DTU Library

Identifying oxysterol interacting proteins through specific degradation

He, Nianzhe

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
He, N. (2022). *Identifying oxysterol interacting proteins through specific degradation*. DTU Chemistry.

Identifying oxysterol interacting proteins through specific degradation

PhD Thesis

Nianzhe He

November 2022

DTU Chemistry

Technical University of Denmark

Kemitorvet, building 207

2800 Kgs. Lyngby

Denmark

PhD Supervisor:

Associate Professor Luca Laraia

Contents

Preface

The work presented in this PhD thesis is the result of my PhD studies carried out at the Department of Chemistry, Technical University of Denmark from December 2019 to November 2022 under the supervision of Associate Professor Luca Laraia.

The thesis is divided into five chapters. Chapter 1 introduces the concept and application of target protein degradation, oxysterols and target identification methods. Chapter 2 discusses the proteomics data of (oxy)sterols-based PROTACs. Chapter 3 and 4 focus on validation of the potential targets of (oxy)sterols. Chapter 5 introduces the design, synthesis and biological evaluation of Aster-A PROTACs. All the experiments and results are provided in the attached document in detail.

Acknowledgement

Foremost, I would like to express my sincere gratitude to my supervisor, Associate Professor Luca Laraia for giving the opportunity to pursue my PhD in his group. Thank you for your invaluable guidance and support on my project. You are always kind, optimistic and give me confidence to go forward.

I would also like to extend my thanks to all the current and former group members for creating an amazing working environment. Dr. Cecilia Rossetti and Dr. Elisa Sassetti are thanked for their expertise and help on cell culture, western blot and proteomics. Dr. Laura Depta is thanked for her support on protein purification and fluorescence polarization. Dr. Thomas Whitmarsh-Everiss, Joseph Hoock and Ruwei Yao are thanked for their insightful suggestions on my project. Oliver Rafn Dan is thanked for his contributions to the synthetic side of the project. Frederik simonsen bro is thanked for revising the abstract.

I would also like to thank my best friends Xurui, Hua Chen, Cheng Xiang for a cherished time spent together in Denmark. My sincere thanks also goes to Deyuan Cong, my former roommate and best friend for discussion about my projects.

Finally, I would like to thank my previous supervisor Professor Zhichao Zhang and my family for their encouragement and support throughout my studies.

Abstract

Proteolysis Targeting Chimeras (PROTACs) are bifunctional small molecules that recruit E3 ligases to proteins of interest (POIs) and realize their degradation via the ubiquitin-proteasome system (UPS). This strategy has been widely applied to degrade specific proteins including kinases, nuclear receptors, transcriptional regulators and regulatory proteins.

Oxysterols are derivatives of cholesterol but whose new functions continue to be discovered. Their potential targets have been partially explored via chemoproteomics and thermal proteome profiling. However, it is still an urgent demand to expand the toolbox to investigate the targets of oxysterols and study how the changes on the cholesterol scaffold including the position and absolute configuration of the hydroxyl groups leads to their target selectivity.

In the first part of this thesis, cholesterol and 4β-hydroxycholesterol were applied into PROTACs by introducing the E3 ligase cereblon (CRBN)-binding ligand, pomalidomide via alkyl and PEG linkers. Based on the hypothesis that the proteins they degrade may be regarded as their targets, the affected proteins were explored by tandem mass tag (TMT)-based proteomics and the most significantly affected protein, Golgi Integral Membrane Protein 4 (GOLIM4), was chosen for further target validation.

The other part of this thesis focused on the synthesis and evaluation of Aster-A PROTACs. Aster-A exerts crucial functions in both cholesterol transporting and autophagy. However, current small-molecule tools targeting Aster-A can only bind to the steroidogenic acute regulatory protein-related lipid transfer (StART)-like domain of Aster-A and the functions of other domains within Aster-A are poorly interrogated by small molecules. Compared with inhibitors, PROTACs can degrade proteins in a catalytic manner which exerts improved efficiency and destroy the whole protein which will remove all functions related to this protein by degradation.

In this part, Aster-A PROTACs were designed to realize the degradation of Aster-A and explore whether the degradation of Aster-A would result in better autophagy inhibition and other phenotypes.

A series Aster-A PROTACs were generated by choosing different E3 ligands and linkers. Among them, **NGF3** was obtained as the only weak Aster-A degrader with DC_{50}=4.8 μM and D_{max}= 41 % in HeLa cells. Further experiments are needed to optimize **NGF3** and investigate the best cell models to study Aster-A degradation.

Danish abstract

Proteolysis Targeting Chimeras (PROTACs) er bifunktionelle små molekyler, som rekrutterer E3-ligaser til proteiner af interesse (POIs), og realiserer deres nedbrydning via ubiquitin-proteasomsystemet (UPS). Denne strategi er blevet anvendt bredt til at nedbryde specifikke proteiner, herunder kinaser, nukleare receptorer, transkriptionelle regulatorer og regulatoriske proteiner.

Oxysteroler er derivater af kolesterol, hvis nye funktioner fortsat opdages. Deres potentielle mål er blevet delvist udforsket via kemoproteomik og termisk proteomprofilering. Det er dog stadig et presserende krav at udvide værktøjskassen til at undersøge *targets* for oxysteroler og undersøge, hvordan ændringerne på kolesterolstrukturen, inklusive positionen og den absolutte konfiguration af hydroxylgrupperne, fører til deres *target*-selektivitet.

I den første del af denne afhandling blev kolesterol og 4β-hydroxycholesterol anvendt til PROTACs ved at introducere den E3 ligase cereblon (CRBN)-bindende ligand, pomalidomid, via alkyl- og PEG-linkere. Baseret på hypotesen om, at de proteiner, som de nedbryder, kan betragtes som deres mål, blev de berørte proteiner udforsket af tandem mass tag (TMT) -baseret proteomik, og det mest signifikant påvirkede protein, GOLIM4, blev valgt til yderligere målvalidering.

Den anden del af denne afhandling fokuserede på syntese og evaluering af Aster-A PROTACs. Aster-A udfører afgørende funktioner i både kolesteroltransport og autofagi. Imidlertid kan nuværende små "værktøjsmolekyler" rettet mod Aster-A kun binde til det steroidogenic acute regulatory protein-related lipid transfer (StART)-lignende domæne af Aster-A, og funktionerne af de andre domæner i Aster-A er ringe undersøgt af små molekyler. Sammenlignet med inhibitorer kan PROTACs nedbryde proteiner på en katalytisk måde, som giver forbedret effektivitet og ødelægger hele proteinet, hvilket vil fjerne alle funktioner relateret til dette protein ved nedbrydning.

I denne del blev Aster-A PROTACs designet til at realisere nedbrydningen af Aster-A og undersøge, om nedbrydningen af Aster-A ville resultere i bedre autofagi-hæmning og andre fænotyper.

En serie Aster-A PROTACs blev genereret ved at vælge forskellige E3-ligander og linkere. Blandt dem blev **NGF3** opnået som den eneste svage Aster-A-nedbryder med DC_{50} = 4,8 μM og D_{max} = 41 % i HeLa-celler. Yderligere eksperimenter er nødvendige for at optimere **NGF3** og undersøge de bedste cellemodeller til at undersøge Aster-A-nedbrydningen.

Abbreviations

20α-HC	20α-hydroxycholesterol
25-HC	25-hydroxycholesterol
4β-HC	4β-hydroxycholesterol
7-KC	7-ketocholesterol
7TM	seven-transmembrane domain
ABPP	activity-based protein profiling
ABPs	activity-based probes
A*f*BPs	affinity-based probes
Akt	protein kinase B
ANK	ankyrin
AR	androgen receptor
ASGPR	asialoglycoprotein receptor
AURORA-A	Aurora kinase A
AUTACs	autophagy-targeting chimeras
B4GALT1	beta-1,4-galactosyltransferase 1
BafA1	bafilomycin A1
BDNF	brain-derived neurotrophic factor
Boc	*tert*-butoxycarbonyl protecting group
BRD4	bromodomain-containing protein 4
c-Abl	Abelson tyrosine kinase
CD71	transferring receptor-1
CDG	congenital disorder of glycosylation
CDK9	cyclin-dependent kinase 9
CE	cellular cholesteryl ester
CETSA	cellular thermal shift assay
CHO	Chinese hamster ovary
cIAP	cellular inhibitor of apoptosis
CI-MPR	cation-independent mannose 6-phosphate receptor

COG	conserved oligomeric Golgi
COPI	coat protein complex I
CQ	chloroquine
CRABP1	cellular retinoic acid binding protein 1
CRABP2	cellular retinoic acid binding protein 2
CRBN	cereblon
CRD	cysteine-rich domain
CRLs	cullin ring ubiquitin ligases
CT	cholestane-3β,5α,6β-triol
CuAAc	copper (I)-catalyzed azide-alkyne cycloaddition
CXCR2	C-X-C motif chemokine receptor 2
DCAF11	DDB1 and CUL4 associated factor 11
DCAF15	DDB1 and CUL4 associated factor 15
DCAF16	DDB1 and CUL4 associated factor 16
EBI2	epstein barr virus-induced G protein–coupled receptor 2
EGFR	epidermal growth factor receptor
ER	estrogen receptor
ER	endoplasmic reticulum
ERBB3	receptor tyrosine-protein kinase erbB-3
ERC	endocytic recycling compartment
ERGIC	ER-Golgi intermediate compartment
ERK	extracellular signal–regulated kinase
ERRα	estrogen related receptor
ERRα	estrogen-related receptor alpha
FAK	Focal adhesion kinase
FBnG	*p*-fluorobenzylguanine
FF	diphenylalanines
FFAT	diphenylalanines in an acidic tract

FKBP12	FK506-binding protein 12
FP	fluorescence polarization
Fz	Frizzled
GALNT2	polypeptide N-acetylgalactosaminyltransferase 2
GARP	Golgi-associated retrograde protein
Gb3	globotriaocylceramide
GlcNAc phosphotransferase	*N*-acetylglucosamine-1-phosphate transferase
GOLIM4	Golgi Integral Membrane Protein 4
GPP130	Golgi phosphoprotein of 130 kDa
HATU	hexafluorophosphate azabenzotriazole tetramethyl uronium
HECT	homologous to the E6-AP carboxyl terminus
HePTP	PTPPBS family of tyrosine phosphatase
HER2	human epidermal growth factor receptor
Hh	Hedgehog
HyT	hydrophobic tagging
IMiDs	immunomodulatory imide drugs
IMM	inner mitochondrial membrane
INSIG	insulin-induced gene
ITDRF	isothermal dose–response fingerprinting
KEAP1	Kelch-like ECH-associated protein 1
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LpL	lipoprotein lipase
LTRs	lysosome-targeting receptors
LXRα	Liver X receptor α
LXRβ	Liver X receptor β
LYTACs	lysosome-targeting chimeras
Man II	α-mannosidase II
MDM2	mouse double minute 2 homologue

MetAP2	methionine aminopeptidase 2
MQ	MaxQuant
MTCs	multi-subunit tethering complexes (MTCs)
NAE	NEDD8-activating enzyme
NCA	*N*-carboxyanhydride
NGF	nerve growth factor-b
NP-40	Nonidet P-40
NPC1	Niemann-Pick disease intracellular cholesterol transporter 1
OMM	outer mitochondrial membrane
ORD	OSBP-related domain
ORPs	oxysterol binding protein -related proteins
OSBP	oxysterol binding protein
PAL	photo-affinity labeling
PBS	phosphate buffered saline
PCs	phosphatidylcholines
PD	Proteome Discoverer
pERK	phosphorylated ERK
PEs	phosphatidylethanolamines
PH	pleckstrin homology
PM	plasma membrane
PNGase F	peptide-N-glycosidase F
POI	proteins of interest
PP2A	serine/threonine phosphatase
PPIs	protein-protein interactions
PROTACs	proteolysis-targeting chimeras
PSs	phosphatidylserines
PTM	post-translational modification
PVDF	polyvinylidene difluoride

RAP	receptor-associated protein
RING	really interesting new gene
RIPK2	serine/threonine kinase
RNF114	RING finger protein 114
RNF4	RING finger protein 4
RORs	Retinoic acid receptor-related orphan receptors
RXR	retinoic X receptor
SCAP	SREBP cleavage activating protein
SERM	selective estrogen receptor modulator
SILAC	stable isotope labeling with amino acids
Sirt2	Sirtuin 2
Smo	Smoothened
SMs	sphingomyelins
SNARE	soluble *N*-ethylmaleimide-sensitive factor activating protein receptor
SREBPs	sterol regulatory element binding proteins
STARDs	steroidogenic acute regulatory-related lipid transfer domain proteins
STAT3	Signal transducer and activator of transcription 3
STEC	Shiga toxin (Stx)-producing Escherichia coli
STIM1	stromal interaction molecule 1
STPs	Sterol transport proteins
STs	sialyltransferases
STx1	Shiga toxin 1
SWG	schweinfurthin G
TACC3	transforming acidic coiled-coil containing protein 3
TFA	trifluoroacetic acid
TGN	*trans*-Golgi network
TM	transmembrane domain
TMEM97	transmembrane protein 97

TMT	tandem mass tags
TPD	target protein degradation
TPP	thermal proteome profiling
tri-GalNAc	triantenerrary *N*-acetylgalactosamine
Ub	ubiquitin
UPS	ubiquitin-proteasome system
v-ATPase	vacuolar type H^+-ATPase
VHL	von Hippel-Lindau tumor suppressor protein

1 Chapter 1 Introduction

1.1 Introduction to proteolysis pathways.

Almost all proteins in living cells are repeatedly being degraded and re-synthesized to maintain crucial homeostatic functions[1]. Protein degradation is also an important process to remove abnormal proteins, such as mis-folded, mutated and damaged proteins to maintain the quality control of proteins. There are two main pathways responsible for protein degradation: the ubiquitin-proteasome system (UPS) and the autophagy-lysosome system. Normally, the UPS is responsible for the removal of short-lived or soluble misfolded proteins. In contrast, lysosomes mediate the elimination of long-lived proteins, protein aggregates or even organelles.

The UPS contains a series of enzymes catalyzing the connection or removal of ubiquitin (Ub)[2], which serves as an important marker for protein degradation by proteasome[3] (Figure 1.1 A). The protein will be marked for degradation by the connection of a single Ub or Ub chains to a lysine residue on its side chain as a post-translational modification (PTM). The modification process involves three enzymes: E1 Ub-activating enzymes, E2 Ub-conjugating enzymes and E3 ubiquitin-protein ligases. Ub is first conjugated to E1 by forming the thioester bond between its C terminus and the catalytic cysteine on E1 in an ATP-dependent manner. Next, Ub is transferred from E1 to a catalytic cysteine on E2. Then, Ub will be transferred to substrates in two ways[4]: E3 catalyzes the transfer of Ub from E2 to substrates directly (Really interesting new gene, RING E3s) or Ub is transferred from E2 to E3 and then further transferred to the substrate (Homologous to the E6-AP Carboxyl Terminus, HECT E3s or RING-in-between-RING, RBR E3s). Finally, the ubiquitinated substrates will be recognized by the proteasome and degraded by it.
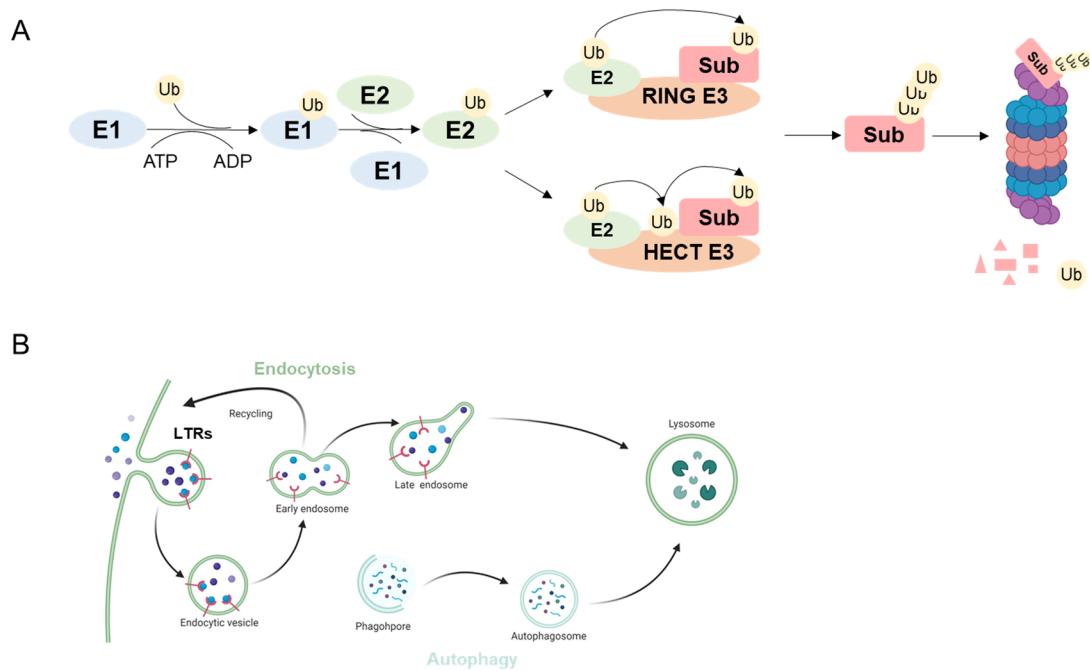

Figure 1.1 The two main protein degradation pathways. (A). Overview of UPS-mediated protein degradation pathways. (B). Lysosomes engage in the recycling and removal of proteins and organelles by endocytosis or autophagy. Created with BioRender.com

Lysosomal degradation is responsible for the removal of extracellular proteins, among others[5]. The lysosome is a membrane-bound organelle containing more than 50 proteases including aspartic, cysteine and serine proteinases[6]. Proteins and even entire organelles can be recruited to lysosomes by various lysosome-targeting receptors (LTRs) and degraded via two main ways: endocytosis or autophagy (Figure 1.1 B). Endocytosis is a crucial way for cells to internalize a series of molecules. This process starts at the plasma membrane (PM) and cell surface proteins can be trafficked either to the PM or other organelles for recycling or to endosome for further lysosomal degradation. Autophagy is a highly conserved process to remove dysfunctional proteins or organelles in a lysosome-dependent manner[7]. Targeted organelles and abnormal proteins (misfolded or aggregated proteins) are wrapped into double-membraned vesicles called autophagosomes. Then, autophagosomes fuse with lysosomes and the digestive enzymes in lysosomes will degrade these contents.

1.2 Target Protein Degradation (TPD)

The UPS and autophagy-lysosome system control the fate of proteins or organelles to maintain cellular homeostasis and normal physiological functions. Recently, these two systems have also been applied to induce targeted protein degradation (TPD), a new paradigm in drug development to degrade disease-associated proteins. A series of bifunctional molecules have been designed to remove proteins of interest (POI). These molecules can be divided into two groups, proteolysis-targeting chimeras (PROTACs)[8] and hydrophobic tagging (HyT)[9], which are mediated by proteasome; as well as lysosome-targeting chimeras (LYTACs)[10] and autophagy-targeting chimeras (AUTACs)[11], which are mediated by lysosome. In this section, the development of these molecules and their applications will be summarized.

1.2.1 Proteolysis Targeting Chimeras (PROTACs)

PROTACs are a series of bifunctional molecules that are able to recruit proteins to the UPS for degradation[12] (Figure 1.2 A). These bifunctional molecules are comprised of three components: a target protein ligand responsible for binding the protein of interest (POI), a ligand for recruiting an E3 ligase and a linker coupling these two parts. When the PROTAC binds to the POI and the E3 ligase together, a ternary complex can be formed and ubiquitin can be transferred from an E2 to the POI. Finally, the poly-ubiquitinated POI is recognized and subsequently degraded by the 26S proteasome (Figure 1.2 B). As such, PROTACs hijack the UPS system to degrade specific cellular proteins. Compared with traditional knockdown methods such as siRNA[13], which removes a specific protein at the RNA level, PROTACs mediate the degradation of a specific protein at the posttranslational level, which emerges as a promising chemical knockdown method.
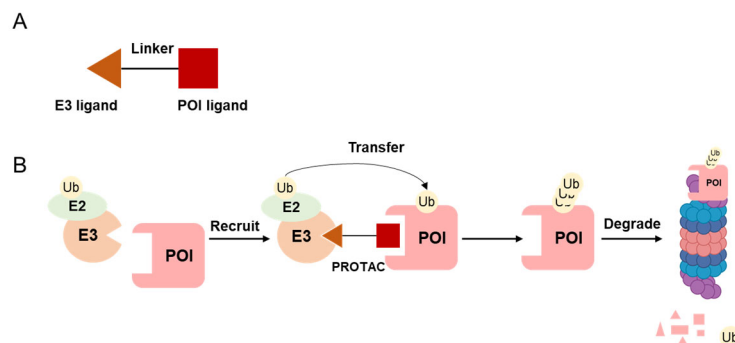
Figure 1.2 PROTACs and their mode of action. (A). The general structure of a PROTAC. (B). The mode of action of a PROTAC.

1.2.1.1   E3 ligases and ligands for PROTACs

There are more than 600 E3s encoded in the human genome. However, only a limited number of E3 ligases have been successfully applied to PROTAC design (Figure 1.3). Initially, nutlin-3a[14] was chosen as the ligand to recruit mouse double minute 2 homologue (MDM2) to design a PROTAC targeting androgen receptor (AR)[15]. However, the degradation efficiency of MDM2-based PROTACs is relatively low: the degradation of AR could only be observed at 10 μM. Following that, bestatin was also selected as the ligand to hijack the E3 ligase cellular inhibitor of apoptosis (cIAP)[16] and a series of PROTACs targeting the estrogen receptor (ER)[17], cellular retinoic acid binding protein 1 and 2 (CRABP1 and 2)[18] and transforming acidic coiled-coil containing protein 3 (TACC3)[19] were designed. Unfortunately, the poor selectivity of bestatin and the self-degradation of the cIAP limit the further application of cIAP1-based PROTACs.

Nowadays, cereblon (CRBN) and von Hippel-Lindau tumor suppressor protein (VHL) are the most commonly used E3 ligases in the design of PROTACs. In 2010, Hiroshi Handa's group found that the primary target of series of immunomodulatory imide drugs (IMiDs) including lenalidomide, pomalidomide and thalidomide is CRBN, a member of the CRL4^{CRBN} E3 ubiquitin ligase class[20]. Later in 2015, the Bradner lab reported the CRBN-based PROTAC, **dBET1** and realized the degradation of bromodomain-containing protein 4 (BRD4)[21]. In 2012, the Crews group published the first class of small molecule inhibitors targeting VHL with high affinity and high selectivity[22]. Later in 2015, they also reported one of the first series of VHL-based PROTACs targeting estrogen related receptor (ERRα) and serine/threonine kinase (RIPK2)[23]. Until now, numerous proteins are successfully targeted for degradation by the two E3 ligases and two PROTACs targeting the AR and ER, **ARV-110** (NCT03888612) and **ARV-471** (NCT04072952) have entered into phase II clinical trials (Figure 1.4)[12].

Moreover, more molecules have been found to target different E3 ligases and further designed into PROTACs (Figure 1.3) including ligands for RING finger protein 4[24] and 114[25] (RNF4 and 114), DDB1 and CUL4 associated factor 11[26], 15[27] and 16[28] (DCAF11, 15 and 16) and Kelch-like ECH-associated protein 1[29] (KEAP1). However, the majority of newly published PROTACs still recruit either CRBN or VHL ligases.
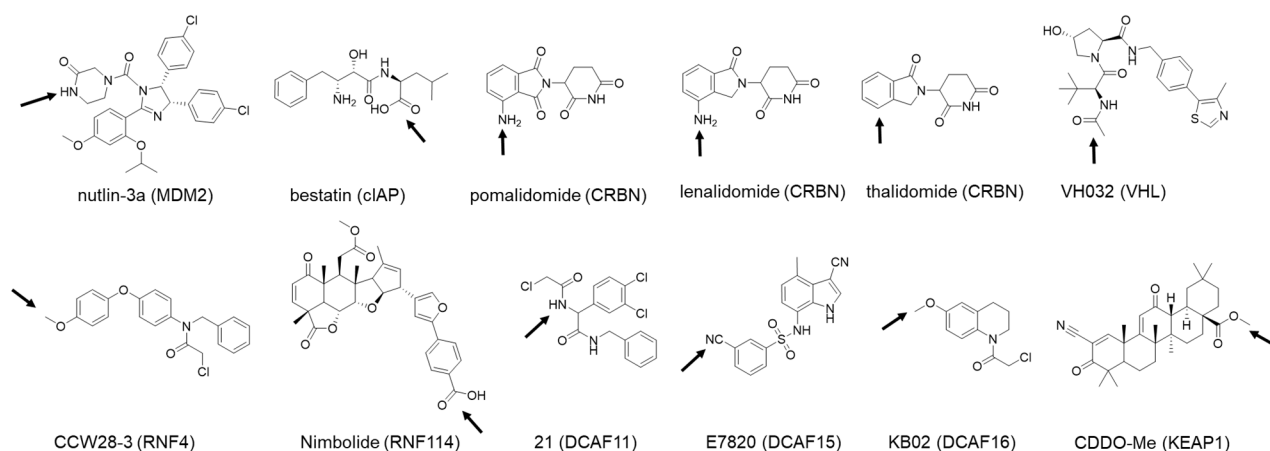
Figure 1.3 Structures of different E3 ligands. Black arrows indicate the position for further linker attachment.

Until now, over 50 proteins have been targeted for degradation by various PROTACs including kinases (CDK9 (Cyclin-dependent kinase 9)[30], BCR-Abl[31], RIPK2[23], c-Abl (Abelson tyrosine kinase)[31], Akt

(PKBalpha)[32]), nuclear receptors (AR[33], ER[17]), transcriptional regulators (BRD4[21], Sirt2 (Sirtuin 2)[34]) and regulatory proteins (CRABP1/2[18], ERRα (Estrogen-related receptor alpha)[23]).
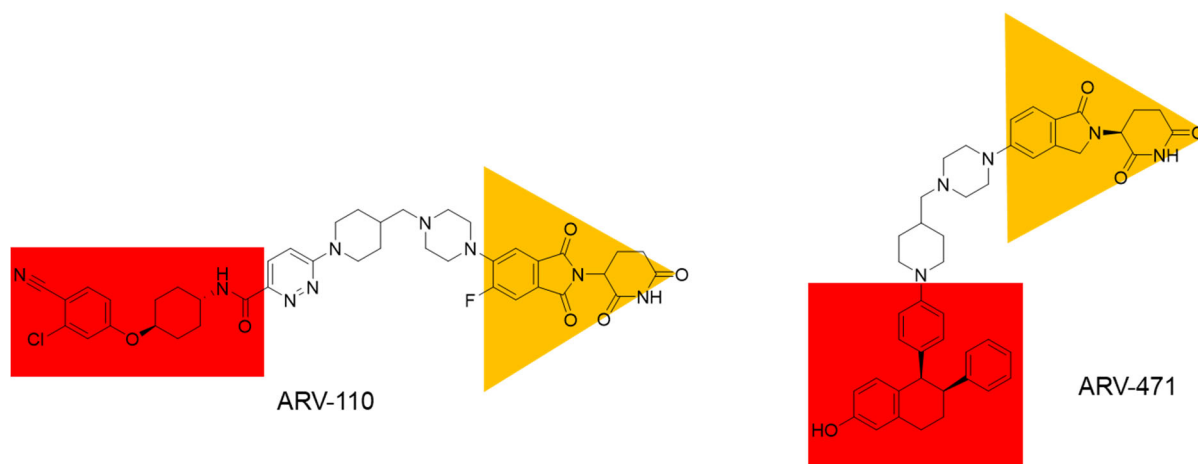
Figure 1.4 Two representative structures of PROTACs in clinical development (ARV-110 and ARV-471). Red rectangles indicate ligands for the POIs, while orange triangles indicate E3 ligase binding ligands.

1.2.1.2   Advantages and disadvantages of PROTACs

Compared with traditional inhibitors, the ability of PROTACs to degrade disease-related proteins offers numerous potential advantages.

First, PROTACs can remove non-catalytic functions of enzymes by degradation. Traditionally, inhibitors can only block the enzymatic activity of a protein but many proteins also exert non-catalytic activities. PROTACs completely degrade the whole protein, which means all functions related to this protein can be blocked. For example, PROTACs targeting Aurora kinase A (AURORA-A) successfully degrade this kinase and the degradation of AURORA-A results in an S-phase defect, a phenotype independent of kinase inhibition[35]. Moreover, AURORA-A degradation successfully induced cellular apoptosis in cancer cell lines. Similar removal of kinase-independent functions by PROTACs has also been applied to Focal adhesion kinase (FAK)[36] and signal transducer and activator of transcription 3 (STAT3)[37].

Second, a PROTAC from a non-selective inhibitor can turn into a selective degrader. Numerous examples demonstrate that turning promiscuous binders into PROTACs results in selective degradation profiles. For example, foretinib is a promiscuous kinase inhibitor that can bind to 133 different kinase at the concentration of 10 μM[38]. When converting foretinib into a VHL-based PROTAC, this molecule can still bind to 52 kinases but can only degrade 36 proteins. This is because the linker part within a PROTAC confers another layer of selectivity. The degradation ability of a given PROTAC is not simply dependent on the tight binding towards the target protein or the E3 ligase but heavily relies on the stability of the ternary complex formed by POI, PROTACs and E3 ligase, which is crucial for the degradation efficiency and selectivity of a PROTAC[39]. By connecting the POI ligand and E3 ligand, the linker part not only enables the POI and E3 ligase to form *de novo* protein-protein interactions (PPIs) from different positions but also engages in forming specific PPIs. Additionally, only the available PPIs mediated by a suitable linker can avoid steric clashes between the POI and E3 ligase, form the stable ternary complex, and further realize the ubiquitination and degradation of the POI. Therefore, although a promiscuous ligand can bind to various

4

proteins, only the proteins that can form the stable ternary complex can be selectively degraded by a PROTAC.

However, there are also disadvantages of PROTACs. PROTACs as bifunctional molecules exhibit quite high molecular weight, which may result in poor cell permeability and stability. In addition, the synthesis of PROTACs is still difficult. As previously mentioned, the ternary complex determines the degradation efficiency of a PROTAC. However, due to the lack of structural information on ternary complex, the design of PROTACs is still an empirical process.

1.2.2 Hydrophobic tagging (HyT)

Under native conditions, the hydrophobic side chains of a protein prefer to be buried inside the core for protein folding and stabilization. Therefore, the exposure of the hydrophobic residues is regarded as a signal of a misfolded protein and the misfolded protein will be further degraded by the quality control system mediated by the UPS[40]. Based on this principle, the HyT strategy was put forward to mimic the misfolded state of a protein by connecting a hydrophobic moiety to the surface of POI and this protein will be further degraded by the UPS[41]. Similarly, HyT consists of three components, a POI binding ligand, a hydrophobic group to mimic the denaturing state of a protein and a spacer connecting these two parts (Figure 1.5 A). The most commonly used hydrophobic groups are adamantane[41] and Boc-protected arginine (Boc$_3$Arg)[42] and this strategy has been successfully applied to the degradation of protein kinases (receptor tyrosine-protein kinase erbB-3, ERBB3)[43], nuclear receptors (AR, Figure 1.5 B)[44] and protein associated with Alzheimer's disease, Tau[45].
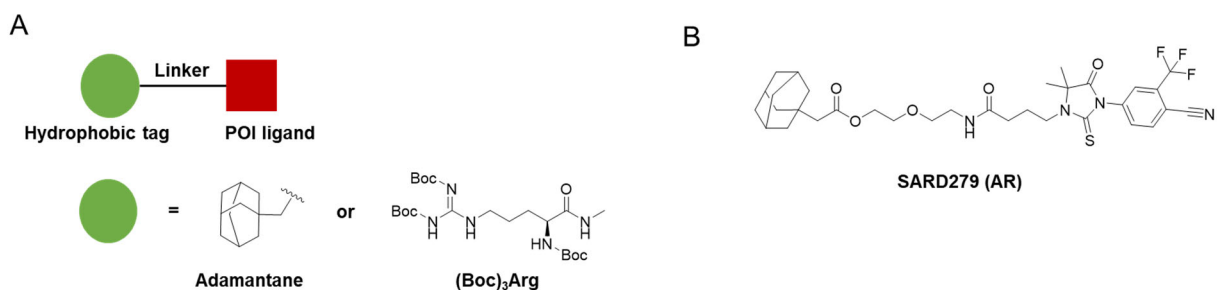
Figure 1.5 Hydrophobic tagging (HyT). (A). The structure of a HyT and the representative hydrophobic moieties. (B). A representative example of AR-targeted HyT, **SARD279**.

1.2.3 Lysosome Targeting Chimeras (LYTACs)

Lysosomal degradation can also be applied to TPD and lead to the degradation of extracellular proteins, aggregated proteins and even organelles. The Bertozzi group designed the first generation of LYTACs by connecting antibodies of POIs with *N*-carboxyanhydride (NCA)-derived glycopolypeptides, a moiety binding to the cation-independent mannose 6-phosphate receptor (CI-MPR) which is a member of LTRs responsible for trafficking proteins to lysosomes (Figure 1.6 A). These molecules have successfully realized the degradation of epidermal growth factor receptor (EGFR) and transferring receptor-1 (CD71)[46]. To realize the tissue-specific degradation, the Bertozzi group designed the second generation of LYTACs by connecting the antibodies of POIs with triantenerrary *N*-acetylgalactosamine (tri-GalNAc), a motif targeting the asialoglycoprotein receptor (ASGPR), which is a liver-specific LTR. GalNAc-LYTACs successfully realize the cell-type specific degradation of human epidermal growth factor receptor 2 (HER2) and EGFR[47].

1.2.4 Autophagy-Targeting Chimeras (AUTACs)

Autophagy-Targeting Chimeras (AUTACs) recruit POIs to autophagosomes for degradation[11]. Hirokazu's group designed the first AUTAC by connecting fumagillol, a ligand targeting methionine aminopeptidase 2 (MetAP2) with *p*-fluorobenzylguanine (FBnG) which serves a tag for autophagic degradation and realized the degradation of MetAP2 via autophagy (Figure 1.6 B). AUTACs have been successfully applied to the degradation of cytoplasmic proteins including FK506-binding protein 12 (FKBP12), MetAP2 and BRD4. Interestingly, AUTAC also realize the degradation of organelles. By linking a guanine tag to 2-phenylindole-3-glyoxyamides, ligands of the translocator protein located on mitochondrial, this AUTAC leads to the removal of damaged mitochondria in trisomy 21-derived fibroblasts.
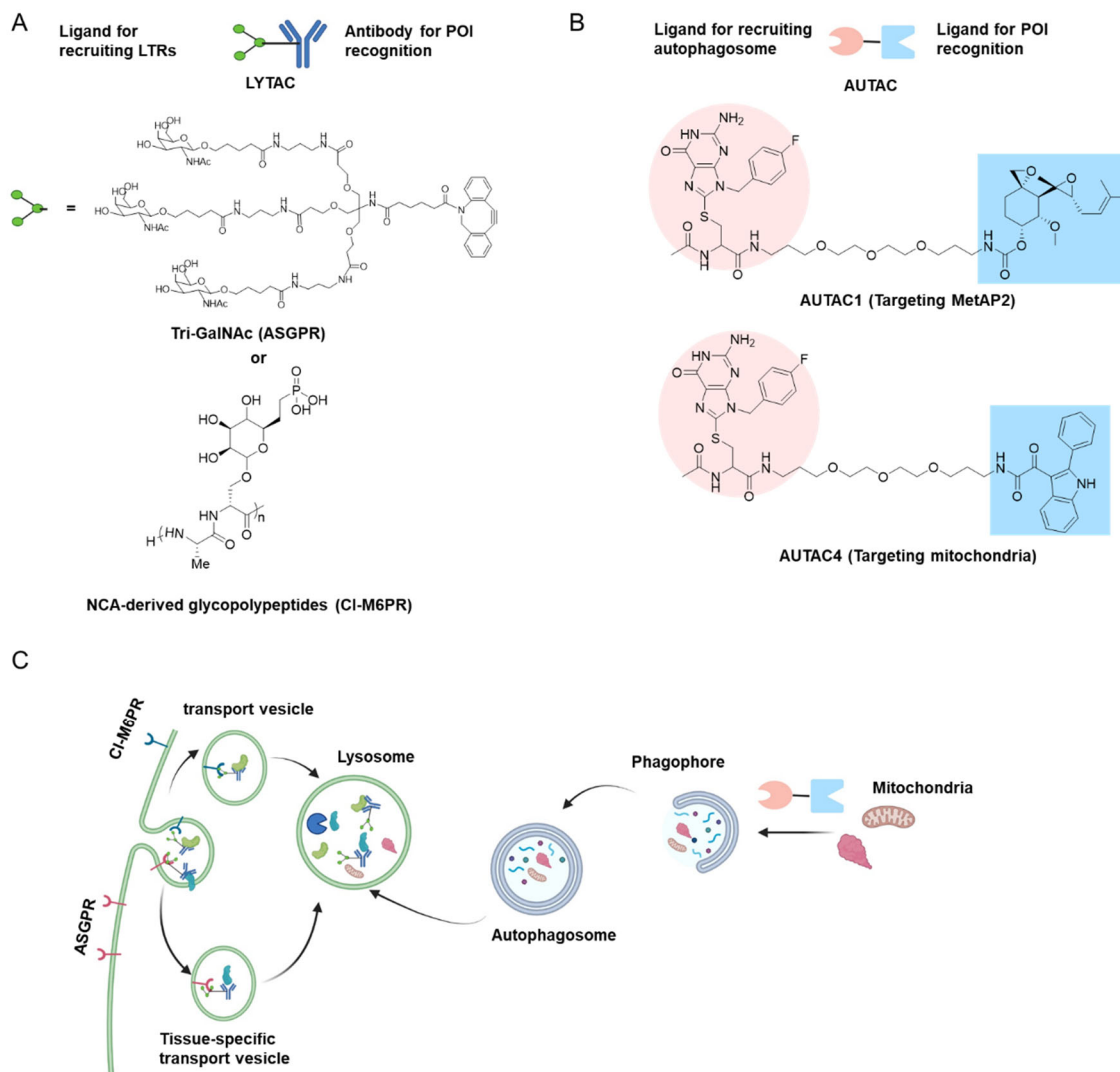
Figure 1.6 TPD strategies recruiting lysosomal-autophagy degradation pathway. (A). Structures of LYTACs. (B). Structures of AUTACs. (C). Mode of actions of LYTACs and AUTACs. Created with BioRender.com

1.3 Cholesterol and oxysterols

Oxysterols are derivatives of cholesterol generated by a series of enzymatic or radical oxidation reactions. Historically, their function as regulators of cholesterol metabolism has been well established[48]. However, recently, their roles in membrane fluidity regulation and several signaling pathways have been uncovered. They are also involved in a series of diseases such as cancer and neurodegenerative diseases[49]. Therefore, there is an urgent demand to understand their potential targets and their modes of action. This section will discuss their structures and (patho)physiological effects.

1.3.1 Derivatives of oxysterols.

Oxysterols are 27-carbon molecules made up of a steroid scaffold and a 6-methylheptan-2-yl side chain. The most common modifications of them involves the addition of hydroxyl, keto, hydroperoxy, epoxy, and carboxyl moieties and their structures are illustrated in Figure 1.7. Among them, the most abundant in human serum are 27-, 24(S)-, 7α-, and 4β-hydroxycholesterol (OHC)[50].
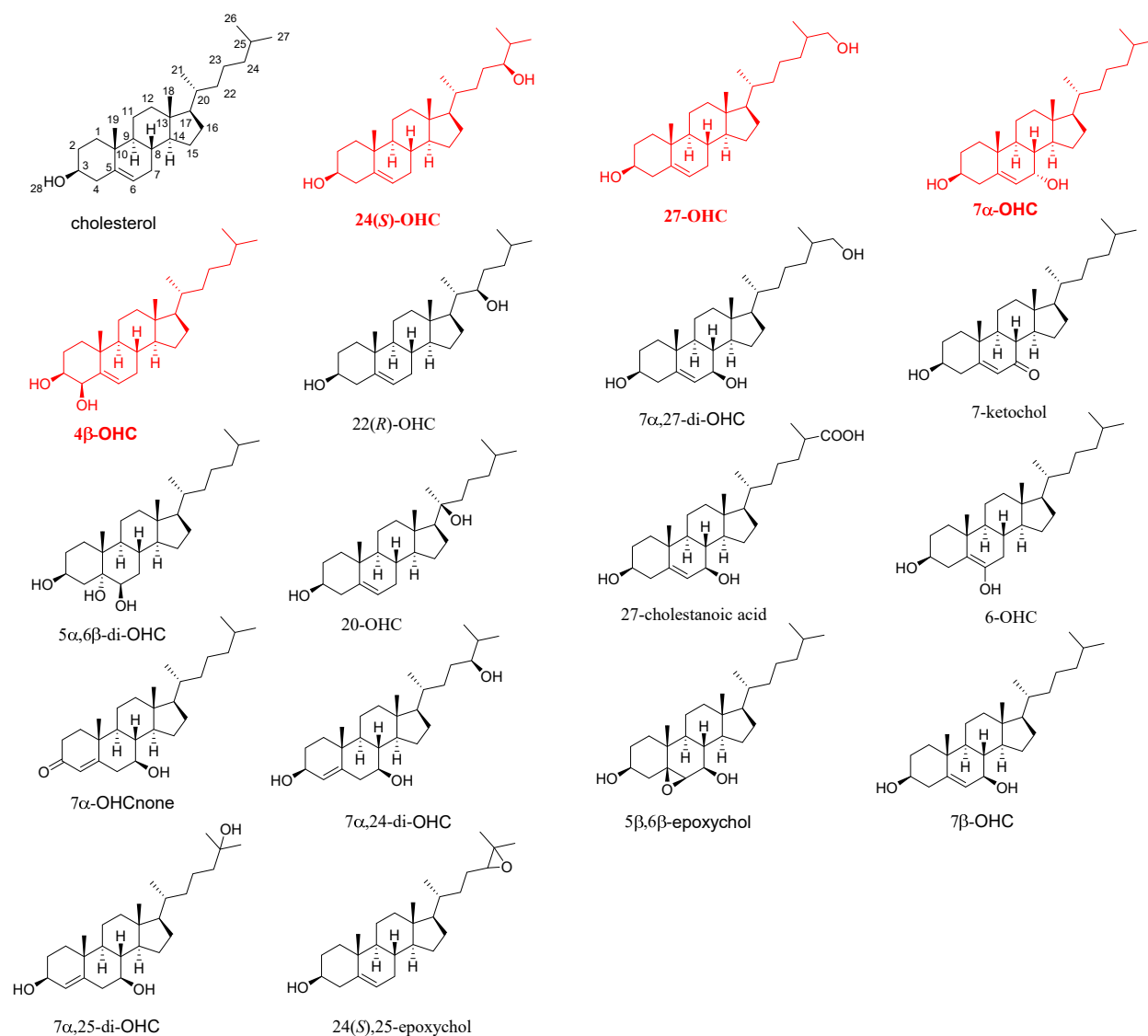
Figure 1.7 Cholesterol and its derivatives. (The most abundant oxysterols in human serum are labelled in red.)

1.3.2 Known targets of oxysterols

Initially, the role of oxysterols was believed to regulate cholesterol metabolism by transforming cholesterol into more polar molecules to favor its elimination. However, more functions of oxysterols have been discovered recently, including regulating gene expression, signaling development and apoptosis and numerous targets of oxysterols have also been identified[51]. The targets of oxysterols can be categorized into two main groups, receptors (nuclear receptors and cell membrane receptors) and transport proteins.

1.3.2.1 Nuclear receptors

Liver X receptor α and β (LXRα and β) are two transcription factors of the nuclear receptor family that can bind to a series of oxysterols[52] including 24(S), 25-epoxychol, 24(S)-OHC[53] and 4β-OHC[54]. LXRs exert their transcriptional activity by forming a heterodimer with retinoic X receptor (RXR). This heterodimer can be activated by oxysterols and further activate the expression of genes involved in lipid metabolism (ATP-binding cassette transporters, sterol regulatory element binding proteins, apolipoproteins)[55], inflammation (apoptosis inhibitor of macrophages) and apoptosis (LPS binding proteins)[56] (Figure 1.8).
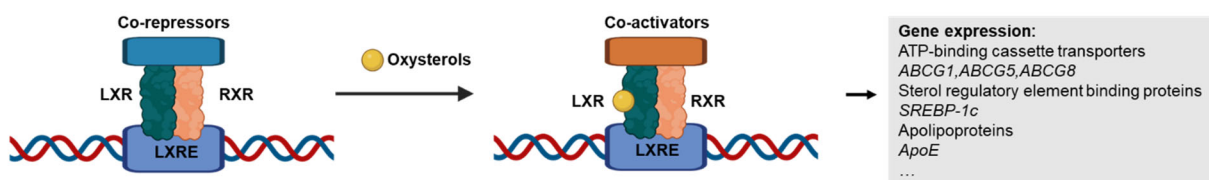
Figure 1.8 Mechanism of LXRs and their modulation by oxysterols. Created with BioRender.com

Retinoic acid receptor-related orphan receptors (RORs) could also be targeted by oxysterols[57]. These receptors are mainly expressed in the liver and responsible for the transcriptional control of lipid metabolism. Among them, RORα and RORγ can bind and further activate by 24(S)-OHC and 7-ketochol but no oxysterol is found to target RORβ yet[58].

The ERs are also the targets oxysterols. ERs involve in reproductive biology and also regulates a series of physiologic activities. 25-OHC is found to activate ERα by direct binding and further promote the expression of ER target genes including cathepsin D, cyclin D1 and progesterone receptor[59]. Additionally, 27-OHC can also bind to ER as a selective estrogen receptor modulator (SERM)[60].

1.3.2.2 Cell membrane receptors

Sterol regulatory element binding proteins (SREBPs) are membrane-bound receptors could be targeted by oxysterols[61]. These transcription factors are responsible for the biosynthesis and uptake of cholesterol and other fatty acids. These proteins will form the complex with SREBP cleavage activating protein (SCAP) to act as sensors for cholesterol levels in cells (Figure 1.9). When the level of cholesterol is low, the SREBP-SCAP complexes will transfer to the Golgi apparatus where SREBPs will be cleaved to generate a soluble NH_2-terminal fragment that can enter the nucleus to regulate the expression of relevant genes. When the cellular level of cholesterol is elevated, this complex will be retained in the membrane of the endoplasmic reticulum (ER) and interact with insulin-induced gene (INSIG) proteins, preventing SERBP trafficking to the Golgi. Both oxysterols and cholesterol regulate SREBP activation: cholesterol binds to SCAP and oxysterols, especially 25-OHC, could bind to INSIG[62].
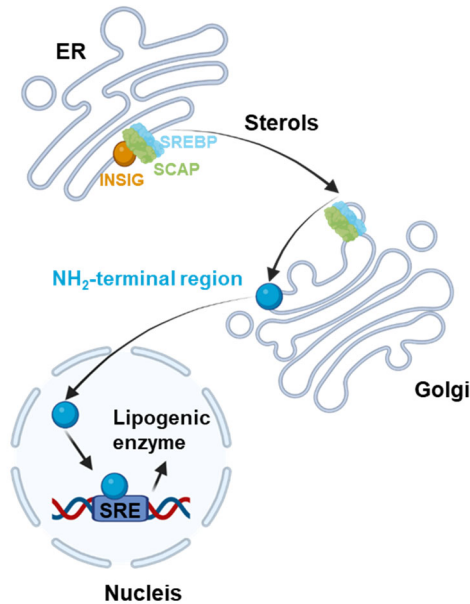
Figure 1.9 Mechanisms of SREBP sensing oxysterols. Created with BioRender.com

Oxysterols could also target the oncoprotein Smoothened (Smo)[63], a member of the Frizzled (Fz) family of seven-transmembrane domain (7TM) proteins responsible for the modulation of Hedgehog (Hh) signal pathway across the PM. The Hh signaling pathway is responsible for numerous aspects of metazoan embryonic development. The insufficient Hh activity will lead to birth defects while the excessive Hh activity is involved in many cancers[64]. When the Hh ligands (Sonic, Indian or Desert Hedgehogs) are absent, Smo is inhibited by the tumor suppressor membrane protein Patched (Ptch) in order to ensure the suppression of Hh signaling pathway (Figure 1.10 A). Upon Hh stimulation, oxysterols (20(*S*)-OHC, 25-OHC, 7-keto-25-OHC and 7-keto-27-OHC) and cholesterol could bind to the cysteine-rich domain (CRD) of SMO, allosterically activate SMO and trigger the downstream signal transduction events of Hh pathway and finally lead to the activation of target gene transcription (Figure 1.10 B).
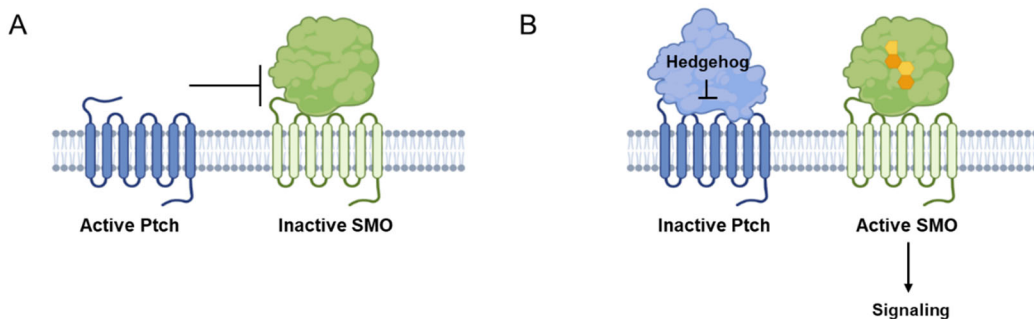
Figure 1.10 Cholesterol and oxysterols can activate the Hh signaling by binding to SMO. (A). When the Hh signaling is absent, Ptch can inihibt SMO. (B). Upon Hh stimulation, cholesterol and oxysterols can bind to SMO and activate the following signal pathways. Created with BioRender.com

C-X-C motif chemokine receptor 2 (CXCR2) is another receptor targeted by oxysterols[65]. It is also a receptor involved in immunity. 22(*R*)-OHC could target and activate CXCR2, resulting in the migration of neutrophil towards tumors and promoting tumor growth.

Epstein Barr virus-induced G protein–coupled receptor 2 (EBI2, also known as GPR183) is also a member of membrane receptors that could be targeted by oxysterols[66]. This receptor is responsible for the migration of follicular B-cell and the production of T-cell dependent antibodies. A series of oxysterols can bind to EBI2 and 7α, 25-di-OHC exhibits the highest affinity towards EBI2. The binding of 7α, 25-di-OHC to EBI2 could promote the migration of activated $CD44^+CD4^+$ T cells and therefore exhibits pronounced pro-inflammatory ability[67].

1.3.2.3 Sterol transport proteins (STPs)

Sterol transport proteins are a series of proteins responsible for vesicular or non-vesicular trafficking cholesterol between different organelles. There are three groups of STPs responsible for non-vesicular sterol trafficking, the oxysterol binding protein (OSBP) -related proteins (ORPs), the steroidogenic acute regulatory-related lipid transfer domain proteins (STARDs) and Aster family proteins[68] (Figure 1.12).

OSBP and ORPs are mainly responsible for trafficking cholesterol from ER to other organelles and maintaining cellular lipid homeostasis. This protein family consists of at least 16 members in humans[69] and most ORPs share a characteristic conserved OSBP-related domain (ORD), which can bind to one lipid molecule[70]. Additionally, except for ORP1S and ORP4S, other OSBP and ORPs usually contain a pleckstrin homology (PH) domain[71] and/or a diphenylalanines (FF) in an acidic tract (FFAT) motif[72] which is responsible for targeting different membranes. Moreover, ORP1L, a longer N-terminally extended version of ORP1, carries a three ankyrin (ANK) repeats, which are responsible for targeting proteins to late endosome. In addition, ORP5[73] and ORP8[74] bear an additional transmembrane domain (TM), which is responsible for targeting to the ER (Figure 1.11).
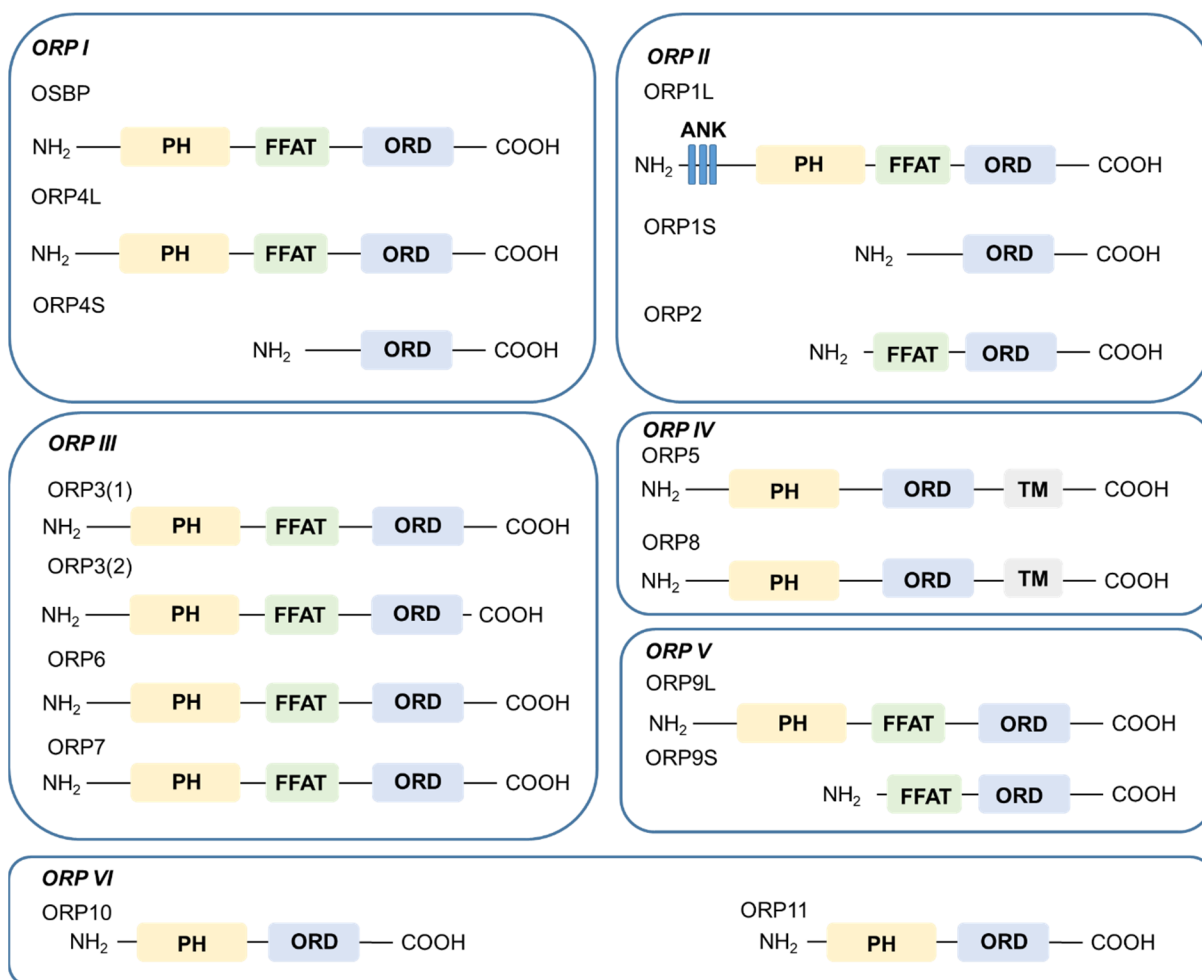
Figure 1.11 Structures of OSBP and ORPS. The ORP proteins can be classified into six groups based on their gene structures and amino acid homology. PH= pleckstrin homology domain, ORD= OSBP-related domain, FFAT= Diphenylalanines (FF) in an acidic tract, TM= transmembrane domain. Figure reproduced from reference[75].

