acid-Metabolites. *EMBO J.* **2010**, *29* (19), 3395–3407. https://doi.org/10.1038/emboj.2010.197.

(103) Artis, D. R.; Lin, J. J.; Zhang, C.; Wang, W.; Mehra, U.; Perreault, M.; Erbe, D.; Krupka, H. I.; England, B. P.; Arnold, J.; et al. Scaffold-Based Discovery of Indeglitazar, a PPAR Pan-Active Anti-Diabetic Agent. *Proc. Natl. Acad. Sci.* **2009**, *106* (1), 262–267. https://doi.org/10.1073/pnas.0811325106.

(104) Wang, J.; Yazdani, S.; Han, A.; Schapira, M. Structure-Based View of the Druggable Genome. *Drug Discov. Today* **2020**, *25* (3), 561–567. https://doi.org/10.1016/j.drudis.2020.02.006.

(105) Chen, X.; Liu, M.; Gilson, M. BindingDB: A Web-Accessible Molecular Recognition Database. *Comb. Chem. High Throughput Screen.* **2001**, *4* (8), 719–725. https://doi.org/10.2174/1386207013330670.

(106) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem. Rev.* **2013**, *113* (4), 2343–2394. https://doi.org/10.1021/cr3003533.

(107) Choi, J. H.; Banks, A. S.; Kamenecka, T. M.; Busby, S. A.; Chalmers, M. J.; Kumar, N.; Kuruvilla, D. S.; Shin, Y.; He, Y.; Bruning, J. B.; et al. Antidiabetic Actions of a Non-Agonist PPARγ Ligand Blocking Cdk5-Mediated Phosphorylation. *Nature* **2011**, *477* (7365), 477–481. https://doi.org/10.1038/nature10383.

(108) Ribeiro Filho, H. V.; Bernardi Videira, N.; Bridi, A. V.; Tittanegro, T. H.; Helena Batista, F. A.; de Carvalho Pereira, J. G.; de Oliveira, P. S. L.; Bajgelman, M. C.; Le Maire, A.; Figueira, A. C. M. Screening for PPAR Non-Agonist Ligands Followed by Characterization of a Hit, AM-879, with Additional No-Adipogenic and Cdk5-Mediated Phosphorylation Inhibition Properties. *Front. Endocrinol. (Lausanne).* **2018**, *9* (FEB). https://doi.org/10.3389/fendo.2018.00011.

(109) Stefaniak, J.; Huber, K. V. M. Importance of Quantifying Drug-Target Engagement in Cells. *ACS Medicinal Chemistry Letters*. 2020, pp 403–406. https://doi.org/10.1021/acsmedchemlett.9b00570.

(110) Chakravarthy, M. V; Lodhi, I. J.; Yin, L.; Malapaka, R. R. V; Xu, H. E.; Turk, J. Identification of a Physiologically Relevant Endogenous Ligand for PPAR a in Liver. *Cell* **2009**, *138* (3), 476–488. https://doi.org/10.1016/j.cell.2009.05.036.

(111) Sauer, S. Ligands for the Nuclear Peroxisome Proliferator-Activated Receptor Gamma. *Trends in Pharmacological Sciences*. 2015, pp 688–704. https://doi.org/10.1016/j.tips.2015.06.010.

(112) Druart, C.; Bindels, L. B.; Schmaltz, R.; Neyrinck, A. M.; Cani, P. D.; Walter, J.; Ramer-Tait, A. E.; Delzenne, N. M. Ability of the Gut Microbiota to Produce PUFA-Derived Bacterial Metabolites: Proof of Concept in Germ-Free versus Conventionalized Mice. *Mol. Nutr. Food Res.* **2015**, *59* (8), 1603–1613. https://doi.org/10.1002/mnfr.201500014.

(113) Druart, C.; Neyrinck, A. M.; Vlaeminck, B.; Fievez, V.; Cani, P. D.; Delzenne, N. M. Role of the Lower and Upper Intestine in the Production and Absorption of Gut Microbiota-Derived PUFA Metabolites. *PLoS One* **2014**, *9* (1). https://doi.org/10.1371/journal.pone.0087560.