Among all the OSBP and ORP members, OSBP, ORP1, ORP2, ORP4 and ORP9 have already been found to bind and deliver cholesterol or oxysterols[76–78]. OSBP is responsible for trafficking cholesterol from ER to *trans*-Golgi network (TGN)[79] and recently its role in delivering cholesterol from the ER to lysosome has been discovered[80]. ORP4 shares the highest structural similarity with OSBP and one of its isoform, ORP4L, is located at ER-Golgi contact sites and mediates cholesterol homeostasis[81]. ORP1 and ORP2 are two structurally related proteins. ORP1S, the short isoform of ORP1, regulates cholesterol trafficking from late endosome to PM and from PM to ER[82]. ORP1L mediates cholesterol trafficking from ER to late endosome but its function in sterol sensing and trafficking has not been fully elucidated[83]. ORP2 also delivers sterol from late endosome or ER to PM[84]. ORP9 also involves cholesterol transportation between ER and the TGN and the knockdown of this protein will lead to the loss of Golgi integrity[85]. However, oxysterols exhibits various binding affinities and selectivity towards these ORPs and leads to various functional outcomes. For example, OSBP, the prototypic protein of this family, can competitively bind to cholesterol and 25-OHC[86]. The binding of cholesterol towards OSBP will lead to the association of serine/threonine phosphatase (PP2A), PTPPBS family of tyrosine phosphatase (HePTP), OSBP and cholesterol. This

11

complex controls the extracellular signal–regulated kinase (ERK) signaling pathway and exert specific phosphatase activity towards phosphorylated ERK (pERK). However, the binding of 25-OHC towards OSBP will result in the dissociation of the complex among OSBP, PP2A and HePTP and the dephosphorylation of pERK will be inhibited.

STARD family proteins are also responsible for lipid transfer. This protein family contains 15 members and can be categorized into six groups based on their abilities for ligand binding and sequence similarities[79,87]. STARD1/3 and STARD4/5/6 are two subfamilies responsible for cholesterol and oxysterols binding. STARD2/7/10/11 is responsible for phospholipid/sphingolipid binding. The remaining three subfamilies lack identified ligands. STARD8/12/13 subfamily is putative to activate Rho GTPases[88]; STARD14/15 subfamily is type II acetyl-CoA thioesterases[89] and STARD9 is a novel mitotic kinesin[90].

As mentioned previously, STARD1/3/4/5/6 have been confirmed to bind and transport sterols[91–93]. STARD1 is located in mitochondrial membranes and responsible for trafficking cholesterol from the outer mitochondrial membrane (OMM) to inner mitochondrial membrane (IMM)[93]. STARD3 involves in delivering cholesterol from ER to the late endosome and mediates the contact between these two organelles[91]. STARD4 is responsible for trafficking cholesterol between PM and ER as well as PM to endocytic recycling compartment (ERC) and STARD5 is also putative to transport cholesterol between PM and ER[94]. STARD6 is putative to function as a testosterone transfer protein[95]. However, the detailed functions of STARD4/5/6 have not been discovered yet.

Aster proteins are a family of membrane proteins located at ER[96]. Recently, their functions in binding and transporting cholesterol have been elucidated. Cholesterol can target to their StART-like domain and be transferred from PM to ER. Additionally, the binding of cholesterol to Asters' StART domain can be blocked by oxysterols such as 25-OHC[97]. The detailed structure and functions of Aster proteins will be illustrated in Chapter 5.
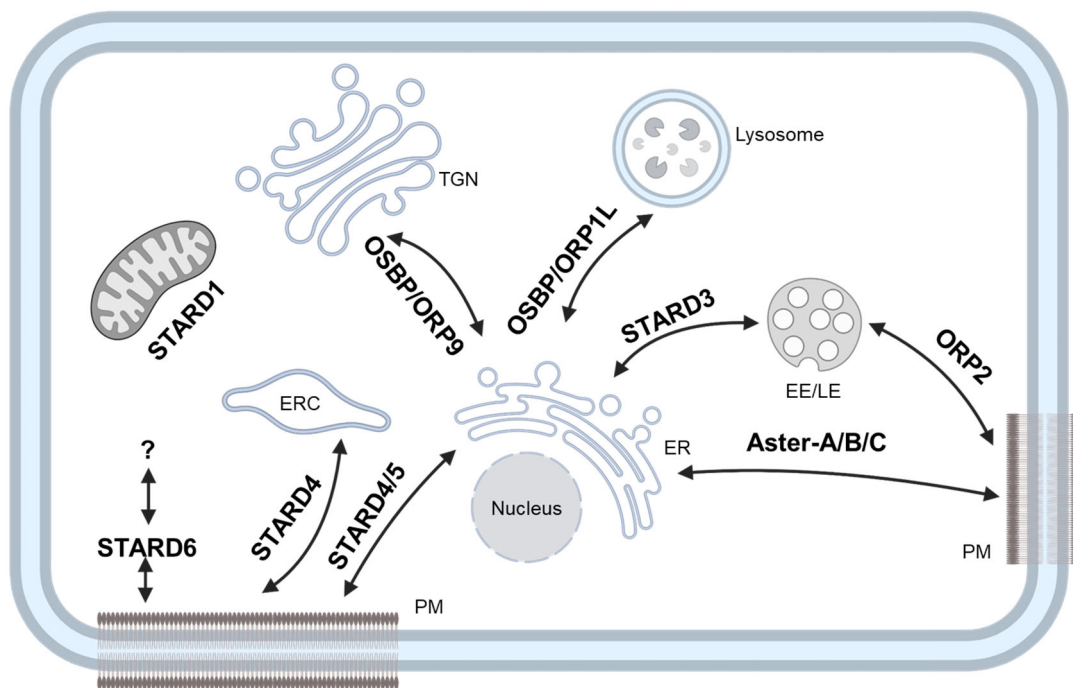
Figure 1.12 Cholesterol transportation in cells. Black arrows indicate the trafficking of cholesterols. ERC= endocytic recycling compartment, ER= endoplasmic reticulum, EE= early endosome, LE= late endosome, TGN=trans-Golgi network, PM=plasma membrane. Figure adapted from reference[75]. Created with BioRender.com

1.3.3   Selectivity of oxysterols.

As summarized before, the targets and functions of oxysterols are diverse. Interestingly, changes on the cholesterol scaffold including the position and absolute configuration hydroxyl groups will generate various oxysterols, which exhibit selectivity towards different targets and exert different functions. For example, the change of absolute configuration of hydroxyl groups generates 22(*S*) and 22(*R*)-OHC. These two oxysterols exhibit opposite effects when they bind to LXRs[98]. 22(*R*)-OHC is characterized as LXRs antagonist but 22(*S*)-OHC behaves as a LXRs agonist. Similarly, 22(*R*)-OHC could bind to and activate CXCR2 while 22(*S*)-OHC exhibits no binding ability towards it. Additionally, the position change of the hydroxyl group also affect the functions of oxysterols. For example, 20(*S*)-OHC could allosterically activate SMO while 22 (*R*)-OHC, 22(*S*)-OHC and 7β-OHC are inactive on the SMO receptor[99].

1.4   Target identification methods

Target identification is a crucial step in the field of chemical biology as it helps to uncover the mechanism of a specific molecule. However, target identification is also a major challenge in this field, especially for the targets of natural products, which are often isolated in small quantities and are challenging to functionalize. With the aid of chemical proteomics[100], different target identification methods have been put forward to discover the targets of a small molecule. In this section, various techniques for target identification will be introduced and discussed.

1.4.1 Activity-based protein profiling (ABPP)

Activity-Based Protein Profiling (ABPP) is one of the most commonly used methods for target identification[101]. This method uses a modified bioactive molecule as a probe and proteomics to label, qualify and quantify the potential targets of a bioactive molecule and uncover its mode(s) of action. According to the different interaction manners between the bioactive molecules and proteins, ABPP can be categorized into two classes: activity-based probes (ABPs) which are based on covalent bioactive molecules[102] and affinity-based probes (AfBPs), where usually an additional photo-affinity labeling (PAL) moiety will be incorporated to the bioactive molecule to form the covalent bond between the targets and this molecule[103].

ABPs consist of three parts: a reactive group for targeting and labelling proteins, a reporter group for enriching or visualizing the proteins and a linker connecting these two parts (Figure 1.13 A).

The reactive group could bind to the active site of the targets covalently. It can be classified into two groups. For ABPs, their molecular structures usually contain electrophilic groups including Michael-acceptors or alpha-halo ketones to form the covalent bond with the conserved active site nucleophiles[104–106]. However, some bioactive molecules, such as natural products, interact with their targets in a non-covalent way. In order to improve the binding affinity of these molecules towards their potential targets, an additional reactive group, PAL, is introduced into these molecules. PAL moieties could be activated by UV-light and then generate highly reactive species to covalently connect to a specific protein. Aliphatic[107], aromatic diazirines[108] as well as benzophenones[109] are the most universally applied PAL moieties (Figure 1.13 C).

The reporter group could facilitate the target identification by visualizing, isolating and quantifying of the labelled proteins. The commonly used reporter groups are fluorophores for visualizing the labeled targets in-gel analysis[110], or affinity tags (usually biotin)[111] for enriching the labelled targets using the biotin-streptavidin system and the enriched proteins can be analyzed by MS (Figure 1.13 B). However, the introduction of such bulky groups may hamper the binding ability of the probes towards the potential targets. Therefore, a series of smaller-sized ABPs were designed by incorporating the bio-orthogonal handle into the molecules[112,113]. These handle could be modified by bio-orthogonal reactions including click chemistry for further visualizing protein targets. Among them, the most common used pair of handles was an azide and a terminal alkyne (Figure 1.13 D)[114]. After copper (I)-catalyzed azide-alkyne cycloaddition (CuAAc), these handles could link to a fluorophore or a biotin group for further analysis.
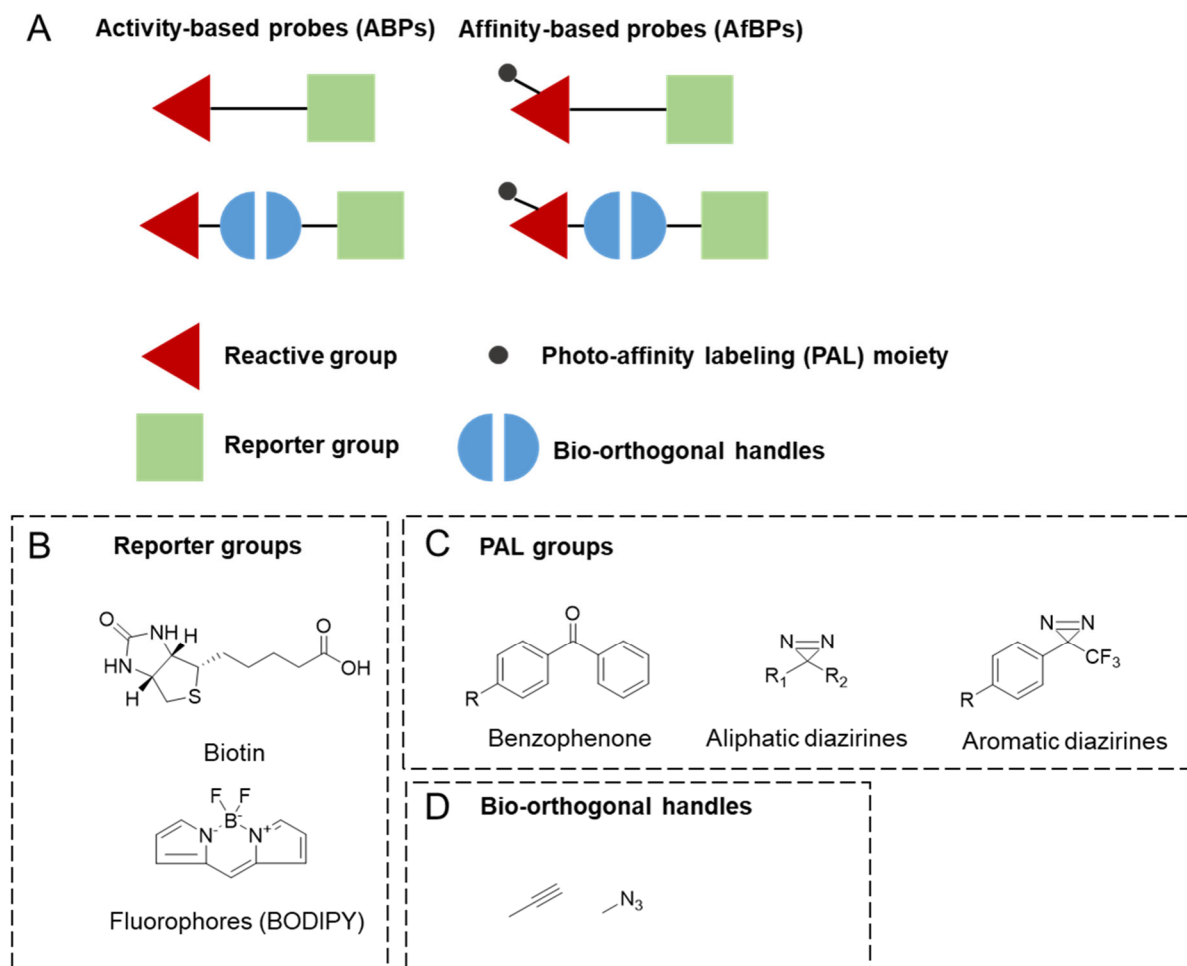
Figure 1.13 Components of ABPs and A*f*BPs. (A). The structures of ABPs and A*f*BPs. (B). Examples of representative reporter groups. (C). Examples of commonly used PAL groups. (D). Examples of commonly used bio-orthogonal handles. Figure reproduced from reference.[115]

Commonly, there are three methods to study ABPs-labelled target proteins: gel-based, proteomics-based and fluorescent microscopy imaging-based platforms (Figure 1.14). After treating the proteomes with ABPs, SDS-PAGE can be performed and the proteins may be detected by in-gel fluorescence[110]. Although it could be widely applied to the analysis of proteins in parallel, the accuracy of this method is not ideal as the bands in gels may contain several proteins and the low-abundance proteins may be omitted because the lack of resolving ability. With the aid of mass spectrometry-based proteomics, ABPP could be utilized in a complex biological system[116]. After treating the whole proteomes with biotinylated ABPs, the probe-target pair will be incubated with streptavidin beads to enrich them and finally, enriched proteins will be digested by trypsin and analyzed by LC-MS. By combining ABPP and chemical proteomics techniques, the potential targets can be identified in an unbiased way and the accuracy and resolvability of target identification are can also be improved. By incorporating a fluorescent group into the ABPs[117], ABPP could also be employed to visualize and localized the activity and distribution of specific proteins using fluorescence microscopy (Figure 1.14).
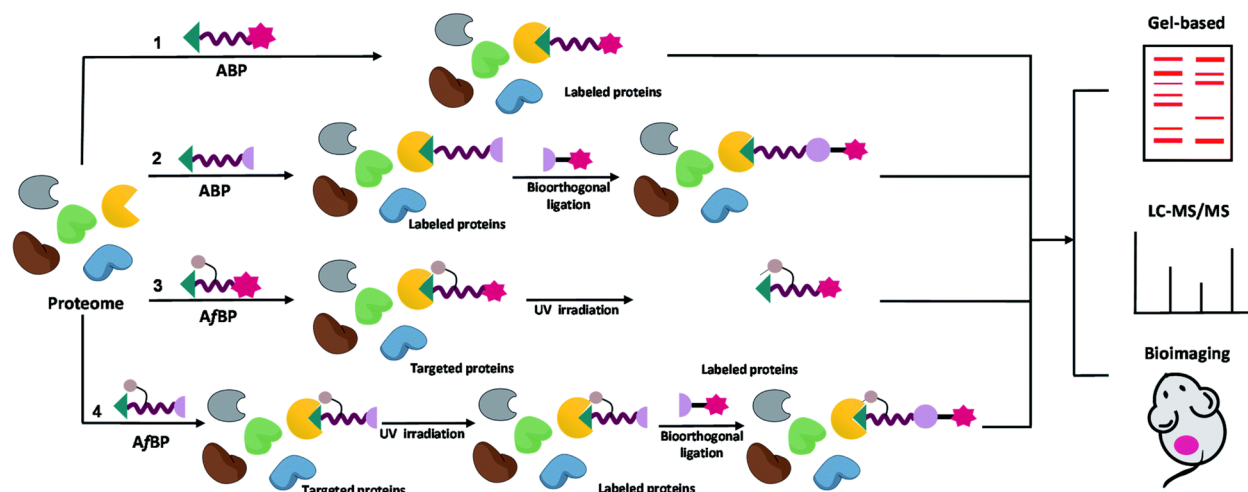
Figure 1.14 The workflow of ABPP for protein identification, enrichment and quantification. Figure reproduced from reference.[115]

1.4.2 Thermal proteome profiling (TPP)

Based on the principle that the binding of a bioactive molecule can affect the melting behavior of a specific protein, thermal proteome profiling (TPP) was developed enabling the target identification of a molecule in living cells without modifying its structure[118].

In this experiment, cells treated with DMSO or a bioactive molecule are heated to various temperatures (Figure 1.15). Then the precipitated proteins are removed by ultracentrifugation and the remaining proteins are analyzed and quantified using quantitative mass spectrometry. This method does not require any modification of the bioactive molecule and enables the assessment of drug binding to more than 7000 proteins in living cells simultaneously. However, there are still some disadvantages of this method. First, the thermostability of some proteins, especially large proteins, may not be affected by the binding of a bioactive molecule, which results in an insignificant shift, or no shift at all. Additionally, a certain number of membrane-bound proteins will still be missing, if the experiment is carried out in cell lysates.
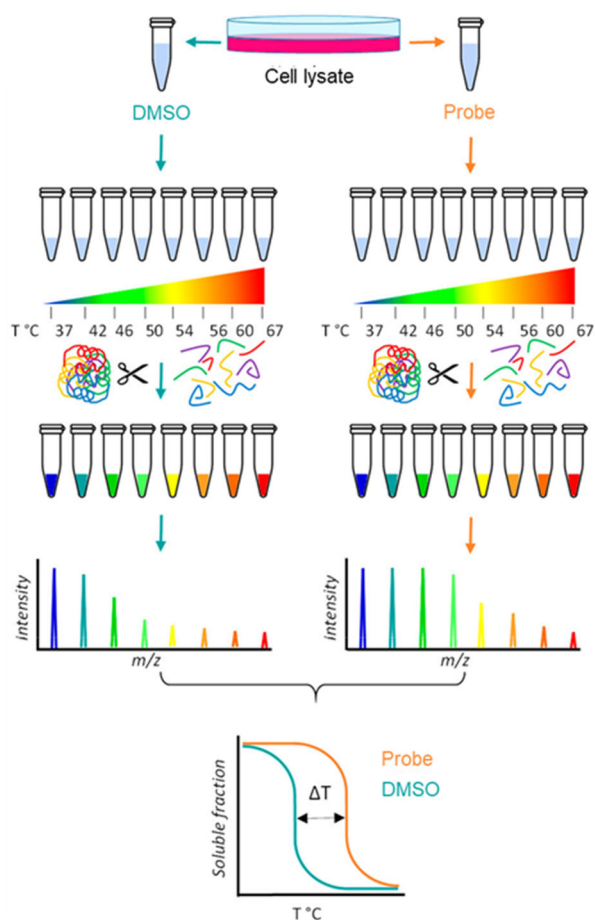
Figure 1.15 The workflow of the thermal proteome profiling (TPP). Figure reproduced from reference.[119]

1.4.3 Cell painting

Apart from target-based screening, phenotypic screening also provides a useful tool for target identification. Compared with target-based screening approaches, phenotypic screening analyzes the cell as a whole and provides a higher degree of physiological relevance. The cell painting assay as a representative phenotypic profiling assay obtains the biological information based on the cellular morphology[120]. This method 'paints' eight cellular components (nucleus, nucleoli, mitochondria, ER, Golgi, plasma membrane, actin cytoskeleton, cytosolic RNA) using different fluorescent dyes. After painting the cells with dyes, automated image analysis will be conducted and more than 1500 morphological features from each stained cell will be extracted to generate the morphological profiles[121] (Figure 1.16). By comparing the similarities and differences of the profiles between the known compounds and unannotated compounds, the target information of the unannotated compounds can be obtained.
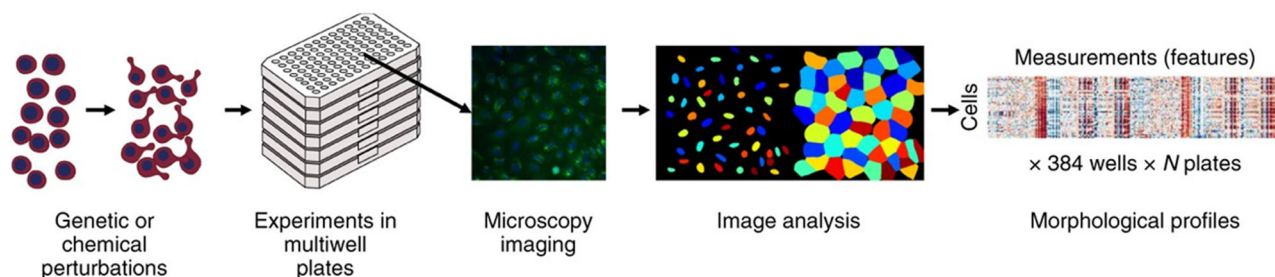
Figure 1.16 The workflow of cell painting. Figure reproduced from reference.[122]

1.4.4 Current methods to identify the targets of (oxy)sterols

1.4.4.1 Identifying the targets of cholesterol and 20(*S*)-hydroxycholesterol by ABPP

To study the targets of cholesterol globally, Cravatt's group applied cholesterol into a series of photo-reactive probes with different stereochemistry by adding a photo-reactive group at the B-ring of the cholesterol scaffold and replacing the side chain with an affinity handle (Figure 1.17). Among them, *trans*-sterol probe shares the most similarity to cholesterol in respect of stereochemistry and molecular topology. Further proteomics data demonstrate that *trans*-sterol probe can identify 265 proteins as the targets of cholesterol including receptors, channels and enzymes involved in protein glycosylation and vesicular transportation.
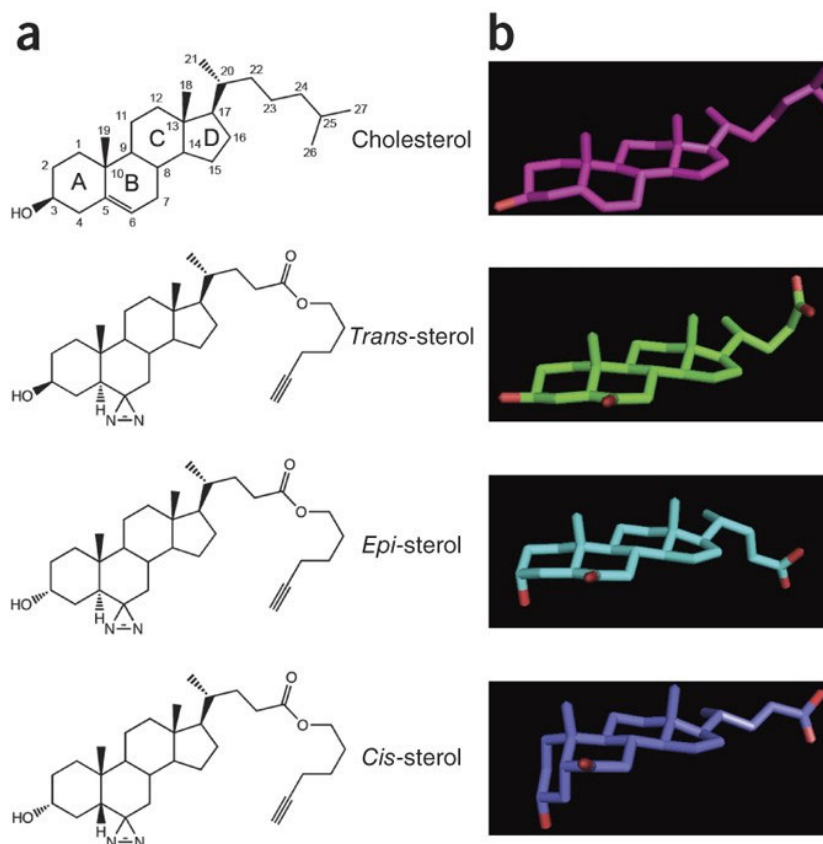
Figure 1.17 The chemical and X-ray structures and of cholesterol and cholesterol-based photo-affinity reactive probes. Figure reproduced from reference[123].

The Ondrus Lab applied the similar strategy to design 20(*S*)-hydroxycholesterol into a photo-affinity probe by introducing an affinity handle at the C-19 position and replacing the side chain with a photo-affinity group (Figure 1.18). This probe keeps the biological activity of 20(*S*)-OHC to active Smo and is further applied to detect the potential targets of 20(*S*)-OHC. Proteomics data indicated this probe can identify a series of known oxysterol targets. Additionally, competition assay further proves this probe can selectively bind to the protein transmembrane protein 97 (TMEM97) and western blotting demonstrates 20(*S*)-OHC can enhance the interaction between TMEM97 and Niemann-Pick disease intracellular cholesterol transporter 1 (NPC1) which indicates its functions in NPC1-mediated cholesterol homeostasis.
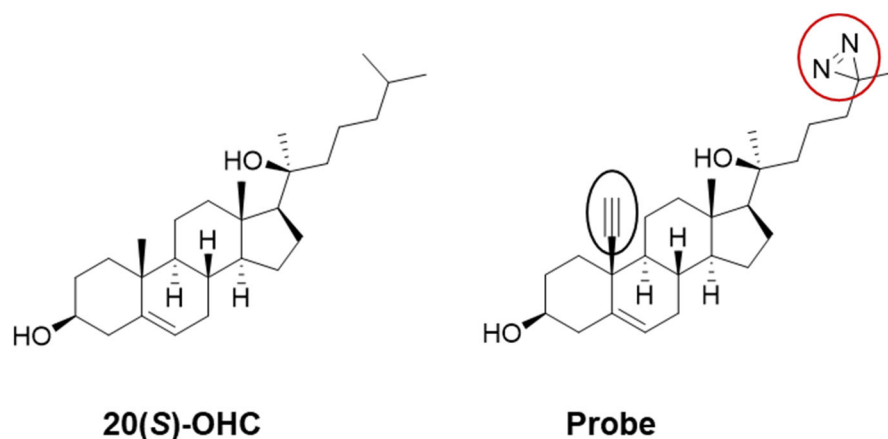
Figure 1.18 The structures of 20(*S*)-hydroxycholesterol (20(*S*)-OHC) and 20(*S*)-OHC-based ABPP probe. Red circle indicates the PAL group and black circle indicates the bio-orthogonal handle.

1.4.4.2   Identifying the targets of β-hydroxycholesterol (4β-HC), cholestane-3β,5α,6β-triol (CT), 7-ketocholesterol (7-KC) and 25-hydroxycholesterol (25-HC) by TPP

As mentioned in 1.3.3, the modification including the position and absolute configuration hydroxyl groups on the cholesterol will generate various oxysterols and these oxysterols will exert different functions. To study the target selectivity of different oxysterols, our group utilized TPP to explored the potential targets of three A- and B-ring oxidized sterols (4β-hydroxycholesterol (4β-HC), cholestane-3β,5α,6β-triol (CT), 7-ketocholesterol (7-KC)) and 25-hydroxycholesterol (25-HC) (Figure 1.19). Proteomics data indicate that each oxysterol exhibits distinct target profiles which further systematically and conclusively proves oxysterols exhibit selectivity towards different targets and exert various functions.
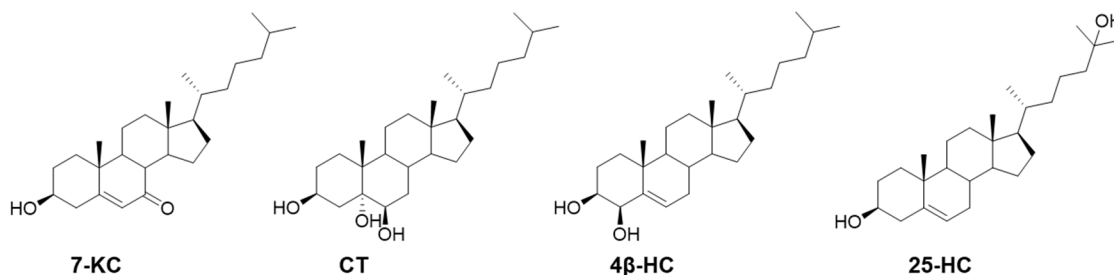
Figure 1.19 The structures of the tested oxysterols by TPP.

1.5   Aims and objectives

As previously mentioned, different target identification methods including ABPP and TPP have been applied to study the targets of (oxy)sterols. In this thesis, the toolbox for exploring the targets of (oxy)sterols will be further expanded.

The first aim of this thesis will be to establish the methodology of target identification via PROTACs by choosing cholesterol as an example and further study how modifications of the sterol scaffold affect its biological activity.

Currently, PROTACs only focus on the degradation of specific proteins using known target ligands. However, in principle, PROTACs could also be applied to target identification. By converting a ligand of unknown target(s) into a PROTAC, the proteins affected by the treatment of this PROTAC can be regarded as potential targets of this molecule and the affected proteins can be revealed by MS-based proteomics. This method has many benefits compared to the existing target identification method. First, this methodology could be conducted in living cells or tissues directly and degrade a series of proteins including both cytosolic and membrane proteins. Additionally, the identification and degradation of a protein can be achieved simultaneously, which means the functions of this protein and the effects of the removal of this protein can also be studied simultaneously without the need for genetic knockdown techniques including siRNA.

This methodology will first be applied to study the targets of (oxy)sterols. As previously mentioned, although several targets of (oxy)sterols have already been identified, the role of (oxy)sterols and their mode of action are still ambiguous. In addition, how the modifications on the cholesterol scaffold at various sites leads to target selectivity and biological activity remains unclear. Therefore, identifying the targets of (oxy)sterols and understanding their roles in biological processes are still outstanding questions in the field.

As a proof-of-concept, cholesterol will first be converted into cholesterol (CHO)-bearing PROTACs. The affected proteins under the treatment of CHO-PROTACs will be explored via MS-based proteomics and the functions of these altered proteins will further explored by a series of molecular and cellular biology experiments. Cholesterol is chosen as a proof-of-concept because its targets and functions have been extensively studied. Then, to study how the modifications on the cholesterol scaffold at various sites affects the functions of oxysterols, 4beta-hydroxycholesterol will also be applied into 4beta-hydroxycholesterol (4βHC)-bearing PROTACs and the degradation profiles of CHO-PROTACs and 4βHC-PROTACs will be compared and analyzed.

The second aim of this thesis will focus on designing Aster-A PROTACs.

Aster-A is a multi-functional protein responsible for both cholesterol and lipid trafficking and autophagy. There has been an established Aster-A inhibitor, **Autogramin-2**, which binds to the StART-like domain of Aster-A, can successfully block the transportation of cholesterol and also inhibit autophagy. However, the functions of other domains within Aster-A are poorly interrogated by small molecules. This project will convert **Autogramin-2** into Aster-A PROTACs to explore (1) whether Aster-A can be degraded (2) whether the autophagy inhibition effects of **Autogramin-2** can be improved by degradation (3) whether there are new phenotypes generated by the degradation of Aster-A.

2 Chapter 2 Synthesis and degradation profile of (oxy)sterol-bearing PROTACs

In this chapter, a series of (oxy)sterols-bearing PROTACs will be designed and the affected proteins will be analyzed by proteomics. To establish the methodology for applying PROTACs to the target identification challenge, I will choose cholesterol, a well-studied bioactive molecule to investigate whether known cholesterol targets can be degraded by cholesterol-bearing PROTACs and then I will apply this method to identify the potential targets of 4β-hydroxycholesterol.

2.1 Design and synthesis of (oxy)sterol-bearing PROTACs

2.1.1 Design and synthesis of cholesterol (CHO)-bearing PROTACs

To apply cholesterol into a PROTAC for further target identification, the CHO-PROTACs should bear these features: 1) a cholesterol scaffold acting as the warhead for potential cholesterol-binding proteins without disrupting critical binding interactions and providing a handle for further linker attachment. 2) an E3 ligand recruiting E3 ligase. 3) flexible linkers with various length, composition and hydrophobicity/hydrophilicity connecting cholesterol and E3 ligand. Based on these considerations, we designed four CHO-PROTACs with different alkyl, PEG or extended glycol linkers, namely **C1-C4** (structures shown in Figure 2.1 B). We kept the steroid scaffold because it is the main moiety for recognizing most cholesterol-binding proteins and replaced the alkyl side chain with a carboxylic acid group (compound **4**) for further linker attachment via the amide coupling. The similar strategy has also been applied to the design of a series of photo-labelling sterol probes (structures shown in Figure 2.1 A)[123]. For the E3 ligase recruiter, we chose pomalidomide, a commonly used E3 ligand in PROTAC design to hijack cereblon, which is a member of Cullin Ring Ubiquitin Ligases (CRLs). Additionally, the introduction of the amide-bond will increase the polarity of the whole molecule and further affect its cell permeability. To exclude the proteins affected by cholesterol scaffold and the additional amide bond, *N*-methylated analogue **C5** was designed as a negative control (Figure 2.1 B).
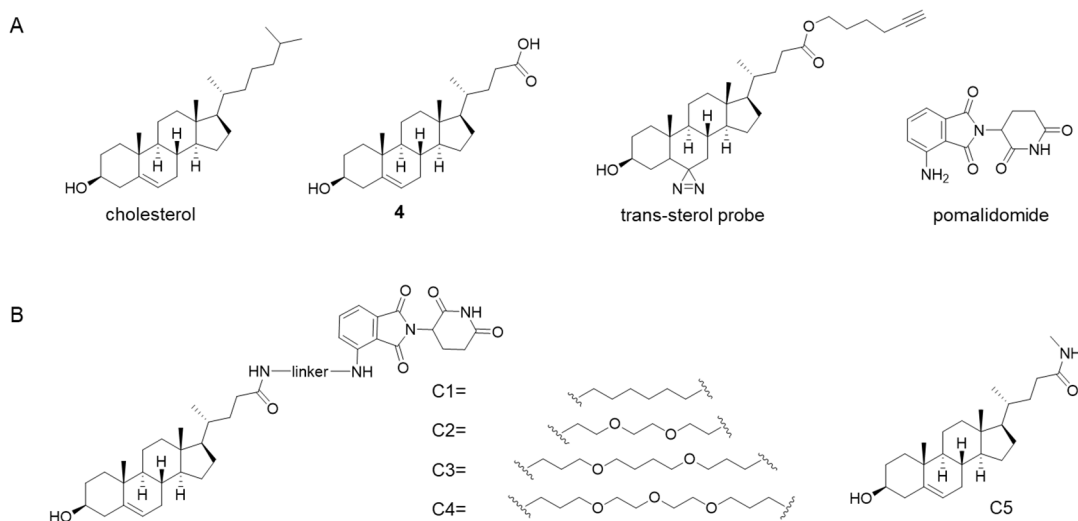
Figure 2.1 Rational design of cholesterol-bearing PROTACs. (A). Chemical structures of cholesterol, compound **4**, trans-sterol probe and the E3 ligand-pomalidomide. (B). Chemical structures of CHO-bearing PROTACs, **C1-C4** and negative control **C5**.

The synthesis of CHO- PROTACs was carried out as follows. As shown in Figure 2.2, for the cholesterol part, compound **4** was prepared via a four-step reactions reported previously[124]. Briefly, hyodeoxycholic acid was chosen as the starting material, followed by acid-catalyzed esterification of the carboxylic acid group, tosylation of hydroxy groups, KOAc mediated substitution and elimination and the final base-catalyzed ester hydrolysis.
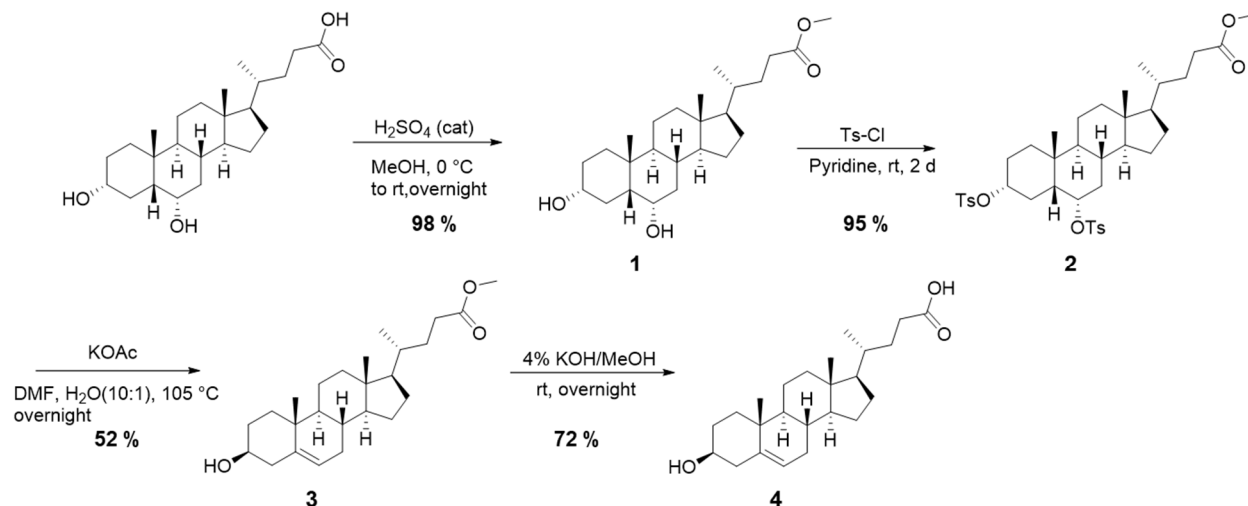
Figure 2.2 Synthesis routine of **4**.

For the pomalidomide part, as shown in Figure 2.3, F-pomalidomide (**5**) was obtained by acetylation between 3-fluorophthalic anhydride and 3-aminopiperidine-2,6-dione hydrochloride. For the linker part, various diamines were mono-protected by (Boc)$_2$O. Then, linkers bearing a free amine group at one end and Boc-protected amine group at the other end were attached to F-pomalidomide via SN$_{Ar}$ reactions to obtain compounds **11-13** and **15-18**.

Finally, the Boc protecting group was removed by TFA and the free amine was reacted with the carboxylic acid group from **4** via HATU-mediated amide coupling to obtain CHO-PROTACs **C1-C4**. Negative control **C5** was obtained by HATU-mediated amide-coupling between **4** and methylamine (Figure 2.4).
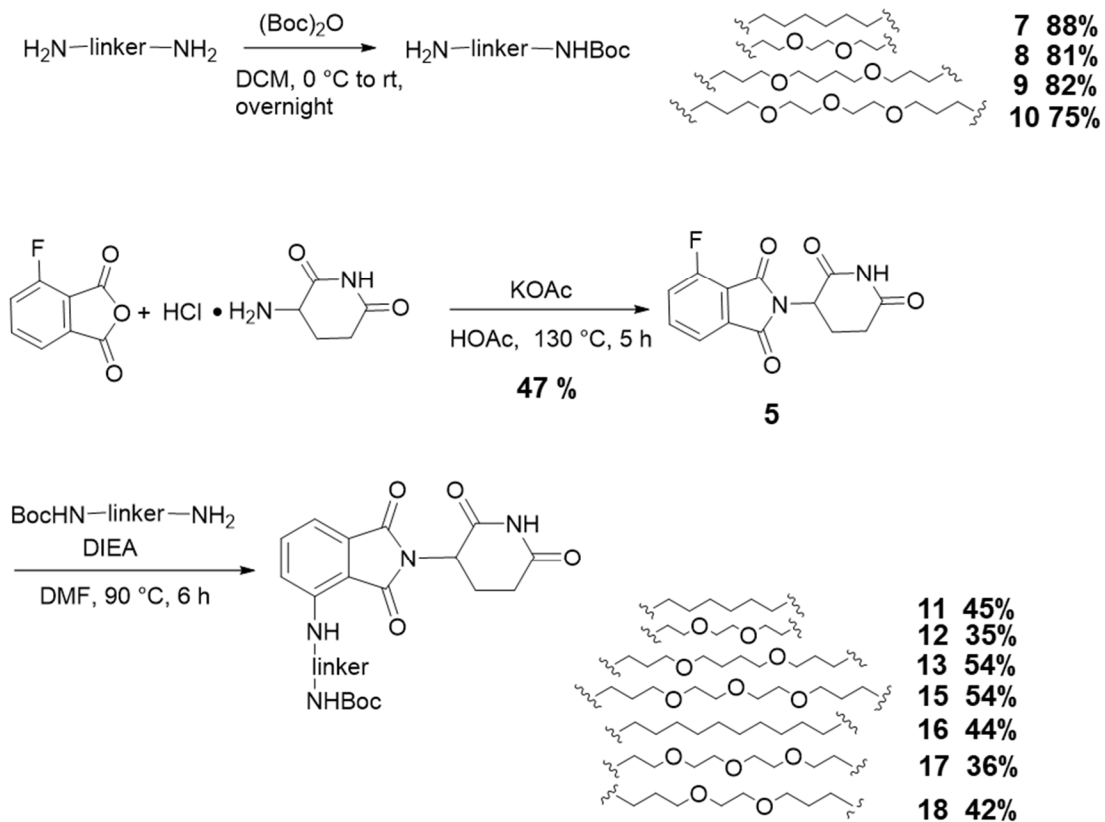
Figure 2.3 Synthesis routes of **11-13** and **15-18**.
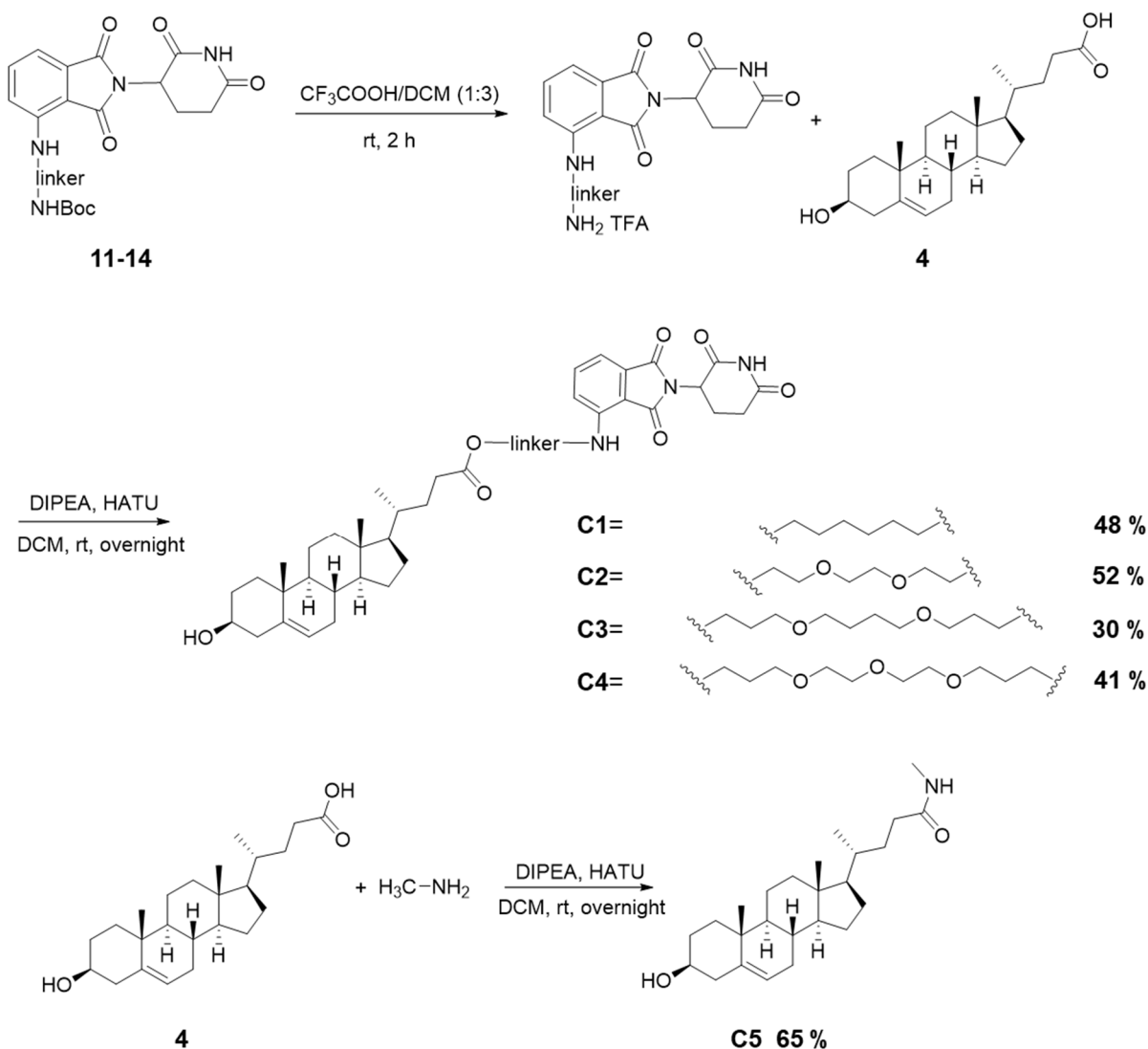
Figure 2.4 Synthesis routes to access **C1-C5**.

2.1.2 Design and synthesis of 4β-hydroxycholesterol (4βHC)-bearing PROTACs

The 4βHC-bearing PROTACs share the similar design rules with CHO-PROTACs by choosing the same E3 ligand and linkers. Similarly, *N*-methylated-4beta-hydroxycholesterol **H5** is also designed as a negative control. The synthesis of 4βHC-PROTACs was also similar with CHO-PROTACs and the only difference being that **4** was further oxidized by SeO₂ to obtain **18** for final amide coupling (Figure 2.5).
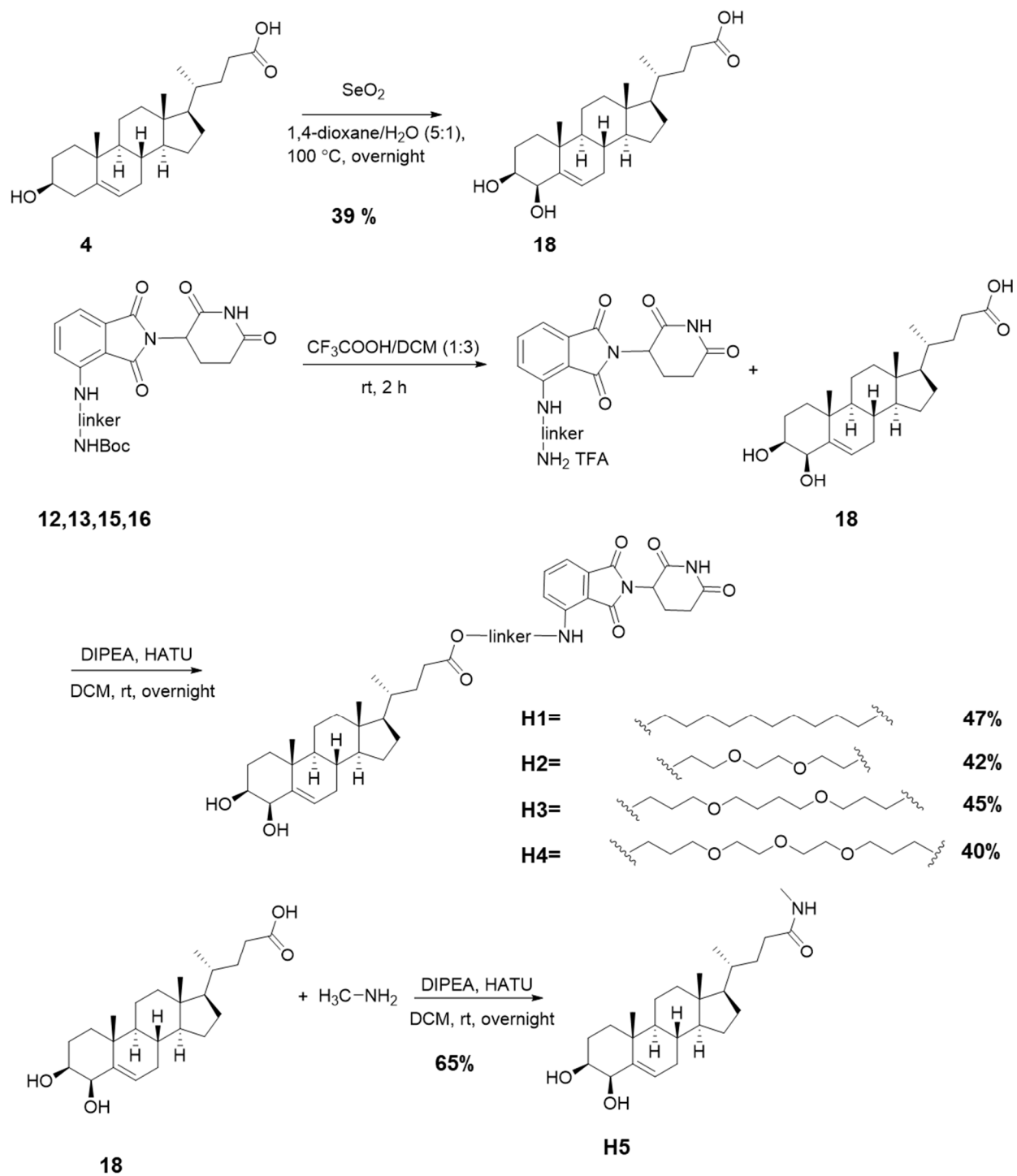
Figure 2.5 Synthesis route towards **H1-H5**.

2.2 Degradation profile of (oxy)sterols-bearing PROTACs via proteomics

2.2.1 Introduction to proteomics

The proteome is the entire set of proteins present in a genome, cell, tissue or organism and proteomics is the method to study proteomes comprehensively[125]. Initially, proteomics relied on two-dimensional gel electrophoresis to separate and identify proteins, which could only cover the highest-abundance proteins in a mixed sample. With developments in mass spectrometry, liquid chromatography-tandem mass spectrometry (LC-MS/MS) has now become the most comprehensive technique in proteomics with the advantages of high sensitivity, high throughput and deep proteome coverage. LC-MS/MS-based proteomics has been successfully applied to investigate the qualification (protein identification), quantification (protein synthesis and degradation) and modification (post-translational modifications) of proteins. It has also been applied to study the interactions between different proteins.

2.2.1.1 Work flow of "bottom-up" proteomics.

Normally, MS-based proteomics for protein identification and quantification can be categorized into two main approaches, namely the "bottom-up" and "top-down" methods, depending on whether digested peptides or intact proteins are analyzed[126]. At the moment, "bottom-up" method is still the most widely used approach. The workflow of "bottom-up" is depicted in Figure 2.6 A. First, proteins are extracted from a given cell line or tissue under native or denaturing conditions. After that, proteins are further digested into peptides by proteolytic enzymes. Then, digested peptides are separated by LC and ionized by the specific ion source. Finally, ionized peptides are sprayed into the mass spectrometer with two readouts, mass-to-charge (*m/z*) ratio values at MS level and the corresponding intensity of each ion at MS/MS level. With the aid of MaxQuant (MQ)[127] or Proteome Discoverer (PD), proteins can be identified by comparing the experimental MS/MS spectra of peptides with *in silico* MS/MS spectra from theoretical peptides and the relative changes in abundance of each peptide across different samples can be obtained by different quantification methods (Figure 2.6 B).
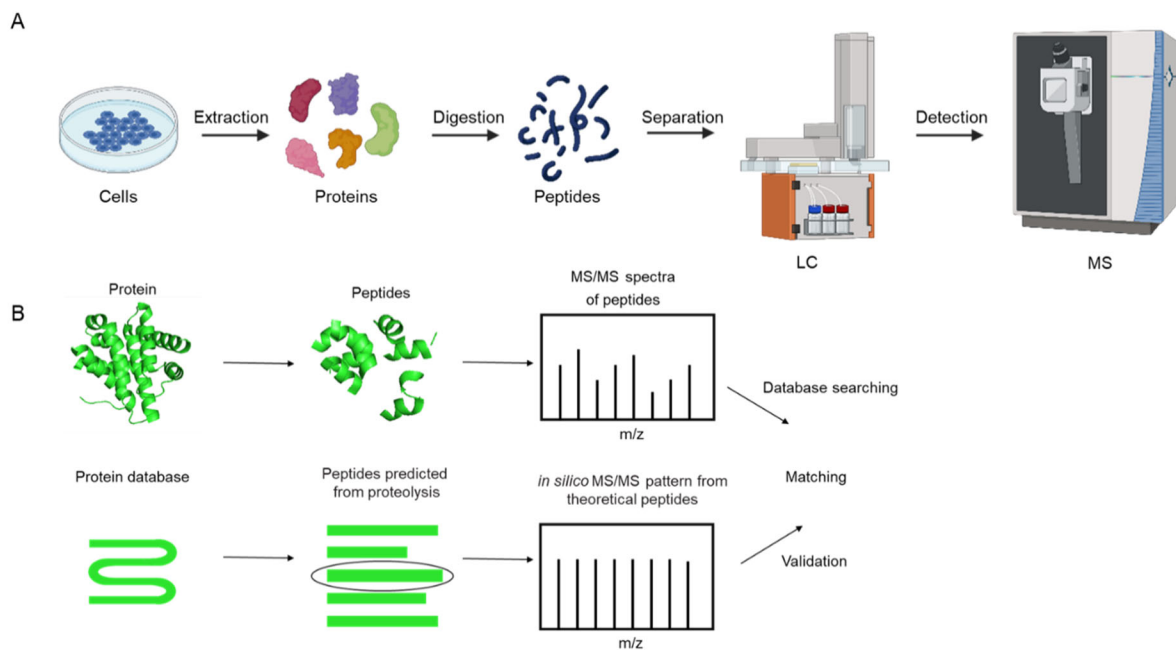
Figure 2.6 Schematic representation of the workflow of "bottom-up" proteomics. (A). Workflow of sample preparation and detection. (B). Workflow of data analysis.