(114) Devillard, E.; McIntosh, F. M.; Duncan, S. H.; Wallace, R. J. Metabolism of Linoleic Acid by Human Gut Bacteria: Different Routes for Biosynthesis of Conjugated Linoleic Acid. *J. Bacteriol.* **2007**, *189* (6), 2566–2570. https://doi.org/10.1128/JB.01359-06.

(115) Gorissen, L.; Raes, K.; Weckx, S.; Dannenberger, D.; Leroy, F.; De Vuyst, L.; De Smet, S. Production of Conjugated Linoleic Acid and Conjugated Linolenic Acid Isomers by Bifidobacterium Species. *Appl. Microbiol. Biotechnol.* **2010**, *87* (6), 2257–2266. https://doi.org/10.1007/s00253-010-2713-1.

(116) Proschak, E.; Heitel, P.; Kalinowsky, L.; Merk, D. Opportunities and Challenges for Fatty Acid Mimetics in Drug Discovery. *J. Med. Chem.* **2017**, *60* (13), 5235–5266. https://doi.org/10.1021/acs.jmedchem.6b01287.

(117) Carter, A. J.; Kraemer, O.; Zwick, M.; Mueller-Fahrnow, A.; Arrowsmith, C. H.; Edwards, A. M. Target 2035: Probing the Human Proteome. *Drug Discov. Today* **2019**, *24* (11), 2111–2115. https://doi.org/10.1016/j.drudis.2019.06.020.

(118) Paiva, S. L.; Crews, C. M. Targeted Protein Degradation: Elements of PROTAC Design. *Curr. Opin. Chem. Biol.* **2019**, *50*, 111–119. https://doi.org/10.1016/j.cbpa.2019.02.022.

(119) Naito, M.; Ohoka, N.; Shibata, N. SNIPERs—Hijacking IAP Activity to Induce Protein Degradation. *Drug Discov. Today Technol.* **2019**, *31*, 35–42. https://doi.org/10.1016/j.ddtec.2018.12.002.

(120) Burslem, G. M.; Crews, C. M. Proteolysis-Targeting Chimeras as Therapeutics and Tools for Biological Discovery. *Cell* **2020**, *181* (1), 102–114. https://doi.org/10.1016/j.cell.2019.11.031.

(121) Huang, H. T.; Dobrovolsky, D.; Paulk, J.; Yang, G.; Weisberg, E. L.; Doctor, Z. M.; Buckley, D. L.; Cho, J. H.; Ko, E.; Jang, J.; et al. A Chemoproteomic Approach to Query the Degradable Kinome Using a Multi-Kinase Degrader. *Cell Chem. Biol.* **2018**, *25* (1), 88-99.e6. https://doi.org/10.1016/j.chembiol.2017.10.005.

(122) Powell, C. E.; Du, G.; Che, J.; He, Z.; Donovan, K. A.; Yue, H.; Wang, E. S.; Nowak, R. P.; Zhang, T.; Fischer, E. S.; et al. Selective Degradation of GSPT1 by Cereblon Modulators Identified via a Focused Combinatorial Library. *ACS Chem. Biol.* **2020**, *15* (10), 2722–2730. https://doi.org/10.1021/acschembio.0c00520.

(123) Medvar, B.; Raghuram, V.; Pisitkun, T.; Sarkar, A.; Knepper, M. A. Comprehensive Database of Human E3 Ubiquitin Ligases: Application to Aquaporin-2 Regulation. *Physiol. Genomics* **2016**, *48* (7), 502–512. https://doi.org/10.1152/physiolgenomics.00031.2016.

(124) Liu, L.; Damerell, D. R.; Koukouflis, L.; Tong, Y.; Marsden, B. D.; Schapira, M. UbiHub: A Data Hub for the Explorers of Ubiquitination Pathways. *Bioinformatics* **2019**, *35* (16), 2882–2884. https://doi.org/10.1093/bioinformatics/bty1067.

(125) Maple, H. J.; Clayden, N.; Baron, A.; Stacey, C.; Felix, R. Developing Degraders: Principles and Perspectives on Design and Chemical Space. *Medchemcomm* **2019**, *10* (10), 1755–1764. https://doi.org/10.1039/C9MD00272C.

(126) Lee, J. H.; Jang, J. K.; Ko, K. Y.; Jin, Y.; Ham, M.; Kang, H.; Kim, I. Y. Degradation of Selenoprotein S and Selenoprotein K through PPARγ-Mediated Ubiquitination Is Required for Adipocyte Differentiation. *Cell Death Differ.* **2019**, *26* (6), 1007–1023. https://doi.org/10.1038/s41418-018-0180-x.

(127) Zhang, C. PPARγ E3 Ubiquitin Ligase Regulates MUC1-C Oncoprotein Stability. *Oncogene* **2014**, *33* (49), 5619–5625. https://doi.org/10.1038/onc.2013.504.

(128) Hou, Y.; Moreau, F.; Chadee, K. PPARγ Is an E3 Ligase That Induces the Degradation of NFκB/P65. *Nat. Commun.* **2012**, *3* (1), 1300. https://doi.org/10.1038/ncomms2270.

(129) Garcia-Vallvé, S.; Guasch, L.; Tomas-Hernández, S.; Del Bas, J. M.; Ollendorff, V.; Arola, L.; Pujadas, G.; Mulero, M. Peroxisome Proliferator-Activated Receptor γ (PPARγ) and Ligand Choreography: Newcomers Take the Stage. *J. Med. Chem.* **2015**, *58* (14), 5381–5394. https://doi.org/10.1021/jm501155f.

(130) Chevrier, S.; Levine, J. H.; Zanotelli, V. R. T.; Silina, K.; Schulz, D.; Bacac, M.; Ries, C. H.; Ailles, L.; Jewett, M. A. S.; Moch, H.; et al. An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* **2017**, *169* (4), 736-749.e18. https://doi.org/10.1016/j.cell.2017.04.016.

(131) Lavin, Y.; Kobayashi, S.; Leader, A.; Amir, E. ad D.; Elefant, N.; Bigenwald, C.; Remark, R.; Sweeney, R.; Becker, C. D.; Levine, J. H.; et al. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell* **2017**, *169* (4), 750-765.e17. https://doi.org/10.1016/j.cell.2017.04.014.

(132) Kiss, M.; Czimmerer, Z.; Nagy, L. The Role of Lipid-Activated Nuclear Receptors in Shaping Macrophage and Dendritic Cell Function: From Physiology to Pathology. *J. Allergy Clin. Immunol.* **2013**, *132* (2), 264–286. https://doi.org/10.1016/j.jaci.2013.05.044.

(133) Guerriero, J. L. Macrophages: The Road Less Traveled, Changing Anticancer Therapy. *Trends Mol. Med.* **2018**, *24* (5), 472–489. https://doi.org/10.1016/j.molmed.2018.03.006.

(134) Lamotte, Y.; Martres, P.; Faucher, N.; Laroze, A.; Grillot, D.; Ancellin, N.; Saintillan, Y.; Beneton, V.; Gampe, R. T. Synthesis and Biological Activities of Novel Indole Derivatives as Potent and Selective PPARγ Modulators. *Bioorg. Med. Chem. Lett.* **2010**, *20* (4), 1399–1404. https://doi.org/10.1016/j.bmcl.2009.12.107.

(135) Asteian, A.; Blayo, A. L.; He, Y.; Koenig, M.; Shin, Y.; Kuruvilla, D. S.; Corzo, C. A.; Cameron, M. D.; Lin, L.; Ruiz, C.; et al. Design, Synthesis, and Biological Evaluation of Indole Biphenylcarboxylic Acids as PPARγ Antagonists. *ACS Med. Chem. Lett.* **2015**, *6* (9), 998–1003. https://doi.org/10.1021/acsmedchemlett.5b00218.

(136) Ward, C. C.; Kleinman, J. I.; Brittain, S. M.; Lee, P. S.; Chung, C. Y. S.; Kim, K.; Petri, Y.; Thomas, J. R.; Tallarico, J. A.; McKenna, J. M.; et al. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem. Biol.* **2019**, *14* (11), 2430–2440. https://doi.org/10.1021/acschembio.8b01083.