2.2.1.2 Relative quantification methods in proteomics

The quantification of MS-based proteomics can be divided into absolute quantification method and relative quantification method. The absolute quantification method applies to determine the absolute expression level of a given protein under specific conditions while relative quantification method applies to compare changes in protein expression level under two or more different conditions. The relative quantification method can mainly be categorized into stable isotope labelling method[128] and label-free method[129]. Stable isotope labelling can also be further divided into metabolic labelling and chemical labelling.

Label-free quantification method

Label-free quantification is quite simple and cheap compared with other methods because special reagents or culturing conditions are not needed[130]. One of the label-free quantification method is to compare the number of peptides or spectra of a given protein, based on the empirical principle that the abundance of proteins correlates with the number of their digested peptides and the identified MS/MS spectra. Another method is to compare the intensities of peptide ions based on the fact that the signal intensity of each ion correlates with its concentration. The intensity of each ion can be measured by the height or area of its chromatographic peak at a given retention time (Figure 2.7).
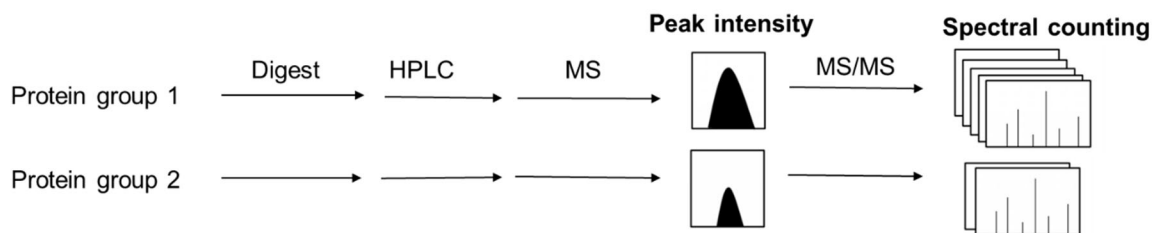
Figure 2.7 Workflow of label-free quantification method.

Metabolic labelling method

Metabolic labelling introduces stable isotope labels into the amino acids of all proteins during the cellular growth. Stable isotope labeling with amino acids (SILAC) is one of the most representative metabolic labelling methods[128]. Two group of cells are separately cultured in normal media with normal amino acids or heavy media containing amino acids (usually Lys or Arg) with heavy stable isotopes (^{13}C or ^{15}N). For example, normal media uses $^{12}C_6$-Lys while heavy media uses $^{13}C_6$-Lys (Figure 2.8). After a few passages, cells will incorporate the normal or heavy Lys into all their proteins, which means all peptides containing a heavy $^{13}C_6$-Lys are 6 Da heavier than the normal peptides. Then cells cultured in two media are mixed in equal proportions, followed by cell lysis and protein digestion. The mixed peptides are then analyzed by proteomics and the relative abundance of proteins under different conditions can be compared at MS level by quantifying the ion intensities of each peptide's isotope pairs. Apart from its application in cell culture, SILAC can also applied to organisms by feeding food with different isotope labels. Since SILAC induces stable isotope labels at an early stage, it can minimize variance during sample preparation process. However, the conditions can be tested by SILAC is limited because of the limited number of isotope labels for heavy amino acids.
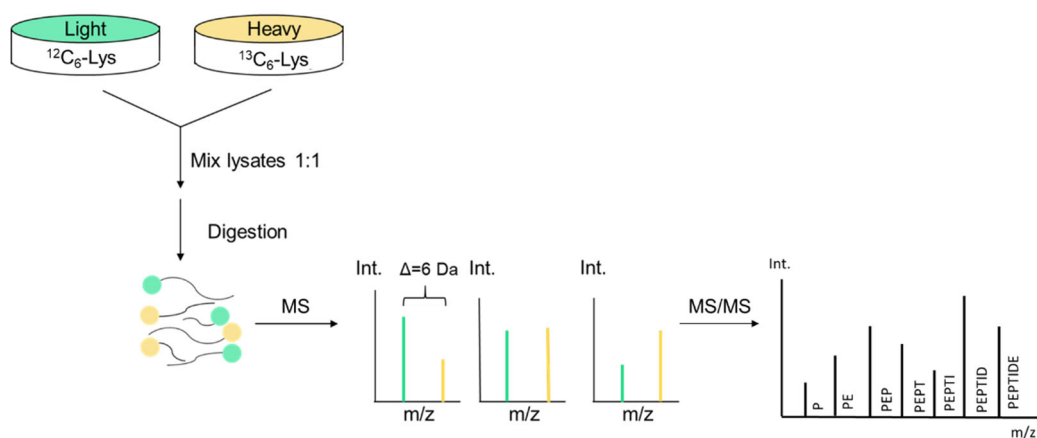
Figure 2.8 Workflow of SILAC.

Chemical labelling method

In contrast to metabolic labelling, chemical labelling introduces stable isotope labels after protein digestion and the quantification information can be obtained at MS/MS level[131]. One of the representative methods of chemical labelling is tandem mass tags (TMT). The structure of a TMT label comprises three components, a reporter group, a balanced group and a NHS-ester group reacting with NH_2 from the peptides (Figure 2.9 A). All the TMT labels share the identical chemical structure but have various isotopes substituted at different positions, which means the reporter group and the balance group have different molecular weights within each tag but the total molecular weights of the reporter group and the balanced group are identical among all TMT labels. At MS level, all the precursor ions share the same molecular weight but at MS/MS level, reported ions can be released during peptides fragmentation and the relative abundance of each peptides can be obtained by comparing the intensities of each reporter ion in low mass region (Figure 2.9 B). TMT labels enable compare more than 10 (10-plex or 11-plex TMT labels) even 16 (16-plex TMT labels) different conditions at one time and it more precise than label-free quantification method.
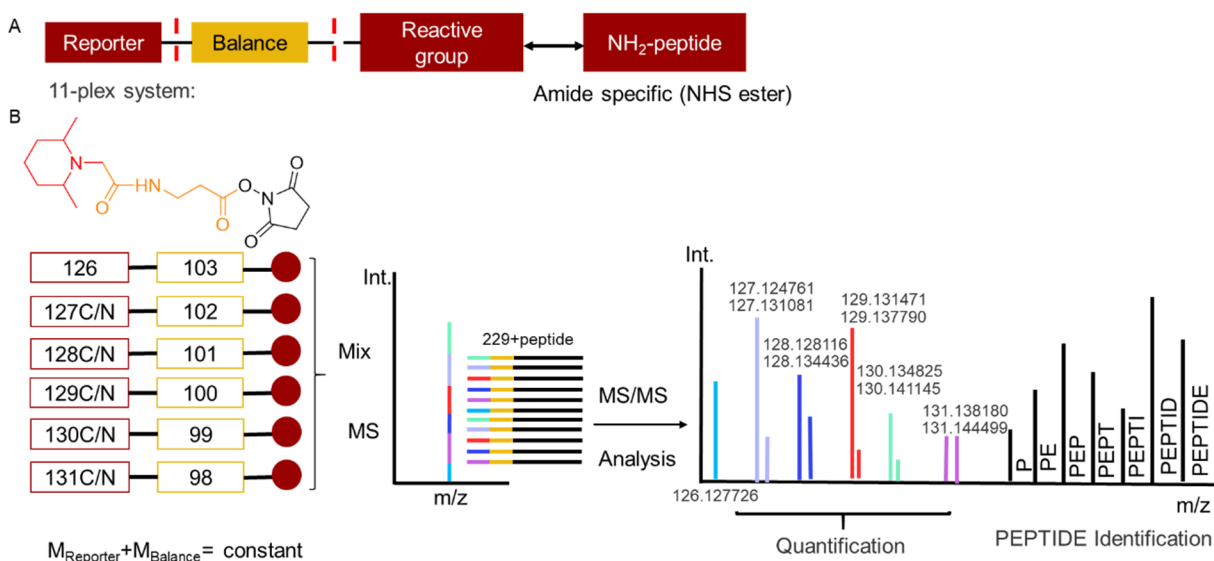
Figure 2.9 Workflow of 11-plex TMT.

2.2.1.3   Experimental design

Based on the mode of action of PROTACs, it is reasonable to assume that proteins reduced in abundance are degraded via the established mechanism, where the proteins bind to **C1-C4** and CRBN to form a ternary complex, leading to polyubiquitination and finally degradation by proteasome. However, we cannot exclude the possibility that the removal of the proteins are due to transcriptional regulation and adaptation. Therefore, after obtaining the proteomics data, further target validation assays are needed to confirm the degradation mechanism.

To realize efficient degradation but also avoid "Hook effect" induced by a higher concentration PROTAC treatment[132], the highest treatment concentration was set to 1 μM and a lower concentration, 100 nM was also set to observe whether the degradation is concentration-dependent or not. The treatment time of all PROTACs was set to 18 h to make sure a pronounced and sustained degradation could be observed. Since there are 10 conditions being tested at a time, we used 11-plex TMT labels to study the degradation profiles of all (oxy)sterols-PROTACs via quantitative proteomics and the detailed workflow is demonstrated as follows.

HeLa cells were treated with all PROTACs at two concentrations (100 nM and 1 μM), negative control **C5** or **H5** at 1 μM and DMSO for 18 h. After that, cells were lysed using PBS containing NP-40 because this detergent has been successfully applied to identifying membrane proteins where most known sterol targets are located[133]. After cell lysis, proteins were digested into peptides by trypsin and the peptides were labelled with the corresponding TMT labels. Then, the labelled peptides were combined and analyzed by LC-MS/MS. Finally, the results were analyzed by MQ or PD (Figure 2.10). All the experiments have been performed in three independent technical replicates to obtain reliable results.
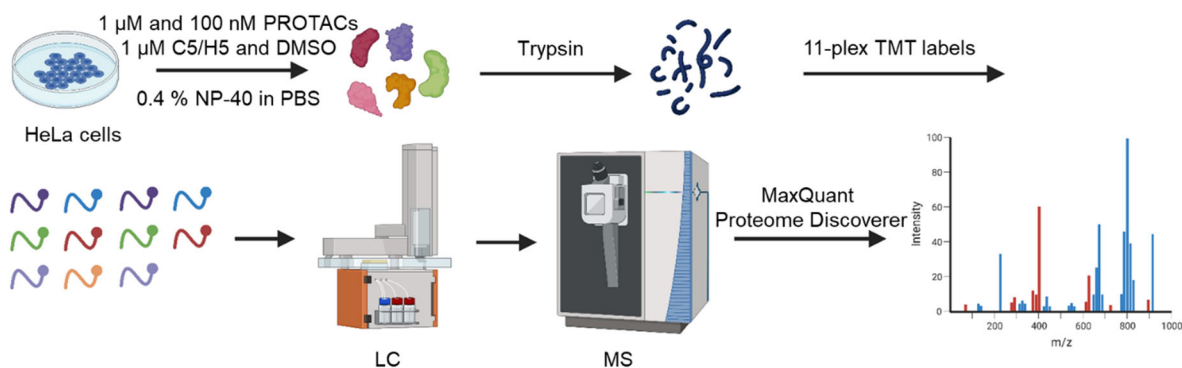
Figure 2.10 Workflow of TMT-based proteomics for target identification of (oxy)sterols.

2.2.2   Degradation profiles of cholesterol (CHO)-bearing PROTACs

In total, we identified 2454 proteins by expression proteomics. No protein is significantly affected under the treatment of 100 nM **C1-C4** or 1 μM **C5** while 21 proteins in total are significantly altered-14 being degraded and 7 being up-regulated after treating with 1 μM **C1-C4**, by setting the threshold: fold change > 2 and p value < 0.05 (Figure 2.11). The detailed information of the affected proteins are exhibited in Table 2.1. Among all the altered proteins, five of them have also been identified as the targets of cholesterol by the photo-affinity cholesterol probes[123], which provides a first validation of this method for target

identification. Theoretically, PROTACs should function as protein degraders, which means protein levels should be decreased. Therefore, the up-regulated proteins are considered as the downstream effects of the degradation of specific proteins and will not be further analyzed.

Comparing the degradation profiles among **C1-C4** at 1 μM, they exhibit distinct degradation profiles (Figure 2.12 A). **C1**, the only alkyl linker containing PROTAC, can degrade most proteins (12 proteins in total) while PEG linker containing PROTACs, **C2-C4**, are more selective and less than 6 proteins in total can be removed by them. However, **C1-C4** also share some similarities (Figure 2.12 B): all of them can significantly degrade QSOX2, E9PAM4, and GOLIM4. For further target validation, we choose GOLIM4, a protein degraded by all CHO-PRTOACs and a known target of cholesterol, for further study (*please see chapter 3 for further details of individual targets and validation efforts*).
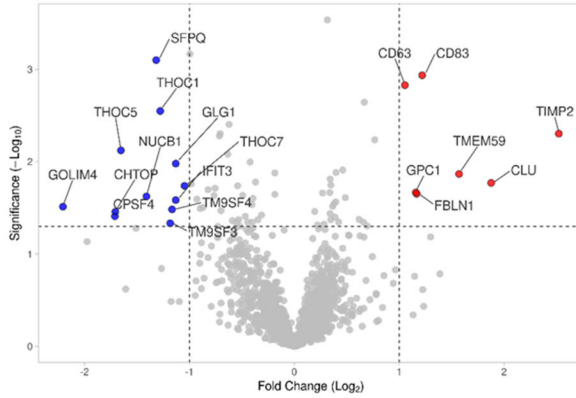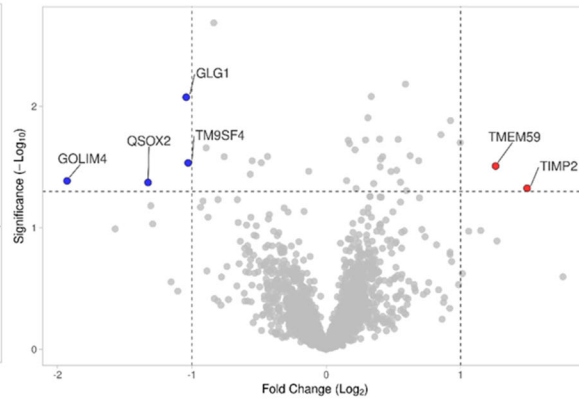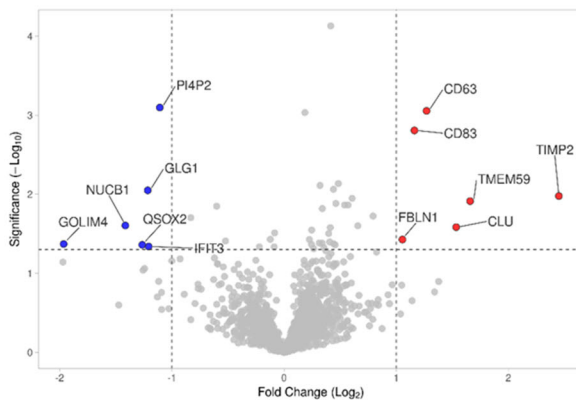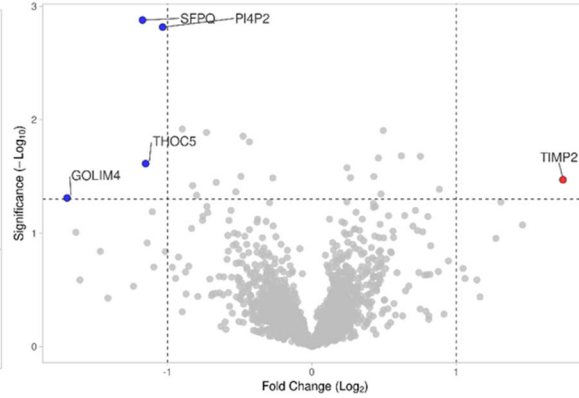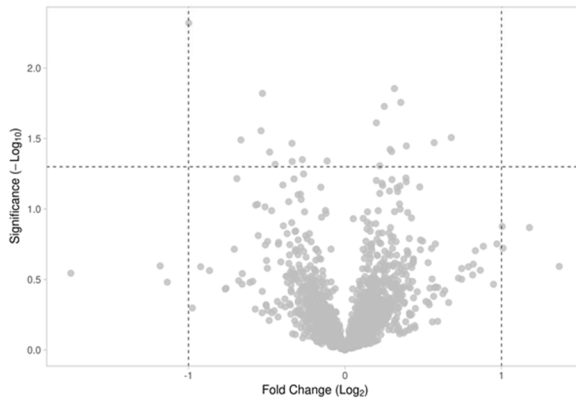
Figure 2.11 Volcano plots indicate the affected proteins under the treatment of CHO-PROTACs compared with DMSO-treated samples. HeLa cells were treated with 1 μM **C1**-**C5** for 18 h. Blue dots indicate the down-regulated proteins and red dots indicate the up-regulated proteins by setting the threshold: fold change > 2 (vertical dashed lines) and p value < 0.05 (horizontal dashed lines). The data were obtained from three independent technical replicates.

Figure 2.12 Degradation profiles of **C1-C4**. (A). Heat map of proteomics profiles of **C1-C4**. Each column represents different conditions and each row represents the altered protein. The color red or blue of each region represents this protein is up-regulated or down-regulated under this condition and the white region represents the protein level remains unchanged or the change is not statically significant. (B). Venn diagram comparing the degradation profiles of **C1-C4**.

Table 2.1 The altered proteins under the treatment of **C1-C4** at 1 μM after 18 h. The italicized *values* represent the observation is statically significant (p < 0.05) but the protein level change is not significant (fold change < 2), and the shading values represents the observation is not statically significant (p > 0.05). Star (*) represents this protein is also identified as the target of cholesterol by the photo-affinity probe.

Gene name	Uniprot ID	Protein name	C1-1 μM $\log_2$FC	C2-1 μM $\log_2$FC	C3-1 μM $\log_2$FC	C4-1 μM $\log_2$FC
GOLIM4*	F8W785	Golgi integral membrane protein 4	-2.21	-1.93	-1.96	-1.70
CPSF4	B7Z7B0	Cleavage and polyadenylation specificity factor subunit 4	-1.71	-0.78	-0.28	-0.21
CHTOP	X6R700	Chromatin target of PRMT1 protein	-1.71	-1.29	-1.26	-0.14
THOC5	F8WCP5	THO complex subunit 5 homolog	-1.65	-0.81	-0.53	-1.15
NUCB1	H7BZI1	Nucleobindin-1	-1.41	-0.57	-1.41	-0.45
SFPQ	H0Y9K7	Splicing factor, proline- and glutamine-rich	-1.32	*-0.89*	-0.45	-1.17
THOC1	A0A087WWS1	THO complex subunit 1	-1.28	-0.46	-0.20	*-0.73*
TM9SF3*	Q5TB53	Transmembrane 9 superfamily member 3	-1.18	-0.88	-1.00	-0.83
TM9SF4*	A0A0C4DFM1	Transmembrane 9 superfamily member 4	-1.17	-1.03	-0.93	-0.76
IFIT3	A0A7P0T855	Interferon-induced protein with tetratricopeptide repeats 3	-1.13	-0.94	-1.21	*-0.80*
GLG1	H3BM42	Golgi apparatus protein 1	-1.13	-1.04	-1.22	*-0.90*
THOC7	A0A5S6STF9	THO complex subunit 7 homolog	-1.05	-0.56	-0.37	-0.75
PI4K2	E9PAM4	Phosphatidylinositol 4-kinase type 2	-1.00	*-0.84*	-1.11	-1.03
QSOX2*	H0Y430	Sulfhydryl oxidase	*-0.89*	-1.33	-1.26	-1.11
CD63*	F8VNT9	CD63 antigen	1.05	1.00	1.27	0.74
GPC1	H7C410	Glypican-1	1.16	*0.92*	*0.79*	1.46
FBLN1	B1AHL2	Fibulin-1	1.17	0.61	1.05	0.70
CD83	A0A087WX61	CD83 antigen	1.22	*0.85*	1.16	*0.88*
TMEM59	Q5T6Z8	Transmembrane protein 59	1.57	1.26	1.66	1.28
CLU	H0YC35	Clusterin	1.88	1.06	1.53	1.31
TIMP2	K7EIX4	Metalloproteinase inhibitor 2	2.52	1.49	2.45	1.74

2.2.3 Degradation profile of 4β-hydroxycholesterol (4βHC)-bearing PROTACs

In total, we identified 4922 proteins by expression proteomics. Almost no changes of protein levels can be observed under the treatment of 100 nM of **H1-H4**, except CAMKMT, which can only be upregulated by 100 nM **H2**, and CCDC9, which can only be upregulated by 100 nM **H4**. Interestingly, THSD1 can be significantly degraded by the negative control, 1 μM **H5**. Although this protein can also be degraded by 1 μM **H1** or **H3**, the degradation is not statistical significant (p value > 0.05).

Since no protein is significantly regulated under the treatment of 1 μM **H2-H4** based on the current threshold: fold change > 2 and p value < 0.05, to study the degradation profiles of **H1-H4** more comprehensively, the threshold is relaxed to fold change > 1.4 and p value < 0.05 (Figure 2.13). The detailed information of the affected proteins are listed in Table 2.2 (**H1**) and Table 2.3 (**H2-H5**). Comparing the degradation profiles among **H1-H4** at 1 μM (Figure 2.14), they exhibit completely distinct degradation profiles. **H1**, the only alkyl linker containing PROTAC, can still degrade the most proteins (21 proteins in total). In contrast, **H2** can only degrade one protein-CSPG4 and **H3** can only degrade two proteins-DGCR8 and CSPG4. **H4** does not exhibit any protein removal ability under the current testing conditions. The different degradation profiles between **H1** and **H2-H4** may be attributed to their various cell permeability. The alkyl linker within **H1** may exhibit increased cell permeability and lead to more proteins being degraded. Therefore, we chose **H1** for further study. First, the degradation profiles of **C1** and **H1** were compared (Figure 2.15 A). **C1** and **H1** exhibit similar degradation profiles: 4 proteins (GOLIM4, GLG1, QSOX2 and TM9SF4) can be degraded by both of them and interestingly, these four proteins are all localized at Golgi apparatus. Second, the degraded proteins of **H1** also exhibit similarities. STRING functional enrichment analysis of the potential targets of **H1** exhibits an enrichment of the Golgi apparatus (14 of 21) and 4 of them are associated with hydrolase activity especially acting on glycosyl bonds (Figure 2.15 B).
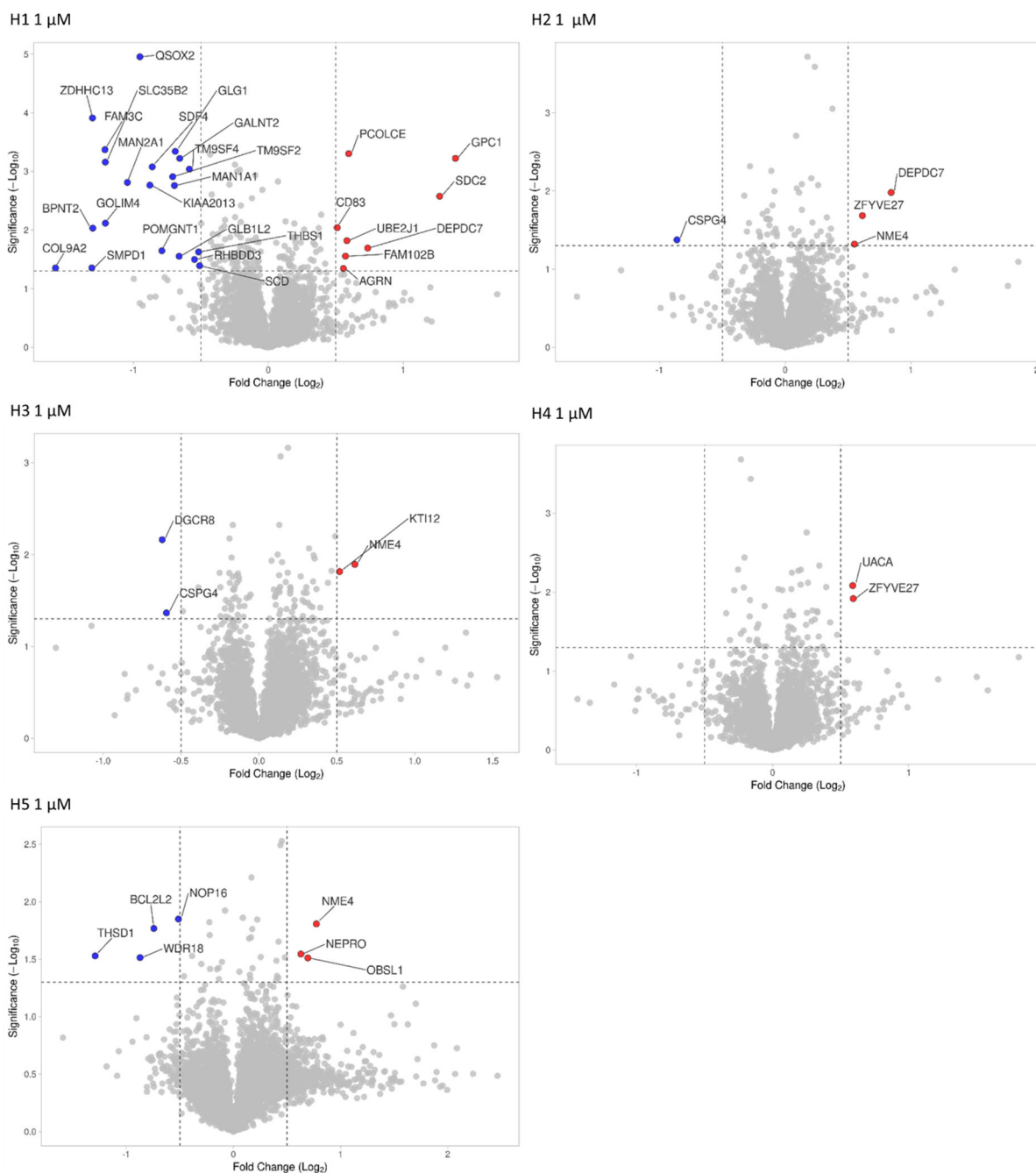
Figure 2.13 Volcano plots indicate the affected proteins under the treatment of 4βHC-PROTACs comapred with DMSO. HeLa cells were treated with 1 μM **H1-H5** for 18 h. Blue dots indicate the down-regulated proteins and red dots indicate the up-regulated proteins by setting the threshold: fold change > 1.4 (vertical dashed lines) and p value < 0.05 (horizontal dashed lines). The data were obtained from three independent technical replicates.
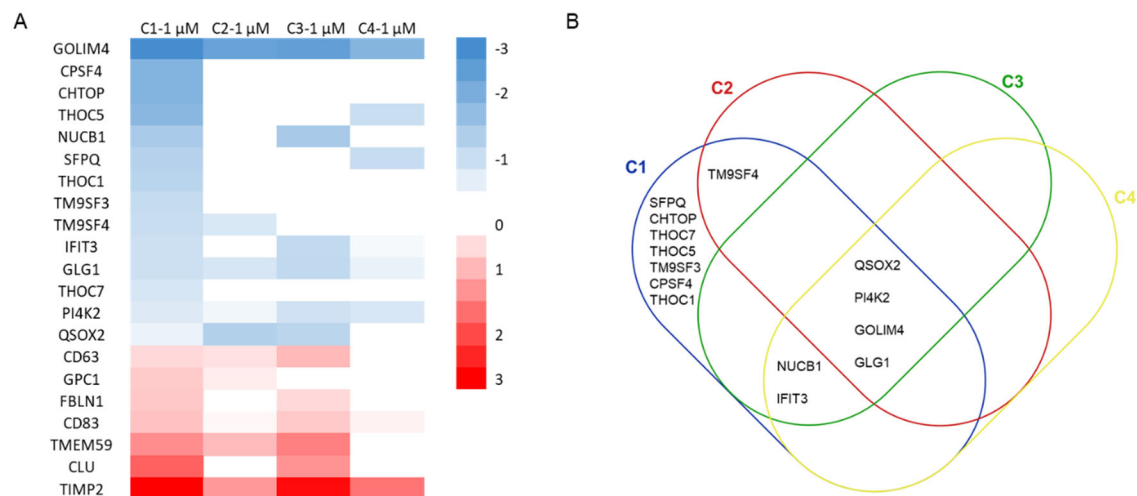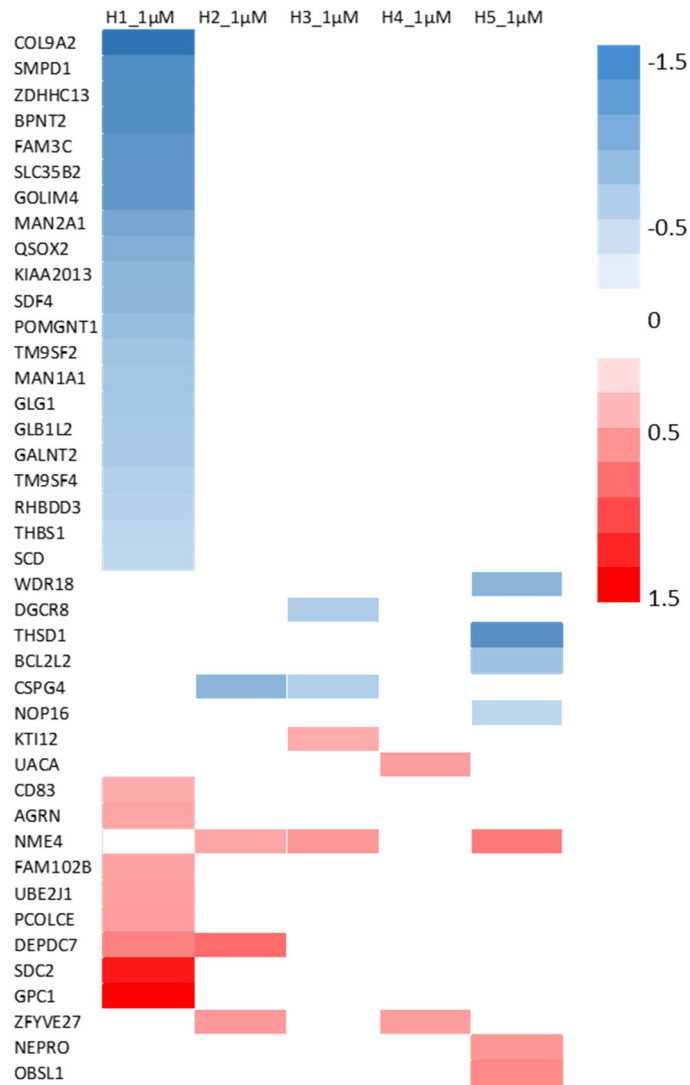
Figure 2.14 Degradation profiles of **H1-H5**. Heat map of proteomics profiles of **H1-H5** at 1 μM. Each column represents different conditions and each row represents the altered protein. The color red or blue of each region represents this protein is up-regulated or down-regulated (Fold change > 1.4) under this condition and the white region represents the protein level remains unchanged or the change is not statically significant.

Table 2.2 The altered proteins under the treatment of **H1** at 1 μM after 18 h. **Values** in bold represent this protein is also degraded by **C1** (1 μM). Dash (-) represents this protein is not identified in the proteomics.

Gene name	Uniprot ID	Protein name	H1-1 μM $\log_2$FC	C1-1 μM $\log_2$FC
COL9A2	Q14055	Collagen alpha-2(IX) chain	-1.58	-
SMPD1	P17405	Sphingomyelin phosphodiesterase	-1.31	-
ZDHHC13	Q8IUH4	Palmitoyltransferase ZDHHC13	-1.30	-
BPNT2	Q9NX62	Golgi-resident adenosine 3,5-bisphosphate 3-phosphatase	-1.30	-
FAM3C	C9JP35	Protein FAM3C	-1.21	-
SLC35B2	Q8TB61	Adenosine 3-phospho 5-phosphosulfate transporter 1	-1.21	-
GOLIM4	F8W785	Golgi integral membrane protein 4	-1.21	-2.21
MAN2A1	Q16706	Alpha-mannosidase 2	-1.05	-
QSOX2	Q6ZRP7	Sulfhydryl oxidase 2	-0.95	-0.89
KIAA2013	Q8IYS2	Uncharacterized protein KIAA2013	-0.88	-
SDF4	A0A5F9UP49	45 kDa calcium-binding protein	-0.86	-
POMGNT1	Q8WZA1	Protein O-linked-mannose beta-1,2-N-acetylglucosaminyltransferase 1	-0.79	-
TM9SF2	Q99805	Transmembrane 9 superfamily member 2	-0.71	-
MAN1A1	P33908	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	-0.70	-
GLG1	Q92896	Golgi apparatus protein 1	-0.69	-1.13
GLB1L2	Q8IW92	Beta-galactosidase-1-like protein 2	-0.66	-
GALNT2	Q10471	Polypeptide N-acetylgalactosaminyltransferase 2	-0.66	-
TM9SF4	A0A0C4DFM1	Transmembrane 9 superfamily member	-0.59	-1.17
RHBDD3	Q9Y3P4	Rhomboid domain-containing protein 3	-0.55	-
THBS1	P07996	Thrombospondin-1	-0.52	-
SCD	O00767	Acyl-CoA desaturase	-0.51	-
CD83	A0A087WX61	CD83 antigen	0.51	1.22
AGRN	A0A494C0G5	Agrin	0.56	-

FAM102B	Q5T8I3	Protein FAM102B	0.57	-
UBE2J1	Q9Y385	Ubiquitin-conjugating enzyme E2 J1	0.58	-
PCOLCE	Q15113	Procollagen C-endopeptidase enhancer 1	0.60	-
DEPDC7	Q96QD5	DEP domain-containing protein 7	0.74	-
SDC2	E7ESK6	Syndecan	1.27	-
GPC1	H7C410	Glypican-1	1.39	1.16
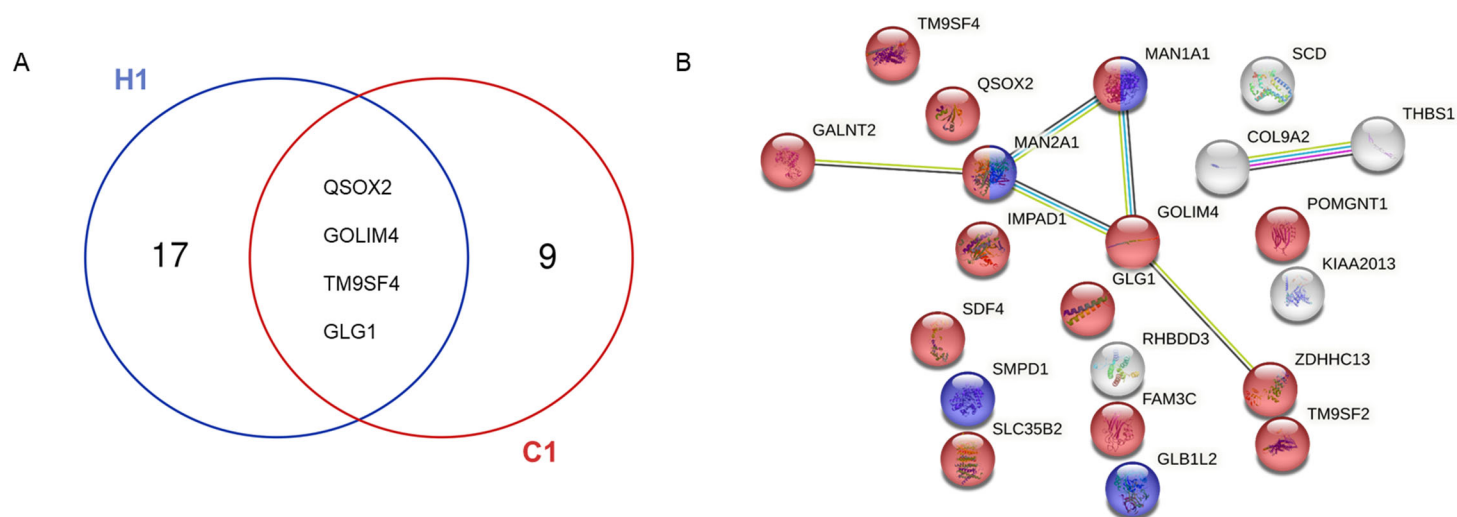
Figure 2.15 Analysis of the potential targets of **H1**. (A). Venn diagram comparing the degradation profiles of **C1** and **H1** at 1 μM. (B). STRING functional analysis of the degraded proteins under the treatment of 1 μM **H1**. Proteins localized at Golgi apparatus are highlighted in red (GO:0005794; FDR: 6.26e-08) and proteins involved in hydrolase activity, acting on glycosyl bonds are highlighted in blue (GO:0016798; FDR: 0.0382).

Table 2.3 The altered proteins under the treatment of **H2-H5** at 100 nM and 1 μM after 18 h. The italicized *values* represent the observation is statically significant (p < 0.05) but the protein level change is not significant (fold change < 1.4), and the shading values represents the observation is not statically significant (p > 0.05).

Gene name	Uniprot id	Protein name	H2-1 μM	H2-100 nM	H3-1 μM	H3-100 nM	H4-1 μM	H4-100 nM	H5-1 μM
			log$_2$FC	log2FC	log2FC	log2FC	log2FC	log2FC	log2FC
BCL2L2	G3V3B7	Bcl-2-like protein 2 (Fragment)	0.15	0.30	-0.13	-0.11	-0.12	0.03	-0.74
CAMKMT	B5MC79	Calmodulin-lysine N-methyltransferase	1.04	1.50	0.63	1.11	0.81	0.76	1.21
CCDC9	Q9Y3X0	Coiled-coil domain-containing protein 9	1.85	1.64	1.33	0.38	1.50	1.61	2.09
CSPG4	Q6UVK1	Chondroitin sulfate proteoglycan 4	-0.86	-0.23	-0.59	-0.48	-0.51	-0.36	0.03
DEPDC7	Q96QD5	DEP domain-containing protein 7	0.84	0.63	0.75	0.63	0.84	0.62	0.21
DGCR8	Q8WYQ5	Microprocessor complex subunit	-0.31	0.26	-0.62	-0.68	-0.39	*-0.30*	*-0.39*
HAUS7	Q99871	HAUS augmin-like complex subunit 7	-0.08	0.56	0.02	-0.30	0.06	-0.04	-0.18
HROB	Q8N3J3	Homologous recombination OB-fold protein	-0.01	0.23	-0.01	0.09	0.13	0.64	-0.16
KTI12	Q96EK9	Protein KTI12 homolog	*0.26*	0.11	0.52	0.08	*0.16*	-0.01	*0.44*
NEPRO	Q6NW34	Nucleolus and neural progenitor protein	0.25	0.39	0.53	0.11	-0.06	0.05	0.63
NEURL4	I3L100	Neuralized-like protein 4	0.41	0.51	0.29	0.20	0.24	0.43	-0.19
NME4	A0A087WVT9	Nucleoside diphosphate kinase	0.55	0.54	0.62	0.44	0.42	0.03	0.77
NOP16	Q9Y3C1	Nucleolar protein 16	0.00	0.02	-0.21	-0.07	-0.24	-0.10	-0.51
OBSL1	A6NN50	Obscurin-like protein 1	0.28	0.36	0.19	0.23	0.33	0.06	0.69
OSBPL5	Q9H0X9	Oxysterol-binding protein-related protein 5	0.28	0.62	0.13	0.22	*0.39*	*0.44*	0.53
RPP25L	Q8N5L8	Ribonuclease P protein subunit p25-like protein	0.52	-0.06	0.88	0.69	0.63	-0.10	0.25
THSD1	Q9NS62	Thrombospondin type-1 domain-containing protein 1	-0.32	-0.30	-1.07	-0.95	-0.26	-0.30	-1.29
UACA	H0YNH8	Uveal autoantigen with coiled-coil domains and ankyrin repeats	-0.04	0.32	0.23	0.33	0.59	0.21	0.28
WDR18	U3KQC1	WD repeat-containing protein 18	*-0.30*	-0.20	*-0.25*	-0.22	-0.10	-0.14	-0.87
ZADH2	Q8N4Q0	Prostaglandin reductase 3	0.35	0.46	0.27	0.17	0.43	0.63	-0.04
ZFYVE27	Q5T4F4	Protrudin	0.61	0.35	0.55	*0.28*	0.59	*0.23*	*0.45*

2.3 Summary and discussion

In this chapter, cholesterol and 4β-hydroxycholesterol-bearing PROTACs were designed and synthesized and their degradation profiles were obtained via TMT-based quantitative proteomics.

For cholesterol-bearing PROTACs, a series of cholesterol-binding protein were identified (such as NPC2, OSBPL8) but only limited numbers of known cholesterol-binding proteins can be degraded by **C1-C4**. In addition to the putative loss of binding towards these targets, the linker within a PROTAC also plays an important role in the degradation selectivity. As mentioned in section 1.2.1.2, the linker can affect the stability of a ternary complex and further affect the degradation selectivity. Cholesterol as a promiscuous binder can undoubtedly bind to numerous proteins. However, when converting cholesterol into a PROTAC, this molecule will form ternary complexes between E3 ligase and different target proteins with various stabilities and only the ternary complex with positive cooperativity can be further degraded by a PROTAC. Additionally, different substrates also exhibit their own preference to different E3 ligases, which may also explain the lack of degradation of cholesterol-binding proteins.

In addition, **C1-C4** also exhibit common but also diverse degradation profiles. This can be explained in two perspectives. First, **C1** may exhibit better cell permeability compared with **C2-C4** and this is because the alkyl linker exhibits lower total polar surface area compared with PEG linkers by reducing the number of polarized atoms. Second, various linkers also affect the stability of the formed ternary complex. An unsuitable linker will also lead to steric clashes between the E3 ligase and the potential target and further leads to lack of degradation.

For 4β-hydroxycholesterol-bearing PROTACs, the number of degraded proteins is even more limited and the overlap of affected proteins between CHO-PROTACs and 4βHC-PROTACs is remarkably low. Except the selectivity conferred by the linker part, the modifications on the cholesterol scaffold also contribute to its degradation selectivity.

Additionally, as illustrated in section 1.3, 4β-hydroxycholesterol, as an endogenous LXR ligand, itself could active the transcription factor, LXRα and further activate the expression of proteins including SREBP1c. However, no known LXR target gene is identified from **H1-H5** proteomics profiles.

It is noteworthy to mention that the concentration and time for the compound treatment will also affect the degradation efficiency of a given substrate and further affect the degradation profile of a PROTAC. A common situation observed for a PROTAC is the Hook effect, whereby the binary complex between a PROTAC and the E3 ligase or a PROTAC and the target protein will be predominated at a high concentration and it will compete with the ternary complex, which means the degradation efficiency will be impaired at high concentration. Therefore, the Hook effect may also be the reason for the lack of degradation of some proteins. Additionally, the treatment time of a PROTAC is also crucial. We set the incubation time to 18 h to realize the complete degradation of all potential target proteins. However, we cannot exclude the possibility that some proteins are affected due to the removal of specific proteins. Moreover, it is worth mentioning that pomalidomide as a molecular glue can also degrade proteins via CRBN[134] and there is an example presenting that a MDM2-PROTAC, **MG-277** fails to degrade the intended target proteins but function as a molecular glue[135]. However, no known substrate of pomalidomide is identified from the proteomics profiles of **C1-C4** and **H1-H4**.

In the next chapter, I choose GOLIM4, which can be significantly degraded by both (oxy)sterols-bearing PROTACs, for further target validation. GOLIM4 has been identified as a putative cholesterol binding protein in a PAL experiment[123] however the functional consequence of this interaction remains to be explored.

3 Chapter 3 Hypothesis I: **C3** directly binds to GOLIM4 and leads to its degradation

In chapter 2, proteomics data indicated that GOLIM4 was downregulated by both CHO- and 4βHC-bearing PROTACs, which suggests that GOLIM4 might be a sterol-interacting protein. In this chapter, the detailed functions of GOLIM4 will be further introduced. Then, a series of experiments will be conducted to explore the mechanism of C3-induced GOLIM4 degradation. Finally, whether GOLIM4 is a sterol-interacting protein would be investigated via target engagement assays.

3.1 Golgi Integral Membrane Protein 4 (GOLIM4)

3.1.1 Structure and biological functions

Golgi Integral Membrane Protein 4 (GOLIM4, or Golgi phosphoprotein of 130 kDa, GPP130) is a type II transmembrane protein mainly localized at the *cis*-Golgi[136]. It consists of three domains, a cytosolic tail (residues 1-12), a transmembrane domain (residues 13-33), and a coiled-coil lumenal stem domain (residues 34-696)[137]. The lumenal domain can be divided into two parts. The coiled-coil luminal domain (residues 34-244) is mainly responsible for its Golgi localization and trafficking, and the acidic C terminus (residues 245-696) is rich in acidic amino acids[138] (Figure 3.1). Under normal physiological conditions, GOLIM4 mediates the transportation of proteins such as GP73 and TGN46 from the early endosome to Golgi via the bypass pathway and the knockdown of GOLIM4 will lead to the accumulation of proteins in the endosome[139]. Interestingly, different from other Golgi resident proteins, the localization, and trafficking of GOLIM4 is pH-sensitive[136]. Weak bases (such as monensin and chloroquine) treatment could redistribute GOLIM4 from Golgi to the endosome, and it relocalises to the Golgi after drug washout[139]. Based on the cycling nature of GOLIM4, the primary function of GOLIM4 is to transport protein cargos between Golgi and endosome.
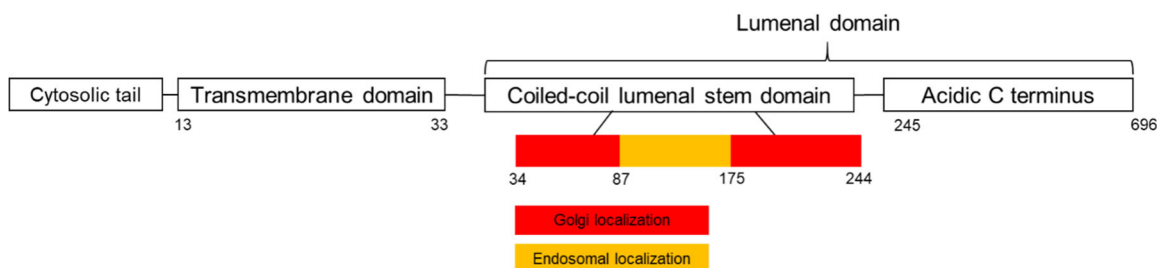
Figure 3.1 Schematic of GOLIM4 structure.

Additionally, GOLIM4 is involved in cellular manganese homeostasis[140]. Mn^{2+} can directly bind to the lumenal domain of GOLIM4[141], leading to its oligomerization[142] and further degradation by lysosome through GGA1 and clathrin-mediated *trans*-Golgi network (TGN) to lysosome pathway[143]. The oligomerized GOLIM4 is recognized by sortilin, a known TGN to lysosome sorting receptor[144].

3.1.2 Functions in diseases.

As mentioned in 3.1.1, GOLIM4 is responsible for trafficking cargos between Golgi and endosomes. Shiga toxin 1 (STx1), which is one of the most lethal toxins produced by *Shigella dysenteriae* 1 and Shiga toxin (Stx)-producing *Escherichia coli* (STEC)[145]*,* also relies on its direct binding to GOLIM4 to exert its toxicity [146].

STx1 is a soluble protein, consisting of a single A subunit and five identical B subunits[147]. STx1 first binds to the cell membrane glycolipid globotriaocylceramide (Gb3) to enter the cell via binding sites located on the B subunit. Then STx1 is transported from the early endosome to Golgi by directly binding to GOLIM4 and further entering the endoplasmic reticulum to active ribosomes via its active A subunit. Retrograde trafficking from early endosomes to Golgi allows STx1 to bypass late endosomes and lysosomes and avoid its degradation[146] (Figure 3.2 A). The trafficking of STx1 can be successfully inhibited by deleting the stem domain (residues 88-175), which is responsible for endosome to Golgi trafficking[140] or Mn^{2+} treatment[148], which leads to GOLIM4 lysosomal degradation. If the retrograde trafficking pathway is impaired, STx1 is re-routed to late endosomes and lysosomes for further degradation (Figure 3.2 B). Therefore, targeting GOLIM4 degradation has been regarded as a useful method to protect cells from STx1 infections.
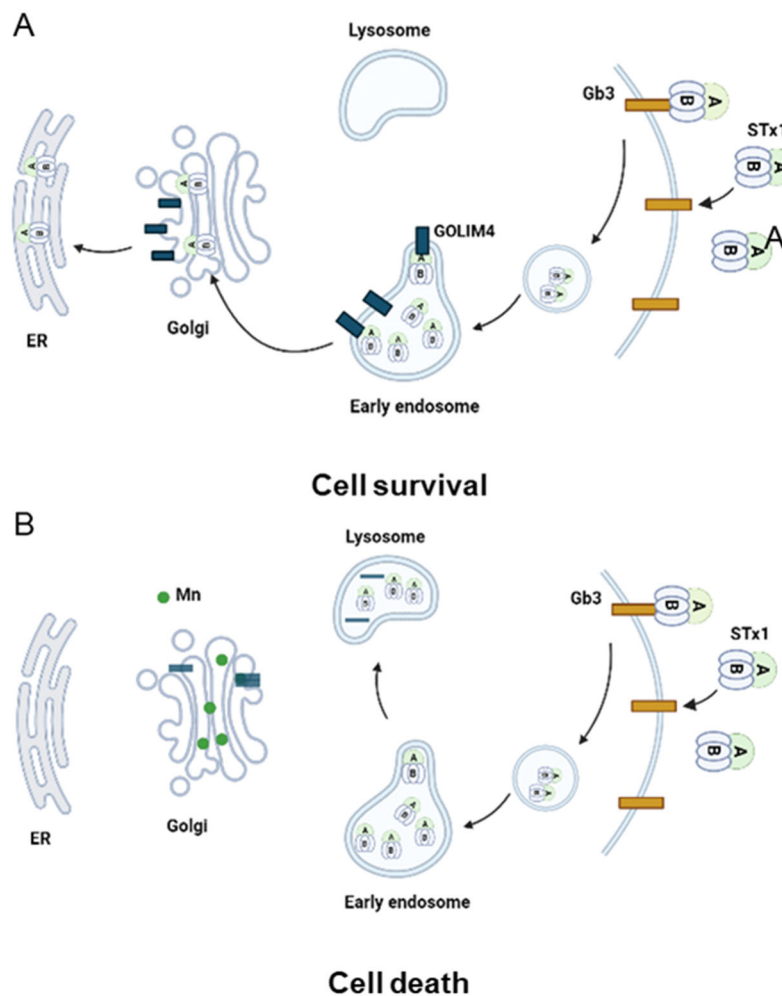
Figure 3.2 GOLIM4 is responsible for STx1 retrograde trafficking. (A).Under normal conditions, STx1 will traffic from early endosome to Golgi via binding to GOLIM4 and finally enter ER to exert its toxic action. (B).Manganese will induce GOLIM4 oligomerization and lysosomal degradation. Under Mn treatment, retrograde trafficking of STx1 is inhibited, and STx1 will enter the lysosome for further degradation. Created with BioRender.com.

Additionally, *GOLIM4*, as a target gene of stromal interaction molecule 1 (STIM1), is found to be significantly overexpressed in head and neck tumor cells[149]. The removal of GOLIM4 could effectively inhibit the proliferation of FaDu cells (human pharyngeal squamous carcinoma cell) and Tca-8113 cells (human tongue squamous carcinoma cell) by blocking cell cycle in the G1 phase and preventing cells from entering the S-phase.

3.2 Target validation of GOLIM4

3.2.1 **C1-C4** could degrade GOLIM4 in two different cell lines.

As mentioned in section 2.3, GOLIM4 was chosen for further target validation. First, western blotting was performed to validate the proteomics data. Western blotting is a commonly used analytical method in molecular and cell biology to separate and identify specific proteins from cell mixtures[150]. Briefly, proteins from a particular mixture are separated based on their molecular weights via electrophoresis. Then proteins are transferred to a solid phase (PVDF or nitrocellulose membrane), and the proteins of interest are visualized using proper primary and secondary antibodies. HeLa cells were treated with **C1-C4** at 100 nM or 1 µM for 18 h, and the GOLIM4 protein level change was analyzed by western blotting. DMSO and **C5** (1 µM) were also included as a control. Consistent with the proteomics data, significant dose-dependent degradation of GOLIM4 (> 90%) could be observed in HeLa cells after treatment with 1 µM **C1-C4** while no obvious degradation of GOLIM4 was observed at 100 nM of **C2-C4**. As a control, no significant depletion of GOLIM4 was observed in HeLa cells treated with 1 µM **C5** (Figure 3.3).
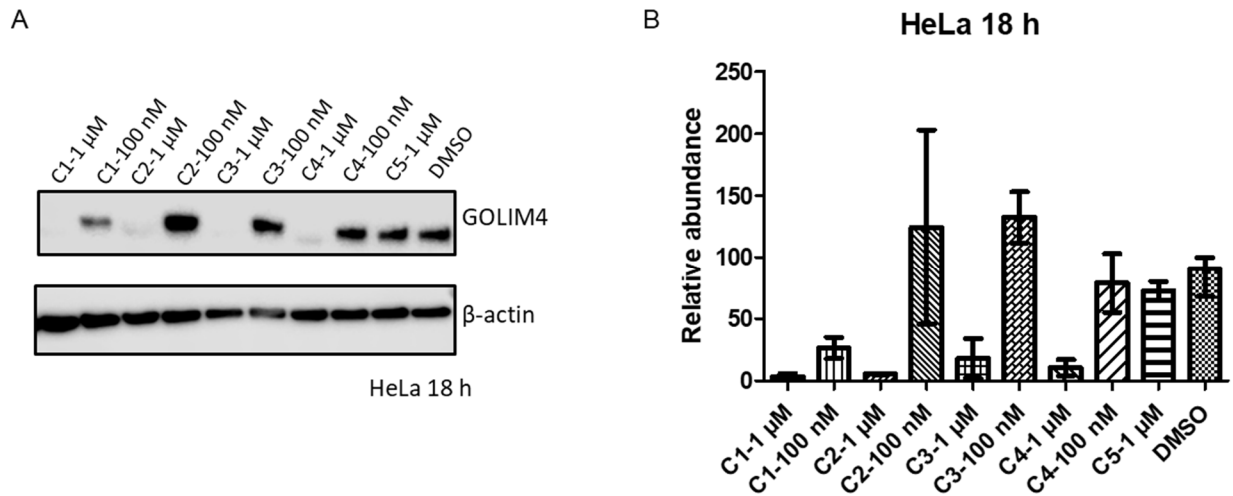
Figure 3.3 **C1-C4** lead to GOLIM4 degradation in HeLa cells. (A). HeLa cells were treated with 100 nM or 1 µM of **C1-C4** for 18 h. 1 µM **C5** and DMSO were treated as control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n = 2). (B) Quantification of the western blot shown in (A) and protein band intensity was normalized to β-actin from DMSO control samples.