(137) Zhang, X.; Crowley, V. M.; Wucherpfennig, T. G.; Dix, M. M.; Cravatt, B. F. Electrophilic PROTACs That Degrade Nuclear Proteins by Engaging DCAF16. *Nat. Chem. Biol.* **2019**, *15* (7), 737–746. https://doi.org/10.1038/s41589-019-0279-5.

(138) Spradlin, J. N.; Hu, X.; Ward, C. C.; Brittain, S. M.; Jones, M. D.; Ou, L.; To, M.; Proudfoot, A.; Ornelas, E.; Woldegiorgis, M.; et al. Harnessing the Anti-Cancer Natural Product Nimbolide for Targeted Protein Degradation. *Nat. Chem. Biol.* **2019**, *15* (7), 747–755. https://doi.org/10.1038/s41589-019-0304-8.

(139) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62* (12), 5673–5724. https://doi.org/10.1021/acs.jmedchem.8b01153.

(140) Tsukidate, T.; Li, Q.; Hang, H. C. Targeted and Proteome-Wide Analysis of Metabolite–Protein Interactions. *Current Opinion in Chemical Biology*. Elsevier Ltd 2020, pp 19–27. https://doi.org/10.1016/j.cbpa.2019.10.008.

(141) Jost, M.; Weissman, J. S. CRISPR Approaches to Small Molecule Target Identification. *ACS Chem. Biol.* **2018**, *13* (2), 366–375. https://doi.org/10.1021/acschembio.7b00965.

(142) Charron, G.; Zhang, M. M.; Yount, J. S.; Wilson, J.; Raghavan, A. S.; Shamir, E.; Hang, H. C. Robust Fluorescent Detection of Protein Fatty-Acylation with Chemical Reporters. *J. Am. Chem. Soc.* **2009**, *131* (13), 4967–4975. https://doi.org/10.1021/ja810122f.

(143) Zhu, K.; Borrelli, K. W.; Greenwood, J. R.; Day, T.; Abel, R.; Farid, R. S.; Harder, E. Docking Covalent Inhibitors: A Parameter Free Approach to Pose Prediction and Scoring. *J. Chem. Inf. Model.* **2014**, *54* (7), 1932–1940. https://doi.org/10.1021/ci500118s.

(144) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant Computational Platform for Mass Spectrometry-Based Shotgun Proteomics. *Nat. Protoc.* **2016**, *11* (12), 2301–2319. https://doi.org/10.1038/nprot.2016.136.

(145) Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* **2015**, *43* (7), e47. https://doi.org/10.1093/nar/gkv007.

(146) Choi, M.; Chang, C. Y.; Clough, T.; Broudy, D.; Killeen, T.; MacLean, B.; Vitek, O. MSstats: An R Package for Statistical Analysis of Quantitative Mass Spectrometry-Based Proteomic Experiments. *Bioinformatics* **2014**, *30* (17), 2524–2526. https://doi.org/10.1093/bioinformatics/btu305.

(147) Lazar, C.; Gatto, L.; Ferro, M.; Bruley, C.; Burger, T. Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *J. Proteome Res.* **2016**, *15* (4), 1116–1125. https://doi.org/10.1021/acs.jproteome.5b00981.

(148) Reiner, A.; Yekutieli, D.; Benjamini, Y. Identifying Differentially Expressed Genes Using False Discovery Rate Controlling Procedures. *Bioinformatics* **2003**, *19* (3), 368–375. https://doi.org/10.1093/bioinformatics/btf877.

(149) Ritz, C.; Baty, F.; Streibig, J. C.; Gerhard, D. Dose-Response Analysis Using R. *PLoS One* **2015**, *10* (12), 1–13. https://doi.org/10.1371/journal.pone.0146021.

(150) Ye, H.; Liu, R.; Li, D.; Liu, Y.; Yuan, H.; Guo, W.; Zhou, L.; Cao, X.; Tian, H.; Shen, J.; et al. A Safe and Facile Route to Imidazole-1-Sulfonyl Azide as a Diazotransfer Reagent. *Org. Lett.* **2013**, *15* (1), 18–21. https://doi.org/10.1021/ol3028708.