To further assess the generality of **C1-C4** among different cell lines, the human lung cancer cell line, A459, was also chosen to test the GOLIM4 degradation ability of **C1-C4**. Similarly, A549 cells were treated with **C1-C5** at 100 nM or 1 µM for 18 h, and the GOLIM4 abundance was detected by western blotting. **C1** and **C3** still exhibited effective GOLIM4 degradation ability; over 90 % of GOLIM4 was removed by both at 1

μM. Differently, the degradation abilities of **C2** and **C4** were impaired in A459 cell lines; less than 50 % of GOLIM4 could be degraded by **C2** and **C4** (Figure 3.4).
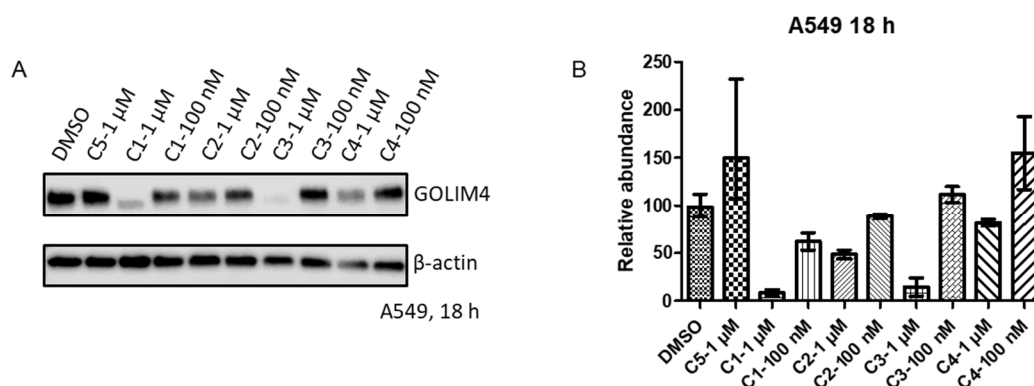
Figure 3.4 **C1-C4** lead to GOLIM4 degradation in A549 cells. (A) A549 cells were treated with 100 nM or 1 μM of **C1-C4** for 18 h. 1 μM **C5** and DMSO were treated as control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2). (B) Quantification of the western blot shown in (A) and protein band intensity was normalized to β-actin from DMSO control samples.

In summary, **C3** exhibited the best degradation ability towards GOLIM4 in both HeLa and A549 cell lines. Therefore, **C3** was selected for further validation experiments.

3.2.2 Degradation of GOLIM4 by **C3** is dose- and time-dependent.

To study the kinetics of GOLIM4 degradation, a time-dependent assay was conducted with **C3**. GOLIM4 levels in HeLa cells were detected over a 48-hour time course treatment with 1 μM **C3**. Progressive degradation of GOLIM4 over 48 hours was observed for **C3** (Figure 3.5). 1 μM **C3** started to significantly degrade GOLIM4 at 4 h, and the maximal GOLIM4 degradation was reached at 8 h when approximately 70 % of GOLIM4 was degraded, and the degradation of GOLIM4 could last 48 h.
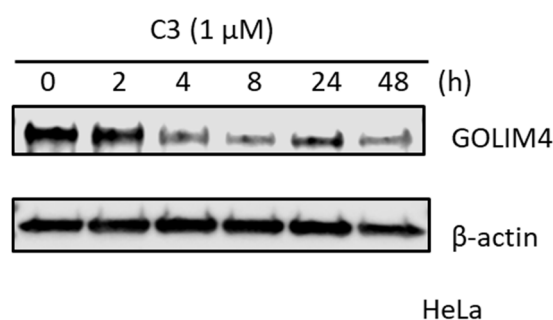
Figure 3.5 Degradation of GOLIM4 induced by **C3** is time-dependent. HeLa cells were treated with 1 μM **C3** at different time points. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=4).

Reassuringly, **C3** showed the concentration-dependent of GOLIM4 degradation with DC_{50} = 0.28 μM (Figure 3.6). 1 μM **C3** could degrade 80 % GOLIM4, and the maximal GOLIM4 degradation was reached at 2 μM.
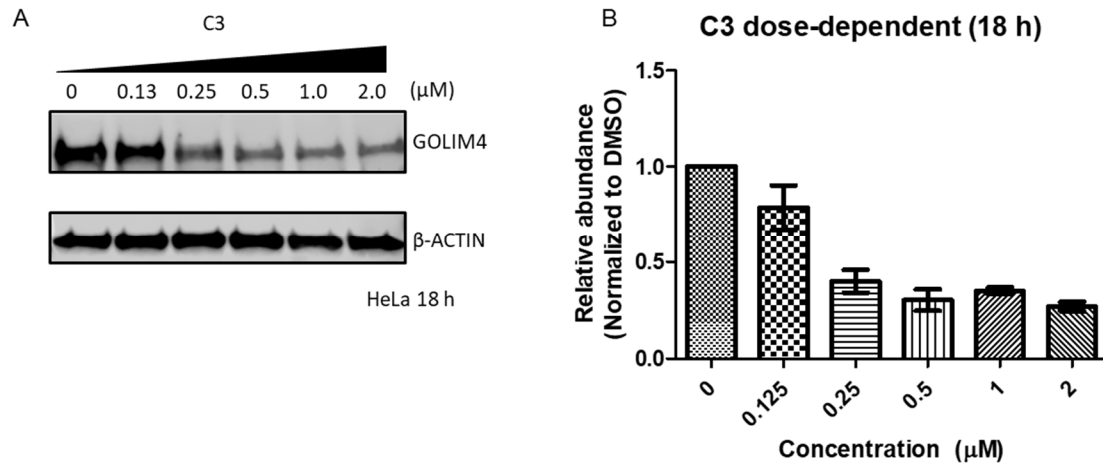
Figure 3.6 Degradation of GOLIM4 induced by **C3** is dose-dependent. (A) HeLa cells were treated with **C3** at different concentrations for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=4). (B) Quantification of the western blot shown in (A) and protein band intensity was normalized to β-actin from DMSO control samples.

3.2.3   Degradation of GOLIM4 by **C3** is not mediated by UPS

Next, the mechanism of **C3**-mediated GOLIM4 degradation was explored. First, we treated cells with 10 μM **C5** to compete with **C3** (1 μM) for its binding with GOLIM4. As shown in Figure 3.7, 10 μM **C5** cannot block GOLIM4 degradation induced by **C3**, which indicates that **C3** and **C5** do not bind to the same position of GOLIM4. Next, to study whether cereblon involves in GOLIM4 degradation, HeLa cells were treated with an excess amount of pomalidomide (20 μM) to compete with **C3** (1 μM) for its binding with cereblon. As shown in Figure 3.7, 20 μM pomalidomide's treatment cannot block GOLIM4 degradation induced by **C3**, which preliminarily indicates cereblon is not involved in **C3**-mediated GOLIM4 degradation.
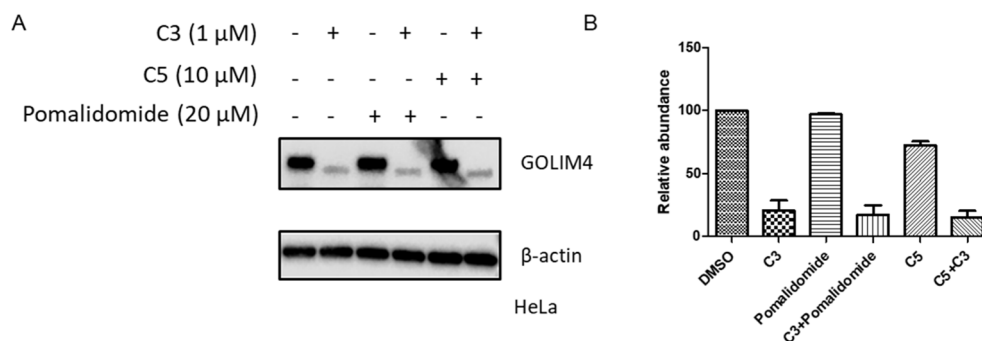
Figure 3.7 Degradation of GOLIM4 induced by **C3** cannot be blocked by excessive **C5** and pomalidomide. (A) HeLa cells were treated with **C5**, **C3,** and pomalidomide for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2). (B) Quantification of the western blot shown in (A). Protein band intensity was normalized to β-actin from DMSO control samples.

To further explore whether cereblon is needed for GOLIM4 degradation, **C3_Me** was synthesized by replacing pomalidomide with *N*-methylated pomalidomide, which blocks its binding to cereblon (Figure 3.8 A)[151]. HeLa cells were treated with 100 nM, 1 μM, and 10 μM **C3_Me** for 18 h. **C3_Me** exhibits similar GOLIM4 degradation ability to **C3** (both can remove > 80 % of GOLIM4 at 1 μM), and its treatment could still degrade GOLIM4 in a concentration-dependent manner after 18 h (Figure 3.8 B).

Taken together, the pomalidomide competition assay and *N*-methylated **C3** degradation assay strongly suggest that the degradation of GOLIM4 induced by **C3** is not mediated by cereblon.
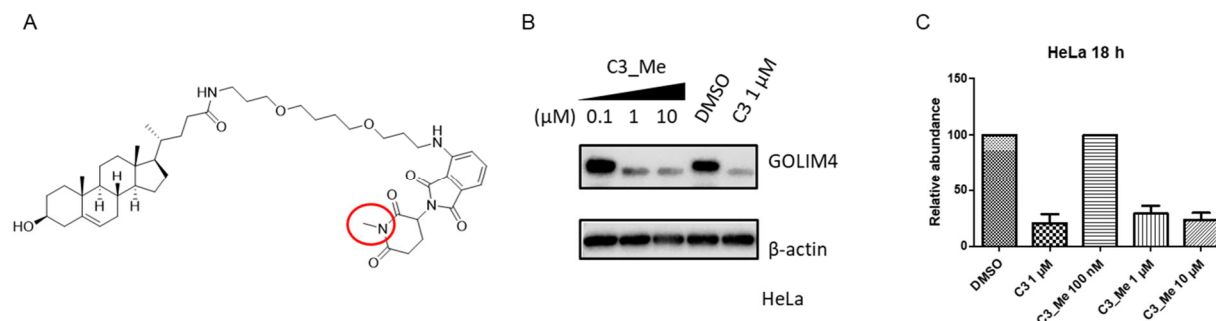
Figure 3.8 **C3_Me** could still degrade GOLIM4 in a dose-dependent manner. (A) Chemical structure of **C3_Me** (Red circle indicates the methyl group). (B) HeLa cells were treated with **C3_Me** at 100 nM, 1 μM, and 10 μM for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2). (C) Quantification of the western blot shown in (A) and protein band intensity was normalized to β-actin from DMSO control samples.

Since cereblon is not involved in GOLIM4 degradation, whether the degradation of GOLIM4 is still mediated by Cullin-RING E3 ubiquitin Ligases (CRLs) was investigated. Because cullin neddylation is necessary for the activity of CRLs, we chose **MLN4924**, an inhibitor of the NEDD8-activating enzyme (NAE), to block the cullin neddylation process and the activity of CRLs (Figure 3.9 A)[152]. We treated HeLa cells with **MLN4924** (5 μM) and **C3** (1 μM) for 18 h but the level of GOLIM4 does not affect by the

treatment of **MLN4924** itself, but **C3** could still degrade GOLIM4 in the presence of **MLN4924** which suggests that CRLs arenot involved in GOLIM4 degradation (Figure 3.9 B).

Finally, whether the degradation of GOLIM4 was induced by UPS was investigated. HeLa cells were treated with 10 μM **MG132**, a well-established proteasome inhibitor (Figure 3.9 A), and 1 μM **C3** for 18 h. Interestingly, the treatment of **MG132** itself could lead to the partial removal of GOLIM4 (> 70 %) and this is because **MG132** can induce caspase activation and further leads to apoptosis-associated Golgi fragmentation[153]. Moreover, the abundance of GOLIM4 was even less (10 % remaining) by co-treatment with **MG132** and **C3** (Figure 3.9 B). These results demonstrate that the proteasome inhibition can affect the stability of GOLIM4 itself, and the removal effect of **C3** cannot be blocked by UPS inhibition.
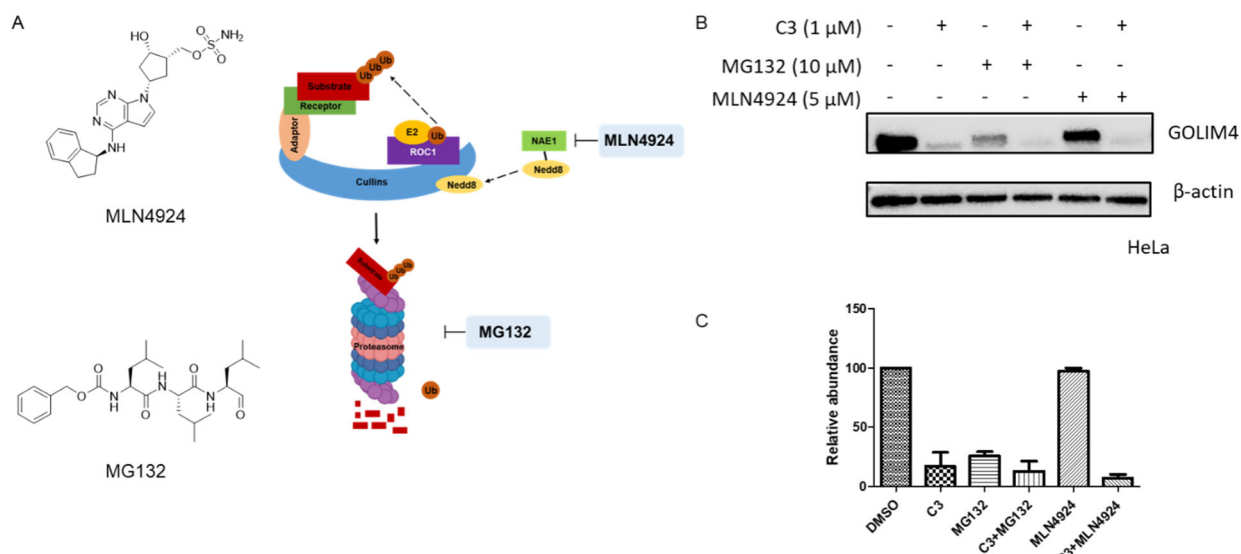
Figure 3.9 Degradation of GOLIM4 by **C3** is not mediated by CRLs and UPS. (A) Chemical structures and mechanism of actions of **MLN4924** and **MG132**. (B) HeLa cells were treated with **C3** (1 μM), **MLN4924** (5 μM), and **MG132** (10 μM) for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2). (C) Quantification of the western blot shown in (A) and protein band intensity was normalized to β-actin from DMSO control samples.

3.2.4 Degradation of GOLIM4 by **C3** is mediated by the lysosomal degradation pathway

In section 3.2.3, we have shown that CRLs and UPS are not involved in **C3**-induced GOLIM4 degradation by pomalidomide competitive assay and **MG132** and **MLN4924** rescue assays. Therefore, we explored the autophagy–lysosome pathway, the other main pathway to mediate protein degradation. In addition, as mentioned in section 3.1, Mn^{2+} could degrade GOLIM4 via a lysosomal pathway. Therefore, Mn^{2+} was also included as a positive control in the following experiments.

First, we treated cells with lysosomal degradation inhibitors, chloroquine (CQ), and bafilomycin A1 (BafA1). Chloroquine can inhibit the function of lysosomes by increasing the pH of lysosomes by accumulation within lysosomes as a weak base and bafilomycin A1 can inhibit the ability of the vacuolar type H^+-ATPase (v-ATPase) to transfer protons into the lysosomes[154] (Figure 3.10 A). Obviously the level of GOLIM4 was not affected by the treatment of CQ (20 μM) itself or cotreatment with CQ (20 μM) and

C3 (1 μM), which strongly suggests that the degradation of GOLIM4 is mediated by the lysosome (Figure 3.10 B). As a positive control, MnCl$_2$ (140 μM) could successfully degrade GOLIM4, and the GOLIM4 degradation induced by MnCl$_2$ could also be blocked by CQ. Interestingly, BafA1 (100 nM) itself could degrade GOLIM4 and this is because the prolonged inhibition of lysosome activity may re-direct protein degradation to the proteasome. However, the cotreatment with BafA1 (100 nM) and **C3** (1 μM) could stabilize GOLIM4.
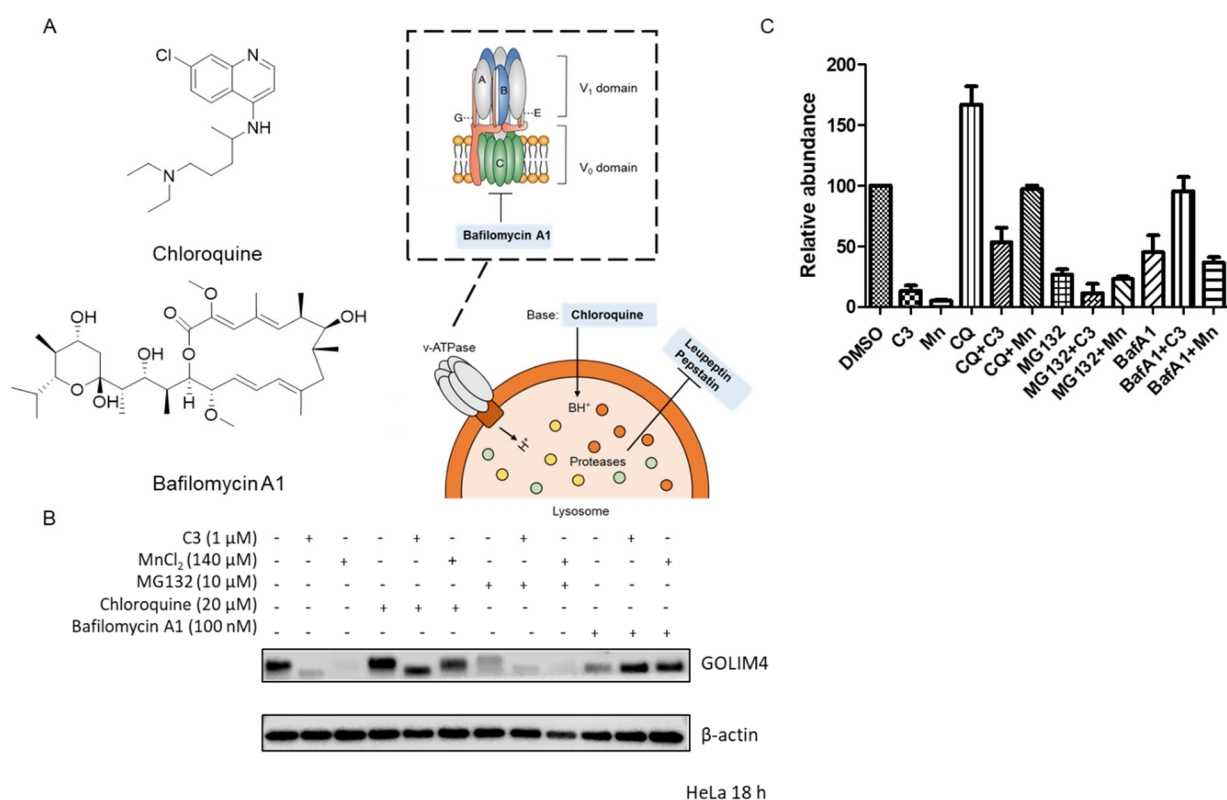
Figure 3.10 CQ and Baf A1 can inhibit **C3** (or MnCl$_2$)-induced GOLIM4 degradation. (A) Chemical structures and mechanism of actions of chloroquine, bafilomycin A1, and lysosomal proteases inhibitors (leupeptin and pepstatin). Figure adapted from Whitmarsh-Everiss *et al.*[155] (B) HeLa cells were treated with **C3** (1 μM), Chloroquine (20 μM), Bafilomycin A1 (100 nM), MnCl$_2$ (140 μM) and MG132 (10 μM) for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2). (C) Quantification of the western blot shown in (A) and protein band intensity was normalized to β-actin from DMSO control samples.

Then, we also treated cells with lysosomal protease inhibitors, leupeptin, which inhibits the activity of serine and cysteine proteases, and pepstatin, which strongly inhibits acid proteases such as renin and pepsin (Figure 3.11 A)[156]. HeLa cells were pretreated with leupeptin (100 μg/mL) and pepstatin (50 μg/mL) for 24 h and then were treated with **C3** (1 μM) or MnCl$_2$ (140 μM) for a further 18 h. As shown in Figure 3.11, the degradation of GOLIM4 could be partially blocked by these lysosomal protease inhibitors, which further proves that the degradation of GOLIM4 was mediated by the lysosomal degradation pathway. Similarly, Mn^{2+}-induced GOLIM4 degradation could also be blocked by these lysosomal protease inhibitors.
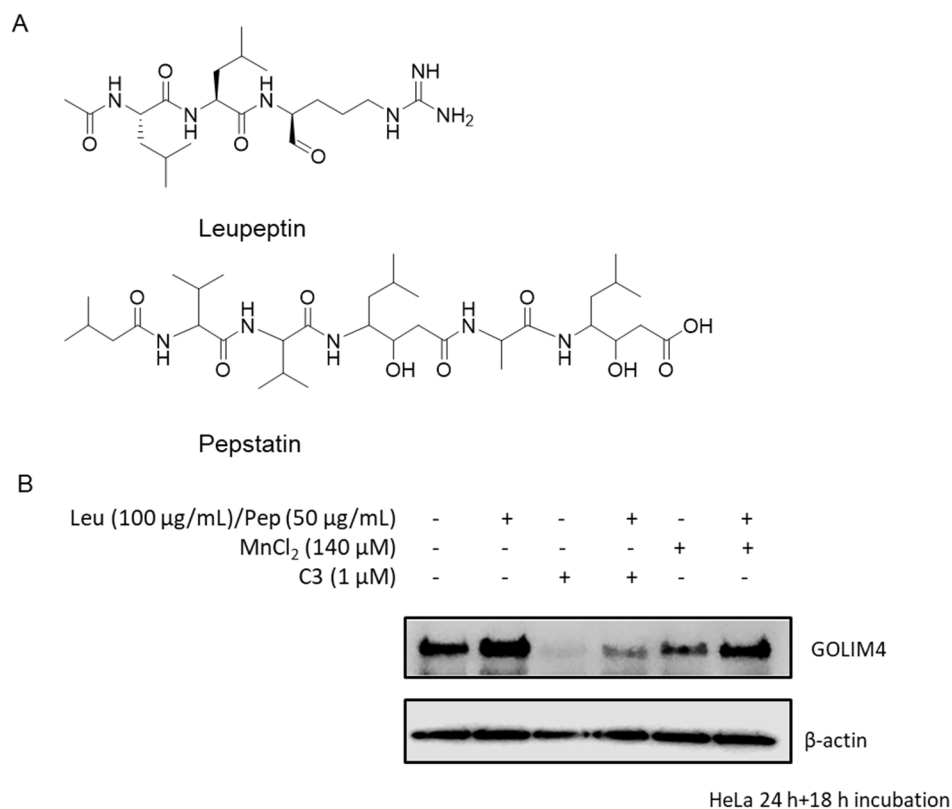
Figure 3.11 **C3**/Mn^{2+} -induced GOLIM4 degradation can be blocked by lysosomal protease inhibitors. (A) Chemical structures of leupeptin and pepstatin. (B) HeLa cells were treated with leupeptin (100 μg/mL) and pepstatin (50 μg/mL) for 24 h and then were treated with **C3** (1 μM) or MnCl$_2$ (140 μM) for a further 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2).

Finally, to study whether autophagy is involved in GOLIM4 degradation, we treated HeLa cells with **SAR405**, a selective inhibitor of VPS34, to inhibit vesicle trafficking from late endosomes to lysosomes and further inhibit autophagy.[157] **SAR405** (1 μM) cannot block the degradation of GOLIM4 which proves autophagy is not involved in GOLIM4 degradation (Figure 3.12).

In summary, similar to Mn^{2+}, GOLIM4 degradation induced by **C3** is mediated by a lysosomal degradation pathway, but autophagy is not involved.

C3 (1 µM)	-	+	-	+	-	+	-	+	-	+
SAR405 (1 µM)	-	-	+	+	-	-	-	-	-	-
MG132 (10 µM)	-	-	-	-	-	-	-	-	+	+
Chloroquine (20 µM)	-	-	-	-	+	+	-	-	-	-
Bafilomycin A1 (100 nM)	-	-	-	-	-	-	+	+	-	-

GOLIM4

β-actin

HeLa 18 h
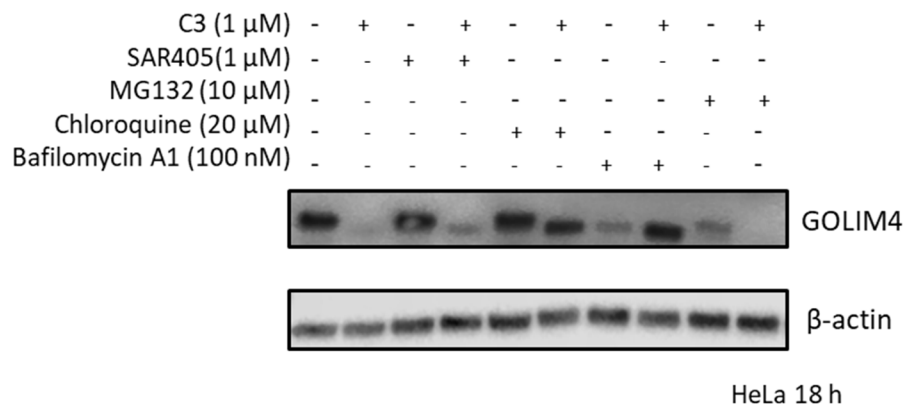
Figure 3.12 Autophagy is not involved in **C3**/Mn^{2+} -induced GOLIM4 degradation. HeLa cells were treated with SAR405 (1 µM), **MG132** (10 µM), CQ (20 µM) and BafA1 (100 nM) for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2).

3.2.5 Degradation of GOLIM4 cannot be blocked by the sortilin inhibitor

As we have shown that both **C3** and Mn^{2+} could degrade GOLIM4 via the lysosomal degradation pathway, the detailed degradation mechanism was further explored. As mentioned in section 3.1, Mn^{2+} binding could lead to GOLIM4 oligomerization, and the oligomerized GOLIM4 is recognized and trafficked to the lysosome by sortilin. Therefore, we will explore whether **C3**-induced GOLIM4 degradation is also mediated by sortilin.

AF38469, a sortilin inhibitor, was chosen as the tool compound to block the function of sortilin (Figure 3.13 A)[158]. We treated HeLa cells with 10 or 50 µM **AF38469** in the presence of **C3** (1 µM) and MnCl$_2$ (140 µM) for 18 h. However, both concentrations of **AF38469** were not able to block the degradation of GOLIM4 induced by **C3** or even Mn^{2+} (Figure 3.13 B). This may be because sortilin, a member of sorting receptors, is responsible for sorting different vesicles by forming different PPIs on other surfaces. From the X-ray crystal structures of sortilin bound to **AF38469**[158] (Figure 3.13, A, PDB ID: 4N7E), **AF38469** could only mimic the C-terminal of neurotensin (a native sortilin ligand) and interfere with limited interactions between sortilin and specific ligands (nerve growth factor-b (NGF), brain-derived neurotrophic factor (BDNF), lipoprotein lipase (LpL), apo lipoprotein AV1, the receptor-associated protein (RAP))[159]. Therefore, GOLIM4 may interact with sortilin at another site where **AF38469** cannot bind. To thoroughly study whether the GOLIM4 degradation is mediated by sortilin, the complete removal of sortilin by knockdown or knockout is necessary.
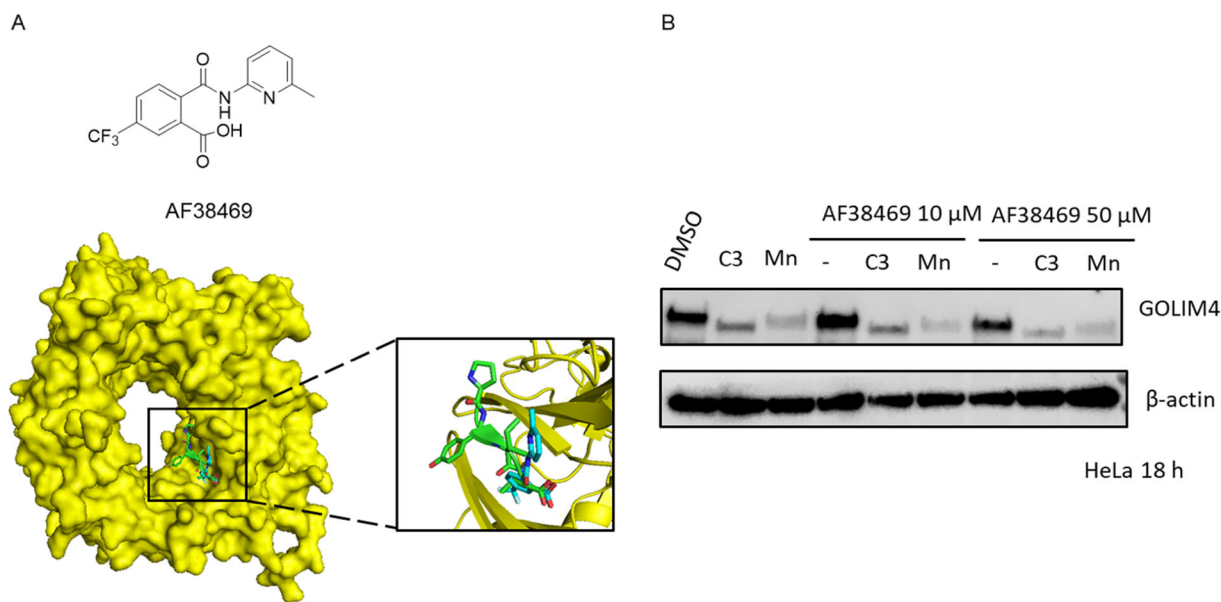
Figure 3.13 Degradation of GOLIM4 cannot be blocked by **AF38469**. (A) Chemical structure of **AF38469** and overlay of X-ray crystal structures of sortilin bound **AF38469** (cyan, PDB: 4N7E) and Neurotensin (C-terminus) (green, PDB: 3F6K). (B) HeLa cells were treated with 10 or 50 μM **AF38469** in the presence of **C3** (1 μM) and $MnCl_2$ (140 μM) for 18 h. DMSO was treated as a control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n=2).

3.2.6 **C3** does not bind to GOLIM4.

To study whether the observed degradation is mediated by direct binding of **C3** to GOLIM4, a cellular thermal shift assay (CETSA) with western blotting read-out was carried out to study the target engagement of **C3**. This assay is based on the principle that the direct binding of a ligand could alter the thermal stability (stabilization or destabilization) of a given target protein. We exposed HeLa cells to **C3** (10 μM) for 30 min, and the harvested cells were subjected to a thermal denaturation gradient from 37.5 °C to 63.4 °C, and the protein level changes were analyzed by western blot. **C3** only exhibits a slight destabilization effect on GOLIM4 (Figure 3.14 A). To address this discrepancy, isothermal dose–response fingerprinting (ITDRF) was conducted at 54.3 °C (Figure 3.14 B) or 63.4 °C (Figure 3.14 C), where the largest differences between control and C3 treated samples were observed. However, **C3** failed to destabilize GOLIM4 in a dose-dependent manner at both temperatures. Based on the CETSA and ITDRF, **C3** does not bind to GOLIM4 directly.

51
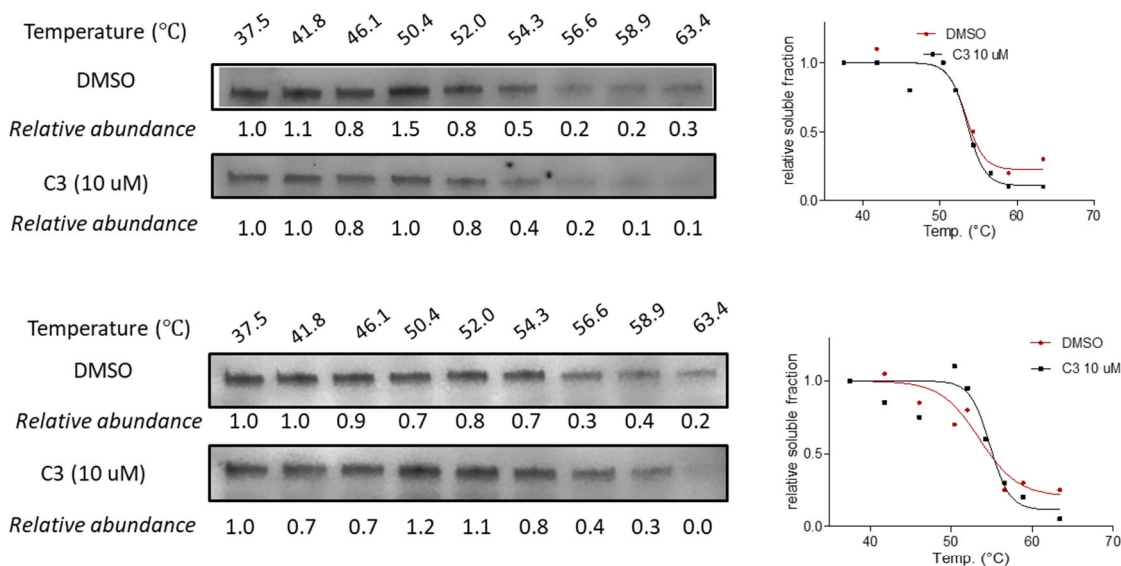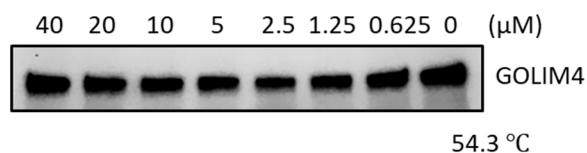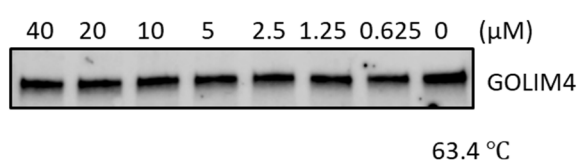
Figure 3.14 **C3** does not bind to GOLIM4 directly. (A) CETSA experiments for **C3** binding to GOLIM4. Two replicates are shown separately, and the protein band intensities were normalized to samples heated to 37.5 ℃. (B) ITDRF experiments for **C3** binding to GOLIM4 to 54.3 ℃. The protein band intensity was normalized to the sample treated with DMSO. (n=3) (C) ITDRF experiments for **C3** binding to GOLIM4 at 63.4 ℃. The protein band intensity was normalized to the sample treated with DMSO. (n=1)

3.3 Summary and discussion

In this chapter, we chose GOLIM4 as the potential target to validate the proteomics data and investigate the mechanism of **C3**-induced GOLIM4 degradation. First, I assumed **C3** functions as a GOLIM4-PROTAC based on: 1) GOLIM4 degradation induced by **C3** has been validated by proteomics and western blotting. 2) GOLIM4 has been identified as a target of cholesterol by another chemoproteomics strategy. However, this assumption was proved incorrect by a pomalidomide competition assay, proteasome inhibitors rescue experiments, and methylated-**C3** degradation assay. Taken together, these experiments demonstrated that cereblon, CRLs, and UPS are not involved in GOLIM4 degradation and further prove that **C3** does not function as a GOLIM4-PROTAC. Then the mechanism of GOLIM4 degradation was established to be lysosome-mediated degradation via lysosomal degradation inhibitors and autophagy inhibitor rescue assay. Since Mn^{2+} also degrades GOLIM4 via the lysosomal degradation pathway, I assumed that, similar to Mn^{2+}, **C3** leads to GOLIM4 oligomerization and further lysosomal degradation. First, I studied the role of sortilin in **C3**-induced GOLIM4 degradation using the sortilin inhibitor **AF38469**. Since sortilin has been proved responsible for Mn^{2+}-induced GOLIM4 depletion, I also included $MnCl_2$ as a positive control. However, this inhibitor cannot rescue the degradation induced by both **C3** and Mn^{2+}. Given the role of sortilin as a sorting receptor, it is reasonable to assume sortilin interacts with different substrates by forming PPIs at various sites, and **AF38469** can only block a limited number of them. Since the interfaces between GOLIM4 and sortilin are still unclear, the best way to investigate the role of sortilin is to remove sortilin completely via knockdown or knockout methods. The purified protein is needed to study whether **C3** also leads to GOLIM4 aggregation. While the purification work was ongoing in our group, whether **C3** directly binds to GOLIM4 was explored by target engagement assay, and ITDRF$_{CETSA}$ suggested that **C3** failed to destabilize GOLIM4 in a dose-dependent way. Eventually, I concluded that the GOLIM4 degradation effect of **C3** is not a direct binding event but a secondary effect based on 1) the GOLIM4 destabilization effect of **C3** is not concentration-dependent. 2) Another downregulated protein induced by **C3**, QSOX2, exhibits a similar degradation pattern to GOLIM4 (will be explained in Chapter 4).

Having established that **C3** does not act as a PROTAC by recruiting CRBN, I asked myself whether it is possible to optimize **C3** into a smaller GOLIM4 degrader by removing the pomalidomide part, which serves as the binding moiety for CRBN. Additionally, because **C3** does not bind to GOLIM4 directly, I sought to investigate the direct targets of **C3**, which are responsible for GOLIM4 degradation. These two questions will be further explored in Chapter 4.

4 Chapter 4 Further exploring the targets of **C3**

In Chapter 3, a series of experiments proved that **C3** does not bind to GOLIM4 directly suggesting that the degradation of GOLIM4 is most likely a secondary effect. In this chapter, the potential target(s) of **C3** will be further explored, and the mechanism of GOLIM4 degradation will be further studied.

4.1 Smaller GOLIM4 degraders

4.1.1 Minimization of **C3**

Since it has been proved that the GOLIM4 degradation does not rely on cereblon, we can assume that the pomalidomide part of **C3** is not necessary for GOLIM4 degradation. Therefore, we wanted to optimize **C3** into smaller GOLIM4 degraders. The minimization of **C3** would be carried out in two ways. First, the fact that **C3** successfully degrades GOLIM4 but **C5** does not indicates that the linker part still plays a role in GOLIM4 degradation. Therefore, we removed the pomalidomide part of **C3** and only kept the linker part to obtain **C3_Boc** (Structure shown in Figure 4.1 A). Second, as mentioned in section 1.2.2, adamantane, as a representative moiety used in hydrophobic tagging (HyT) strategies, has been successfully applied to degrade various endogenous proteins[41]. Therefore, we also replaced pomalidomide with adamantane and kept the same linker to obtain **C3_HyT** (Structure shown in Figure 4.1 A).

The synthesis of **C3_Boc** and **C3_HyT** was straightforward. **C3_Boc** was obtained by HATU-mediated amide coupling between **4** and **9** (Figure 4.1 B). The **C3_HyT** was obtained via a three-step reaction. First, 1-adamantaneacetic acid was coupled to **9** via HATU-mediated amide coupling, and then the Boc-protecting group was removed by TFA and finally the free amine and **4** were coupled via HATU-mediated amide coupling to give **C3_HyT** (Figure 4.1 B).
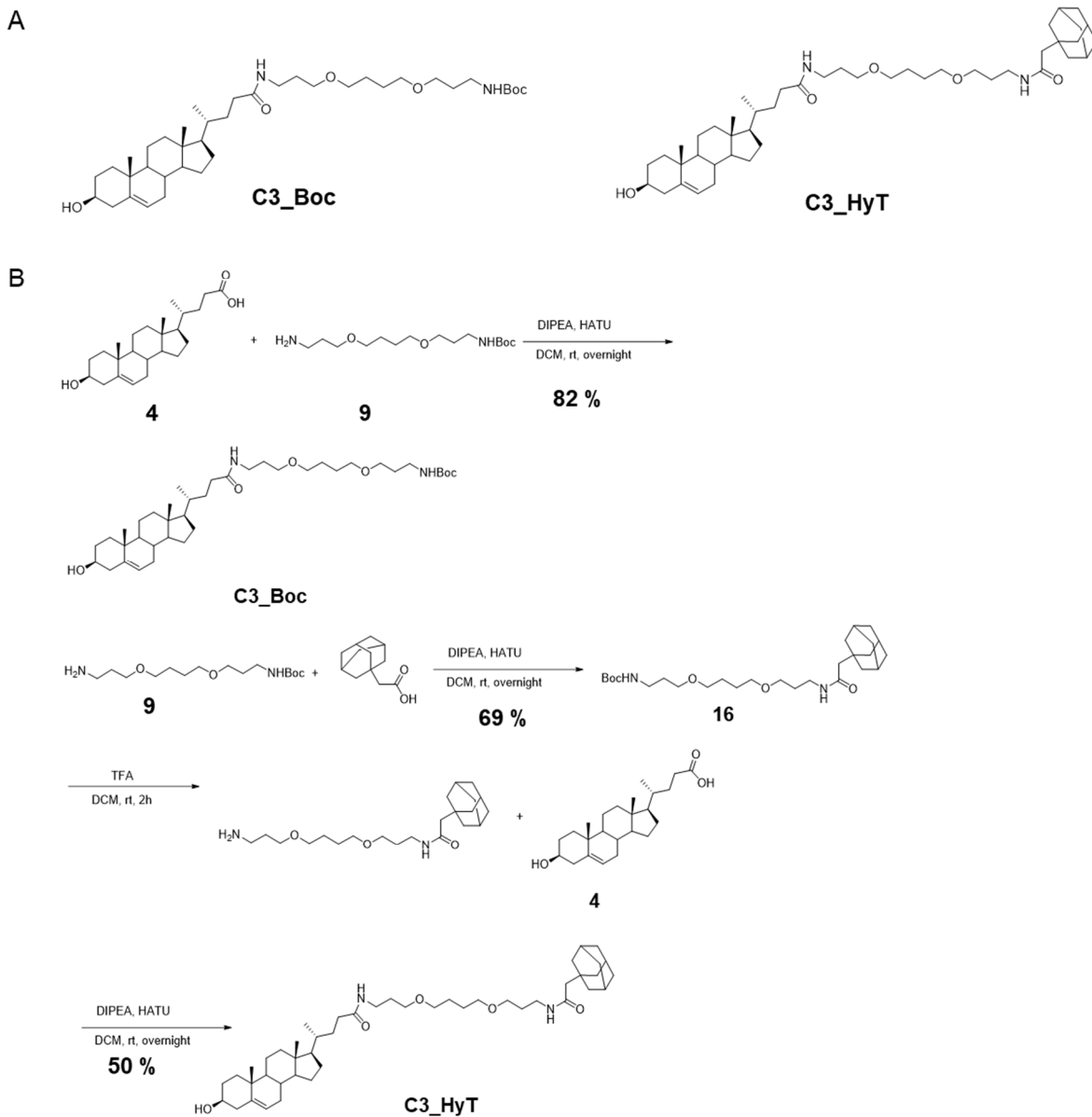
Figure 4.1 Design and synthesis routine of **C3_Boc** and **C3_HyT** (A). The structures of **C3_Boc** and **C3_HyT.** (B). The synthesis route to access **C3_Boc** and **C3_HyT**.

After obtaining **C3_Boc** and **C3_HyT**, we first tested their degradation abilities towards GOLIM4. By treating HeLa cells with different concentrations of **C3_Boc** and **C3_HyT** for 18 h, both **C3_Boc** (Figure 4.2 B) and **C3_HyT** (Figure 4.2 A) exhibited GOLIM4 removal abilities in a dose-dependent manner. Among all the GOLIM4 degraders (**C3**, **C3_Me**, **C3_Boc** and **C3_HyT**), **C3_HyT** exhibits the best GOLIM4 degradation ability; nearly 20% of GOLIM4 can be removed at 100 nM while no noticeable GOLIM4 removal effect can be observed for other degraders at this concentration.
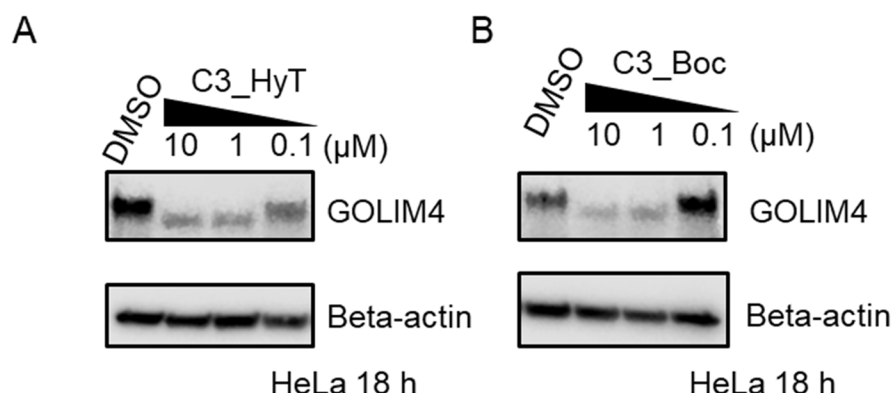
Figure 4.2 Degradation of GOLIM4 by **C3_Boc** and **C3_HyT**. (A). HeLa cells were treated with 100 nM, 1 μM and 10 μM of **C3_HyT** for 18 h. DMSO was treated as control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n = 2). (B). HeLa cells were treated with 100 nM, 1 μM and 10 μM of **C3_Boc** for 18 h. DMSO was treated as control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n = 2).

4.1.2   Golgi resident proteins and glycosylation enzymes are degraded by **C3** and its derivatives.

4.1.2.1   QSOX2 exhibits a similar degradation pattern with GOLIM4.

In the meantime, we focused on validating QSOX2, another down-regulated protein of CHO-bearing PROTACs. As shown in Table 4.1, proteomics data showed that QSOX2 can be degraded by 1 μM **C2** and **C3** with a weaker degradation efficiency compared with GOLIM4. When we studied the mechanism of QSOX2 degradation, we surprisingly found QSOX2 exhibits a similar degradation pattern to GOLIM4. **C3** could also degrade QSOX2 in a dose-dependent manner with DC_{50}=0.5 μM (Figure 4.3 A). QSOX2 degradation is not mediated by cereblon or even UPS but relies on the lysosome (Figure 4.3 B and C). Additionally, **C3_Boc**, and **C3_HyT** could also degrade QSOX2 in a dose-dependent manner (Figure 4.3 D). Taken together, we confirmed that neither GOLIM4 nor QSOX2 degradation is induced by the direct binding of **C3**. Based on the fact that these two proteins are all localized at Golgi[160], we hypothesized that **C3** might cause the destabilization of the Golgi apparatus and then lead to the degradation of Golgi-associated proteins.

Table 4.1 Both GOLIM4 and QSOX2 can be significantly degraded by 1 μM **C2** and **C3**

Gene name	Uniprot ID	Protein name	C1	C2	C3	C4	C5
			log_2FC	log_2FC	log_2FC	log_2FC	log_2FC
GOLIM4	F8W785	Golgi integral membrane protein 4	-2.20	-1.93	-1.96	-1.69	-0.006
QSOX2	H0Y430	Sulfhydryl oxidase	-0.88	**-1.32**	**-1.26**	-1.11	-0.07
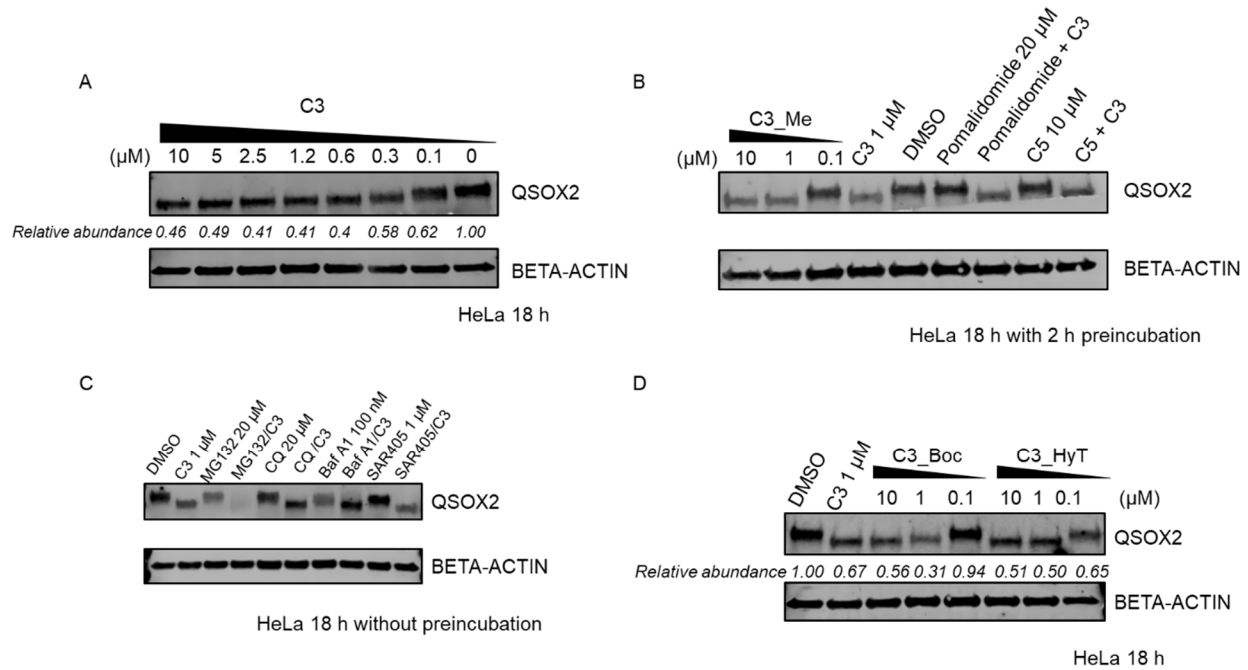
Figure 4.3 QSOX2 exhibits a similar degradation pattern with GOLIM4. (A). **C3** degrades QSOX2 in a dose-dependent way. HeLa cells were treated with different concentrations of **C3** for 18 h. DMSO was treated as control. The protein levels of QSOX2 and β-actin were analyzed by western blot (n = 1). (B). CRBN does not involve in QSOX2 degradation. HeLa cells were treated with **C5**, **C3**, **C3_Me** and pomalidomide for 18 h. DMSO was treated as a control. The protein levels of QSOX2 and β-actin were analyzed by western blot (n = 1). (C). The degradation of QSOX2 was mediated by lysosome but not UPS or autophagy. HeLa cells were treated with **C3** (1 μM), **MG132** (20 μM), Chloroquine (20 μM), Bafilomycin A1 (100 nM) and SAR405 (1 μM) for 18 h. DMSO was treated as a control. The protein levels of QSOX2 and β-actin were analyzed by western blot (n=1). (D). QSOX2 can be degraded by **C3_Boc** and **C3_HyT**. HeLa cells were treated with 100 nM, 1 μM and 10 μM of **C3_Boc** or **C3_HyT** for 18 h. DMSO was treated as control. The protein levels of GOLIM4 and β-actin were analyzed by western blot (n = 1).

4.1.2.2 Degradation profile of **C3_Me**, **C3_Boc** and **C3_HyT**.

To study whether other Golgi-resident proteins are affected by **C3** and its derivatives, we investigated the degradation profiles of **C3_Me**, **C3_Boc** and **C3_HyT** via TMT-based proteomics. We totally identified 6556 proteins by proteomics. Among them, 37 proteins are significantly altered: 23 being degraded and 14 being up-regulated by setting the threshold: fold change > 2 and p-value < 0.05 (Figure 4.4 and Table 4.2).
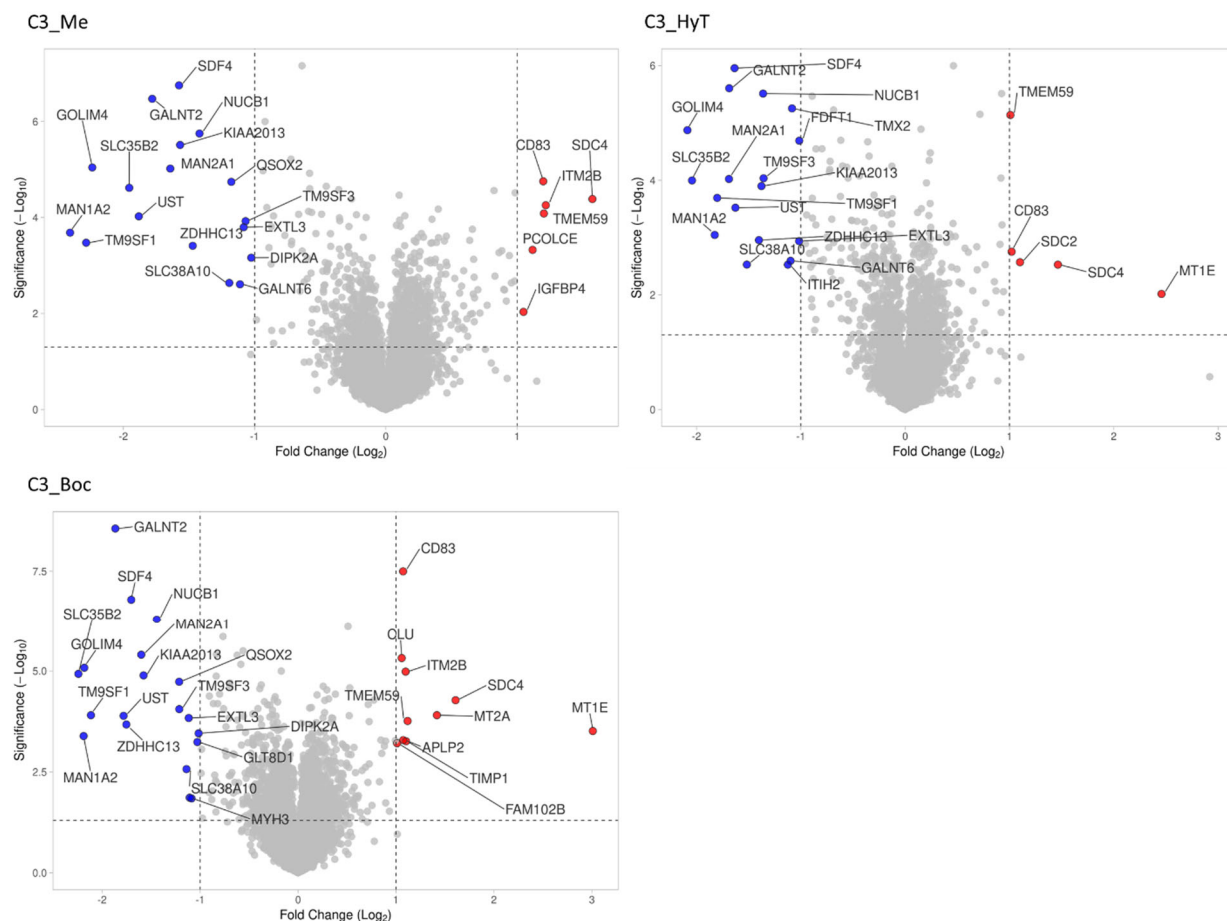
Figure 4.4 Volcano plots indicate the affected proteins under the treatment of **C3** derivatives compared with DMSO-treated samples. HeLa cells were treated with 1 μM **C3_Me**, **C3_Boc** and **C3_HyT** for 18 h. Blue dots indicate the down-regulated proteins and red dots indicate the up-regulated proteins by setting the threshold: fold change > 2 (vertical dashed lines) and p value < 0.05 (horizontal dashed lines). The data were obtained from three independent technical replicates.

This time **C3_HyT**, **C3_Boc,** and **C3_Me** exhibit almost the same degradation profiles with almost all significantly downregulated proteins removed by **C3_HyT**, **C3_Boc** and **C3_Me** (Figure 4.5 A). Comparing the degradation profiles of **C3** with its derivatives, it is also clear that **C3** exhibits a similar degradation profile with **C3_HyT**, **C3_Boc** and **C3_Me** (Table 4.2). Therefore, it is reasonable to postulate that they serve the same functions and may bind to the same targets. Additionally, the same degradation profiles of the positive PROTAC, **C3** and the negative control PROTAC (**C3_Me**) indicate **C3** does not exert its function as a PROTAC.

Additionally, the downregulated proteins induced by **C3** and its derivatives also exhibit similarities. STRING functional enrichment analysis of the degraded proteins reveals an enrichment of proteins localized to the Golgi apparatus (16 of 23) which means **C3** and its derivatives mainly affect the abundance of Golgi-resident proteins (Figure 4.5 B). Moreover, among them, several glycosyltransferases are also affected, including MAN1A2, MAN2A1, GALNT2, EXTL3, GLT8D1 and GALNT6, which reflects the glycosylation process is also affected by **C3** and its derivatives. In addition, several proteins (11 of 23) are

58

also found in coat protein complex I (COPI) vesicles[161]. COPI is mainly responsible for trafficking cargos within Golgi and retrograde trafficking from the Golgi to ER[162]. Therefore, the proteomics data indicates that **C3** and its derivatives also affect COPI-mediated retrograde trafficking.

In summary, proteomics data further indicates that **C3** and its derivatives can lead to the degradation of Golgi-resident proteins, glycosyltransferases and affect the trafficking within Golgi. Therefore, we hypothesized that **C3** interacts with the component of Golgi and then leads to degradation of Golgi proteins and affects the normal functions of Golgi apparatus.
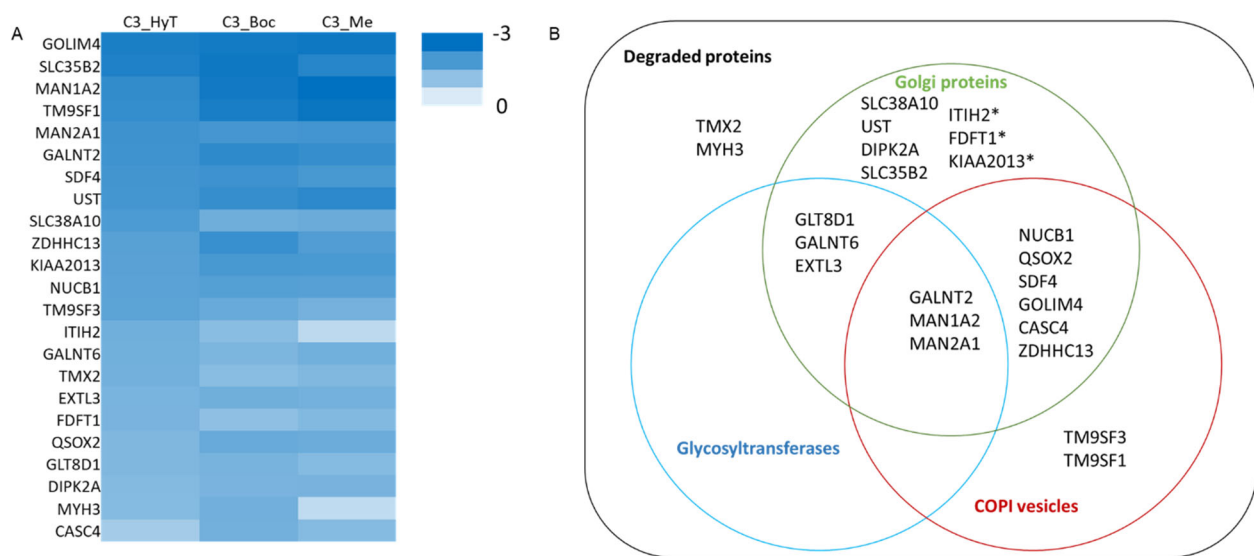
Figure 4.5 Degradation profiles of **C3_Me**, **C3_Boc** and **C3_HyT**. (A). Heat map of proteomics profiles of **C3_Me**, **C3_Boc** and **C3_HyT** at 1 μM. Each column represents different conditions and each row represents the altered protein. The color blue of each region represents this protein is down-regulated under this condition and the white region represents the protein level remains unchanged or the change is not statically significant. (B). The classification of the degraded proteins. The degraded proteins can be classified in to three groups, Golgi proteins (green circle, GO: 0005794; FDR: 8.76e-10. Golgi proteins not being enriched by STRING functional analysis are labelled with star*), glycosyltransferases (blue circle) and COPI vesicles (red circle).

Table 4.2 The significantly altered proteins under the treatment of **C3_HyT**, **C3_Boc** and **C3_Me** at 1 µM after 18 h. The shading values represent that the observation is not statistically significant (p > 0.05). The star * represents proteins can be removed by **C3** at 1 µM after 18 h. The cross X represents this protein is not identified by protemics. The italicized *values* represent the observation is statically significant (p < 0.05) but the protein level change is not significant (fold change < 2).

Gene name	Uniprot ID	Protein name	C3_HyT log$_2$FC	C3_Boc log$_2$FC	C3_Me log$_2$FC	C3 log$_2$FC
GOLIM4*	O00461	Golgi integral membrane protein 4	-2.09	-2.18	-2.24	-1.96
SLC35B2	Q8TB61	Adenosine 3'-phospho 5'-phosphosulfate transporter 1	-2.04	-2.24	-1.95	X
MAN1A2	O60476	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IB	-1.83	-2.19	-2.41	X
TM9SF1	O15321	Transmembrane 9 superfamily member 1	-1.80	-2.11	-2.28	X
MAN2A1	Q16706	Alpha-mannosidase 2	-1.69	-1.60	-1.64	X
GALNT2	Q10471	Polypeptide N-acetylgalactosaminyltransferase 2	-1.69	-1.87	-1.78	X
SDF4	Q9BRK5	45 kDa calcium-binding protein	-1.64	-1.70	-1.58	-1.97
UST	Q9Y2C2	Uronyl 2-sulfotransferase	-1.63	-1.78	-1.88	X
SLC38A10	Q9HBR0	Putative sodium-coupled neutral amino acid transporter 10	-1.52	-1.14	-1.19	X
ZDHHC13	Q8IUH4	Palmitoyltransferase ZDHHC13	-1.40	-1.75	-1.47	X
KIAA2013	Q8IYS2	Uncharacterized protein KIAA2013	-1.38	-1.58	-1.57	X
NUCB1*	Q02818	Nucleobindin-1	-1.36	-1.44	-1.42	-1.41
TM9SF3*	Q9HD45	Transmembrane 9 superfamily member 3	-1.36	-1.21	-1.07	-1.00
ITIH2	P19823	Inter-alpha-trypsin inhibitor heavy chain H2	-1.12	-0.83	-0.26	X
GALNT6	Q8NCL4	Polypeptide N-acetylgalactosaminyltransferase 6	-1.10	-0.99	-1.11	X
TMX2	Q9Y320	Thioredoxin-related transmembrane protein 2	-1.08	-0.83	-0.94	X
EXTL3	O43909	Exostosin-like 3	-1.02	-1.12	-1.08	X
FDFT1	P37268	Squalene synthase	-1.01	-0.77	-0.92	*-0.60*
QSOX2*	Q6ZRP7	Sulfhydryl oxidase 2	-0.96	-1.21	-1.18	-1.26
GLT8D1	Q68CQ7	Glycosyltransferase 8 domain-containing protein 1	-0.94	-1.03	-0.87	X
DIPK2A	Q8NDZ4	Deleted in autism protein 1	-0.88	-1.02	-1.03	X
MYH3	P11055	Myosin-3	-0.87	-1.11	-0.23	X
CASC4	Q6P4E1	Protein CASC4	-0.53	-1.09	-0.87	X
TIMP1	P01033	Metalloproteinase inhibitor 1	0.76	1.10	0.96	X

APLP2	Q06481	Amyloid-like protein 2	0.85	1.07	0.94	X
FAM102B	Q5T8I3	Protein FAM102B	0.91	1.01	0.90	X
CLU	P10909	Clusterin	0.92	1.06	0.99	1.53
IGFBP4	P22692	Insulin-like growth factor-binding protein 4	0.92	0.89	1.05	X
PCOLCE	Q15113	Procollagen C-endopeptidase enhancer 1	0.92	0.99	1.12	X
ITM2B	Q9Y287	Integral membrane protein 2B	0.93	1.10	1.22	X
MT2A	P02795	Metallothionein-2	0.94	1.42	0.24	X
TMEM59	Q9BXS4	Transmembrane protein 59	1.01	1.12	1.20	1.65
CD83	Q01151	CD83 antigen	1.02	1.07	1.20	1.16
SDC2	P34741	Syndecan-2	1.10	0.93	0.90	X
SDC4	P31431	Syndecan-4	1.46	1.61	1.57	X
MT1E	P04732	Metallothionein-1E	2.46	3.01	0.52	X
ITPK1	Q13572	Inositol-tetrakisphosphate 1-kinase	2.92	0.89	1.15	-0.16

4.2 Complexes associated with tethering containing helical rods (CATCHR)

4.2.1 Golgi and its functions

The Golgi apparatus, as the intracellular membranous organelle, functions as a central hub for protein trafficking and the main site of post-translational modifications, especially the glycosylation of proteins and lipids[163]. A large amount of cargos, especially newly synthesized proteins, coming from endoplasmic reticulum (ER) and the ER-Golgi intermediate compartment (ERGIC) will enter the Golgi for future maturation by trafficking through different cisternae and they will be sorted and transported to different destinations (endosome or plasma membrane) when they reach the *trans*-Golgi network (TGN)[164]. This route is defined as anterograde trafficking. Additionally, a large proportion of trafficking also happens in an opposite direction, from endosome to Golgi, which is defined as retrograde trafficking. Retrograde trafficking is essential for receptor transportation, Golgi quality control and cell signaling[165].

Apart from being transported to different destinations, proteins are also subject to post-translational modifications within Golgi, including glycosylation and phosphorylation[166]. There are many glycosylation enzymes located at Golgi, including α-mannosidase II (Man II), polypeptide N-acetylgalactosaminyltransferase 2 (GALNT2) and sialyltransferases (STs)[167]. Glycosylation is essential in protein stability, trafficking, and other normal functions.

4.2.2 General introduction of CATCHR

Complexes associated with tethering containing helical rods (CATCHR), as a member of multi-subunit tethering complexes (MTCs), are a group of proteins and complexes responsible for transporting cargos to different parts within cells[168]. The complexes contain four components (Figure 4.6), the Dsl1 complex, which is responsible for vesicles trafficking from Golgi to ER, the Golgi-associated retrograde protein (GARP) complex, which is responsible for trafficking between endosome and the *trans*-Golgi network (TGN), the conserved oligomeric Golgi (COG) complex which is mainly accountable for retrograde trafficking through the Golgi and the exocyst complex which is responsible for fusion vesicles to the plasma membrane.
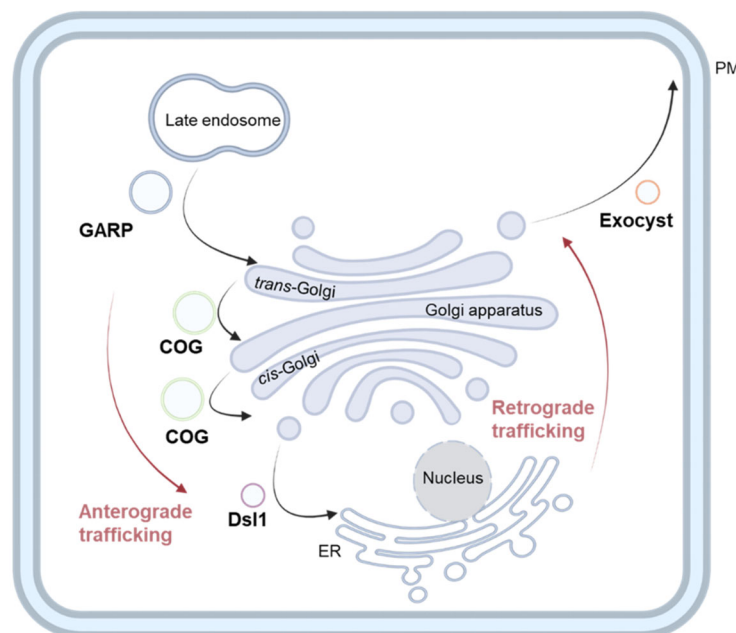

Figure 4.6 CATCHR controls cargo trafficking within cells. Created with BioRender.com.

4.2.3 Conserved oligomeric Golgi (COG) complex

4.2.3.1 Structure and biological functions

The conserved oligomeric Golgi (COG) complex interacts with nearly all trafficking receptors within Golgi[169]. Therefore, it plays an essential role in maintaining the integrity of the Golgi structure, maturation of glycans and trafficking cargos[170]. The complex consists of eight subunits, named COG1-COG8 and these subunits can be categorized into two lobes, COG1-4 (lobe A) and COG5-8 (lobe B), respectively. Among them, COG1 and COG8 form the "bridge" between two lobes by mutual interaction[171]. Lobe A and lobe B are separately localized within Golgi: lobe A is mainly localized on Golgi stacks while lobe B preferentially presents on vesicle-like structures[172].

As previously mentioned, the COG complex is involved in cargo trafficking by direct interaction with Rab proteins[173], which are essential for vesicle formation and tethering, soluble N-ethylmaleimide-sensitive factor activating protein receptors (SNAREs), which are responsible for vesicular fusion[174] and most importantly, COPI-coated vesicles, which transport vesicles via retrograde trafficking pathway. Additionally, the COG complex mediates glycosylation within Golgi by sorting and trafficking glycosyltransferases or affecting the stability of glycosyltransferases[175].

4.2.3.2 Phenotypes of COG depletion

Given the crucial roles of the COG complex, the phenotypes of COG complex defects, including mutations or depletion, will be introduced in this section.

Misglycosylation.

Misglycosylation (deglycosylation or hypoglycosylation) is one of the most representative phenotypes of COG depletion. Actually, several members of the COG complex were discovered because they led to defects in Golgi-associated glycosylation reactions[176]. In addition, mutations in a COG member can lead to human congenital disorder of glycosylation (CDG)[177].

The Lupashin lab found that COG 1 and 2 knockout in Chinese hamster ovary (CHO) cells significantly altered the N-linked glycosylation of plasma membrane-localized glycoconjugates and the glycosylation pattern of GOLIM4 (GPP130)[178]. Similar glycosylation defective patterns could also be observed in COG3[179], 5 and 7-depleted[171] HeLa cells. In COG3 KD HeLa cells, mislocalization of various glycosyltransferases, such as MAN2A1, MGAT1, and GALNT2, can be observed, which may be the reason for abnormal glycosylation patterns in COG-depletion cells.

Protein instability.

In COG 1 and 2 knockout CHO cells, seven Golgi membrane proteins (named GEARs), including MAN2A1 and GOLIM4, were found to be down-regulated[153]. Further experiments proved that their intracellular localizations had been changed. For example, GOS-28, typically located with Golgi, was mislocalized to the cytoplasm and ER.

GOLIM4 and several key glycosylation enzymes, such as beta-1,4-galactosyltransferase 1 (B4GALT1), were also down-regulated in *ΔCOG4* HEK293T cell lines[180].

Trafficking abnormalities.

COG depletion could affect retrograde trafficking, but anterograde trafficking and sorting were not affected. In COG 1 and 2 knockout CHO cells, although the stability of COPI subunits was not affected, the distribution of β and ε –COPI subunits has been significantly altered, which means the retrograde trafficking was affected[153]. Additionally, in COG-depleted HeLa cells, the redistribution of enzymes localized at *trans-*

Golgi to ER was dramatically delayed upon the treatment of Brefeldin A, which also indirectly indicated that the retrograde trafficking was partially blocked[176]. Also, the retrograde of Shiga-toxin was significantly impaired in *ΔCOG3* HeLa cell lines[179].

Golgi fragmentation.

Previous studies demonstrated that Golgi structure was altered in fibroblasts from CDG patients. Additionally, Golgi fragmentation can also be observed in *ΔCOG2* HEK293T cell lines and *ΔCOG3* HeLa cell lines[179].

4.2.4 Golgi-associated retrograde protein (GARP) complex

4.2.4.1 Structure and biological functions
The Golgi-associated retrograde protein (GARP) complex contains four subunits, vacuolar protein sorting 51 (VPS51), VPS52, VPS53 and VPS54[181]. This complex is mainly located in TGN and mediates the fusion of vesicles from late endosome to the TGN.

4.2.4.2 Phenotypes of GARP depletion
GARP depletion will generate similar phenotypes as with COG depletion, including the abnormal glycosylation and degradation of Golgi proteins including GOLIM4, the mislocalization and disability of key glycosylation enzymes including MGAT1 and B4GalT1[181], the change of Golgi morphology and the mislocalization of COPI accessory proteins. Additionally, the mutations of GARP complex will lead to neurodevelopmental diseases in human[182].

4.3 **C3** generates similar phenotypes with COG/GARP knockdown
After summarizing the phenotypes of COG/GARP knockdown, it is interesting to find that **C3** and its derivatives generate similar phenotypes with COG/GARP knockdown. Therefore, it is reasonable to postulate that **C3** may interact with COG/GARP. In this section, more phenotypes of **C3** and its derivatives will be explored and the interaction of **C3_Boc** with COG complex will be further studied.

4.3.1 Glycosylation of Golgi resident proteins is impaired by the treatment of **C3** and its derivatives.
By comparing the western blotting results of GOLIM4 and QSOX2 degradation, it is noticeable that the electrophoretic mobilities of GOLIM4 and QSOX2 are increased when these two proteins are degraded by **C3** and its derivatives (Figure 4.7). The observed band shift on western blotting usually indicates the change of PTMs of the target protein. Based on the phenotype that **C3** and its derivatives lead to the mislocalization and degradation of glycosyltransferases, it is reasonable to postulate that **C3** and its derivatives lead to the misglycosylation of GOLIM4 and QSOX2 and further lead to their degradation.
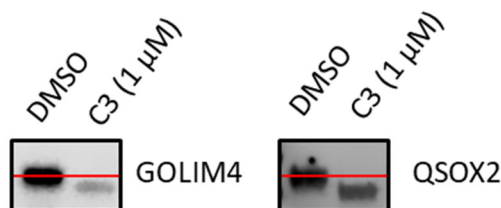

Figure 4.7 The electrophoretic mobilities of GOLIM4 and QSOX2 are increased under the treatment of **C3**. Red line indicates the normal height of GOLIM4 or QSOX2.

Then, the glycosylation state of GOLIM4 and QSOX2 under the treatment of **C3** and its derivatives was further studied. To explore whether the N-linked glycosylation state was affected by **C3** and its derivatives, tunicamycin, a widely used *N*-glycosylation inhibitor and peptide-*N*-glycosidase F (PNGase F), an enzyme capable of removing almost all N-linked glycans were selected. Tunicamycin is a mixture of natural antibiotics but it also acts as an inhibitor of *N*-acetylglucosamine-1-phosphate transferase (GlcNAc phosphotransferase) which is responsible for the transfer of *N*-acetylglucosamine-1-phosphate to dolichol phosphate[183]. The inhibition of this step can block the biosynthesis of N-linked glycoproteins. PNGase F as an amidase can break the bond between the GlcNac and the asparagine residue which cleaves almost all types of N-linked oligosaccharides from glycoproteins but it failed to remove the GlcNac linked to a α1-3-fucose residue (Figure 4.8)[184].
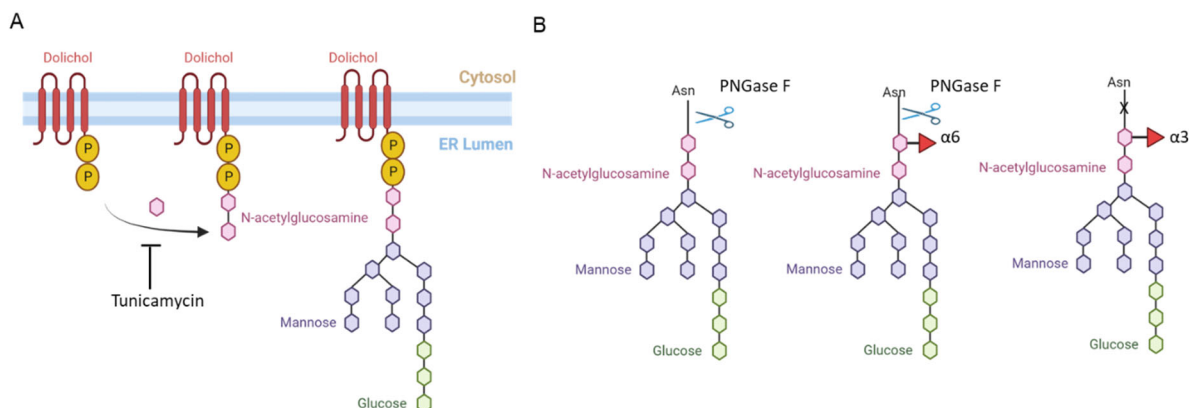
Figure 4.8 The mechanism of tunicamycin and PNGase F in blocking *N*-glycosylation. (A). Tunicamycin inhibits the transfer of *N*-acetylglucosamine-1-phosphate to dolichol phosphate. (B). PNGase F breaks the bond between the GlcNac and the asparagine residue but cannot remove the GlcNac linked to a α1-3-fucose residue. Adapted from "Protein Glycosylation in the ER", by BioRender.com (2022). Retrieved from https://app.biorender.com/biorender-templates

The deglycosylation assay of GOLIM4 and QSOX2 was conducted in two ways. First, HeLa cells were treated with tunicamycin at 5 μg/mL, **C3** at 1 μM and **C3_HyT** at 1 μM for 18 h. Second, denatured whole cell lysate of HeLa was treated with PNGase F for 2 h. To comprehensively compare whether **C3** and its derivatives generate the same deglycosylation type with PNGase F, the lysate of HeLa treated with **C3_HyT** was also treated with PNGase F under the denaturing conditions. Clearly, reduced bands of QSOX2 can be observed under test conditions which indicates the glycosylation state of QSOX2 was successfully affected by tunicamycin and PNGase F (Figure 4.9). Interestingly, each condition generates bands at different heights indicating they do not affect the same type of glycosylation of QSOX2. First, the similar band shift under the treatment of **C3** and **C3_HyT** reconfirms **C3** and its derivatives affect the same glycosylation type of QSOX2 and serve the same functions. Second, compared with treating **C3_HyT** along, the further deglycosylation of QSOX2 under the treatment of **C3_HyT** and PNGase F indicates not all the *N*-glycosylation is affected by the treatment of **C3** and its derivatives. Third, comparing the bands under the treatment of tunicamycin and PNGase F indicates PNGase F failed to remove all *N*-glycosylation of QSOX2. Unfortunately, PNGase F and tunicamycin failed to remove the *N*-glycosylation of GOLIM4 which proves GOLIM4 is insensitive to these tool compounds (Figure 4.9).

Figure 4.9 The glycosylation state of QSOX2 but not GOLIM4 can be affected by tunicamycin (Tuni) and PNGase F. HeLa cells were treated with tunicamycin at 5 μg/mL, **C3** at 1 μM and **C3_HyT** at 1 μM for 18 h. HeLa lysates treated with **DMSO** or **C3_HyT** were further treated with PNGase F for 2 h. The protein levels of GOLIM4 and QSOX2 were analyzed by western blot (n=1).

In summary, based on the phenotype that **C3** and its derivatives lead to the degradation of glycosyltransferases, it is reasonable to postulate that the band shift of GOLIM4 and QSOX2 are due to 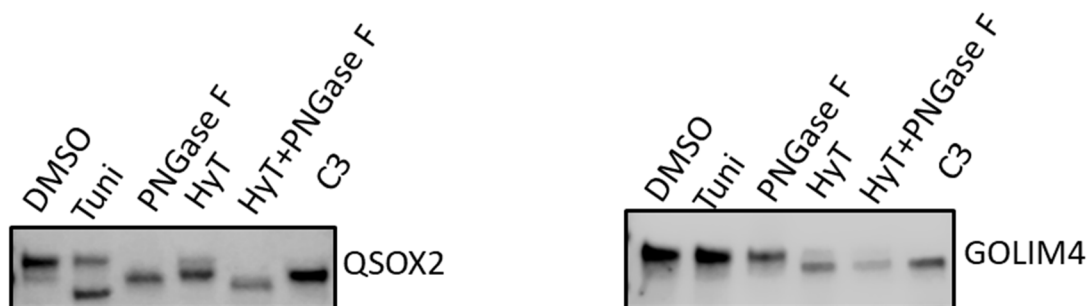misglycosylation. Additionally, the successful removal of the *N*-glycosylation of QSOX2 indicates QSOX2 is a glycoprotein. However, the detailed deglycosylation type led by **C3** and its derivatives cannot be simply determined by the treatment of tunicamycin or PNGase F and the reasons are various. First, the *N*-linked glycosylation of QSOX2 is partially but not completely affected by the treatment of **C3** and its derivatives. Therefore, the bands shift of **C3** and its derivatives is less pronounced compared with tunicamycin or PNGase F. Second, it is also possible that **C3** and its derivatives affect the *O*-glycosylation rather than *N*-glycosylation of QSOX2. Therefore, to study the detailed affected glycosylation type of QSOX2 under the treatment of **C3** and its derivatives, more types of endoglycosidases are needed to generate more types of deglycosylation state of QSOX2. Additionally, glycoproteomics is another useful tool to study the detailed glycosylation types of a glycoprotein[185]. Moreover, although PNGase F and tunicamycin failed to remove N-linked glycosylation of GOLIM4, GOLIM4 has been reported as a heavily glycosylated-protein[136]. Therefore, it is reasonable to postulate GOLIM4 is insensitive to PNGase F and tunicamycin.

4.4 **C3** may interact with COG complex.

In section 4.3, the detailed phenotypes of **C3** and its derivatives have been studied and as mentioned in section 4.2, these phenotypes could also be observed in the depletion of COG complex or GARP complex. Therefore, we postulated **C3** and its derivatives may interact with COG or GARP complex directly. In this section, whether **C3** and its derivatives directly bind to COG complex will be further explored.

Due to the lack of purified proteins from COG complex, the binding of **C3** towards COG complex will be detected via the pull down assay. As mentioned in Chapter 1, the pull down assay is a method for affinity purification and it utilizes an immobilized bait (small molecules, proteins, antibodies) to capture the proteins that can bind to this bait when exposed to the target proteins or whole cell lysate[186]. Therefore, if we fix **C3** onto an immobilized phase and expose this phase to the whole cell lysate, the potential targets of **C3** will be captured. Then the captured proteins will be detected towards a series of COG members' primary antibodies via western blotting to check whether COG members can be enriched by the immobilized phase. The immobilized phase contains two parts. First, a magnetic bead covered with streptavidin and the beads can be simply fixed by a magnet. Second, a biotinylated **C3** probe with biotin on

the one end and the scaffold of **C3** on the other end. The scaffold of **C3** is responsible for the recognition of the potential targets of **C3** and the biotin is responsible for the binding towards streptavidin to fix the probe onto the beads.

The biotinylated **C3** probe, **C3_Bio**, was obtained by replacing the pomalidomide group with biotin and the synthesis of **C3_Bio** was relatively simple and straightforward. Biotin was coupled to **9** via HATU-mediated amide coupling and the the Boc-protection group of **17** was removed by TFA in DCM, and then the free amine was connected to **4** via HATU-mediated amide coupling to generate **C3_Bio** (Figure 4.10).
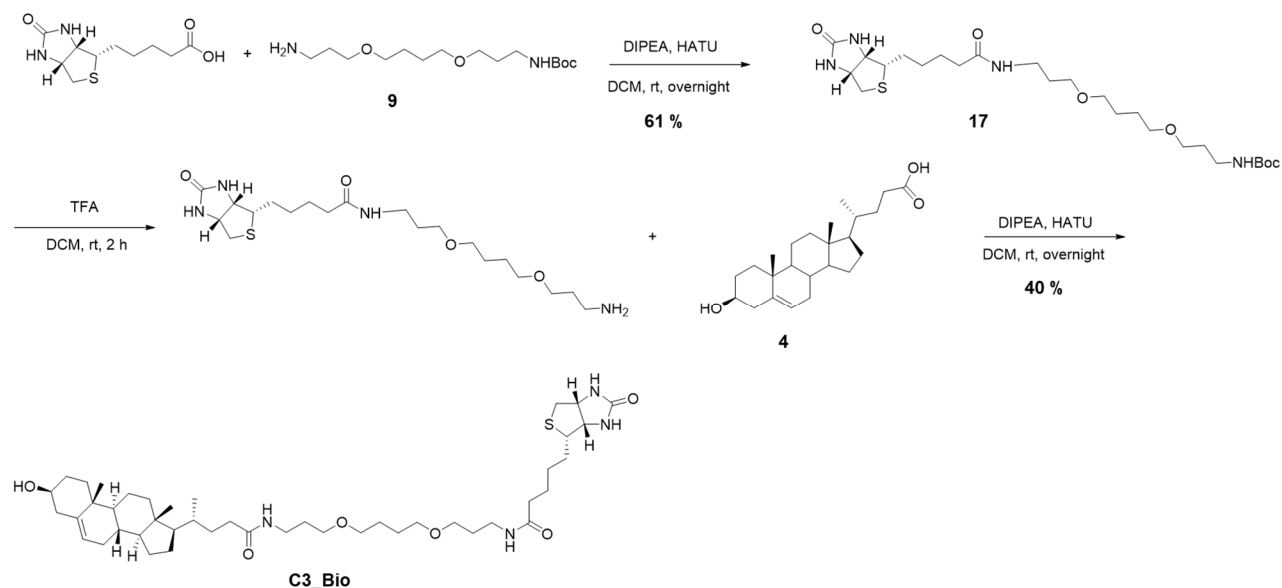
Figure 4.10 The structure and the synthesis routine of **C3_Bio**

The workflow of **C3_Bio** pull down was demonstrated as follows (Figure 4.11). First, the biotinylated probe **C3_Bio** was incubated with the commercially available streptavidin magnetic beads. The binding affinity of biotin towards streptavidin is extremely high with a $K_d \approx 10^{-14}$ mol/L so the interaction of biotin and streptavidin can only be removed by harsh conditions such as boiling in denaturing buffer. Therefore, by incubating **C3_Bio** with the streptavidin beads, the scaffold of **C3** was fixed onto an immobilized phase. Then the immobilized phase was incubated with the whole cell HeLa lysate to capture the potential targets. After the enrichment, the captured proteins were released by boiling in SDS-denaturing buffer and detected by western blotting.
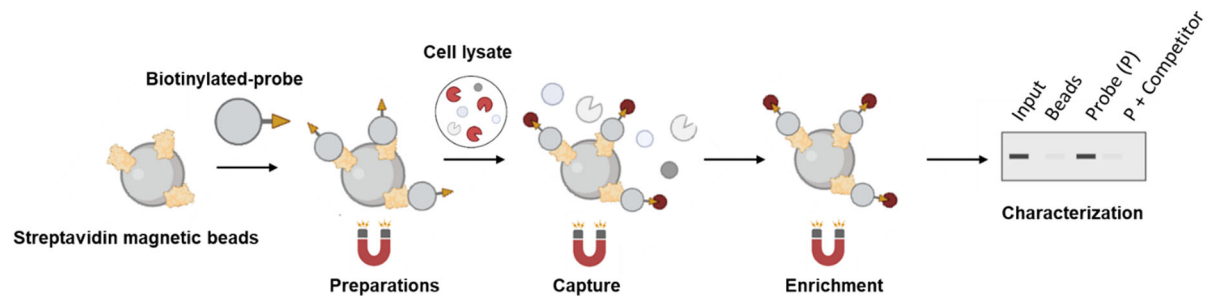
Figure 4.11 The workflow of biotin and streptavidin-based pull down.

For the first screening, we detected whether COG1-4, COG6 and GOLIM4 can be captured by the probe. It is interesting to find that COG6 can be enriched by **C3_Bio** while no binding to COG1-4 can be observed (Figure 4.12 A). Additionally, GOLIM4 cannot be enriched by **C3_Bio** either which reconfirms GOLIM4 cannot bind to **C3** directly. However, the binding of COG 6 towards **C3_Bio** cannot be blocked by 30 μM of **C3_Boc** and this may due to the precipitation issue at the high concentration. Therefore, the concentration of **C3_Boc** was decreased to 5 μM to compete for the binding between **C3_Bio** and COG6. Clearly, OG6 can still be enriched by **C3_Bio** and this time the binding can be blocked by **C3_Boc** (Figure 4.12 B). However, the beads themselves also bind to COG 6. Therefore, it is difficult to exclude the possibility that the COG 6 can be enriched by the bead rather than **C3_Bio**.
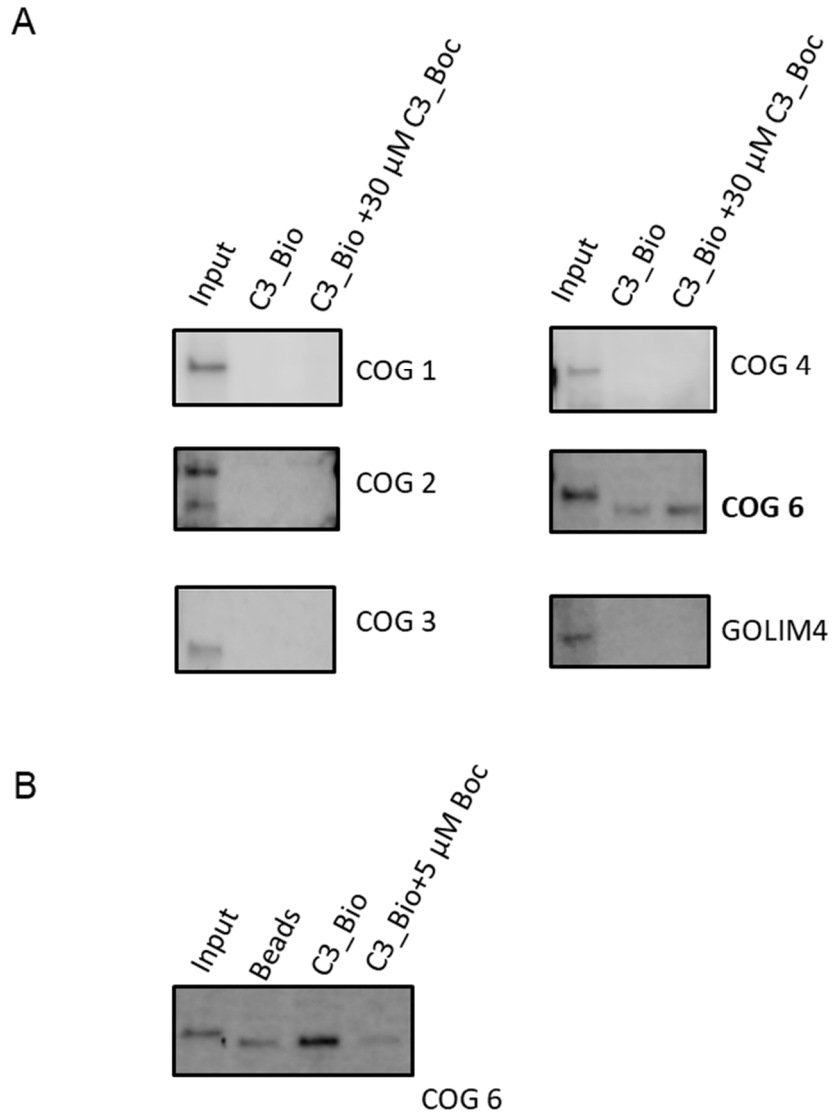
Figure 4.12 Only COG6 can be enriched by **C3_Bio**. (A). Only COG6 can be enriched by **C3_Bio** but the enrichment of COG6 cannot be blocked by 30 μM of **C3_Boc**. (n=2) (B). COG6 can be enriched by **C3_Bio** and the enrichment of COG6 can be blocked by 5 μM of **C3_Boc** but COG6 can also be enriched by the beads. (n=2)

Then, CETSA with western blotting read-out was conducted to check whether **C3** and its derivatives bind to COG6. We exposed whole HeLa cell lysate to **C3_HyT** (10 μM) or DMSO for 10 min. Then, the mixture was subjected to a thermal denaturation gradient from 37.5 °C to 63.0 °C and the protein level changes were analyzed by western blot. Clearly **C3_HyT** could only weakly stabilize COG 6 (Figure 4.13). Therefore, more *in vitro* experiments are needed to confirm the binding of **C3_HyT** towards COG6.

	37.5	43.2	46.1	47.5	49.3	51.6	53.9	58.4	63.0	(°C)
Relative abundance	1.00	0.96	0.67	0.67	0.36	0.22	0.18	0	0	COG6
Relative abundance	1.00	0.65	0.48	0.40	0.30	0.21	0.15	0.07	0	DMSO

HeLa
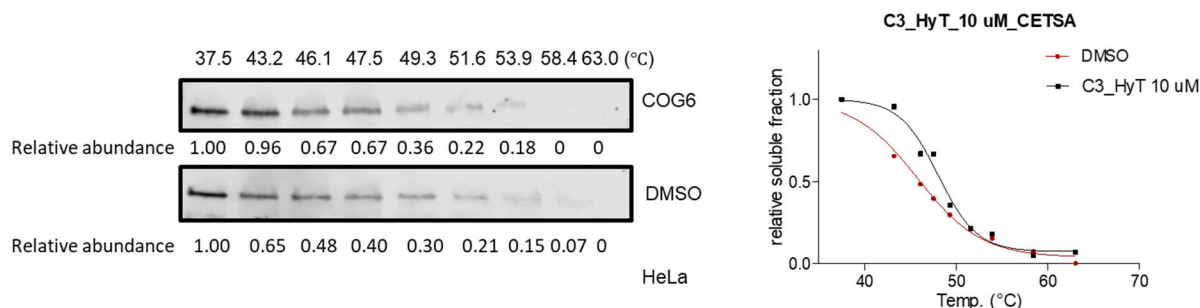
Figure 4.13 **C3_HyT** exhibits a weak stabilization effect on COG6. CETSA experiments for **C3_HyT** binding to GOLIM4. The protein band intensities were normalized to samples heated to 37.5 °C (n=2).

4.5   **C3** derivatives bind to OSBP

While I was studying the interaction between **C3** and COG complex, the binding of **C3** derivatives towards a series of sterol transport proteins (STPs) was also determined to check whether the observed phenotypes are generated by the binding to STPs. This was carried out by Dr. Laura Depta, a post-doc in the group. Due to the intrinsically fluorescence of pomalidomide which will interfere with the FP assay, only the binding abilities of **C5**, **C3_Boc** and **C3_HyT** towards STPs were studied. First, we detected the binding affinity of **C5**, **C3_Boc** and **C3_HyT** towards Aster proteins. As mentioned in Chapter 1, Aster proteins (Aster-A, B and C) are a family of STPs that are responsible for trafficking cholesterol from PM to ER. Obviously, **C5**, **C3_Boc** and **C3_HyT** failed to target Aster proteins at 10 µM (Figure 4.14 A).

Then, we detected the binding affinity of **C3_Boc** and **C3_HyT** towards OSBP. OSBP is also a member of STPs responsible for exchanging cholesterol and PI4P at ER-Golgi membrane contact sites[187]. We studied the dose-response of **C3_Boc** and **C3_HyT** towards OSBP inhibition in a fluorescence polarization (FP) assay monitoring the inhibition of binding of 22-NBD-Cholesterol to the recombinant sterol binding domain of OSBP. Interestingly, both **C3_Boc** and **C3_HyT** exhibit high affinity towards OSBP. **C3_Boc** can inhibit OSBP with IC_{50} of 73 nM and **C3_HyT** can inhibit OSBP with IC_{50} of 141 nM. In this assay, **OSW-1**, an established OSBP inhibitor, can bind to OSBP with IC_{50} of 28 nM and 25-hydroxycholesterol (25-HC), an endogenous OSBP regulator can bind to OSBP with IC_{50} of 195 nM (Figure 4.14 C). Therefore, **C3** its derivatives can target to OSBP and whether the phenotypes generated by C3 are due to OSBP inhibition should also be explored.

Figure 4.14 **C3** derivatives can selectively inhibit OSBP. n = 3, data is mean ± sem. Experiments carried out by Dr. Laura Depta, a post-doc in the group. (A). **C5** and **C3** derivatives cannot inhibit binding of 22-NBD-Cholesterol (20 nM) to the ASTER domain of the Aster proteins. (B). **C3_Boc** and **C3_HyT** exhibit potent inhibition of OSBP.

4.6 Summary, discussion and outlook

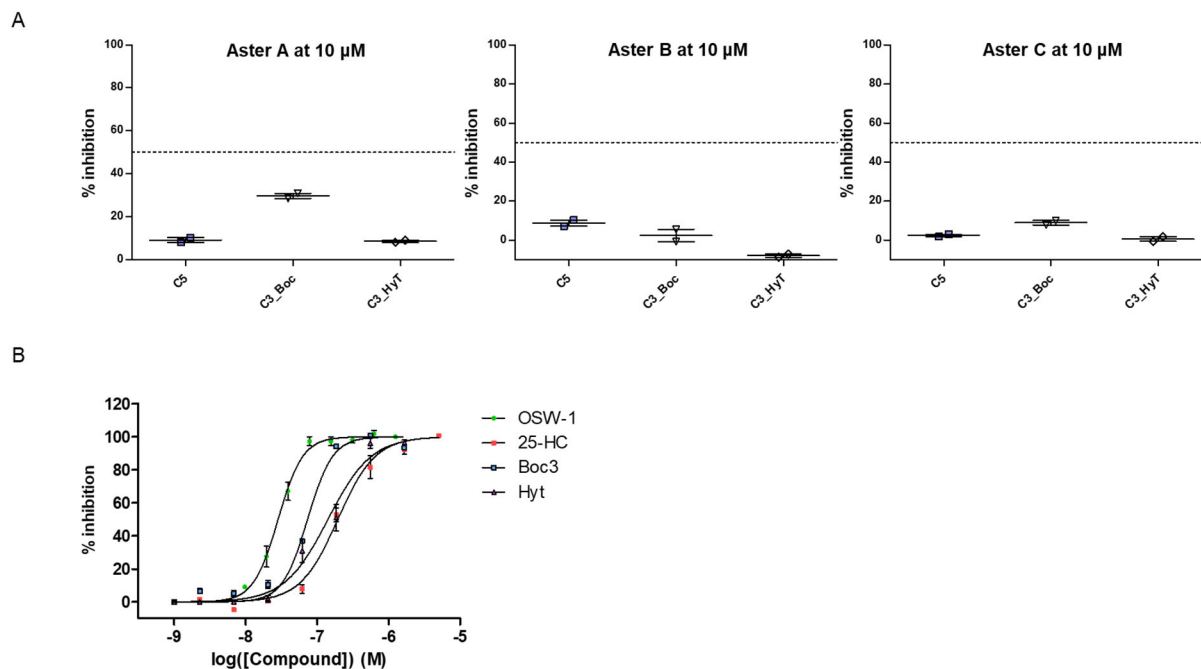In Chapter 3, I have shown that CRBN and UPS are not involved in the degradation of GOLIM4. Therefore, in this chapter, the detailed mechanism of GOLIM4 degradation was explored. First, I optimized **C3** into smaller GOLIM4 degraders by removing the pomalidomide part and obtained two **C3** derivatives, namely **C3_Boc** and **C3_HyT**. Interestingly, we found **C3_Boc** was enough to trigger GOLIM4 degradation. In the meantime, we also discovered that **C3** and its derivatives could also degrade another Golgi resident protein (QSOX2) via the same mechanism. Since we have already proved **C3** does not directly bind to GOLIM4, we postulated that **C3** may interact with the component of Golgi and lead to its destabilization and further lead to the degradation of Golgi resident proteins.

To systemically study the proteins affected by **C3** derivatives, the degradation profiles of **C3** derivatives were also obtained by TMT-based proteomics. The proteomics data indicate that **C3** derivatives exhibit similar degradation profiles which indicates they may serve the same functions. Then the interpretation of the affected proteins indicates that the degraded proteins can be categorized into three groups, Golgi resident proteins, glycosyltransferases and COPI vesicles. Then it is exciting to find that the depletion of COG and GARP complex will generate similar phenotypes with the treatment of **C3**. Therefore, more phenotypes of **C3** were explored.

First, it is noticeable that GOLIM4 and QSOX2 exhibit clear bands shift under the treatment of **C3** which indicates the PTMs of these proteins are affected. Based on the phenotype that **C3** leads to the degradation of glycosyltransferases, we postulated that the glycosylation state of GOLIM4 and QSOX2 are affected under the treatment of **C3** and its derivatives and the affected glycosylation of QSOX2 and GOLIM4 leads to the increase of their electrophoretic mobilities. Then, to further explore the affected glycosylation state of GOLIM4 and QSOX2 under the treatment of **C3** and its derivatives, two established tool compounds-tunicamycin and PNGase F were chosen to study whether N-linked glycosylation of GOLIM4 and QSOX2 was affected. First, it is clear that the glycosylation state of QSOX2 but not GOLIM4 was successfully affected by PNGase F and tunicamycin. Given that GOLIM4 has been proved as a heavily glycosylated protein, this protein may resist the treatment of PNGase F or tunicamycin. Then, it is interesting to find that the glycosylation states are differently affected by **C3_HyT**, PNGase F and tunicamycin which indicates **C3** and its derivatives may partially affect the N-linked glycosylation or even affect the O-linked glycosylation of QSOX2. To further study the affected glycosylation of QSOX2, more types of endoglycosidases are needed to remove different glycosylation modifications of QSOX2 for comparison. Additionally, glycoproteomics will be another useful tool to study the glycosylation of QSOX2 under different conditions.

Since **C3** and COG/GARP depletion generate similar phenotypes, we first studied whether **C3** interacts with COG complex directly. Proteomics data indicate that COG complex can be identified but no degradation of COG complex can be observed. Therefore, to study whether **C3** can interact with COG complex directly, a pull-down assay based on biotin and streptavidin was conducted. First, pull-down assay reconfirmed **C3** cannot bind to GOLIM4. Then, pull-down experiment showed COG6 can be enriched by the biotinylated **C3** probe, namely **C3_Bio** and the enrichment can be inhibited by the co-incubation with 5 μM **C3_Boc**. To further confirm the binding between COG6 and **C3_HyT**, the CETSA assay towards COG6 was conducted. However, the observed stabilization effect of **C3_HyT** towards COG6 was not significant. Based on the fact that the bead itself can enrich COG6 and **C3_HyT** only weakly stabilizes COG6, we cannot certainly conclude that **C3** and its derivatives directly interact with COG6 and therefore more experiments need to be conducted to confirm the binding of **C3** with COG6. Additionally, there are total of 8 members of a COG complex and we only tested 5 of them (COG 1-4, COG 6). Therefore, whether COG 5, 7 or 8 interacts with **C3** and its derivatives will be further studied.

It is noteworthy to mention that although other COG members cannot be enriched by **C3_Bio**, we cannot exclude the possibility that the tested COG members may still interact with **C3** and its derivatives. This is because of the lack of a real 'linker' part within the structure of **C3_Bio** and therefore, the steric clashes between the streptavidin and the potential target may disrupt its binding towards **C3_Bio**. The extended glycol moiety within **C3_Bio** should not be regarded as a linker of the probe but as the scaffold of **C3**. Therefore, in the next step, an extended linker will be attached to **C3_Boc** to avoid these steric clashes between streptavidin and the potential targets.

Finally, because we still fail to figure out the targets of **C3** and its derivatives, we wanted to check whether **C3** derivatives target STPs and generate these phenotypes. Therefore, we detected the binding affinity of **C3_HyT** and **C3_Boc** towards different STPs including Aster proteins and OSBP. A fluorescence polarization assay clearly showed **C3** derivatives cannot target Aster proteins but interestingly, both **C3_HyT** and **C3_Boc** can inhibit OSBP in the nanomolar range, which means it is also possible that the phenotypes generated by **C3** and its derivatives are due to the binding towards OSBP. Currently, it has been well established that the inhibition of OSBP by the complex natural product **OSW-1** can lead to the mislocalization of OSBP from ER/Golgi membrane contact sites to an area localized at Golgi and the degradation of OSBP in HCT116[188] and HUVECs cells[189]. Moreover, **OSW-1** also leads to the fragmentation of Golgi[188] and affects the key lipid pools including elevating the level of cellular cholesteryl ester (CE), ceramides and sphingomyelins (SMs) and the reduction of phosphatidylcholines (PCs), phosphatidylethanolamines (PEs) and phosphatidylserines (PSs)[189]. Additionally, another known OSBP inhibitor, schweinfurthin G (**SWG**), which can bind to OSBP in nanomolar range, also affects the localization of OSBP in RPE-1 cells and further affects the morphology of TGN and Golgi trafficking[190]. Therefore, further experiments are needed to investigate the relationship between OSBP inhibition and the phenotypes of **C3**.
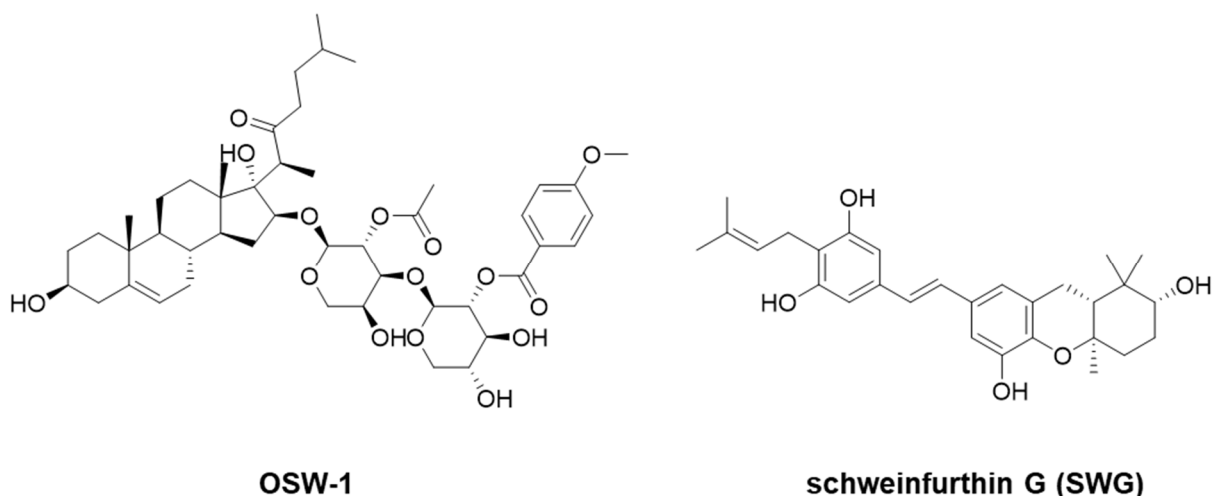
Figure 4.15 Known OSBP inhibitors.

In conclusion, we started from designing cholesterol-bearing PROTACs but eventually found that the detected protein degradation is not due to the direct binding towards these proteins but the downstream effect of inhibition towards COG complex or OSBP. These results demonstrate the pitfalls of applying PROTACs for target identification. First, it is difficult to determine the mechanism of the observed protein degradation. The ideal situation is the PROTACs bind to these proteins and then lead to their degradation

via the established mechanism. However, we cannot exclude the possibility that PROTACs may only function as the inhibitors and the inhibition of a protein further leads to the observed protein degradation. In addition, especially for CRBN-recruiting PROTACs, it is worth mentioning that pomalidomide, lenalidomide and thalidomide themselves function as molecular glues and lead to the degradation of specific proteins such as GSPT1 and ARID2. Therefore, it is necessary to check whether the degraded proteins are the neo-substrates of these immunomodulatory drugs. Moreover, it is also possible that PROTACs affect the transcript expression levels of these proteins and further affect the expression level of these proteins. Therefore, it is difficult to discriminate the observed degradation is a direct binding event towards the PROTAC or a secondary event. However, this shortcoming can be easily circumvented by comparing the degradation profiles of a positive PROTAC, which can still bind to the chosen E3 ligase and a negative PROTAC, which the binding toward E3 ligase has been blocked. Take pomalidomide-based PROTACs as an example. By comparing the degradation profiles of a pomalidomide-based PROTAC (**C3**) and a methylated pomalidomide-based PROTAC (**C3_Me**), the proteins removed by both can be regarded as the downstream events and the proteins only degraded by the positive PROTAC can be regarded as the potential targets.

Second, the intrinsic selectivity of PROTACs further hinders its application in target identification because the binding of a PROTAC does not promise a degradation event. As aforementioned, different linker parts (length or types), different choices of E3 ligases and different linker attachment positions will all affect the stability of the POI-PROTAC-E3 ternary complex and further affect the degradation selectivity and efficiency. Although a PROTAC can bind to a series of targets, only limited numbers of proteins can be degraded and further detected by proteomics. Therefore, to study the targets of a ligand comprehensively, a series of PROTACs bearing different linkers, E3 ligands are needed.

Third, the test conditions, especially the concentration and the incubation time of a PROTAC will also affect the degradation profiles of a PROTAC. Numerous examples demonstrate that PROTACs at a higher concentration will lead to the hook effect which halts the protein degradation so the protein levels remain stable. However, it is difficult to determine the effective concentration of a PROTAC towards each protein. Therefore, to study the targets of a ligand more comprehensively, it is necessary to test different concentrations of a given PROTAC. In addition, incubation time also affects the degradation efficiency of a PROTAC and longer incubation time (18 h) will result in a series of downstream events that we cannot discriminate from the proteomics data. Therefore, a short incubation time seems more suitable to exclude the downstream effect of a PROTAC but the threshold for the significant degradation should be relaxed as the degradation effect may be unnoticeable in a short incubation period.

In the end, applying PROTACs for target identification may provide a convenient way to simultaneously identify targets and study the functions of the targets. However, it is better to combine this method with other methods including cell painting, pull down or photo-affinity labeling for target identification.

5   Chapter 5 Design, synthesis, and biological evaluation of PROTACs targeting Aster-A

5.1   Introduction to Aster family proteins

5.1.1   Structures and functions of Aster family proteins.

Aster/GRAMD1 proteins are a highly conserved family of integral membrane proteins anchored to the ER. This family contains three members, Aster-A (GRAMD1A), Aster-B (GRAMD1B), and Aster-C (GRAMD1C). Their structures comprise an N-terminal GRAM domain, a StART-like domain, and a C-terminal transmembrane domain (Figure 5.1)[96]. Aster proteins are mainly localized in ER during the steady state and will be recruited to the ER-PM contact sites when the levels of accessible cholesterol are elevated in the PM.
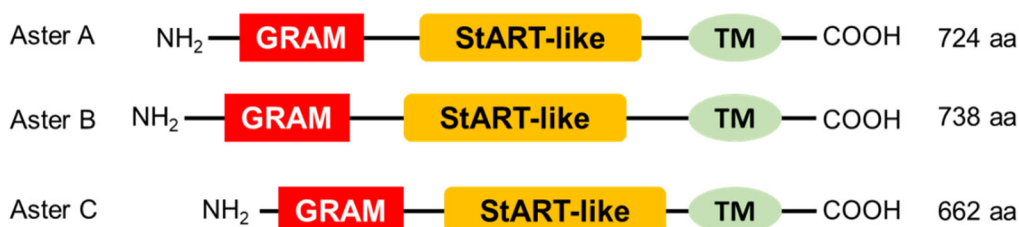
Figure 5.1 Schematics of Aster proteins structures. The amino acid lengths are indicated.

The GRAM domain contains a typical pleckstrin homology (PH) domain, which is usually responsible for regulating organelle contacts. Take the GRAM domain of Aster-B, which is the most studied GRAM domain within Aster family proteins, as an example[96]. Modeling studies based on the Lam6 (homolog of Aster-B in yeast) GRAM domain reveal that the Aster-B GRAM domain contains a basic patch[191], which plays a role in sensing anionic lipids in membranes, and a distinct site near the basic patch, which functions as a sensor for accessible cholesterol. The two distinct sites within the Aster-B GRAM domain collectively contribute to recognizing anionic lipids and accessible cholesterol within the PM.

The StART-like domain is one of the critical components within Aster family proteins, and this domain is responsible for transporting lipids, including cholesterol. Based on the crystal structure of Aster-A StART-like domain (Figure 5.2)[96], there is a hydrophobic groove consisting of a highly curved 7-stranded beta-sheet responsible for sterol binding and a long C-terminal helix, together with two short N-terminal helices close this hydrophobic cavity.
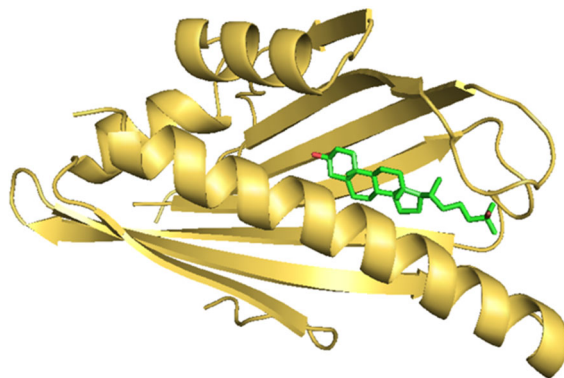
Figure 5.2 Crystal structure of mouse Aster-A StART-like domain with 25-hydroxy cholesterol (PDB: 6GQF[96]).

Besides their common roles in cholesterol sensing and trafficking, each Aster protein also exhibits distinct functions. This can be expected from their distinct tissue expression patterns: Aster-A is highly expressed in the central nervous system, especially the brain, Aster-B is highly expressed in endocrine tissues, especially parathyroid and thyroid, and Aster-C is highly expressed in the liver and testes[96]. Aster-A also plays a role in autophagy and this function is discovered by the identification of an Aster-A selective inhibitor, **Autogramin-2**[97]. This molecule can block the transportation of cholesterol between ER and inhibit autophagosome initiation, suggesting the role of Aster-A in autophagy. Aster-B is responsible for trafficking HDL-derived cholesterol from the PM to ER for further steroidogenesis in steroidogenic tissues and therefore its deficiency in mice will lead to impaired adrenal cholesteryl ester storage, decreasing levels of serum corticosterone. Additionally, Aster-B contains a distinct mitochondrial targeting sequence, which tethers mitochondrial with the ER[192]. Therefore, Aster-B also mediates cholesterol transfer from ER to mitochondrial and the ablation of Aster-B will lead to mitochondrial dysfunction. Aster-C is also found involving in interaction with mitochondria via its GRAM domain recently and the depletion of Aster-C leads to the elevated level of mitochondrial cholesterol and increased mitochondrial bioenergetics[193]. In addition, similar to Aster-A, Aster-C has been reported to also function as a negative regulator of starvation-induced autophagy.

5.1.2 Small molecule regulators of Aster family proteins

Autogramins[97] are a series of autophagy inhibitors discovered by phenotypic screening for inhibition of autophagosome formation. Further proteomics studies demonstrate **Autogramins** exert their autophagy inhibition abilities by targeting Aster-A and these molecules exhibit marked selectivity towards Aster-B, Aster-C and other unrelated sterol binding proteins such as STARD1 or LXR-β. Fluorescent polarization experiments proved that **Autogramin-2** (structure shown in Figure 5.3), one of the derivatives of **Autogramins**, successfully compete with the binding of 22-NBD-cholesterol to Aster-A StART domain with $IC_{50} = 349 \pm 51$ nM. Docking studies further confirm that **Autogramin-2** and cholesterol bind to the same region on Aster-A. Functional studies demonstrate **Autogramin-1**, another derivate of **Autogramins**, impairs the formation of phagophore in response to starvation.

Autogramin-1   Autogramin-2
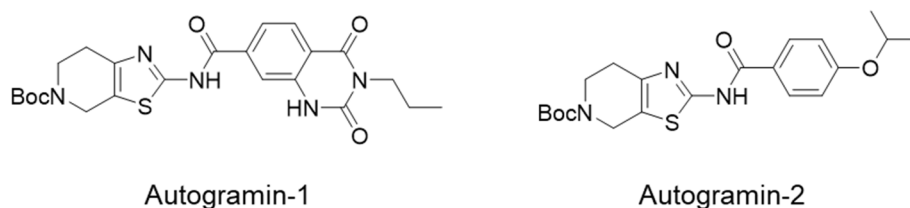
Figure 5.3 Chemical structures of Autogramins.

A series of sterol analogs including 20α-hydroxycholesterol (20α-HC) and **U18666A** also exhibit inhibition activity towards Aster family proteins (structures shown in Figure 5.4 A)[194]. 20α-HC can bind to all Aster proteins in the micromolar range (Aster-A: $IC_{50} = 0.87$ µM, Aster-B: $IC_{50} = 0.21$ µM, and Aster-C: $IC_{50} = 1.41$ µM). **U18666A**, an established inhibitor of Niemann-Pick C1 (NPC1), could also bind to Aster proteins with poor selectivity (Aster-A: $IC_{50} = 2.33$ µM, Aster-B: $IC_{50} = 0.52$ µM, and Aster-C: $IC_{50} = 9.50$ µM) and inhibit the ability of Aster proteins to traffick cholesterol between membranes *in vitro*. Further optimization of **U18666A** and 20α-hydroxycholesterol gave **AI-3d** and **AI-1l**. **AI-3d** acting as a pan Aster proteins inhibitor (Aster-A: $IC_{50} = 0.11$ µM, Aster-B: $IC_{50} = 0.06$ µM, and Aster-C: $IC_{50} = 0.71$ µM) and **AI-1l** could selectively binds to Aster-C with $IC_{50} = 0.85$ µM. However, recently, **AI-1l** has also been reported to inhibit the activity of sterol regulatory element binding protein 2 (SREBP2) and inhibit hedgehog signaling.
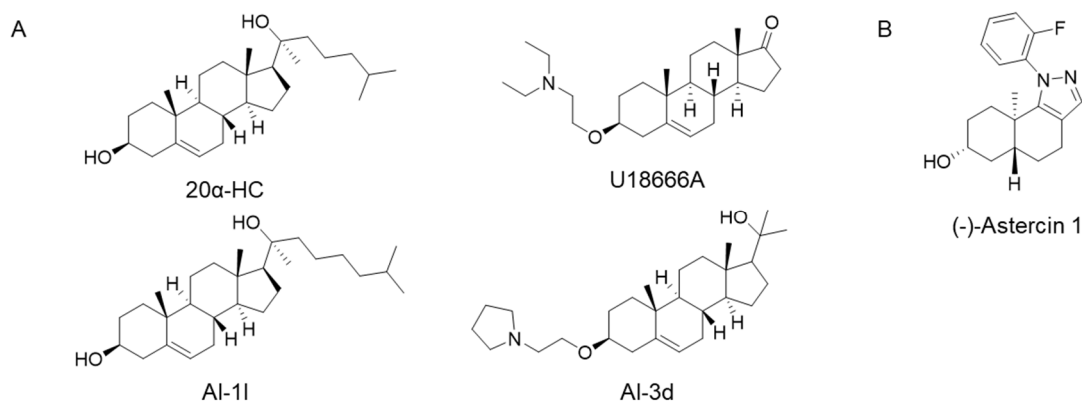
Figure 5.4 Chemical structures other Aster proteins inhibitors. (A) Chemical structures of 20α-hydroxycholesterol and its derivatives. (B) Chemical structures of **(-)-Astercin 1**.

Based on screening a sterol-inspired compound library which is established by fusion between a steroidal scaffold and a series of heterocyclic scaffolds, **(-)-Astercin 1** (Figure 5.4 B) was obtained as a selective Aster-C inhibitor in our lab ($IC_{50} = 1.51 \pm 0.56$ µM)[195]. It exhibits more than 30-fold selectivity over Aster-B with $IC_{50} = 46.8 \pm 5.78$ µM and over 50-fold selectivity over Aster-A with $IC_{50} > 80$ µM. *In vitro* FRET-based cholesterol transfer assay confirmed **(-)-Astercin 1** could block Aster-mediated sterol transfer.

5.2 Project outline

Aster-A exerts crucial functions in both cholesterol transporting and autophagy. However, current small-molecule tools targeting Aster-A all function as inhibitors and only focus on occupying the StART-like domain of Aster-A. The functions of other domains within Aster-A are poorly interrogated by small molecules.

Compared with traditional small-molecule tools, which only inhibit the catalytic function of a specific protein, PROTACs can remove all functions (non-enzymatic, scaffolding, etc.) related to this protein by degradation via the UPS. Additionally, the degradation ability of a PROTAC does not rely on the binding to the functional catalytic/enzymatic domain of this protein but only depends on a transient interaction with this protein at any domains. Moreover, once the proteins are degraded, PROTACs will be released and engaged in a new cycle of protein destruction. This catalytic nature enables PROTACs to potentially exert improved and diverse activity compared with inhibition.

In this chapter, based on the advantages of PROTACs on protein degradation, removal of proteins' non-catalytic functions and catalytic nature, I will focus on applying the established Aster-A inhibitor, **Autogramin-2** into PROTACs to obtain the first series of Aster-A PROTACs. These will be used as tools to study the functions of Aster-A. Based on the above discussion, this project will be conducted in the following steps.

1. Aster-A PROTACs will be obtained by connecting **Autogramin-2**, an Aster-A selective inhibitor with known binding mode towards Aster-A, with pomalidomide, a commonly used E3 ligand targeting CRBN via different linkers and their degradation abilities will be tested in the established HeLa cells.

2. A series of SAR studies of Aster-A PROTACs will be carried out based on exploring the preference of E3 ligases, suitable linkers and the suitable cell lines for Aster-A detecting.

5.3 Design and synthesis of CRBN-recruiting Aster-A PROTACs

5.3.1 Design of CRBN-recruiting Aster-A PROTACs

Aster-A PROTACs will follow the similar design rules introduced in section 2.1. Briefly, the Aster-A PROTACs should consist of three components: an Aster-A inhibitor that binds to the target protein, an E3 ligand, and a linker connecting the two parts.

We chose **Autogramin-2** as the Aster-A ligand. **Autogramin-2** is a selective Aster-A inhibitor with K_i=290 nM determined by fluorescence polarization[97]. The homology model of the Aster-A StART domain with docked **Autogramin-2** indicates that the isopropoxy group of **Autogramin-2** is exposed to the solvent (Figure 5.5 A). Therefore, it is possible to replace this group with a carboxylic acid group for further linker attachment. The similar design strategy has also been applied to the design of BODIPY–autogramin (structure shown in Figure 5.5 B). This probe retains the binding capacity towards the StART domain of Aster-A with a K_d of 49 ± 12 nM determined by fluorescence polarization. We still chose pomalidomide as the E3 ligand and the same linkers used in CHO-PROTACs to design the first series of Aster-A PROTACs (structures shown in Figure 5.5 C).

Compound	Linker	Number of non-H atoms
NGF1		6
NGF2		8
NGF3		12
NGF4		13
NGF5		10
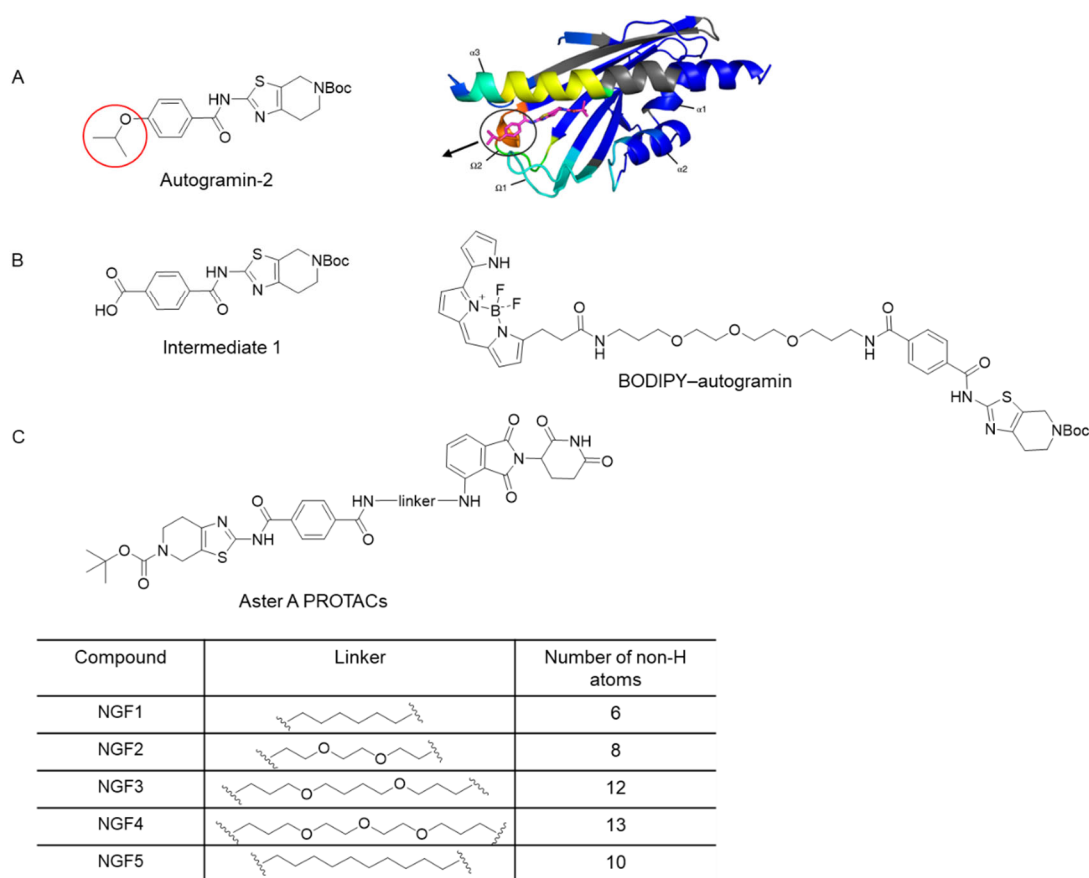
Figure 5.5 Rational design of Aster-A PROTACs. (A) Chemical structure of **Autogramin-2** and the structure of homo models of the Aster-A StART domain with docked **Autogramin-2**. The isopropoxy moiety is circled in red (left) and black (right) and the black arrow indicates the exit vector for further linker connection. (B) Chemical structures of **intermediated 1** (compound **21**) and BODIPY-autogramin. (C) Chemical structures of Aster-A PROTACs, **NGF1-NGF5** and the number of non-H atoms

5.3.2 Synthesis of CRBN-recruiting Aster-A PROTACs

The synthesis of Aster-A PROTACs was carried out as follows. For the **Autogramin-2** part (compound **21**), commercially available N-Boc piperidone was treated with 5,5-dibromobarbituric acid, resulting in the formation of bromopiperidinone. Then condensation reaction was performed between the bromopiperidinone and thiourea to give the desired aminothiazole. Then the amine was coupled to mono-methyl terephthalate via HATU-mediated amide coupling, and the resulting ester was hydrolyzed into **21** by NaOH in H_2O/THF. The synthesis of the pomalidomide part has been demonstrated in section 2.1.1, and the final PROTACs were obtained by HATU-mediated amide coupling (Figure 5.6).
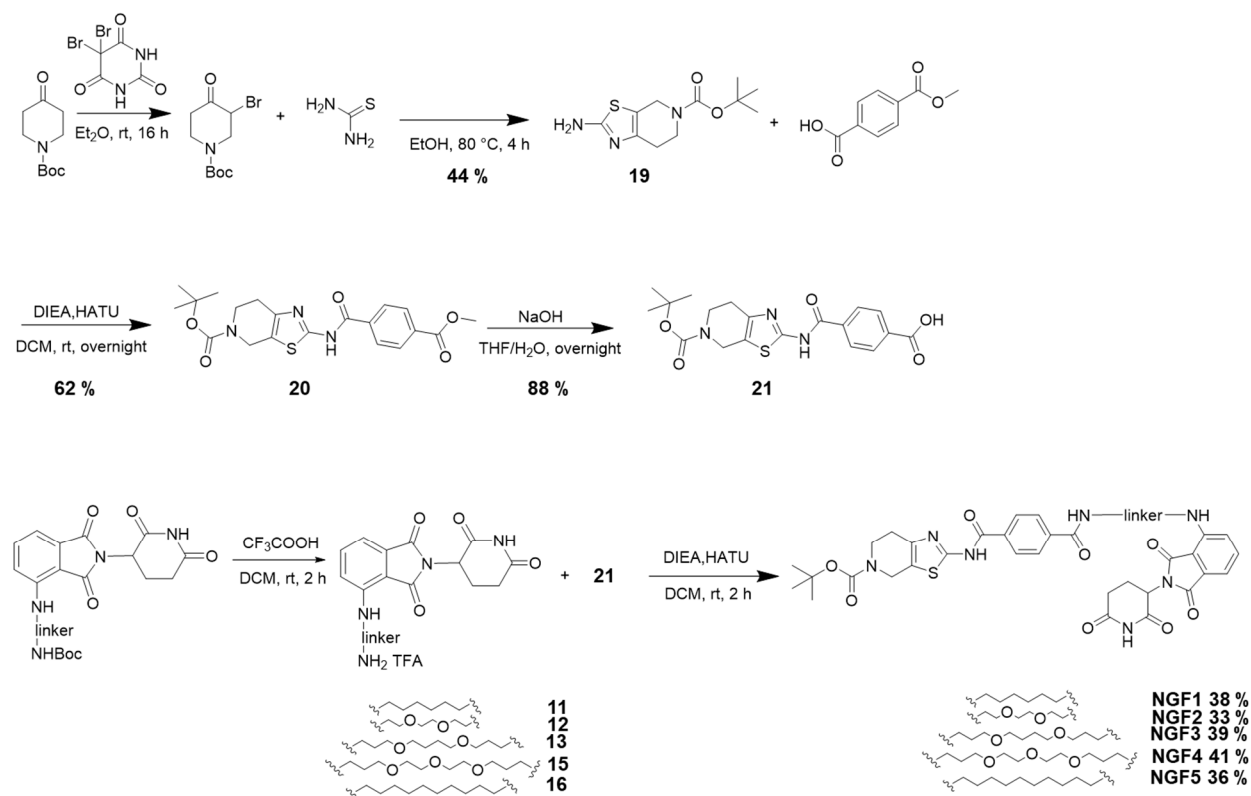
Figure 5.6 Synthesis of key intermediate **21** and Aster-A PROTACs.

5.3.3 Optimization of final HATU-mediated amide coupling

Initially, the conversion in the last step, HATU-mediated amide coupling between the amine and the acid (**21**) was relatively low (20 %, determined by LC-MS, Figure 5.7 B), and the formation of the guanidinium byproduct could be observed by LC-MS. Take the synthesis of **NGF2** as an example (Figure 5.7 A). Typically, the free amine (as a TFA salt) is obtained by removing the Boc-protecting group using TFA, and the remaining TFA is simply removed by vacuum without further purification. Then, **21**, DIPEA and HATU are dissolved in DCM and stirred for 10 min, followed by the addition of amine. However, monitoring this reaction by LC-MS after 2 h, three peaks can be clearly characterized: by product guanidinium (peak 2, Figure 5.7 C), unreacted acid (peak 8, Figure 5.7 C) and the desired product (peak 9, Figure 5.7 C).
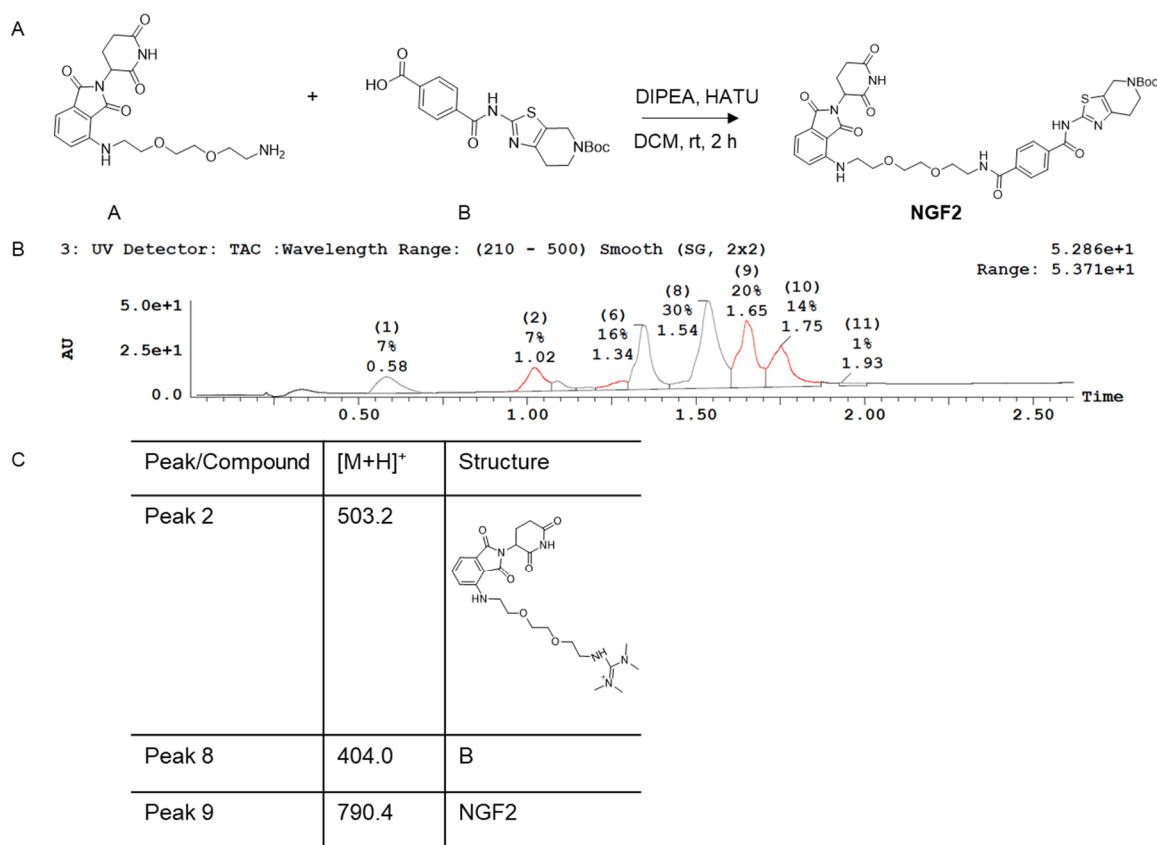
Figure 5.7 Side products formed in the final HATU-mediated amide coupling. (A) Synthesis scheme of **NGF2**. (B) LC-MS spectra (UV absorbance) of the amide-coupling reaction after 2 hours. (C) Structural and m/z information about each LC-MS peak.

Based on the mechanism of HATU-mediated amide coupling (Figure 5.8)[196], the carboxylate anion attacks HATU to form the *O*-acyl(tetramethyl)isouronium salt. Then, OAt anion attacks the isouronium salt to obtain the active OAt ester. Finally, amine reacts with the active ester to give the desired amide. However, a side-reaction could also happen when the amine reacts with HATU to form the guanidinium byproduct[197]. To minimize the formation of guanidinium, the mixture time of the acid, DIPEA, and HATU has been prolonged from 10 min to 1 hour to generate the active ester (Yield was shown in Figure 5.6).
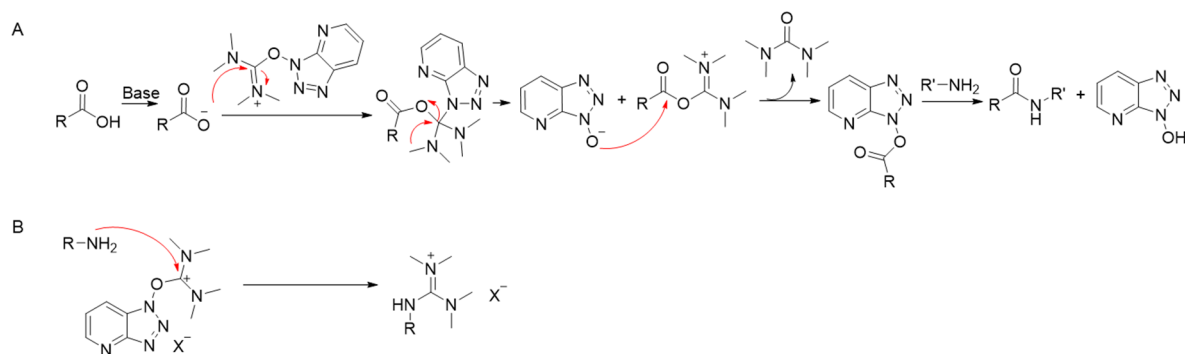
Figure 5.8 Mechanism of HATU-mediated amide coupling (A) and the side reaction between the amine and HATU (B).

5.4 **NGF3** exhibits a weak Aster-A removal effect in HeLa cell lines.

After obtaining five Aster-A PROTACs, namely **NGF1**-**5**, we tested their degradation ability towards Aster-A in HeLa cells. HeLa cells were treated with **NGF1**-**5** at 300 nM and 3 μM for 18 h, and the Aster-A protein level change was analyzed by western blotting. Clearly, only **NGF3** exhibits a weak Aster-A degradation effect; about 30 % of Aster-A can be removed by 3 μM **NGF3,** while the other four PROTACs and **Autogramin-2** cannot degrade Aster-A at all (Figure 5.9 A). Then we treated HeLa cells with different concentrations of **NGF3** for 18 h. Reassuringly, **NGF3** can degrade Aster-A in a dose-dependent manner with DC_{50}=4.8 μM and the maximum Aster-A degradation can be observed at 10 μM with D_{max} of 41% (Figure 5.9 B).
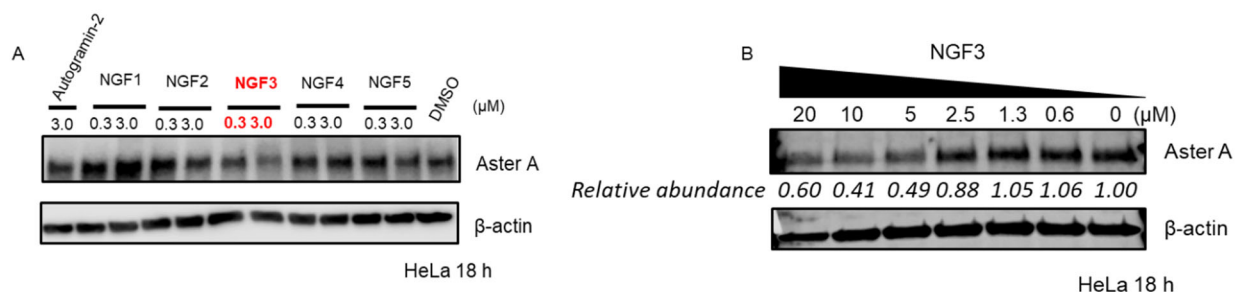
Figure 5.9 **NGF3** leads to Aster-A degradation in HeLa cells. (A) HeLa cells were treated with 300 nM or 3 μM of **NGF1-5** for 18 h. 3 μM **Autogramin-2** and DMSO were treated as control. The protein levels of Aster-A and β-actin were analyzed by western blot (n=2). (B) **NGF3** degrades Aster-A in a dose-dependent manner.

5.5 Optimization of **NGF3**.

After screening the first series of Aster-A PROTACs, we obtained **NGF3** as a weak Aster-A PROTAC and we wanted to optimize this molecule into a more effective Aster-A PROTAC. Therefore, a series of SAR studies were carried out. The degradation efficiency of a PROTAC relies on these three components, the

selection of the POI ligands, E3 ligands, and the linkers. These factors collectively affect the formation of the stable POI-PROTAC-E3 ternary complex[198]. However, because of the lacking structural information on the Aster-A-**NGF3**-CRBN ternary complex, the SAR of **NGF3** could only be explored empirically. Therefore, the optimization of **NGF3** will be conducted in two ways. First, to explore the effect of E3 ligase preference on Aster-A degradation, we retained similar linkers used in the pomalidomide-based PROTACs but employed another commonly used E3 ligase (VHL) to obtain a new series of VHL-based PROTACs. Second, to explore the effect of linkers on the formation of ternary complex and the degradation efficiency of Aster-A PROTACs, we kept pomalidomide as the ligand to recruit CRBN but selected different linkers.

5.5.1   SAR I: Preference of E3 ligases-similar linker but different E3 ligand

It has been demonstrated that hijacking different E3 ligases will lead to different degradation profiles of a given PROTAC. Take dasatinib PROTACs as an example[199]. A CRBN-based dasatinib PROTAC could successfully degrade both c-ABL and BCR-ABL but a VHL-based dasatinib PROTAC could only degrade c-ABL. The reasons for differences in E3 ligase selectivity might be the different affinity towards each E3 ligase or the target protein has its own preference towards different E3 ligases, which leads to the different ubiquitination rates and degradation efficiencies.

Since only **NGF3** exhibits weak Aster-A degradation ability, we decided to hijack another commonly used E3 ligase-VHL to design VHL-recruiting PROTACs and tested whether Aster-A can be degraded by them. Therefore, **VH032**[200] was chosen as the VHL ligand to obtain three VHL-recruiting Aster-A PROTACs, namely **NGV1-3**. **VH032 (22)** was synthesized according to the literature[201]. Briefly, the bromoaryl compound was prepared through reductive amination of 4-bromobenzaldehyde and then followed by Heck coupling with 4-methylthiazole, Boc deprotection and amide coupling. Then, **VH032 (22)** was connected to different linkers via HATU-mediated amide coupling and finally coupled with **21** via HATU-mediated amide-coupling to obtain **NGV1-3** (Figure 5.10 A, Figure 5.11). Similarly, we tested **NGV1-3** at 300 nM and 3 μM for 18 h in HeLa cells. **NGV1-3** does not exhibit Aster-A removal ability (Figure 5.10 B). Although it seems a degradation of Aster-A under the treatment of 3 μM **NGV1**, it is not reproducible (Figure 5.12) and the artificial degradation is because of the poor quality of Aster-A primary antibody.
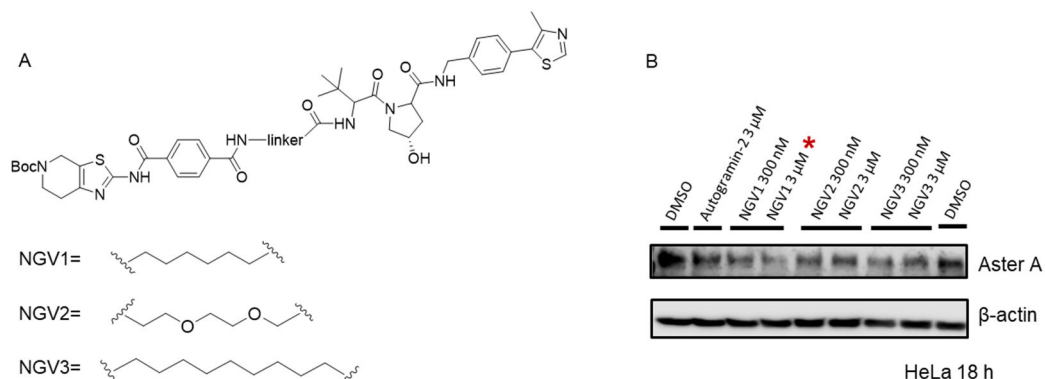
Figure 5.10 VHL-based Aster-A PROTACs failed to degrade Aster-A in HeLa cells. (A) Chemical structures of **NGV1-3**. (B) HeLa cells were treated with 300 nM or 3 μM of **NGV1-3** for 18 h. 3 μM **Autogramin-2** and DMSO were treated as control. The protein levels of Aster-A and β-actin were analyzed by western blot (n=2). * It seems a degradation of Aster-A under the treatment of 3 μM NGV1 but it is not reproducible and the artificial degradation is because of the poor quality of Aster-A primary antibody.
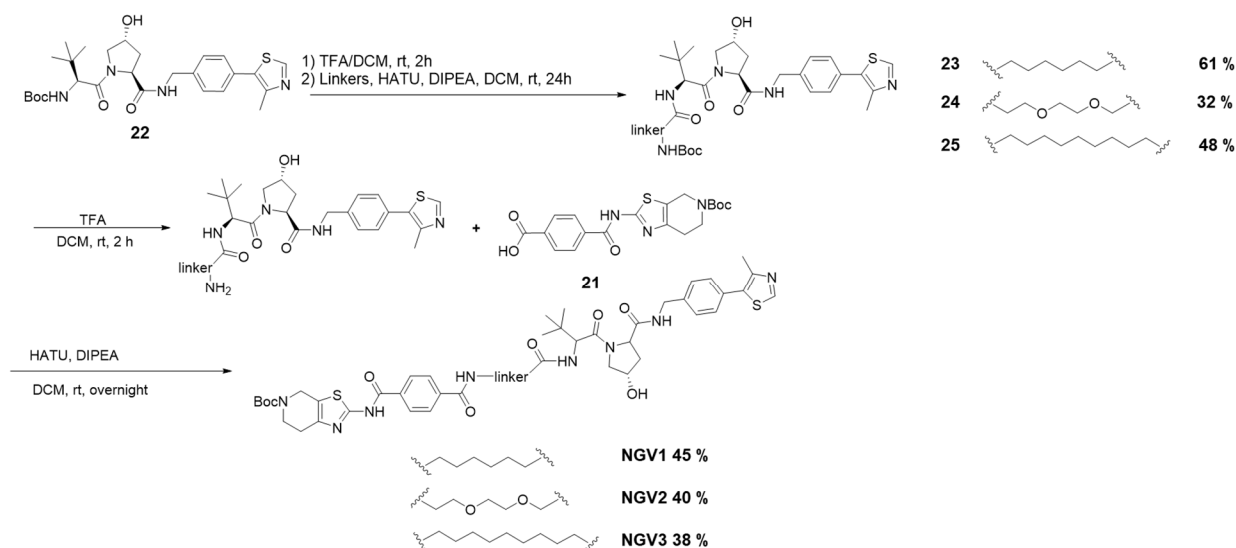
Figure 5.11 The synthesis routine of **NGV1-3**

5.5.2 SAR II: Suitable linkers-same E3 ligand but different linkers.

Since VHL-recruiting PROTACs could not degrade Aster-A either, we chose to continue with CRBN and focused on optimizing the linker of **NGF3**. This is because an efficient PROTAC not only relies on the effective binding of the POI ligands and the E3 ligands but also depends on a suitable inker which could combine the target protein with E3 ligase to form a stable ternary complex[202]. From **NGF2** to **NGF4,** the linkers have 8, 12, and 13 non H-atoms, respectively. Since one more C-O bond turns **NGF3** into a completely ineffective PROTAC, **NGF4**. We wanted to explore the effects of one or two less C-O bonds in the structure of **NGF3**. Therefore, two more linkers with 10 and 11 non H-atoms were introduced between the pomalidomide and **Autogramin-2** to obtain two more **NGF3** derivatives, namely **NGF6** and **NGF7** (structures shown in Figure 5.12 A). Then their degradation activities were also tested in HeLa cell lines at higher concentration, 10 μM for 18 h. However, **NGF6** and **NGF7** still failed to degrade Aster-A (Figure 5.12 B).
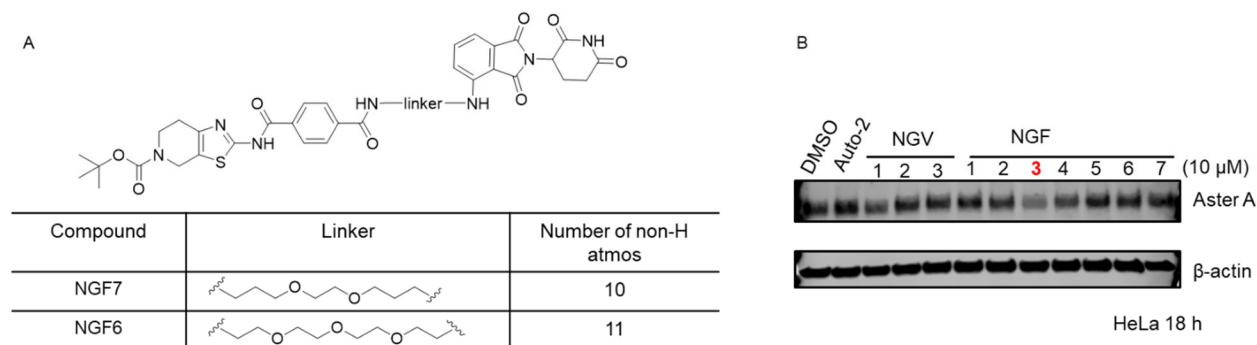
Figure 5.12 **NGF6** and **7** failed to degrade Aster-A in HeLa cells. (A) Chemical structures of **NGF6** and **7**. (B) HeLa cells were treated with 10 μM of **NGV1-3** and **NGF1-7** for 18 h. 3 μM **Autogramin-2** and DMSO were treated as control. The protein levels of Aster-A and β-actin were analyzed by western blot (n=2).

5.6 **NGF3**-induced Aster-A degradation is cell-line dependent

A PROTAC can exhibit various degradation efficiencies across different cell lines[203]. This can be attributed to different reasons, including the expression levels of the target protein and E3 ligases or the mutations on the E3 ligases. We hypothesized that HeLa cell line may not be the best cell model to study Aster-A degradation. Therefore, two additional cell lines, human bone osteosarcoma epithelial cells, U2OS and human lung adenocarcinoma cells, A549 were selected to investigate the effects of Aster-A degradation.

Among three tested cell lines, A549 has the highest Aster-A expression level but the least expression levels of VHL, while U2OS has the comparable levels of Aster-A and VHL with HeLa cells but the least expression level of CRBN (Figure 5.13).
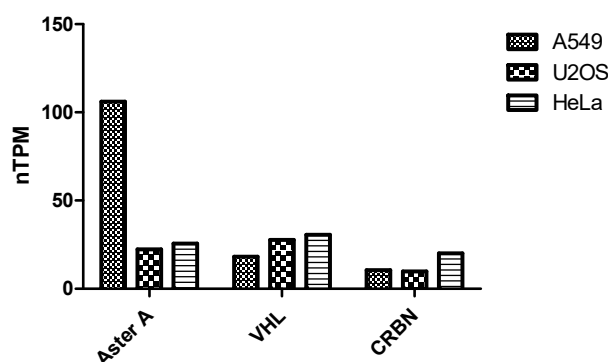
Figure 5.13 Expression levels of Aster-A, VHL and CRBN in A549, U2OS and HeLa cells (Data source: Human protein atlas https://www.proteinatlas.org/).

A459 cell lines were treated with different concentrations of **NGF1-7** and **NGV1-3** for 18 h. However, none of them leads to the degradation of Aster-A (Figure 5.14).

U2OS cell lines were also treated with 10 μM of **NGF1-7** and **NGV1-3** for 18 h. However, none of them including **NGF3** can degrade Aster-A (Figure 5.15).

In conclusion, among all the tested cell lines, Aster-A degradation induced by **NGF3** is cell line-dependent and can only be observed in HeLa cells.
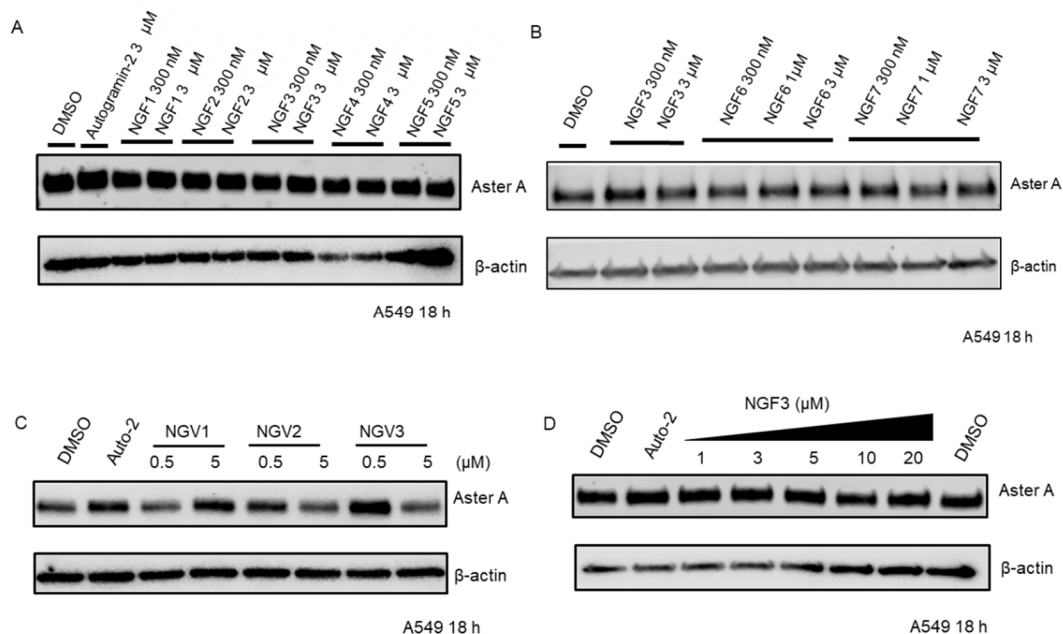
Figure 5.14 All Aster-A PROTACs including **NGF3** failed to degrade Aster-A in A549 cells. (A). A459 cells were treated with 300 nM and 3 μM of **NGF1-5** for 18 h. 3 μM **Autogramin-2** and DMSO were used as controls. The protein levels of Aster-A and β-actin were analyzed by western blot (n=2). (B). A459 cells were treated with 300 nM, 1 μM and 3 μM of **NGF3, 6 and 7** for 18 h and DMSO were used as controls. The protein levels of Aster-A and β-actin were analyzed by western blot (n=2). (C). A459 cells were treated with 500 nM and 5 μM of **NGV1-3** for 18 h. 3 μM **Autogramin-2** and DMSO were used as controls. The protein levels of Aster-A and β-actin were analyzed by western blot (n=2). (D). A459 cells were treated with different concentrations of **NGF3** for 18 h. 3 μM **Autogramin-2** and DMSO were used as controls. The protein levels of Aster-A and β-actin were analyzed by western blot (n=1).
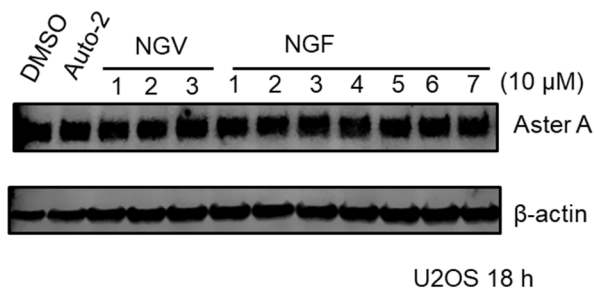
Figure 5.15 All Aster-A PROTACs including **NGF3** failed to degrade Aster-A in U2OS cells. U2OS cells were treated with 10 μM of **NGV1-3** and **NGF1-7** for 18 h. 3 μM **Autogramin-2** and DMSO were used as controls. The protein levels of Aster-A and β-actin were analyzed by western blot (n=1).

86

5.7 Summary and outlook

In this chapter, we first designed a series of pomalidomide-based Aster-A PROTACs with various alkyl, PEG or extended glycerol linkers and screened them in HeLa cells. Among these five PROTACs, only **NGF3** exhibited a weak degradative ability of Aster-A. Therefore, based on **NGF3**, SAR studies were carried out by either shortening the linker part to obtain **NGF6** and **NGF7** or by recruiting another E3 ligand to obtain **NGV1-3**. Unfortunately, among all the PROTACs, **NGF3** is still the only effective Aster-A degrader. Then we started to explore the effects of different cell models. There are numerous factors that affect the degradation ability of a PROTAC in cells, including the synthesis rate of the target protein, the expression level of the target protein, the expression level of the E3 ligase and the different mutations on E3 ligase in different cell lines. A549 was chosen because its high expression level of Aster-A proteins and U2OS was chosen because this cell line is a common cell model to study the function for Aster-A. However, none of our PROTACs exhibited effective Aster-A degradation at a range of concentrations. The reasons of an ineffective PROTAC are also diverse, such as impaired binding affinity towards target protein or E3 ligase, poor cell permeability, fast protein re-synthesis rate or the combination of POI ligand, E3 ligand and linker cannot form productive ternary complex. However, due to lacking structural information, the rational optimization of a PROTAC remains arduous and empirical.

The further optimization of Aster-A PROTACS will be carried out from different aspects. First, we will still focus on **NGF3** as this molecule has already exhibited weak Aster-A degradation ability and we will explore the effects of various exit vectors on Aster-A degradation. As shown in Figure 5.16, the glutarimide moiety of pomalidomide inserts deeply into the hydrophobic pocket of CRBN, which contributes to the majority of affinity between the molecule and the receptor and the phthalimide part is exposed to the solvent. Therefore, besides connecting the linker through C4 position, we also want to connect the linker through C5 position to generate another **NGF3**-like Aster-A PROTAC.
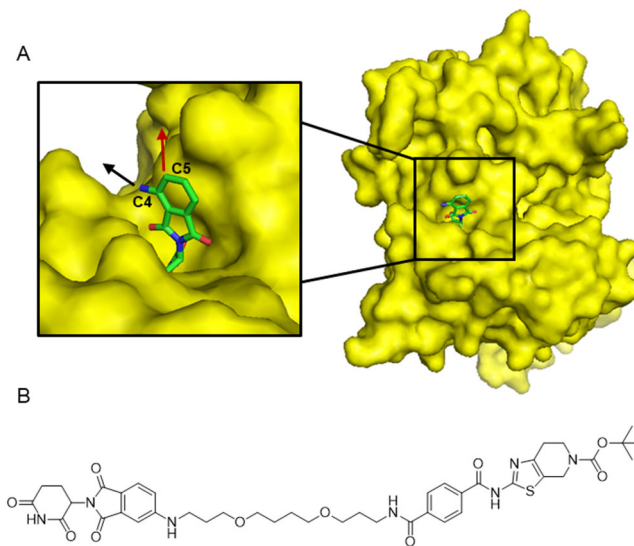
Figure 5.16 Further optimization of **NGF3**. (A) The binding mode of **pomalidomide** with CRBN[204] (PDB ID: 4CI3). Black arrow indicates the exit vector of **NGF3** and red arrow indicates the new exit vector for linker attachment. (B) The structure of **NGF3**-like Aster-A PROTAC.

Additionally, we will further study the effects of the linker part. Due to limited structural information of CRBN-**NGF3**-Aster-A ternary complex, it is difficult to decide the rigidity of the linker part so we will still

choose flexible linkers. Since we have only optimized (poly)ethylene glycol containing linkers and only two lengths of alkyl linkers (6 and 10 carbons) were synthesized and tested. In the following experiments, more alkyl linkers (8 and 12 carbons) will be selected and synthesized to obtain more diverse alkyl linker-based Aster-A PROTACs (Figure 5.17).

Moreover, we will screen more cell lines to find the most suitable model for detecting changed in Aster-A levels.

Compound	Linker	Number of non-H atmos
NGF8		8
NGF9		12
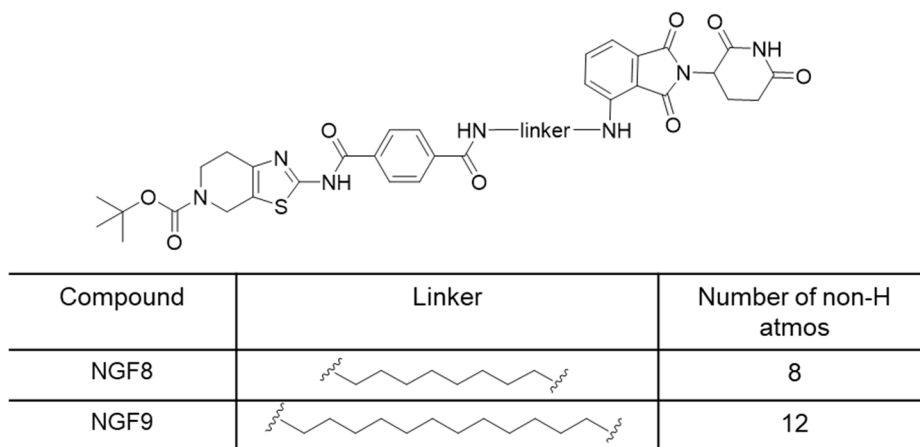
Figure 5.17 New structures of alkyl-based Aster-A PROTACs.

6 Experimental sections

6.1 Synthesis of compounds

6.1.1 General directions

Commercially available reagents were used without further purification and all solvents were of HPLC quality. All reactions were run under a N_2 atmosphere and were monitored by thin layer chromatography (TLC) and/or reversed-phase ultra-performance liquid chromatography mass spectrometry (RP-UPLC-MS)

Analytical TLC was conducted on Merck aluminium sheets covered with silica (C60). The plates were either visualized under UV-light or stained by dipping in a developing agent followed by heating. $KMnO_4$ [3 g in water (300 mL) along with K_2CO_3 (20 g) and 5% aqueous NaOH (5 mL)] was used as developing agents. Flash column chromatography was performed using Merck Geduran® Si60 (40-63 μm) silicagel.

All new compounds were characterized by ^{1}H NMR, ^{13}C NMR, MS (ESI), HRMS (ESI) and optical rotation (byproducts were not fully characterized). For the recording of ^{1}H NMR and ^{13}C NMR a Bruker Ascend with a Prodigy cryoprobe (operating at 400 MHz for proton and 100 MHz for carbon) was used. The chemical shifts (δ) are reported in parts per million (ppm) and the coupling constants (J) in Hz. For spectra recorded in DMSO, signal positions were measured relative to the signal for DMSO (δ 2.50 ppm for 1H NMR and δ 39.43 ppm for ^{13}C NMR). For spectra recorded in CDCl$_3$, signal positions were measured relative to the signal for CHCl$_3$ (δ 7.26 ppm for ^{1}H NMR and δ 77.0 ppm for ^{13}C NMR).

Analytical RP-UPLC-MS (ESI) analysis was performed on a S2 Waters AQUITY RP-UPLC system equipped with a diode array detector using an Thermo Accucore C18 column (d 2.6 μm, 2.1 x 50 mm; column temp: 50 °C; flow: 1.0 mL/min). Eluents A (0.1% HCO_2H in H_2O) and B (0.1% HCO_2H in MeCN) were used in a linear gradient (5% B to 100% B) in 2.4 min and then held for 0.1 min at 100% B (total run time: 2.6 min). The LC system was coupled to a SQD mass spectrometer.

Ultra-high Performance Liquid Chromatography-High Resolution Mass Spectrometry (UHPLC-HRMS) was performed on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 × 150 mm, 1.9 μm) with a linear gradient consisting of water (A) and acetonitrile (B) both buffered with 20 mM formic acid, starting at 10% B and increased to 100% in 10 min where it was held for 2 min, returned to 10% in 0.1 min and remaining for 2 min (0.35 mL/min, 60 °C). An injection volume of 1 μL was used. MS detection was performed in both positive and negative detection on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250 °C, gas flow of 8 L/min, sheath gas temperature of 300 °C and flow of 12 L/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for m/z 85–1700 in MS mode and m/z 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 μL/min using a 1:100 splitter. The solution contained 1 μM tributylamine (Sigma-Aldrich) and 10 μM Hexakis(2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The [M + H]$^+$ ions (m/z 186.2216 and 922.0098 respectively) of both compounds was used.

6.1.2   Synthesis of cholesterol-bearing PROTACs.
Hyodeoxycholic acid methyl ester (1)
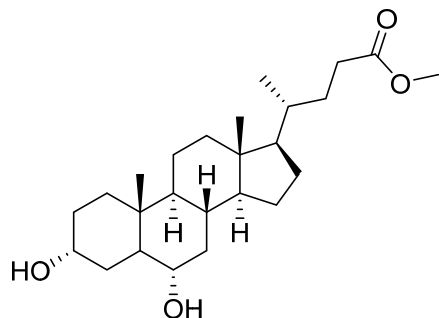
Hyodeoxycholic acid (20 g, 51 mmol) was dissolved in 100 mL methanol and then concentrated H_2SO_4 (2.5 mL) was added dropwise at 0 °C under N_2 atmosphere. The solution was stirred overnight at room temperature. Then, the solvent was removed under vacuum and the crude was dissolved in 40 mL EtOAc, washed by saturated $NaHCO_3$ (2 × 25 mL) and brine (2 × 25 mL). The organic phase was collected, dried over anhydrous sodium sulfate and evaporated under reduced pressure to obtain compound **1**.

White solid (20 g, 98%).

The NMR data matched those reported in the literature[205].

^{1}H NMR (400 MHz, CDCl$_3$) δ 4.09 – 4.01 (m, 1H), 3.66 (s, 3H), 3.64 – 3.55 (m, 1H), 2.40 – 2.31 (m, 1H), 2.27 – 2.17 (m, 1H), 2.02 – 1.30 (m, 17H), 1.23 – 0.98 (m, 7H), 0.95 – 0.87 (m, 6H), 0.64 (s, 3H).

^{13}C-NMR (101 MHz, CDCl$_3$): δ 174.89, 71.80, 68.32, 56.26, 56.05, 51.65, 48.48, 42.98, 40.07, 39.94, 36.10, 35.68, 35.48, 35.09, 34.97, 31.19, 31.09, 30.31, 29.28, 28.25, 24.33, 23.61, 20.88, 18.39, 12.16.

MS *m/z* (ES+) calcd for C$_{25}$H$_{42}$O$_4$ [2M+H]$^+$ 813.6, found: 813.5.

Methyl 3α,6α-ditosyloxy-5β-cholan-24-oate (2)
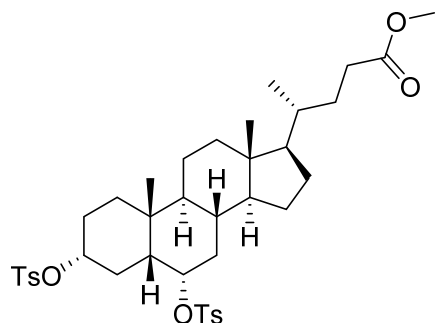
1 (5 g, 12.5 mmol) was dissolved in dry pyridine (30 mL) and then a solution of TsCl (7 g, 37.5 mmol) dissolved in 20 mL dry pyridine was added dropwise. The solution was stirred at room temperature for 2 days. And then, 200 mL 10% HCl was poured into the solution and ice chips were added gradually. The precipate was filtered off and washed by water to give compound **2**.

White solid (8.5 g, 95%).

The NMR data matched those reported in the literature[206].

¹H NMR (400 MHz, CDCl₃) δ 7.83 – 7.72 (m, 4H), 7.36 (t, *J* = 8.0 Hz, 4H), 4.85 – 4.77 (m, 1H), 4.38 – 4.27 (m, 1H), 3.68 (s, 3H), 2.48 (s, 6H), 2.40 – 2.31 (m, 1H), 2.28 – 2.18 (m, 1H), 1.98 – 0.93 (m, 24H), 0.90 (d, *J* = 6.3 Hz, 3H), 0.82 (s, 3H), 0.61 (s, 3H).

¹³C-NMR (101 MHz, CDCl₃): δ 174.81, 144.84, 134.64, 134.63, 129.98, 129.94, 127.75, 127.66, 127.14, 81.92, 79.82, 55.95, 55.90, 51.65, 46.47, 42.96, 39.72, 39.58, 36.29, 35.39, 34.96, 32.22, 31.14, 31.01, 28.10, 27.53, 26.60, 24.06, 23.02, 21.82, 20.65, 18.34, 12.09.

Methyl 3β-hydroxy-5-cholen-24-oate (3)
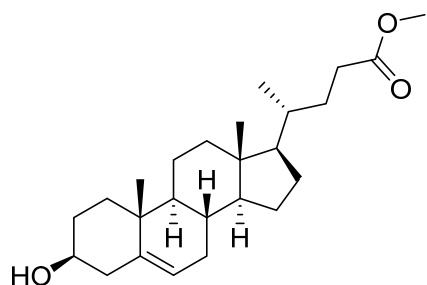
3 (1 g, 1.4 mmol) and KOAc (100 mg, 1 mmol) were dissolved in 10 mL DMF and 1mL H₂O. The solution was refluxed at 105 ℃ overnight. And then, the solution was cooled down to room temperature, extracted by EtOAc (3 × 20 mL) and washed by brine (3 × 20 mL). The organic phase was collected, dried over anhydrous sodium sulfate and evaporated under reduced pressure. The residue was purified by column chromatography (gradient of pentane/EtOAC 8:1 to 2:1).

White solid (282 mg, 52%).

The NMR data matched those reported in the literature[206].

¹H NMR (400 MHz, CDCl₃) δ 5.35 (dt, *J* = 5.5, 1.9 Hz, 1H), 3.66 (s, 3H), 3.57 – 3.46 (m, 1H), 2.41 – 2.16 (m, 4H), 2.05 – 1.74 (m, 8H), 1.65 – 1.23 (m, 9H), 1.20 – 1.02 (m, 4H), 1.00 (s, 3H), 0.92 (d, *J* = 6.5 Hz, 3H), 0.68 (s, 3H).

¹³C-NMR (101 MHz, CDCl₃): δ 174.93, 140.88, 121.82, 71.94, 56.87, 55.91, 51.64, 50.22, 42.51, 42.43, 39.88, 37.39, 36.64, 35.52, 32.03, 32.01, 31.79, 31.21, 31.16, 28.26, 24.40, 21.21, 19.54, 18.46, 12.01.

MS *m/z* (ES+) calcd for C₂₅H₄₀O₃ [M+H]⁺ 389.3, found: 389.1.

3β-hydroxy-5-cholenoic acid (4)
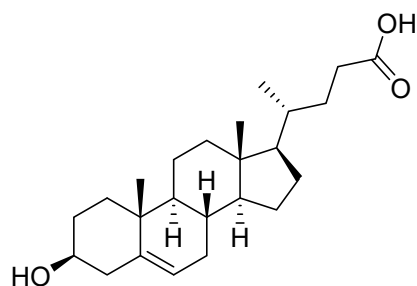
3 (500 mg, 1.28 mmol) was dissolved in 10 mL 4% KOH/MeOH and the solution was stirred overnight at room temperature. Then, 1 M HCl was added gradually and the precipitate was filtered off and washed with H₂O to obtain compound **4**.

White solid (344 mg, 72%).

The NMR data matched those reported in the literature[207].

^{1}H NMR (400 MHz, CDCl$_3$) δ 5.37 (dt, J = 4.9, 1.9 Hz, 1H), 3.57 – 3.47 (m, 1H), 2.51 – 1.08 (m, 25H), 1.03 (s, 3H), 0.96 (d, J = 6.6 Hz, 3H), 0.71 (s, 3H).

^{13}C-NMR (101 MHz, CDCl$_3$): δ 178.17, 142.23, 122.43, 72.44, 58.14, 57.30, 51.70, 43.55, 43.02, 41.12, 38.55, 37.68, 36.71, 33.26, 33.01, 32.34, 32.30, 32.01, 29.14, 25.30, 22.19, 19.86, 18.79, 12.31.

2-(2,6-dioxopiperidin-3-yl)-4-fluoroisoindoline-1,3-dione (5)
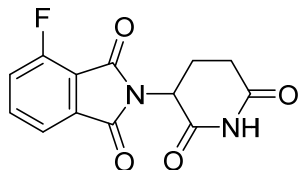
3-Fluorophthalic anhydride (996 mg, 6 mmol), 3-aminopiperidine-2,6-dione hydrogen chloride (658 mg, 4 mmol) and NaOAc (558 mg, 6.8 mmol) were added to HOAc (20 mL) and the solution was heated at 135 °C for 8 h. Then, the reaction was cooled down to room temperature and the residue was suspended in 100 mL ice water. The solid was collected by vacuum filtration without further purification.

Grey solid (512 mg, 47%).

The NMR data matched those reported in the literature[208].

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.16 (s, 1H), 7.96 (td, J = 7.9, 4.5 Hz, 1H), 7.80 (d, J = 7.3 Hz, 1H), 7.75 (t, J = 8.9 Hz, 1H), 5.17 (dd, J = 12.9, 5.4 Hz, 1H), 2.90 (ddd, J = 17.1, 13.9, 5.5 Hz, 1H), 2.69 – 2.53 (m, 2H), 2.07 (dtd, J = 13.0, 5.3, 2.3 Hz, 1H).

MS *m/z* (ES+) calcd for C$_{13}$H$_9$FN$_2$O$_4$ [M+ H]$^+$ 277.1, found 276.9.

4-fluoro-2-(1-methyl-2,6-dioxopiperidin-3-yl)isoindoline-1,3-dione (6)
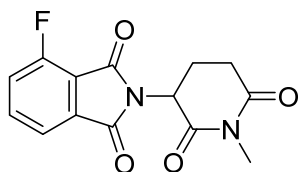
5 (100 mg, 0.36 mmol) and NaH (21 mg, 0.54 mmol) were dissolved in 6 mL DMF and stirred for 15 min. Then, MeI (76.6 mg, 0.54 mmol) was added to the mixture dropwisely and the reaction mixture was stirred at 0-25 °C for 8 h. Then, the solution was diluted with 30 mL water and extracted by EtOAc (3 × 10 mL), washed with brine (3 × 20 mL), dried over anhydrous Na$_2$SO$_4$ and concentrated to obtain **6** as a green oil without further purification.

The data matched those reported in the literature[209].

MS *m/z* (ES+) calcd for C$_{14}$H$_{11}$FN$_2$O$_4$ [M+ H]$^+$ 290.3, found 290.9.

General procedure I: Mono-Boc protection of diamines.

A solution of Di-*tert*-butyl dicarbonate (2.20 g, 10.1 mmol) in CH_2Cl_2 (20 mL) was added dropwisely to a solution of the corresponding diamine (60 mmol) in CH_2Cl_2 (60 mL) at 0 °C. The solution was stirred at 0 °C for 3 h and then at room temperature overnight. And then the mixture was washed by H_2O and brine (each 30 mL). The organic phase was collected, dried over Na_2SO_4 and evaporated under reduced pressure.

tert-butyl (6-aminohexyl)carbamate (7)
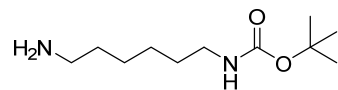
The compound was prepared using General procedure I to give a colorless oil (1.9 g, 88%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 4.59 (s, 1H), 3.11 (q, J = 6.9 Hz, 2H), 2.71 (t, J = 7.0 Hz, 2H), 2.10 (s, 2H), 1.52 – 1.46 (m, 4H), 1.44 (s, 9H), 1.38 – 1.30 (m, 4H).

tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate (8)
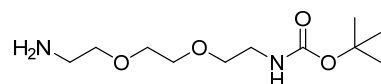
The compound was prepared using General procedure I to give a colorless oil (2.0 g, 81%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 5.21 (s, 1H), 3.59 (s, 4H), 3.54 – 3.48 (m, 4H), 3.29 (q, J = 5.4 Hz, 2H), 2.86 (t, J = 5.2 Hz, 2H), 1.41 (s, 9H).

tert-butyl (3-(4-(3-aminopropoxy)butoxy)propyl)carbamate (9)
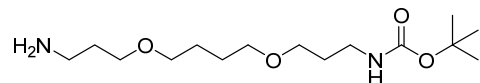
The compound was prepared using General procedure I to give a colorless oil (2.5 g, 82%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 5.01 (s, 1H), 3.52 – 3.35 (m, 8H), 3.23 – 3.16 (m, 2H), 2.80 (t, J = 6.7 Hz, 2H), 1.77 – 1.67 (m, 4H), 1.65 – 1.56 (m, 4H), 1.41 (s, 9H).

tert-butyl (3-(2-(2-(3-aminopropoxy)ethoxy)ethoxy)propyl)carbamate (10)
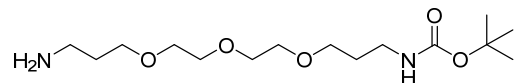
The compound was prepared using General procedure I to give a colorless oil (3.2 g, 75%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 5.39 (t, J = 5.7 Hz, 1H), 3.42 – 3.25 (m, 12H), 2.94 (q, J = 6.3 Hz, 2H), 2.53 (t, J = 6.7 Hz, 2H), 1.54 – 1.44 (m, 4H), 1.18 (s, 9H).

General procedure II: Nucleophilic aromatic substitution (S$_N$Ar).

4-fluoro-pomalidomide (**5**, 1 eq.) or 4-fluoro-2-(1-methyl-2,6-dioxopiperidin-3-yl)isoindoline-1,3-dione (**6**, 1 eq), DIPEA (4 eq) and the corresponding mono-protected diamine (1 eq) were dissolved in dry DMF. The mixture was stirred at 90 °C for 6 h. After cooling to room temperature, the mixture was diluted with H$_2$O (20 mL) and then extracted by EtOAc (3 × 30 mL). The combined organic phase was further washed by brine (3 x 30 mL), dried over anhydrous sodium sulfate and evaporated under reduced pressure. The residue was purified by column chromatography (gradient of pentane/EtOAc 4:1 to 1:4).

***tert*-butyl (6-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)hexyl)carbamate (11)**
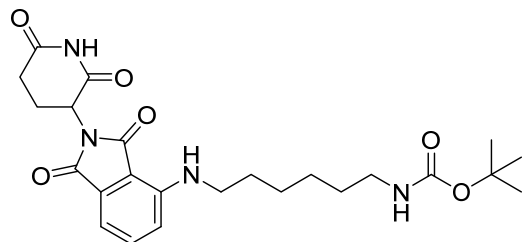
**5** (400 mg, 1.45 mmol) and **7** (313 mg, 1.45 mmol) were combined according to general procedure II. The product was a yellow solid (309 mg, 45%).

The NMR data matched those reported in the literature[210].

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.99 (s, 1H), 7.54 – 7.49 (m, 1H), 7.12 (d, J = 7.1 Hz, 1H), 6.90 (d, J = 8.6 Hz, 1H), 4.97 – 4.90 (m, 1H), 4.53 (s, 1H), 3.29 (t, J = 7.0 Hz, 2H), 3.14 (s, 2H), 2.96 – 2.71 (m, 3H), 2.20 – 2.13 (m, 1H), 1.75 – 1.57 (m, 4H), 1.57 – 1.49 (m, 2H), 1.47 (s, 9H), 1.45(m, 2H).

MS *m/z* (ES+) calcd for C$_{24}$H$_{43}$N$_4$O$_6$ [M+ H$^+$] 473.2, found 473.6.

***tert*-butyl (2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethoxy)ethyl)carbamate (12)**

5 (800 mg, 2.9 mmol) and **8** (720 mg, 2.9 mmol) were combined according to general procedure II. The product was a yellow oil (512 mg, 35%).

The NMR data matched those reported in the literature[210].

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.48 (s, 1H), 7.51 (dd, J = 8.5, 7.1 Hz, 1H), 7.13 (d, J = 7.1 Hz, 1H), 6.93 (d, J = 8.5 Hz, 1H), 5.09 (s, 1H), 4.97 (s, 1H), 3.74 (t, J = 5.3 Hz, 2H), 3.70 – 3.64 (m, 4H), 3.58 (t, J = 5.2 Hz, 2H), 3.49 (t, J = 5.3 Hz, 2H), 3.37 – 3.31 (m, 2H), 2.93 – 2.71 (m, 3H), 2.18 – 2.11 (m, 1H), 1.45 (s, 9H).

MS *m/z* (ES+) calcd for C$_{24}$H$_{32}$N$_4$O$_8$ [M+ H$^+$] 505.2, found 505.1.

tert-butyl (3-(4-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4 yl)amino)propoxy)butoxy)propyl)carbamate (13)

5 (400 mg, 1.45 mmol) and **9** (441 mg, 1.45 mmol) were combined according to general procedure II. The product was a yellow oil (439 mg, 54%).

The NMR data matched those reported in the literature[210].

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.25 (s, 1H), 7.51 (dd, J = 8.5, 7.1 Hz, 1H), 7.11 (d, J = 7.1 Hz, 1H), 6.94 (d, J = 8.5 Hz, 1H), 6.47 (s, 1H), 4.97 – 4.89 (m, 1H), 3.56 (t, J = 5.7 Hz, 2H), 3.52 – 3.39 (m, 8H), 3.23 (q, J = 6.2 Hz, 2H), 2.96 – 2.70 (m, 3H), 2.18 – 2.11 (m, 1H), 1.94 (p, J = 6.2 Hz, 2H), 1.76 (p, J = 6.3 Hz, 2H), 1.72 – 1.61 (m, 4H), 1.46 (s, 9H).

MS m/z (ES+) calcd for C$_{28}$H$_{40}$N$_4$O$_8$ [M+ H$^+$] 561.3, found 561.2.

tert-butyl (3-(4-(3-((2-(1-methyl-2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)butoxy)propyl)carbamate (14)
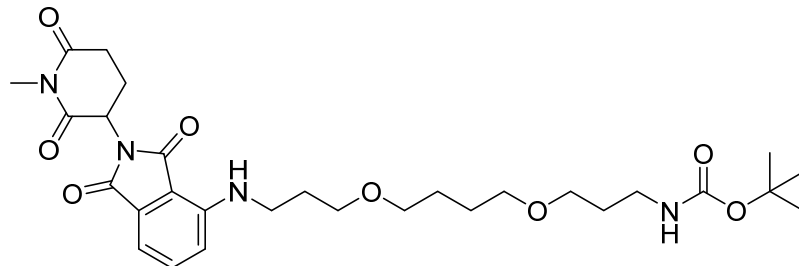
6 (150 mg, 0.52 mmol) and **9** (157 mg, 0.52 mmol) were combined according to general procedure II. The product was a green oil (120 mg, 40%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.50 (dd, J = 8.5, 7.1 Hz, 1H), 7.10 (dd, J = 7.2, 0.6 Hz, 1H), 6.94 (d, J = 8.6 Hz, 1H), 4.96 – 4.91 (m, 1H), 3.54 (q, J = 5.2, 4.7 Hz, 2H), 3.51 – 3.37 (m, 8H), 3.23 (s, 5H), 3.02 – 2.93 (m, 1H), 2.84 – 2.71 (m, 2H), 2.15 – 2.05 (m, 1H), 1.96 – 1.89 (m, 2H), 1.76 (p, J = 6.3 Hz, 2H), 1.70 – 1.64 (m, 4H), 1.46 (s, 9H).

MS m/z (ES+) calcd for C$_{29}$H$_{42}$N$_4$O$_8$ [M+ H$^+$] 575.3, found 575.3.

tert-butyl (3-(2-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)ethoxy)propyl)carbamate (15)
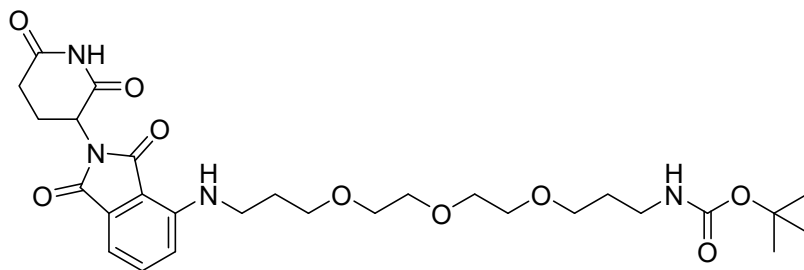
5 (500 mg, 1.81 mmol) and **10** (580 mg, 1.81 mmol) were combined according to general procedure II. The product was a yellow oil (564 mg, 54%).
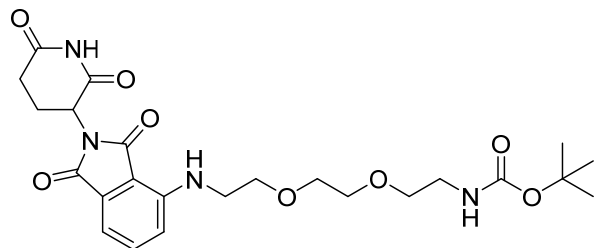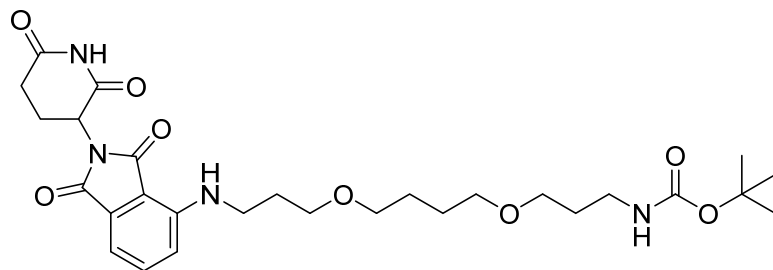
The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 8.17 (s, 1H), 7.51 (dd, J = 8.5, 7.1 Hz, 1H), 7.11 (d, J = 7.1 Hz, 1H), 6.96 (d, J = 8.5 Hz, 1H), 4.96 – 4.90 (m, 1H), 3.74 – 3.59 (m, 10H), 3.55 (t, J = 6.0 Hz, 2H), 3.43 (t, J = 6.6 Hz, 2H), 3.27 – 3.21 (m, 2H), 2.96 – 2.68 (m, 3H), 2.19 – 2.12 (m, 1H), 1.96 (p, J = 6.2 Hz, 2H), 1.77 (p, J = 6.2 Hz, 2H), 1.46 (s, 9H).

MS m/z (ES+) calcd for $C_{28}H_{40}N_4O_9$ [M+ H⁺] 577.3, found 577.7.

tert-butyl (10-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)decyl)carbamate (16)

5 (500 mg, 1.81 mmol) and tert-butyl (10-aminodecyl)carbamate (493 mg, 1.81 mmol) were combined according to general procedure II. The product was a yellow oil (421 mg, 44%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 8.08 (s, 1H), 7.51 (dd, J = 8.5, 7.1 Hz, 1H), 7.11 (d, J = 7.1 Hz, 1H), 6.91 (d, J = 8.5 Hz, 1H), 4.98 – 4.89 (m, 1H), 4.54 (s, 1H), 3.28 (t, J = 7.0 Hz, 2H), 3.12 (t, J = 7.1 Hz, 2H), 2.98 – 2.68 (m, 3H), 2.19 – 2.13 (m, 1H), 1.68 (p, J = 7.1 Hz, 4H), 1.47 (s, 9H), 1.44 – 1.39 (m, 2H), 1.37 – 1.26 (m, 10H).

MS m/z (ES+) calcd for $C_{28}H_{40}N_4O_6$ [M+ H⁺] 529.3, found 529.3.

tert-butyl (2-(2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethoxy)ethoxy)ethyl)carbamate (17)
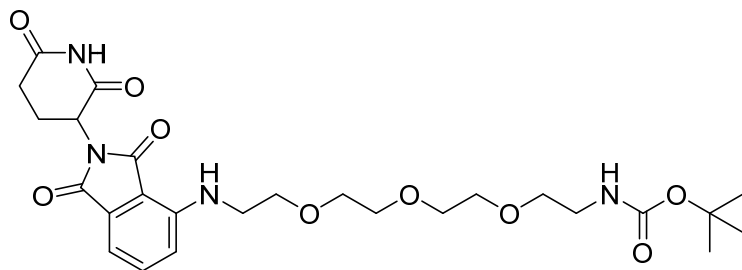
5 (283 mg, 1.02 mmol) and tert-butyl (2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethyl)carbamate (300 mg, 1.02 mmol) were combined according to general procedure II. The product was a yellow oil (200 mg, 36%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 7.50 (dd, J = 8.5, 7.1 Hz, 1H), 7.12 (d, J = 7.1 Hz, 1H), 6.94 (d, J = 8.5 Hz, 1H), 4.97 – 4.90 (m, 1H), 3.74 (t, J = 5.4 Hz, 2H), 3.69 (s, 4H), 3.67 – 3.64 (m, 2H), 3.64 – 3.60 (m, 2H), 3.54 (t, J = 5.2 Hz, 2H), 3.49 (t, J = 5.4 Hz, 2H), 3.31 (t, J = 5.2 Hz, 2H), 2.94 – 2.69 (m, 3H), 2.18 – 2.09 (m, 1H), 1.45 (s, 9H).

MS _m/z_ (ES+) calcd for $C_{26}H_{36}N_4O_9$ [M+ H⁺] 549.2, found 549.2.

tert-butyl (3-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)propyl)carbamate (18)
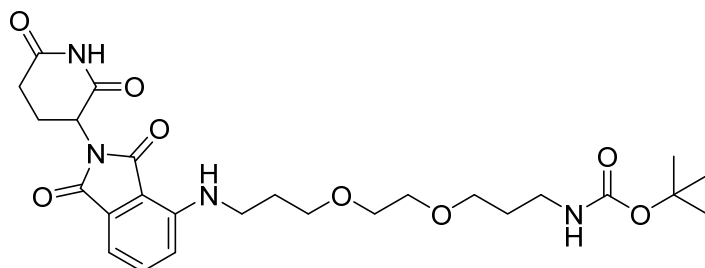

**5** (500 mg, 1.80 mmol) and tert-butyl (3-(2-(3-aminopropoxy)ethoxy)propyl)carbamate (500 mg, 1.80 mmol) were combined according to general procedure II. The product was a yellow oil (403 mg, 42%).

The NMR data matched those reported in the literature[210].

¹H NMR (400 MHz, CDCl₃) δ 7.50 (dd, J = 8.5, 7.1 Hz, 1H), 7.10 (d, J = 7.0 Hz, 1H), 6.94 (d, J = 8.6 Hz, 1H), 5.05 – 4.90 (m, 2H), 3.66 – 3.59 (m, 6H), 3.56 (t, J = 6.1 Hz, 2H), 3.43 (t, J = 6.5 Hz, 2H), 3.23 (q, J = 6.0 Hz, 2H), 2.94 – 2.70 (m, 3H), 2.19 – 2.11 (m, 1H), 1.96 (p, J = 6.2 Hz, 2H), 1.77 (p, J = 6.3 Hz, 2H), 1.45 (s, 9H).

MS _m/z_ (ES+) calcd for $C_{26}H_{36}N_4O_9$ [M+ H⁺] 533.2, found 533.2.

General procedure III: Boc-deprotection and HATU-mediated amide coupling.

The corresponding mono-Boc protected linker coupled pomalidomide was dissolved in CH_2Cl_2 (3 mL) and then trifluoroacetic acid (1 mL) was added dropwise at 0 °C. The mixture was then stirred at rt for 2 h. Then the solvent was removed, washed by $NaHCO_3$ and extracted by CH_2Cl_2. The orange oily residue was further dried under vacuum.

The carboxylic acid (1 eq), DIPEA (2 eq) and HATU (1.5 eq) were dissolved in 10 mL CH_2Cl_2. After 5 min, deprotected amine (1.2 eq) was dissolved in 5 mL CH_2Cl_2 and added to the same flask. The mixture was stirred at rt overnight. Then, the mixture was poured into H_2O (10 mL) and extracted by CH_2Cl_2 (3 x 10 mL). The organic phase was collected, dried over anhydrous sodium sulfate and evaporated under reduced pressure. The residue was purified by column chromatography (gradient of CH_2Cl_2/MeOH 80:1 to 30:1).

(4R)-N-(6-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)hexyl)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)pentanamide (C1)
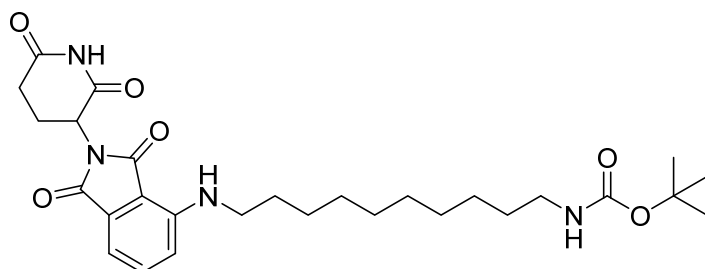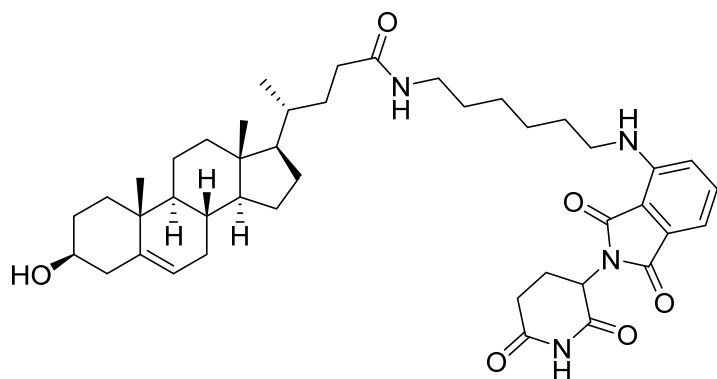

4 (80 mg, 0.21 mmol) and **11** (120 mg, 0.252 mmol) were combined according to general procedure III. The product was a yellow solid (73 mg, 48%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.72 (s, 1H), 7.46 (dd, J = 8.5, 7.1 Hz, 1H), 7.03 (d, J = 7.0 Hz, 1H), ,6.86 (d, J = 8.6 Hz, 1H), 6.72(s, 1H), 6.23 (s, 1H), 5.93 (t, J = 5.7 Hz, 1H), 5.35 – 5.28 (m, 1H), 4.95 – 4.81 (m, 1H), 3.30 – 3.14 (m, 5H), 2.89 – 2.66 (m, 3H), 2.37 (s, 2H), 2.20 – 2.28 (m, 2H), 2.12 – 1.99 (m, 2H), 1.98 – 1.87 (m, 2H), 1.87 – 1.68 (m, 4H), 1.69 – 1.60 (m, 2H), 1.51 – 1.46 (m, 4H), 1.43 – 1.36 (m, 12H), 1.16 – 0.98 (m, 4H), 0.96 (s, 3H), 0.94 – 0.89 (m, 4H), 0.63 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.21, 171.63, 169.55, 168.85, 167.72, 147.01, 140.87, 136.25, 132.46, 121.58, 116.82, 111.41, 109.78, 71.74, 56.75, 55.85, 50.12, 48.90, 43.50, 42.58, 42.41, 42.25, 39.79, 39.54, 37.31, 36.52, 35.58, 33.64, 32.03, 31.91, 31.60, 31.44, 29.48, 29.16, 28.22, 26.65, 24.31, 22.82, 21.11, 19.45, 18.45, 11.93.

HRMS (ESI+) *m/z* found 729.4585 [M+H]$^+$, C$_{43}$H$_{60}$N$_4$O$_6$ calculated 729.4591 (Δ = -0.82 ppm)

Purity: 100 %. **Retention time**: 2.20 min

(4R)-N-(2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethoxy)ethyl)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)pentanamide (C2)

4 (80 mg, 0.21 mmol) and **12** (127 mg, 0.252 mmol) were combined according to general procedure III. The product was a yellow solid (83 mg, 52%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.89 (s, 1H), 7.47 (dd, J = 8.5, 7.1 Hz, 1H), 7.08 (d, J = 7.0 Hz, 1H), 6.89 (d, J = 8.5 Hz, 1H), 6.50 (s, 1H), 6.14 (t, J = 5.4 Hz, 1H), 5.32 – 5.30 (m, 1H), 4.92 – 4.88 (m, 1H), 3.71 (t, J = 5.3 Hz, 2H), 3.65 (d, J = 2.2 Hz, 4H), 3.57 (t, J = 5.1 Hz, 2H), 3.55 – 3.47 (m, 1H), 3.46 – 3.40 (m, 4H), 2.87 – 2.68 (m, 3H), 2.25 – 2.18 (m, 2H), 2.12 – 2.01 (m, 3H), 2.00 – 1.90 (m, 2H), 1.83 – 1.73 (m, 3H), 1.52 – 1.32 (m, 7H), 1.36 – 1.17 (m, 4H), 1.08 – 1.01 (m, 4H), 0.97 (s, 3H), 0.89 – 0.85 (m, 4H), 0.63 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.00, 171.52, 169.48, 168.74, 167.65, 146.80, 140.89, 136.19, 132.62, 121.65, 116.78, 111.85, 110.45, 71.76, 70.72, 70.14, 70.08, 69.32, 56.78, 55.88, 50.15, 48.98, 42.41, 42.38, 42.32, 39.82, 39.33, 37.34, 36.56, 35.58, 33.58, 31.94, 31.81, 31.67, 31.48, 28.23, 24.34, 22.95, 21.14, 19.48, 18.47, 11.95.

HRMS (ESI+) *m/z* found 761.4483 [M+H]$^+$, C$_{43}$H$_{60}$N$_4$O$_8$ calculated 761.4489 (Δ = -0.78 ppm)

Purity: 100 %. **Retention time**: 1.99 min

(4R)-N-(3-(4-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)butoxy)propyl)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)pentanamide (C3)
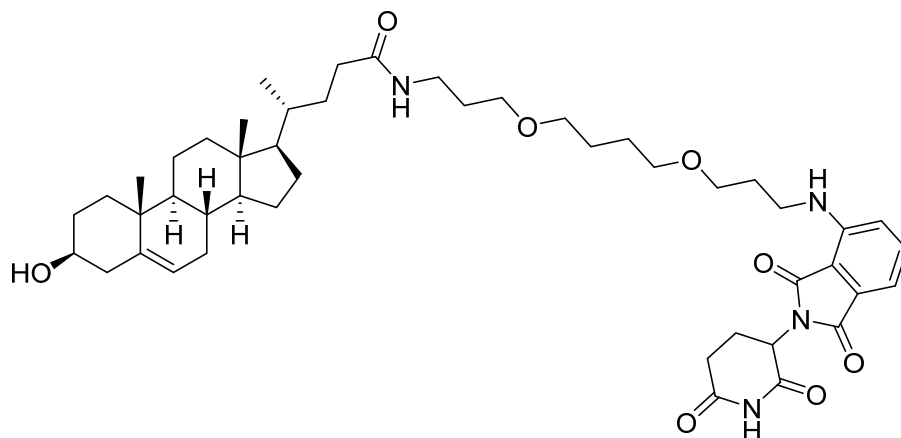
4 (80 mg, 0.21 mmol) and **13** (145 mg, 0.252 mmol) were combined according to general procedure III. The product was a yellow solid (51 mg, 30%).

¹H NMR (400 MHz, CDCl₃) δ 8.62 (s, 1H), 7.45 (dd, *J* = 8.6, 7.1 Hz, 1H), 7.06 (d, *J* = 7.1 Hz, 1H), 6.89 (d, *J* = 8.5 Hz, 1H), 6.17 (t, *J* = 5.4 Hz, 1H), 5.32 – 5.30 (m, 1H), 4.91 – 4.86 (m, 1H), 3.53 – 3.49 (m, 5H), 3.44 – 3.42 (m, 4H), 3.39 – 3.30 (m, 4H), 2.87 – 2.70 (m, 3H), 2.26 – 2.16 (m, 4H), 2.12 – 2.09 (m, 2H), 1.98 – 1.90 (m, 4H), 1.84 – 1.74 (m, 6H), 1.66 – 1.61 (m, *J* = 2.2 Hz, 4H), 1.50 – 1.39 (m, 6H), 1.31 – 1.24 (m, 3H), 1.13 – 1.00 (m, 4H), 0.98 (s, 3H), 0.91– 0.89 (m, 4H), 0.65 (s, 3H).

¹³C NMR (101 MHz, CDCl₃) δ 173.84, 171.44, 169.43, 168.67, 167.75, 147.05, 140.90, 136.17, 132.60, 121.65, 116.69, 111.44, 109.98, 71.79, 71.09, 70.96, 69.78, 68.55, 56.82, 55.90, 50.17, 48.94, 42.45, 42.33, 40.46, 39.85, 38.24, 37.35, 36.57, 35.61, 33.75, 31.97, 31.96, 31.68, 31.50, 29.46, 29.21, 28.26, 26.58, 26.50, 24.35, 22.93, 21.16, 19.49, 18.50, 11.97.

HRMS (ESI+) *m/z* found 817.5110 [M+H]⁺, C₄₇H₆₈N₄O₈ calculated 817.5115 (Δ = -0.61 ppm)

Purity: 100 %. **Retention time**: 2.22 min

(4R)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-(3-(4-(3-((2-(1-methyl-2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)butoxy)propyl)pentanamide (C3_Me)
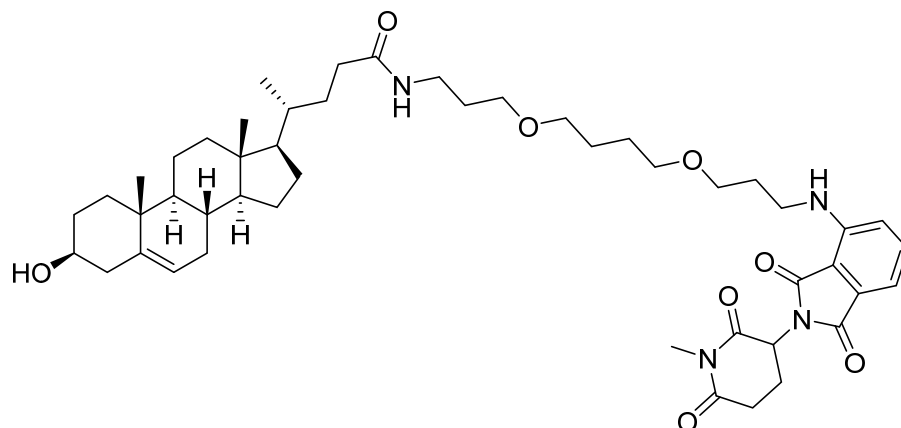
4 (80 mg, 0.21 mmol) and **14** (141 mg, 0.252 mmol) were combined according to general procedure III. The product was a yellow solid. (70 mg, 41%).
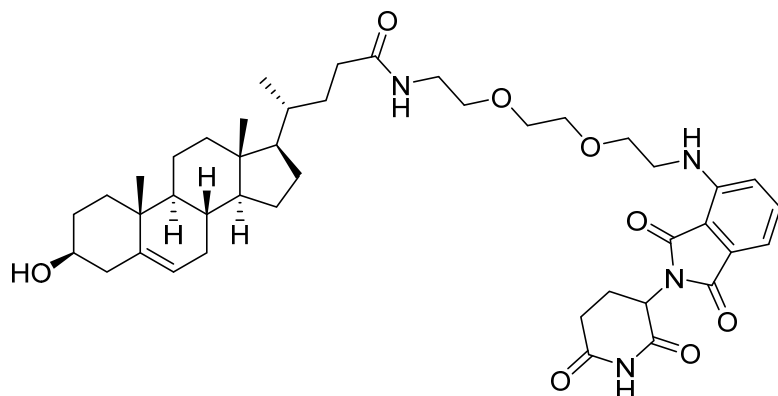
¹H NMR (400 MHz, CDCl₃) δ 7.47 (dd, *J* = 8.5, 7.1 Hz, 1H), 7.08 – 7.06 (m, 1H), 6.89 (d, *J* = 8.5 Hz, 1H), 6.14 (t, *J* = 5.3 Hz, 1H), 5.33 – 5.31 (m, 1H), 4.92 – 4.87 (m, 1H), 3.53 – 3.47 (m, 5H), 3.47 – 3.43 (m, 4H), 3.42 – 3.31 (m, 4H), 3.19 (s, 3H), 2.95 – 2.94 (m, 1H), 2.77 – 2.72 (m, 2H), 2.33 – 2.16 (m, 3H), 2.15 – 2.00 (m, 3H), 1.98 – 1.89 (m, 4H), 1.88 – 1.72 (m, 6H), 1.66 – 1.64 (m, 4H), 1.51 – 1.38 (m, 6H), 1.37 – 1.21 (m, 3H), 1.17 – 1.01 (m, 4H), 0.98 (s, 3H), 0.93 – 0.88 (m, 4H), 0.65 (s, 3H).

¹³C NMR (101 MHz, CDCl₃) δ 173.72, 171.37, 169.64, 169.16, 167.88, 147.05, 140.90, 136.14, 132.66, 121.68, 116.65, 111.44, 110.10, 71.81, 71.03, 69.99, 68.35, 56.84, 55.93, 50.19, 49.71, 42.47, 42.38, 40.29, 39.87, 38.39, 37.37, 36.59, 35.64, 33.77, 32.03, 31.98, 31.73, 29.60, 29.24, 28.29, 27.36, 26.68, 26.53, 24.37, 22.26, 21.17, 19.51, 18.53, 11.99.

HRMS (ESI+) *m/z* found 831.5270 [M+H]⁺, C₄₈H₇₀N₄O₈ calculated 831.5272 (Δ = -0.24 ppm)

Purity: 93 %. **Retention time**: 2.35 min

tert-butyl (3-(4-(3-(2-(adamantan-1-yl)acetamido)propoxy)butoxy)propyl)carbamate (16)
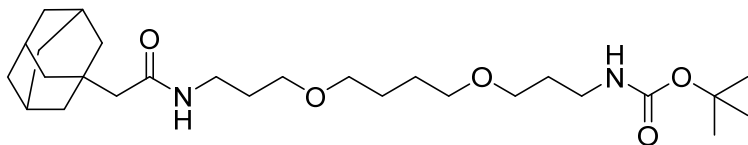
1-Adamantylacetic acid (200 mg, 1.03 mmol) and **9** (376 mg, 1.23 mmol) were combined according to general procedure III. The product was a colorless oil (344 mg, 69%).

¹H NMR (400 MHz, DMSO-d_6) δ 7.64 (t, J = 5.6 Hz, 1H), 6.76 (t, J = 5.7 Hz, 1H), 3.54 – 3.45 (m, 1H), 3.38 – 3.33 (m, 8H), 3.06 (q, J = 6.6 Hz, 2H), 2.96 (q, J = 6.6 Hz, 2H), 1.91 (s, 3H), 1.80 (s, 2H), 1.69 – 1.48 (m, 19H), 1.38 (s, 9H).

(R)-N-(3-(4-(3-(2-(adamantan-1-yl)acetamido)propoxy)butoxy)propyl)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)pentanamide (C3_HyT)
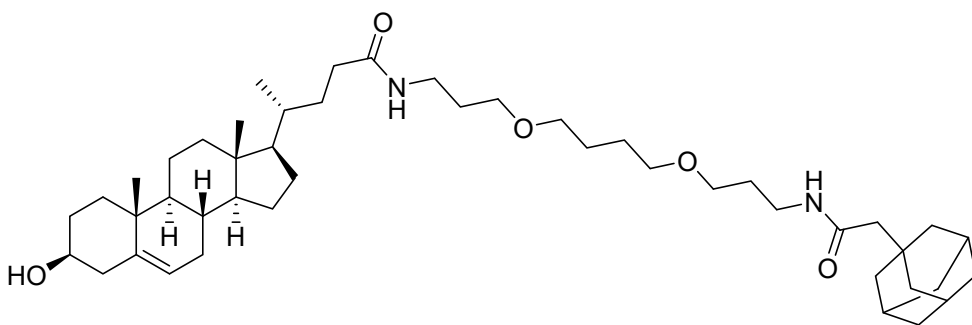
4 (45 mg, 0.12 mmol) and **16** (55 mg, 0.144 mmol) were combined according to general procedure III. The product was a transparent oil (44 mg, 50%).

¹H NMR (400 MHz, CDCl₃) δ 6.15 (t, J = 5.4 Hz, 1H), 6.04 (t, J = 5.3 Hz, 1H), 5.35 – 5.28 (m, 1H), 3.53 – 3.44 (m, 5H), 3.44 – 3.36 (m, 4H), 3.35 – 3.27 (m, 4H), 2.31 – 2.13 (m, 4H), 2.08 – 2.01 (m, 2H), 1.99 – 1.93 (m, 4H), 1.88 (s, 2H), 1.86 – 1.78 (m, 3H), 1.77 – 1.70 (m, 5H), 1.69 (s, 1H), 1.65 (s, 2H), 1.64 – 1.60 (m, 6H), 1.59 – 1.57 (m, 7H), 1.55 – 1.35 (m, 8H), 1.35 – 1.17 (m, 3H), 1.14 – 1.02 (m, 4H), 0.98 (s, 3H), 0.92 – 0.86 (m, 4H), 0.65 (s, 3H).

¹³C NMR (101 MHz, CDCl₃) δ 173.65, 170.99, 140.95, 121.62, 71.74, 70.91, 69.91, 69.84, 56.83, 55.93, 52.03, 50.18, 42.77, 42.46, 42.37, 39.86, 38.18, 38.07, 37.37, 36.87, 36.59, 35.63, 33.77, 32.80, 31.97, 31.71, 29.46, 29.32, 28.74, 28.28, 26.68, 26.64, 24.36, 21.16, 19.50, 18.52, 11.98.

HRMS (ESI+) m/z found 737.5822 [M+H]⁺, $C_{48}H_{70}N_4O_8$ calculated 737.5834 (Δ = -1.62 ppm)

$[\alpha]_D^{20}$: -24° (c=0.9, EtOH)

tert-butyl (3-(4-(3-((R)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)pentanamido)propoxy)butoxy)propyl)carbamate (C3_ Boc)
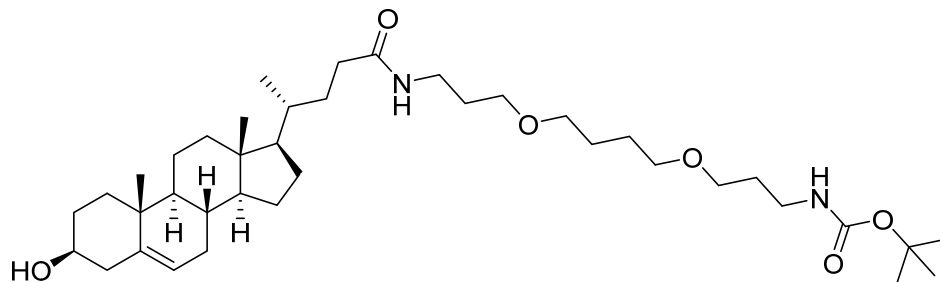
4 (90 mg, 0.24 mmol) and **9** (88 mg, 0.288 mmol) were combined according to general procedure III. The product was a transparent oil (130 mg, 82%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 6.18 (s, 1H), 5.37 – 5.33 (m, 1H), 4.93 (s, 1H), 3.53 – 3.47 (m, 5H), 3.47 – 3.43 (m, 4H), 3.35 (q, J = 6.1 Hz, 2H), 3.21 (q, J = 6.2 Hz, 2H), 2.33 – 2.17 (m, 3H), 2.17 – 1.92 (m, 4H), 1.91 – 1.69 (m, 8H), 1.68 – 1.62 (m, 4H), 1.60 – 1.46 (m, 5H), 1.44 (s, 10H), 1.37 – 1.23 (m, 2H), 1.20 – 1.01 (m, 5H), 1.00 (s, 3H), 0.97 – 0.92 (m, 4H), 0.68 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 173.71, 156.13, 140.92, 121.72, 71.84, 70.99, 70.82, 69.99, 69.31, 56.86, 55.96, 50.21, 42.49, 42.40, 39.89, 38.37, 37.39, 36.61, 35.66, 33.78, 32.00, 31.75, 29.88, 29.22, 28.56, 28.30, 26.65, 26.61, 24.39, 21.19, 19.52, 18.54, 12.01.

MS *m/z* (ES+) calcd for C$_{39}$H$_{69}$N$_2$O$_6$ [M+ H$^+$] 661.5 found 661.5.

$[\alpha]_D^{20}$: -31.2° (c=1.00, EtOH)

(4R)-N-(3-(2-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)ethoxy)propyl)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)pentanamide (C4)
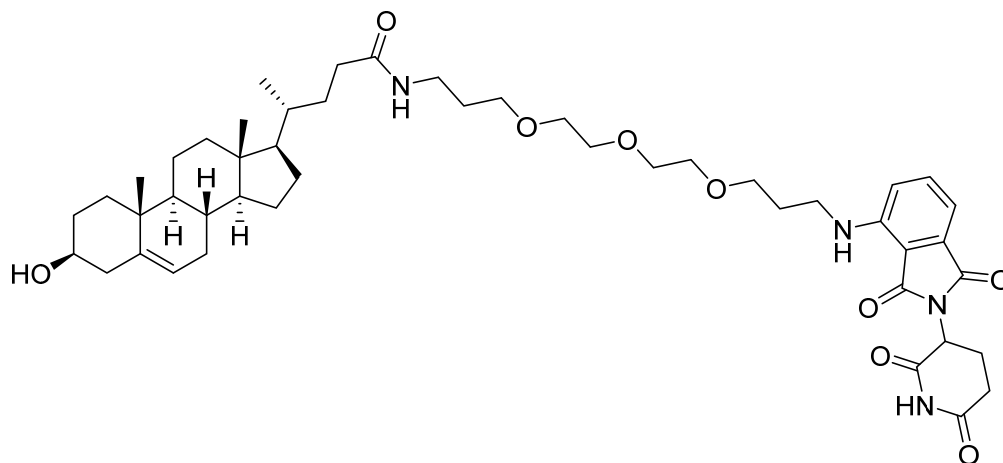
4 (80 mg, 0.21 mmol) and **15** (210 mg, 0.252 mmol) were combined according to general procedure III. The product was a yellow solid (100 mg, 57%).

¹H NMR (400 MHz, CDCl$_3$) δ 8.53 (s, 1H), 7.47 (dd, J = 8.5, 7.1 Hz, 1H), 7.07 (d, J = 7.1 Hz, 1H), 6.92 (d, J = 8.5 Hz, 1H), 6.26 (t, J = 5.5 Hz, 1H), 5.33 – 5.31 (m, 1H), 4.92 – 4.87 (m, 1H), 3.68 – 3.52 (m, 12H), 3.52 – 3.44 (m, 1H), 3.40 – 3.30 (m, 4H), 2.88 – 2.67 (m, 3H), 2.26 – 2.16 (m, 3H), 2.14 – 2.00 (m, 3H), 1.97 – 1.88 (m, 4H), 1.85 – 1.70 (m, 6H), 1.57 – 1.39 (m, 7H), 1.33 – 1.21 (m, 2H), 1.14 – 1.01 (m, 4H), 0.98 (s, 3H), 0.97 – 0.90 (m, 4H), 0.65 (s, 3H).

¹³C NMR (101 MHz, CDCl$_3$) δ 173.98, 171.35, 169.45, 168.63, 167.74, 147.06, 140.90, 136.22, 132.61, 121.68, 116.75, 111.49, 110.00, 71.81, 70.59, 70.56, 70.54, 70.18, 70.09, 69.05, 56.83, 55.95, 50.18, 48.96, 42.46, 42.36, 40.36, 39.86, 38.02, 37.36, 36.59, 35.64, 33.63, 31.97, 31.71, 31.52, 29.37, 29.03, 28.27, 24.37, 22.94, 21.17, 19.51, 18.53, 11.99.

HRMS (ESI+) *m/z* found 833.5055 [M+H]$^+$, C$_{47}$H$_{68}$N$_4$O$_9$ calculated 833.5065 (Δ = -1.19 ppm)

Purity: 98 %. **Retention time**: 2.17 min

(R)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-methylpentanamide (C5)
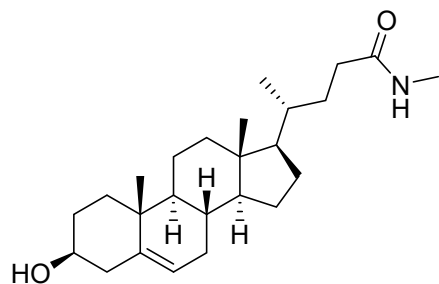
4 (50 mg, 0.133 mmol) and methylamine (5 mg, 0.16 mmol) were combined according to general procedure III. The product was a white powder (34 mg, 65%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 5.65 – 5.49 (m, 1H), 5.39 – 5.35 (m, 1H), 3.59 – 3.50 (m, 1H), 2.84 (s, 3H), 2.34 – 2.22 (m, 2H), 2.11 – 2.01 (m, 4H), 1.91 – 1.84 (m, 3H), 1.62 – 1.43 (m, 7H), 1.37 – 1.27 (m, 4H), 1.19 – 1.06 (m, 4H), 1.03 (s, 3H), 0.98 – 0.93 (m, 4H), 0.71 (d, J = 2.5 Hz, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.63, 140.88, 121.82, 71.94, 56.87, 55.96, 50.22, 42.52, 42.43, 39.90, 37.39, 36.64, 35.70, 33.63, 32.02, 31.79, 28.32, 26.57, 24.41, 21.22, 19.54, 18.56, 12.03.

HRMS (ESI+) *m/z* found 388.3216 [M+H]$^+$, C$_{25}$H$_{41}$NO$_2$ calculated 388.3212 (Δ = -0.07 ppm)

$[\alpha]_D^{20}$: -33.8° (c=0.7, EtOH)

tert-butyl (3-(4-(3-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propoxy)butoxy)propyl)carbamate (17)
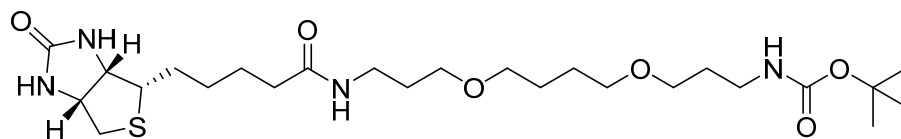
D-biotin (200 mg, 0.82 mmol) and **9** (300 mg, 0.98 mmol) were combined according to general procedure III. The product was a white powder (265 mg, 61%)

The data matched those reported in the literature[211].

^{1}H NMR (400 MHz, CDCl$_3$) δ 4.58 – 4.50 (m, 1H), 4.38 – 4.30 (m, 1H), 3.56 – 3.46 (m, 6H), 3.44 (s, 4H), 3.25 – 3.18 (m, 2H), 3.18 – 3.16 (m, 1H), 2.95 – 2.88 (m, 1H), 2.80 – 2.74 (m, 1H), 2.24 – 2.18 (m, 2H), 1.80 – 1.72 (m, 5H), 1.71 – 1.57 (m, 7H), 1.44 (s, 11H).

MS *m/z* (ES+) calcd for C$_{25}$H$_{46}$N$_4$O$_6$S [M+ H$^+$] 531.3 found 531.1.

(R)-4-((3S,8S,9S,10R,13R,14S,17R)-3-hydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-(3-(4-(3-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)propoxy)butoxy)propyl)pentanamide (C3_Bio)
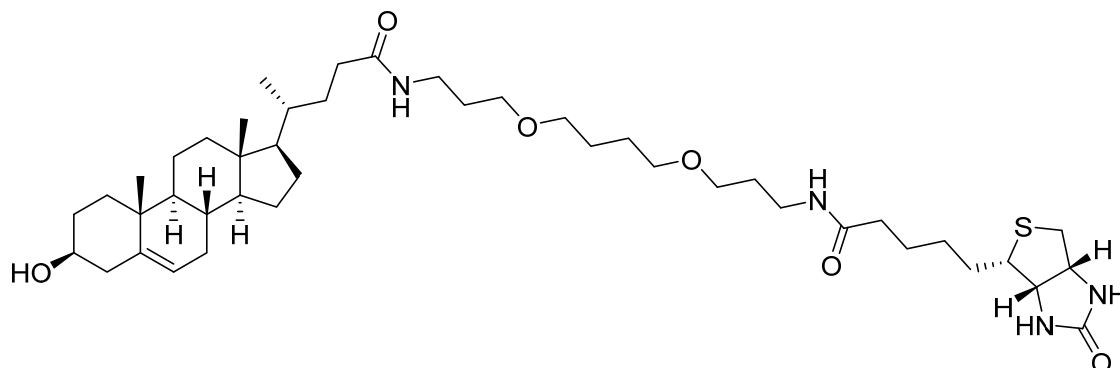
4 (72 mg, 0.20 mmol) and **17** (100 mg, 0.232 mmol) were combined according to general procedure III. The product was a white solid (62 mg, 40%).

¹H NMR (400 MHz, CDCl₃) δ 6.53 (s, 1H), 6.26 (s, 1H), 6.18 (s, 1H), 5.38 – 5.34 (m, 1H), 4.57 – 4.52 (m, 1H), 4.38 – 4.32 (m, 1H), 3.55 – 3.50 (m, 8H), 3.46 (s, 4H), 3.36 (s, 4H), 3.21 – 3.14 (m, 1H), 2.97 – 2.90 (m, 1H), 2.80 – 2.75 (m, 1H), 2.25 – 2.01 (m, 11H), 1.87 – 1.77 (m, 7H), 1.70 – 1.62 (m, 6H), 1.55 – 1.44 (m, 8H), 1.36 – 1.28 (m, 2H), 1.21 – 1.04 (m, 4H), 0.99 (s, 3H), 0.94 – 0.90 (m, 4H), 0.67 (s, 3H).

¹³C NMR (101 MHz, CDCl₃) δ 173.89, 173.18, 163.78, 140.94, 121.73, 71.86, 70.95, 70.90, 69.92, 69.71, 62.05, 60.41, 56.89, 55.96, 55.57, 50.96, 50.23, 42.51, 42.40, 40.65, 39.91, 38.30, 38.10, 37.40, 36.63, 36.09, 35.67, 33.79, 32.02, 31.75, 29.37, 29.24, 28.32, 28.24, 28.19, 26.68, 25.73, 24.41, 21.21, 19.54, 18.58, 12.03.

MS *m/z* (ES+) calcd for C₄₄H₇₄N₄O₆S [M+ H⁺] 787.5 found 787.6.

$[\alpha]_D^{20}$: 7.7° (c=0.3, EtOH)

6.1.3 Synthesis of 4beta-hydroxycholesterol-bearing PROTACs.

3β, 4β-dihydroxy-5-cholenoic acid (18)
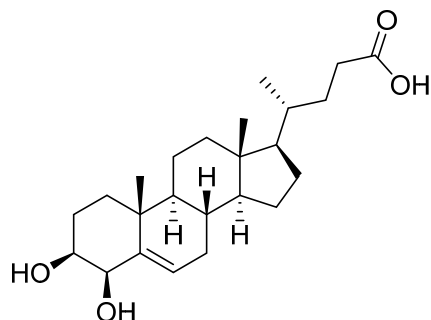
3β-hydroxy-5-cholenoic acid (**4**, 200 mg, 0.53 mmol) and SeO_2 (94 mg, 0.85 mmol) were dissolved in 10 mL 1,4-dioxane and 2 mL H_2O. The solution was refluxed at 100 °C overnight. And then, the solution was cooled down to room temperature. The black precipitate was filtered and the filtrate was dissolved in 30 mL Et_2O and then the filtrate was washed by 40 mL water and 40 mL brine respectively. Then, the organic phase was dried by anhydrous Na_2SO_4, filtered, and the solvent was removed by rotatory-evaporator. The crude was a red-brown crude product. The crude product was purified by flash chromatography on silica gel (40:1 DCM/MeOH) resulting in a yellow powder (81 mg, 39 %).

^{1}H NMR (400 MHz, MeOD): δ 5.67 – 5.60 (m, 1H), 4.05 (d, J = 3.5 Hz, 1H), 3.49 – 3.42 (m, 1H), 2.38 – 2.29 (m, 1H), 2.25 – 1.75 (m, 8H), 1.69 – 1.42 (m, 6H), 1.41 – 1.23 (m, 3H), 1.20 (s, 3H), 1.18 – 0.99 (m, 4H), 0.96 (d, J = 6.5 Hz, 3H), 0.93 – 0.87 (m, 1H), 0.73 (s, 3H).

^{13}C NMR (101 MHz, MeOD): δ 178.18, 144.44, 128.54, 78.55, 73.81, 58.29, 57.26, 51.87, 43.54, 41.05, 38.62, 37.25, 36.71, 33.24, 33.19, 32.33, 32.01, 29.14, 25.84, 25.27, 21.65, 21.50, 18.79, 12.30.

(4R)-4-((3S,4R,8S,9S,10R,13R,14S,17R)-3,4-dihydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-(10-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)decyl)pentanamide (H1)
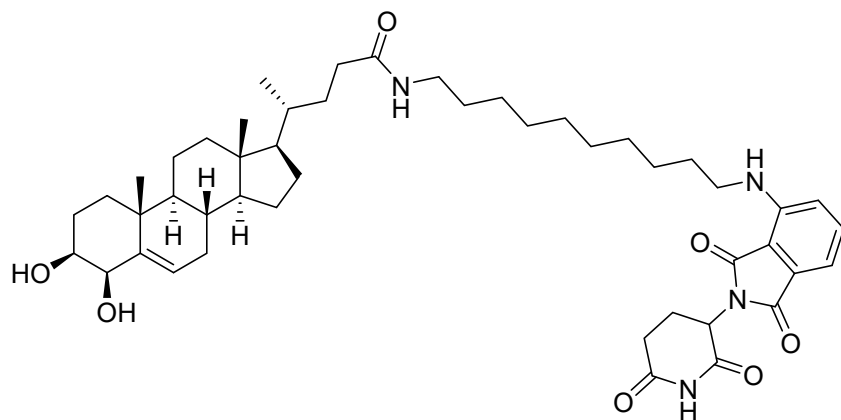
18 (22 mg, 0.056 mmol) and **16** (29 mg, 0.068 mmol) were combined according to general procedure III. The product was a yellow solid (20 mg, 47%).

¹H NMR (400 MHz, CDCl₃) δ 8.30 (s, 1H), 7.51 (dd, J = 8.5, 7.2 Hz, 1H), 7.10 (d, J = 7.1 Hz, 1H), 6.90 (d, J = 8.5 Hz, 1H), 6.24 (s, 1H), 5.71 (s, 1H), 5.70 – 5.65 (m, 1H), 4.98 – 4.90 (m, 1H), 4.14 (d, J = 3.5 Hz, 1H), 3.80 – 3.67 (m, 1H), 3.59 – 3.53 (m, 1H), 3.30 – 3.22 (m, 4H), 3.21 – 3.14 (m, 1H), 2.94 – 2.70 (m, 3H), 2.36 – 2.22 (m, 2H), 2.17 – 1.98 (m, 4H), 1.93 – 1.77 (m, 4H), 1.67 – 1.58 (m, 3H), 1.52 – 1.48 (m, 5H), 1.46 – 1.42 (m, 5H), 1.31 (s, 9H), 1.19 (s, 3H), 1.18 – 0.99 (m, 5H), 0.97 – 0.88 (m, 4H), 0.69 (s, 3H).

¹³C NMR (101 MHz, CDCl₃) δ 174.21, 171.20, 169.64, 168.56, 167.77, 147.15, 142.87, 136.26, 132.61, 128.76, 116.81, 111.49, 109.92, 77.24, 72.59, 57.00, 55.91, 50.27, 48.98, 42.79, 42.50, 39.87, 39.79, 37.07, 36.11, 35.64, 33.69, 32.17, 32.08, 31.93, 31.54, 29.67, 29.54, 29.39, 28.30, 27.04, 27.01, 25.47, 24.36, 22.95, 21.15, 20.65, 18.52, 12.01.

MS *m/z* (ES+) calcd for $C_{47}H_{68}N_4O_7$ [M+ H⁺] 801.5 found 801.2.

(4R)-4-((3S,4R,8S,9S,10R,13R,14S,17R)-3,4-dihydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-(2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethoxy)ethyl)pentanamide (H2)

18 (28 mg, 0.07 mmol) and **12** (35 mg, 0.086 mmol) were combined according to general procedure III. The product was a yellow solid (23 mg, 42%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.77 (s, 1H), 7.58 – 7.45 (m, 1H), 7.13 (d, J = 7.1 Hz, 1H), 6.92 (d, J = 8.5 Hz, 1H), 6.19 (t, J = 5.5 Hz, 1H), 5.72 – 5.62 (m, 1H), 4.98 – 4.87 (m, 1H), 4.14 (d, J = 3.5 Hz, 1H), 3.75 (t, J = 5.3 Hz, 2H), 3.68 (s, 4H), 3.63 – 3.56 (m, 3H), 3.50 – 3.43 (m, 4H), 2.93 – 2.68 (m, 3H), 2.30 – 2.20 (m, 1H), 2.17 – 2.01 (m, 3H), 2.02 – 1.87 (m, 2H), 1.87 – 1.72 (m, 3H), 1.68 – 1.62 (m, 1H), 1.61 – 1.50 (m, 3H), 1.49 – 1.21 (m, 6H), 1.18 (s, 3H), 1.14 – 0.93 (m, 5H), 0.92 – 0.87 (m, 4H), 0.66 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.12, 171.29, 169.40, 168.57, 167.55, 146.72, 142.73, 136.12, 132.55, 128.59, 116.71, 111.80, 110.39, 77.26, 72.46, 70.67, 70.08, 69.91, 69.23, 56.85, 55.76, 50.16, 48.90, 42.34, 42.31, 39.66, 39.37, 36.97, 35.99, 35.51, 33.44, 32.03, 31.78, 31.74, 31.39, 28.14, 25.32, 24.22, 22.89, 21.03, 20.52, 18.37, 11.88.

MS *m/z* (ES+) calcd for C$_{43}$H$_{60}$N$_4$O$_9$ [M+ H$^+$] 777.4 found 777.5.

(4R)-4-((3S,4R,8S,9S,10R,13R,14S,17R)-3,4-dihydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-(3-(4-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)butoxy)propyl)pentanamide (H3)
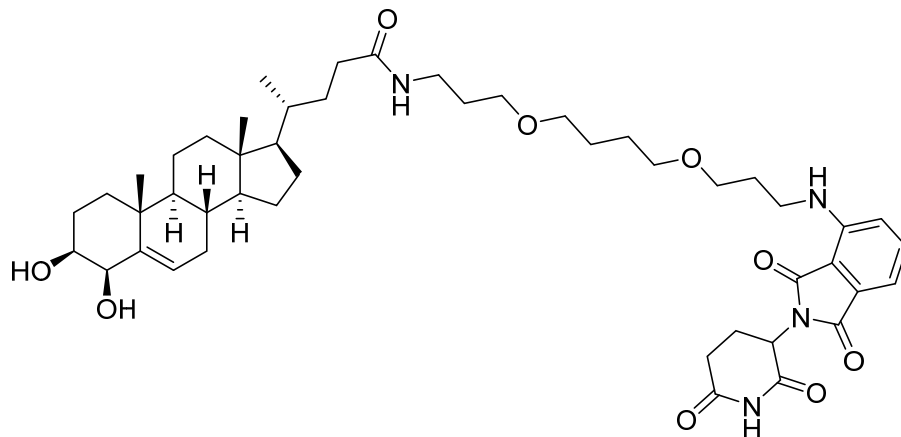
18 (28 mg, 0.07 mmol) and **13** (40 mg, 0.086 mmol) were combined according to general procedure III. The product was a yellow solid (26 mg, 45%).

¹H NMR (400 MHz, Chloroform-*d*) δ 8.52 (s, 1H), 7.50 (dd, *J* = 8.5, 7.2 Hz, 1H), 7.10 (d, *J* = 7.1 Hz, 1H), 6.93 (d, *J* = 8.5 Hz, 1H), 6.30 (t, *J* = 5.3 Hz, 1H), 5.71 – 5.64 (m, 1H), 4.96 – 4.89 (m, 1H), 4.14 (d, *J* = 3.4 Hz, 1H), 3.60 – 3.50 (m, 5H), 3.49 – 3.43 (m, 4H), 3.42 – 3.34 (m, 4H), 2.92 – 2.70 (m, 3H), 2.29 – 2.21 (m, 1H), 2.18 – 2.04 (m, 3H), 2.02 – 1.84 (m, 5H), 1.85 – 1.73 (m, 4H), 1.72 – 1.62 (m, 5H), 1.60 – 1.51 (m, 3H), 1.50 – 1.40 (m, 3H), 1.36 – 1.25 (m, 3H), 1.18 (s, 3H), 1.16 – 0.96 (m, 5H), 0.96 – 0.87 (m, 4H), 0.68 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 173.91, 171.23, 169.34, 168.51, 167.64, 146.95, 142.73, 136.08, 132.52, 128.59, 116.59, 111.38, 109.92, 77.24, 72.44, 71.02, 70.89, 69.73, 68.49, 56.88, 55.76, 50.16, 48.84, 42.36, 40.40, 39.67, 38.30, 36.96, 35.98, 35.53, 33.57, 32.04, 31.88, 31.79, 31.40, 29.36, 29.06, 28.16, 26.50, 26.42, 25.32, 24.22, 22.85, 21.02, 20.52, 18.40, 11.89.

MS *m/z* (ES+) calcd for $C_{47}H_{68}N_4O_9$ [M+ H$^+$] 833.5 found 833.5.

(4R)-4-((3S,4R,8S,9S,10R,13R,14S,17R)-3,4-dihydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-(3-(2-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)ethoxy)propyl)pentanamide (H4)
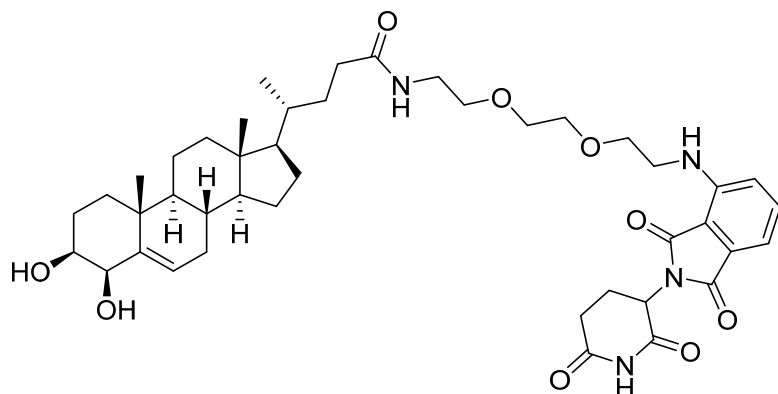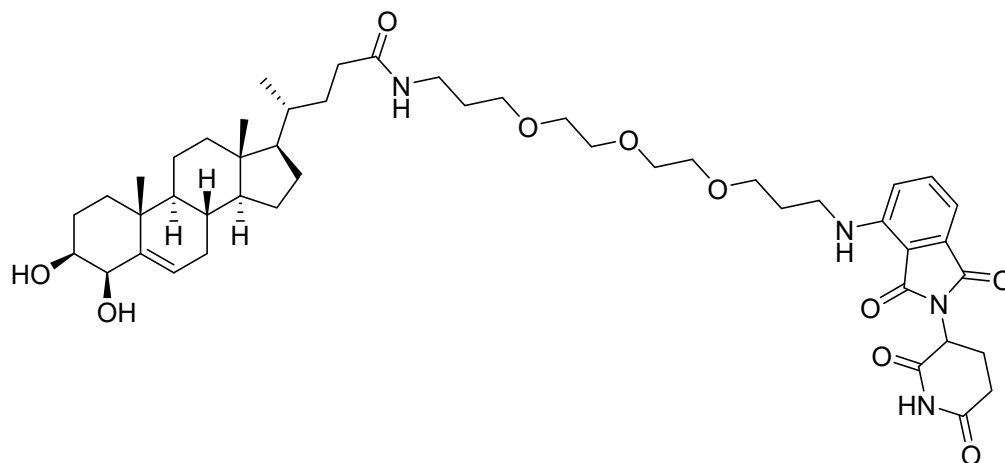

18 (20 mg, 0.05 mmol) and **14** (24 mg, 0.06 mmol) were combined according to general procedure III. The product was a yellow solid (17 mg, 40%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.26 (s, 1H), 7.52 (dd, J = 8.5, 7.2 Hz, 1H), 7.12 (d, J = 7.1 Hz, 1H), 6.96 (d, J = 8.5 Hz, 1H), 6.59 (s, 1H), 5.71 – 5.66 (m, 1H), 4.95 – 4.90 (m, 1H), 4.15 (d, J = 3.6 Hz, 1H), 3.75 – 3.66 (m, 4H), 3.65 – 3.56 (m, 9H), 3.46 – 3.36 (m, 4H), 2.93 – 2.75 (m, 3H), 2.38 – 2.25 (m, 2H), 2.18 – 2.08 (m, 3H), 2.00 – 1.90 (m, 4H), 1.84 – 1.77 (m, 4H), 1.62 – 1.52 (m, 3H), 1.49 – 1.42 (m, 3H), 1.40 – 1.22 (m, 4H), 1.19 (s, 3H), 1.16 – 1.00 (m, 5H), 0.97 – 0.90 (m, 4H), 0.69 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 171.04, 169.32, 168.41, 167.59, 146.87, 142.73, 136.14, 132.50, 130.87, 128.78, 128.67, 116.70, 111.49, 109.96, 72.45, 70.43, 70.41, 70.05, 69.97, 69.00, 68.54, 56.85, 55.75, 50.11, 48.84, 42.35, 40.35, 39.64, 37.08, 36.90, 35.96, 35.53, 33.49, 32.16, 32.02, 31.86, 31.77, 31.38, 31.16, 29.19, 28.76, 28.14, 26.36, 25.35, 24.20, 22.83, 22.61, 21.01, 20.50, 19.82, 18.37, 11.86.

MS *m/z* (ES+) calcd for C$_{47}$H$_{68}$N$_4$O$_{10}$ [M+ H$^+$] 849.5 found 849.9.

(R)-4-((3S,4R,8S,9S,10R,13R,14S,17R)-3,4-dihydroxy-10,13-dimethyl-2,3,4,7,8,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-N-methylpentanamide (H5)
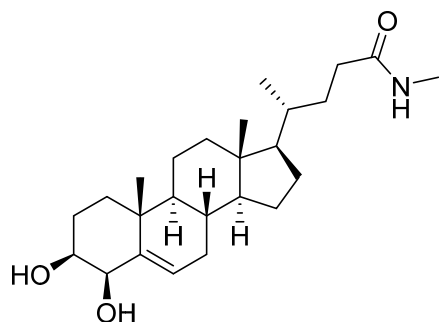
18 (25 mg, 0.065 mmol) and methylamine (3 mg, 0.078 mmol) were combined according to general procedure III. The product was a white powder (17 mg, 65%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 5.69 (dd, *J* = 4.9, 1.7 Hz, 1H), 5.57 (s, 1H), 4.15 (d, *J* = 3.5 Hz, 1H), 3.60 – 3.54 (m, 1H), 2.82 (d, *J* = 4.3 Hz, 3H), 2.29 – 2.23 (m, 1H), 2.18 (s, 2H), 2.13 – 2.04 (m, 2H), 2.03 – 1.97 (m, 1H), 1.95 – 1.76 (m, 4H), 1.70 – 1.52 (m, 4H), 1.52 – 1.40 (m, 3H), 1.37 – 1.29 (m, 2H), 1.20 (s, 3H), 1.14 – 1.06 (m, 3H), 0.98 – 0.89 (m, 4H), 0.70 (s, 3H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.31, 142.74, 128.68, 77.27, 72.46, 56.88, 55.80, 50.16, 42.37, 39.68, 36.94, 35.99, 35.55, 33.53, 32.04, 31.85, 31.81, 28.17, 26.36, 25.37, 24.23, 21.04, 20.53, 18.41, 11.89

MS *m/z* (ES+) calcd for C$_{25}$H$_{41}$NO$_3$ [M+ H$^+$] 404.3 found 404.2.

$[\alpha]_D^{20}$: -42.0° (c=0.45, EtOH)

6.1.4 Synthesis of Aster-A PROTACs.

tert-butyl 2-amino-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (19)
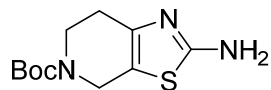
N-BOC piperidone (3.5 g, 17.5 mmol) and dibromo barbituric acid (3.0 g, 10.5 mmol) were dissolved in dry diethyl ether (50 mL) and the reaction mixture was stirred at rt for 16 h. Then the precipitate was filtered off and washed with diethyl ether. The filtrate was concentrated under reduced pressure to yield the bromopiperidinone, which was used without further purification.

Then bromopiperidinone and thiourea (1.5 g, 19.5 mmol) were dissolved in 30 mL ethanol. The mixture was heated at 80 °C for 4 h. The orange precipitate was filtered off and the filtrate was concentrated under reduced pressure. Then the crude was dissolved in EtOAc and saturated NaHCO$_3$ and the aqueous phase was washed with EtOAc. The organic layer was collected, dried with anhydrous Na$_2$SO$_4$ and concentrated under reduced pressure to yield a yellow solid. (960 mg, 44% over two steps).

tert-butyl 2-(4-(methoxycarbonyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (20)
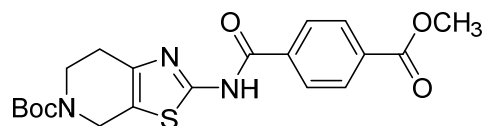
Mono-methyl terephthalate (846 mg, 4.7 mmol), DIPEA (1.35 mL, 7.8 mmol) and HATU (1.78 g, 4.7 mmol) were dissolved in 20 mL CH$_2$Cl$_2$. After 5 min, **19** (1 g, 3.9 mmol) was dissolved in 5 mL CH$_2$Cl$_2$ and added to the same flask. The mixture was stirred at rt overnight. Then, the mixture was poured into H$_2$O (10 mL) and extracted by CH$_2$Cl$_2$ (3 x 10 mL). The organic phase was collected, dried over anhydrous Na$_2$SO$_4$ and evaporated under reduced pressure. The residue was purified by column chromatography (gradient of pentane/EA 8:1 to 2:1).

White solid (1 g, 62 %)

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.19 (d, J = 8.2 Hz, 2H), 8.11 (d, J = 8.2 Hz, 2H), 4.60 (s, 2H), 3.96 (s, 3H), 3.73 (t, J = 5.8 Hz, 2H), 2.74 – 2.62 (m, 2H), 1.49 (s, 9H).

MS *m/z* (ES+) calcd for C$_{20}$H$_{23}$N$_3$O$_5$S [M+ H$^+$] 418.1 found 418.1.

4-((5-(tert-butoxycarbonyl)-4,5,6,7-tetrahydrothiazolo[5,4-c]pyridin-2-yl)carbamoyl)benzoic acid (21)
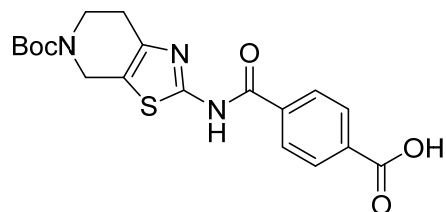
20 (590 mg, 1.41 mmol) was dissolved in 10 mL THF and 5 mL 10 % NaOH (aq) was added to the reaction mixture. The solution was stirred overnight at room temperature. Then, 1 M HCl was added gradually and the precipitate was filtered off and washed with H_2O to obtain compound **21**.

White solid (500 mg, 88%).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.17 (d, J = 8.3 Hz, 2H), 8.09 – 8.05 (m, 2H), 4.55 (s, 2H), 3.68 (t, J = 5.8 Hz, 2H), 2.70 (t, J = 6.0 Hz, 2H), 1.44 (s, 9H).

MS *m/z* (ES-) calcd for $C_{19}H_{21}N_3O_5S$ [M-H]$^-$ 402.1, found: 402.0.

General procedure IV: Boc-deprotection and HATU-mediated amide coupling.

The corresponding mono-Boc protected linker coupled pomalidomide was dissolved in CH_2Cl_2 (3 mL) and then trifluoroacetic acid (1 mL) was added dropwise at 0 °C. The mixture was then stirred at rt for 2 h. Then the solvent was removed, washed by $NaHCO_3$ and extracted by CH_2Cl_2. The orange oily residue was further dried under vacuum.

The carboxylic acid (1 eq), DIPEA (2 eq) and HATU (1.5 eq) were dissolved in 10 mL CH_2Cl_2. After 60 min, deprotected amine (1.2 eq) was dissolved in 5 mL CH_2Cl_2 and added to the same flask. The mixture was stirred at rt overnight. Then, the mixture was poured into H_2O (10 mL) and extracted by CH_2Cl_2 (3 x 10 mL). The organic phase was collected, dried over anhydrous sodium sulfate and evaporated under reduced pressure. The residue was purified by column chromatography (gradient of CH_2Cl_2/MeOH 80:1 to 30:1).

tert-butyl 2-(4-((6-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)hexyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF1)
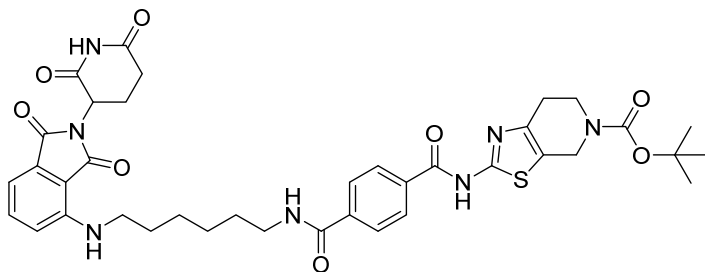

21 (60 mg, 0.165 mmol) and **11** (80 mg, 0.198 mmol) were combined according to general procedure IV. The product was a yellow solid (47 mg, 38%).

¹H NMR (400 MHz, Chloroform-*d*) δ 10.94 (s, 1H), 8.02 (d, *J* = 8.2 Hz, 2H), 7.64 (d, *J* = 8.1 Hz, 2H), 7.48 (dd, *J* = 8.5, 7.1 Hz, 1H), 7.06 (d, *J* = 7.1 Hz, 1H), 6.88 (d, *J* = 8.7 Hz, 1H), 6.58 (t, *J* = 5.8 Hz, 1H), 6.20 (d, *J* = 6.0 Hz, 1H), 5.00 – 4.93 (m, 1H), 4.55 (s, 2H), 3.71 (t, *J* = 5.8 Hz, 2H), 3.47 – 3.37 (m, 2H), 3.31 – 3.24 (m, 2H), 2.89 – 2.71 (m, 5H), 2.18 – 2.12 (m, 1H), 1.68 – 1.59 (m, 4H), 1.47 (s, 9H), 1.47 – 1.39 (m, 4H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 172.18, 170.73, 169.62, 167.70, 166.55, 163.93, 158.09, 154.61, 147.00, 138.61, 136.24, 133.80, 132.39, 128.21, 127.23, 116.80, 111.45, 109.69, 80.4653.45, 48.93, 42.35, 39.99, 31.57, 29.23, 29.05, 28.42, 26.59, 26.54, 22.79.

HRMS (ESI+) *m/z* found 758.2969 [M+H]⁺, $C_{38}H_{43}N_7O_8S$ calculated 758.2972 (Δ = -0.39 ppm)

tert-butyl 2-(4-((2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethoxy)ethyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF2)

21 (100 mg, 0.25 mmol) and **12** (120 mg, 0.3 mmol) were combined according to general procedure IV. The product was a yellow solid (65 mg, 33%).

¹H NMR (400 MHz, CDCl₃) δ 11.03 (s, 1H), 8.09 (d, *J* = 8.1 Hz, 2H), 7.70 (d, *J* = 8.1 Hz, 2H), 7.50 (dd, *J* = 8.5, 7.1 Hz, 1H), 7.08 (dd, *J* = 9.8, 7.6 Hz, 2H), 6.74 (d, *J* = 6.8 Hz, 1H), 6.53 (s, 1H), 4.95 (dd, *J* = 11.9, 5.3 Hz, 1H), 4.62 (s, 2H), 3.82 – 3.73 (m, 4H), 3.70 (d, *J* = 4.4 Hz, 8H), 3.57 – 3.38 (m, 2H), 2.99 – 2.68 (m, 5H), 2.21 – 2.08 (m, 1H), 1.52 (s, 9H).

¹³C NMR (101 MHz, CDCl₃) δ 171.66, 170.89, 169.53, 167.51, 166.22, 163.85, 159.18, 154.50, 147.20, 138.47, 135.84, 133.23, 132.24, 128.43, 127.35, 117.59, 111.72, 110.13, 80.70, 70.61, 70.51, 69.72, 48.93, 42.71, 40.13, 31.61, 28.42, 22.88.

HRMS (ESI+) *m/z* found 790.2864 [M+H]⁺, $C_{38}H_{43}N_7O_{10}S$ calculated 790.2870 (Δ = -0.76 ppm)

tert-butyl 2-(4-((3-(4-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)butoxy)propyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF3)

21 (100 mg, 0.25 mmol) and **13** (120 mg, 0.3 mmol) were combined according to general procedure IV. The product was a yellow solid (83 mg, 39%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 11.09 (s, 1H), 8.08 (d, J = 8.2 Hz, 2H), 7.63 (d, J = 8.2 Hz, 2H), 7.51 (dd, J = 8.5, 7.2 Hz, 1H), 7.10 (d, J = 7.0 Hz, 1H), 7.02 (d, J = 5.2 Hz, 1H), 6.94 (d, J = 8.7 Hz, 1H), 6.31 (s, 1H), 5.02 – 4.93 (m, 1H), 4.60 (s, 2H), 3.79 – 3.70 (m, 2H), 3.66 – 3.56 (m, 4H), 3.52 – 3.46 (m, 4H), 3.44 – 3.39 (m, 2H), 3.38 – 3.35 (m, 2H), 2.93 – 2.75 (m, 5H), 2.23 – 2.08 (m, 1H), 1.92 – 1.88 (m, 2H), 1.88 – 1.82 (m, 2H), 1.71 – 1.63 (m, 2H), 1.51 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 171.91, 171.00, 169.48, 167.64, 166.01, 163.93, 158.25, 154.56, 146.97, 138.58, 136.15, 133.62, 132.35, 128.18, 127.10, 126.90, 116.71, 111.36, 109.65, 80.43, 70.99, 70.66, 70.21, 68.02, 53.42, 48.90, 48.83, 40.11, 39.29, 31.54, 29.43, 28.93, 28.39, 28.38, 26.68, 26.33, 22.90.

HRMS (ESI+) m/z found 846.3497 [M+H]$^+$, C$_{42}$H$_{51}$N$_7$O$_{10}$S calculated 846.3496 (Δ = 0.11 ppm)

Purity: 100 %. **Retention time**: 2.17 min

tert-butyl 2-(4-((3-(2-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)ethoxy)propyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF4)
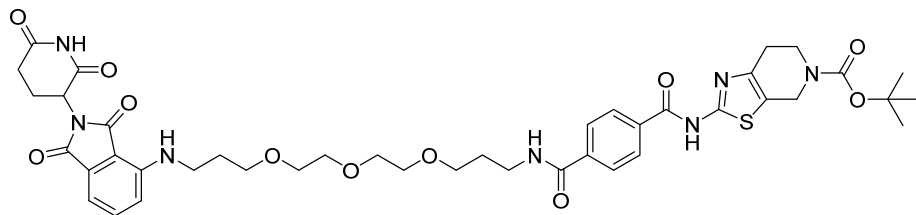
21 (40 mg, 0.1 mmol) and **15** (58 mg, 0.12 mmol) were combined according to general procedure IV. The product was a yellow solid (35 mg, 41 %).

^{1}H NMR (400 MHz, CDCl$_3$) δ 11.60 (s, 1H), 8.08 (d, J = 8.2 Hz, 2H), 7.61 (dd, J = 8.5, 7.1 Hz, 1H), 7.56 (d, J = 8.2 Hz, 2H), 7.15 (s, 1H), 7.06 (d, J = 8.5 Hz, 1H), 6.94 (s, 1H), 6.25 (s, 1H), 5.05 – 4.94 (m, 1H), 4.68 – 4.49 (m, 2H), 3.84 – 3.71 (m, 2H), 3.70 – 3.67 (m, 4H), 3.68 – 3.62 (m, 6H), 3.52 – 3.36 (m, 6H), 3.01 – 2.74 (m, 5H), 2.25 – 2.09 (m, 1H), 1.96 – 1.88 (m, 2H), 1.88 – 1.74 (m, 2H), 1.51 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 171.77, 171.64, 169.58, 167.71, 166.00, 163.98, 158.46, 154.57, 147.10, 138.47, 136.57, 133.50, 132.28, 128.19, 127.13, 117.09, 111.41, 109.50, 80.41, 70.62, 70.41, 70.21, 70.18, 69.78, 68.19, 48.97, 39.62, 38.46, 31.64, 29.67, 28.98, 28.40, 22.87.

HRMS (ESI+) m/z found 862.344 [M+H]$^+$, C$_{42}$H$_{51}$N$_7$O$_{11}$S calculated 862.3445 (Δ = -0.58 ppm)

116

tert-butyl 2-(4-((10-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)decyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF5)
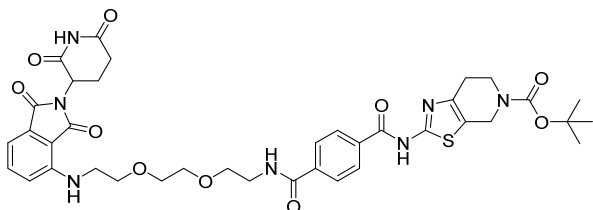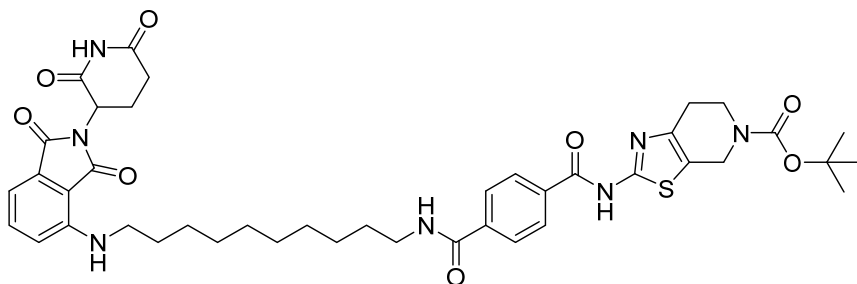

21 (25 mg, 0.063 mmol) and **16** (30 mg, 0.075 mmol) were combined according to general procedure IV. The product was a yellow solid (18 mg, 36%).

^{1}H NMR (400 MHz, Chloroform-d) δ 10.60 (s, 1H), 8.10 (d, J = 8.2 Hz, 2H), 7.74 (d, J = 8.2 Hz, 2H), 7.57 – 7.50 (m, 1H), 7.11 (d, J = 7.1 Hz, 1H), 6.92 (d, J = 8.6 Hz, 1H), 6.32 (t, J = 5.7 Hz, 1H), 6.24 (t, J = 5.6 Hz, 1H), 5.00 – 4.92 (m, 1H), 4.60 (s, 2H), 3.75 (t, J = 5.9 Hz, 2H), 3.47 (q, J = 6.6 Hz, 2H), 3.31 – 3.23 (m, 2H), 2.93 – 2.71 (m, 5H), 2.22 – 2.08 (m, 1H), 1.68 – 1.60 (m, 4H), 1.49 (s, 9H), 1.41 – 1.29 (m, 12H).

^{13}C NMR (101 MHz, Chloroform-d) δ 171.98, 170.24, 169.54, 167.67, 166.27, 163.93, 158.43, 154.55, 147.06, 138.78, 136.14, 133.62, 132.44, 128.32, 127.31, 116.74, 111.32, 109.70, 80.56, 48.90, 42.55, 40.19, 31.56, 29.48, 29.05, 28.99, 28.97, 28.93, 28.88, 28.41, 26.71, 26.55, 22.89.

HRMS (ESI+) m/z found 814.3594 [M+H]$^+$, C$_{42}$H$_{51}$N$_7$O$_8$S calculated 814.3598 (Δ = -0.49 ppm)

tert-butyl 2-(4-((2-(2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethoxy)ethoxy)ethyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF6)
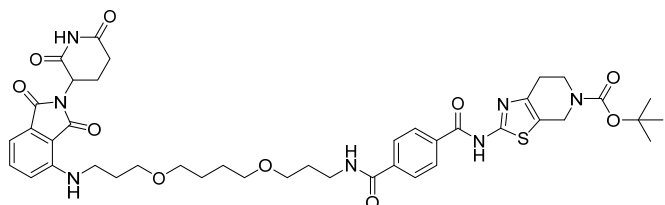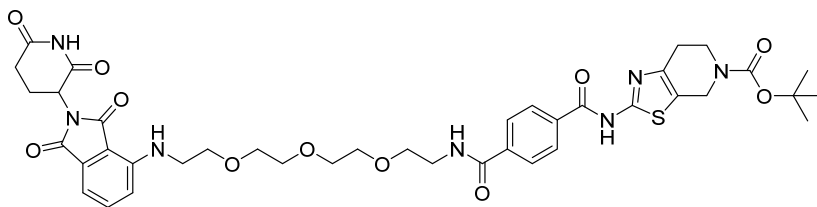

21 (40 mg, 0.1 mmol) and **17** (54 mg, 0.12 mmol) were combined according to general procedure IV. The product was a yellow solid (27 mg, 29%).

^{1}H NMR (400 MHz, Chloroform-d) δ 8.12 – 8.04 (m, 2H), 7.64 (d, J = 8.2 Hz, 2H), 7.55 (dd, J = 8.5, 7.1 Hz, 1H), 7.13 (d, J = 7.1 Hz, 1H), 7.03 (d, J = 8.5 Hz, 1H), 6.98 (t, J = 4.9 Hz, 1H), 6.37 (s, 1H), 5.03 – 4.95 (m, 1H), 4.62 (s, 2H), 3.80 – 3.57 (m, 17H), 3.51 – 3.43 (m, 2H), 3.03 – 2.75 (m, 5H), 2.24 – 2.15 (m, 1H), 1.52 (s, 9H).

^{13}C NMR (101 MHz, Chloroform-d) δ 171.84, 171.51, 169.64, 167.77, 166.81, 164.18, 158.65, 154.73, 147.04, 138.65, 136.40, 133.83, 132.47, 128.34, 127.49, 117.35, 112.01, 110.20, 80.64, 70.61, 70.59, 70.56, 70.04, 69.96, 49.13, 43.32, 42.81, 40.41, 31.75, 29.83, 28.55, 23.06

MS m/z (ES+) calcd for C$_{40}$H$_{48}$N$_7$O$_{11}$S [M+ H$^+$] 834.3 found 834.3.

tert-butyl 2-(4-((3-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)propyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGF7)
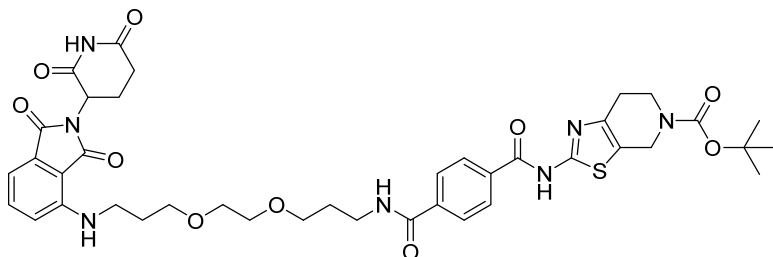
21 (30 mg, 0.07 mmol) and **18** (39 mg, 0.09 mmol) were combined according to general procedure IV. The product was a yellow solid (25 mg, 39%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 11.75 (s, 1H), 8.07 – 8.00 (m, 2H), 7.53 (dd, J = 8.5, 7.1 Hz, 1H), 7.43 (d, J = 8.2 Hz, 2H), 7.13 (d, J = 7.0 Hz, 2H), 6.91 (d, J = 8.6 Hz, 1H), 5.97 (s, 1H), 5.31 (s, 1H), 5.04 – 4.93 (m, 1H), 4.64 – 4.54 (m, 2H), 3.84 – 3.61 (m, 7H), 3.60 – 3.56 (m, 2H), 3.54 – 3.42 (m, 3H), 3.36 – 3.27 (m, 1H), 3.23 – 3.15 (m, 1H), 3.06 – 2.97 (m, 1H), 2.92 – 2.70 (m, 4H), 2.22 – 2.15 (m, 1H), 2.02 – 1.84 (m, 2H), 1.79 – 1.68 (m, 1H), 1.51 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 172.04, 171.71, 169.74, 167.69, 166.19, 164.04, 158.09, 154.67, 146.98, 138.50, 136.46, 133.83, 132.27, 127.97, 127.06, 116.77, 111.72, 109.73, 80.30, 70.75, 70.21, 69.80, 67.80, 49.07, 39.62, 39.26, 31.63, 29.36, 28.79, 28.43, 22.99.

MS *m/z* (ES+) calcd for C$_{40}$H$_{47}$N$_7$O$_{10}$S [M+ H$^+$] 818.9 found 819.1.

tert-Butyl ((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamate (22)
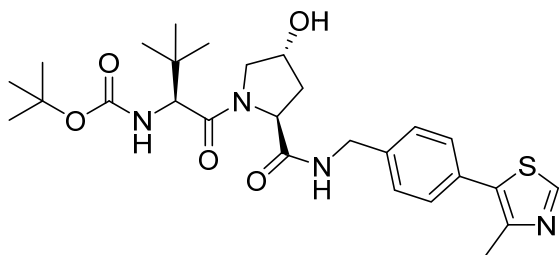
This compound was synthesized according to a procedure described in the literature and the NMR data matched those reported in the literature[201].

^{1}H NMR (400 MHz, CDCl$_3$): δ 8.73 (s, 1H), 7.45 (t, J = 5.9 Hz, 1H), 7.38 – 7.31 (m, 4H), 5.18 (d, J = 9.0 Hz, 1H), 4.76 (t, J = 7.9 Hz, 1H), 4.60 – 4.49 (m, 2H), 4.31 (dd, J = 15.0, 5.1 Hz, 1H), 4.15 (d, J = 9.1 Hz, 1H), 4.07 (d, J = 11.4 Hz, 1H), 3.58 (dd, J = 11.4, 3.6 Hz, 1H), 2.59 – 2.53 (m, 1H), 2.52 (s, 3H), 2.12 (dd, J = 13.6, 7.9 Hz, 1H), 1.40 (s, 9H), 0.91 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 172.85, 170.74, 156.58, 150.63, 138.44, 132.12, 130.74, 129.64, 128.31, 80.59, 70.26, 59.04, 58.37, 56.59, 43.37, 35.80, 34.82, 28.43, 26.47, 15.93.

tert-butyl (6-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-6-oxohexyl)carbamate (23)
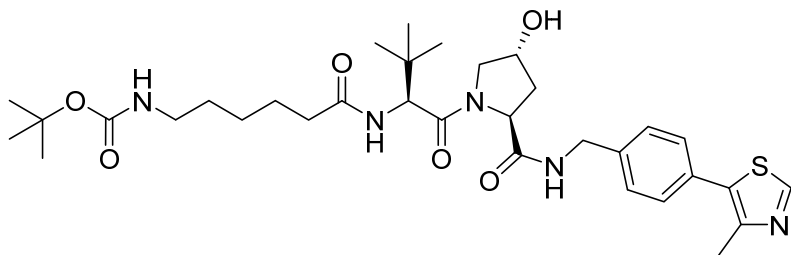
22 (500 mg, 0.94 mmol) and Boc-6-aminohexanoic acid (261 mg, 1.13 mmol) were combined according to general procedure IV. The product was a white solid (363 mg, 61 %).

The NMR data matched those reported in the literature[212].

¹H NMR (400 MHz, MeOD): δ 8.88 (s, 1H), 7.49 – 7.39 (m, 4H), 4.63 (s, 1H), 4.60 – 4.47 (m, 3H), 4.36 (d, J = 15.5 Hz, 1H), 3.91 (d, J = 10.9 Hz, 1H), 3.80 (dd, J = 11.0, 3.9 Hz, 1H), 3.02 (t, J = 7.0 Hz, 2H), 2.48 (s, 3H), 2.35 – 2.18 (m, 3H), 2.12 – 2.04 (m, 1H), 1.68 – 1.56 (m, 2H), 1.52 – 1.44 (m, 2H), 1.42 (s, 9H), 1.38 – 1.28 (m, 2H), 1.04 (s, 9H).

¹³C NMR (101 MHz, MeOD): δ 175.96, 174.49, 172.36, 152.85, 149.04, 140.30, 133.43, 131.52, 130.37, 128.99, 71.09, 60.82, 59.00, 58.01, 43.70, 41.22, 38.92, 36.54, 30.64, 28.79, 27.49, 27.04, 26.70, 15.80.

tert-butyl (2-(2-(2-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethoxy)ethoxy)ethyl)carbamate (24)

22 (300 mg, 0.56 mmol) and Boc-8-amino-3,6-dioxaoctanoic acid (179 mg, 0.68 mmol) were combined according to general procedure IV. The product was a white solid (121 mg, 32 %).

The NMR data matched those reported in the literature[213].

¹H NMR (400 MHz, MeOD): δ 8.98 (s, 1H), 8.57 (t, J = 6.0 Hz, 1H), 7.46 – 7.42 (m, 1H), 7.41 – 7.37 (m, 4H), 6.76 (t, J = 5.8 Hz, 1H), 5.15 (d, J = 3.6 Hz, 1H), 4.57 (d, J = 9.6 Hz, 1H), 4.47 – 4.32 (m, 3H), 4.26 (dd, J = 15.8, 5.7 Hz, 1H), 3.96 (s, 2H), 3.67 (dd, J = 10.7, 4.0 Hz, 1H), 3.63 – 3.57 (m, 3H), 3.57 – 3.50 (m, 2H), 3.44 – 3.36 (m, 2H), 3.08 (q, J = 6.0 Hz, 2H), 2.44 (s, 3H), 2.06 (dd, J = 13.1, 7.7 Hz, 1H), 1.94 – 1.87 (m, 1H), 1.35 (s, 9H), 0.94 (s, 9H).

¹³C NMR (101 MHz, MeOD): δ 172.18, 169.63, 169.07, 156.04, 151.93, 148.22, 139.88, 131.61, 130.19, 129.18, 127.95, 78.05, 70.86, 70.03, 69.75, 69.33, 59.19, 57.04, 56.15, 42.16, 38.38, 36.19, 28.68, 26.66, 16.40.

tert-butyl (10-(((S)-1-((2S,4R)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-10-oxodecyl)carbamate (25)

22 (200 mg, 0.37 mmol) and Boc-10-Aminodecanoic acid (130 mg, 0.45 mmol) were combined according to general procedure IV. The product was a white solid (124 mg, 48%).

The NMR data matched those reported in the literature[213].

¹H NMR (400 MHz, DMSO-d_6) δ 8.98 (s, 1H), 8.56 (t, J = 6.1 Hz, 1H), 7.84 (d, J = 9.3 Hz, 1H), 7.40 (q, J = 8.2 Hz, 4H), 6.74 (t, J = 5.8 Hz, 1H), 5.12 (d, J = 3.4 Hz, 1H), 4.54 (d, J = 9.3 Hz, 1H), 4.47 – 4.38 (m, 2H), 4.37 – 4.33 (m, 1H), 4.21 (dd, J = 15.9, 5.5 Hz, 1H), 3.71 – 3.60 (m, 2H), 2.87 (q, J = 6.6 Hz, 2H), 2.44 (s, 3H), 2.31 – 2.18 (m, 1H), 2.15 – 1.98 (m, 2H), 1.95 – 1.84 (m, 1H), 1.56 – 1.39 (m, 1H), 1.36 (s, 9H), 1.22 (s, 12H), 0.93 (s, 9H).

tert-butyl 2-(4-((6-((1-((4S)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-6-oxohexyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGV1)
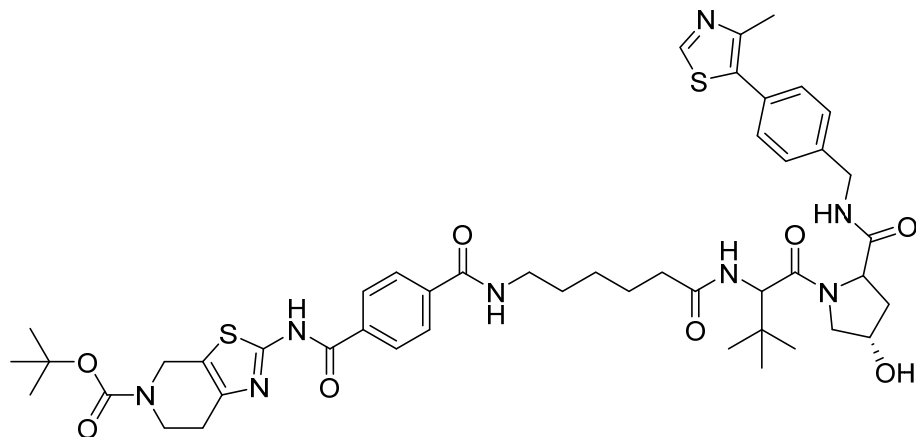

19 (100 mg, 0.21 mmol) and **23** (150 mg, 0.23 mmol) were combined according to general procedure IV. The product was a white solid (88 mg, 45%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.66 (s, 1H), 7.81 (d, J = 8.0 Hz, 2H), 7.73 – 7.67 (m, 2H), 7.21 (s, 4H), 6.93 (s, 1H), 4.72 (d, J = 9.3 Hz, 1H), 4.62 (d, J = 5.7 Hz, 1H), 4.58 – 4.46 (m, 3H), 4.33 (t, J = 8.2 Hz, 1H), 4.18 – 4.01 (m, 2H), 3.83 – 3.66 (m, 3H), 3.24 – 3.15 (m, 1H), 3.14 – 3.09 (m, 1H), 2.72 (d, J = 5.7 Hz, 2H), 2.44 (s, 3H), 2.34 – 2.17 (m, 2H), 2.15 – 2.01 (m, 2H), 1.47 – 1.38 (m, 4H), 1.33 (t, J = 7.3 Hz, 2H), 1.02 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.09, 171.97, 171.04, 166.25, 164.82, 157.42, 154.64, 150.35, 148.32, 137.96, 137.90, 134.69, 131.61, 130.71, 129.20, 128.09, 127.78, 127.43, 80.49, 70.15, 59.59, 58.08, 57.54, 45.72, 42.92, 39.85, 36.80, 35.84, 35.33, 28.47, 28.43, 26.73, 26.59, 16.06.

MS m/z (ES+) calcd for C$_{47}$H$_{60}$N$_8$O$_8$S$_2$ [M+ H$^+$] 929.2 found 929.4.

**tert-butyl 2-(4-((2-(2-(2-((1-((4S)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethoxy)ethoxy)ethyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGV2)**

19 (100 mg, 0.21 mmol) and **24** (135 mg, 0.23 mmol) were combined according to general procedure IV. The product was a white solid (81 mg, 40%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.69 (s, 1H), 8.04 (q, J = 6.2 Hz, 1H), 7.93 (d, J = 8.1 Hz, 2H), 7.83 (d, J = 8.1 Hz, 2H), 7.70 (d, J = 9.4 Hz, 1H), 7.33 (d, J = 7.9 Hz, 2H), 7.25 (d, J = 8.0 Hz, 2H), 6.78 (t, J = 5.8 Hz, 1H), 4.71 (d, J = 9.4 Hz, 1H), 4.61 (s, 1H), 4.56 (s, 2H), 4.53 − 4.47 (m, 1H), 4.27 (t, J = 8.3 Hz, 1H), 4.22 − 4.16 (m, 1H), 4.05 (d, J = 11.3 Hz, 1H), 3.94 (s, 2H), 3.87 − 3.77 (m, 1H), 3.74 − 3.58 (m, 10H), 3.40 − 3.29 (m, 1H), 2.68 (s, 2H), 2.50 (s, 3H), 2.37 − 2.28 (m, 1H), 2.14 − 2.05 (m, 1H), 1.49 (s, 9H), 0.99 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 171.04, 170.83, 170.45, 166.93, 164.26, 157.12, 154.63, 150.41, 148.44, 138.60, 137.83, 134.52, 131.51, 130.96, 129.40, 127.95, 127.91, 127.82, 80.43, 71.28, 70.04, 69.70, 69.53, 59.05, 57.28, 56.96, 53.44, 43.17, 39.64, 36.93, 36.03, 28.42, 26.47, 16.07.

MS m/z (ES+) calcd for C$_{47}$H$_{60}$N$_8$O$_{10}$S$_2$ [M+ H$^+$] 961.4 found 961.3.

tert-butyl 2-(4-((10-((1-((4S)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-10-oxodecyl)carbamoyl)benzamido)-6,7-dihydrothiazolo[5,4-c]pyridine-5(4H)-carboxylate (NGV3)
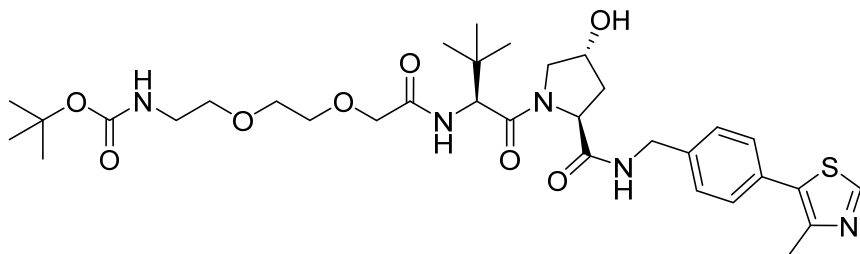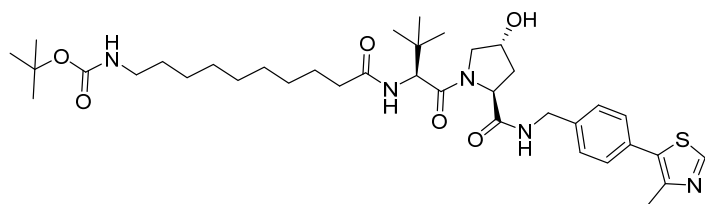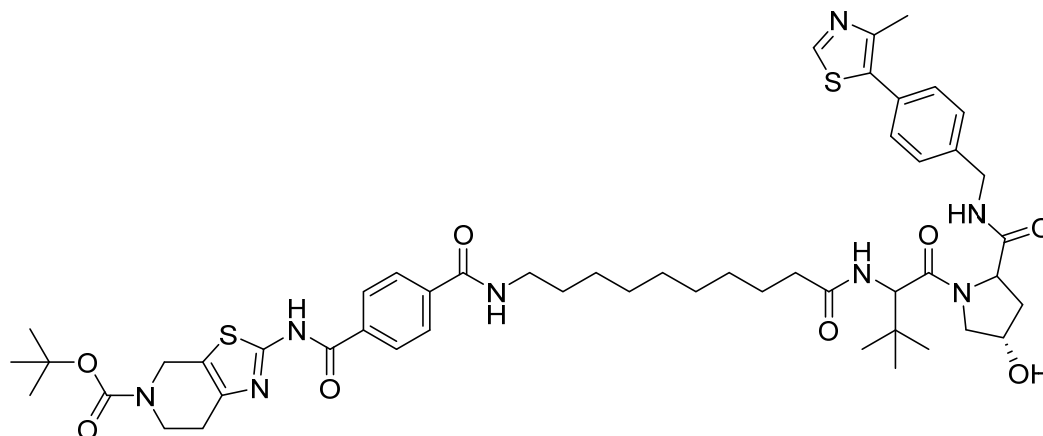

19 (40 mg, 0.083 mmol) and **25** (55 mg, 0.092 mmol) were combined according to general procedure IV. The product was a white solid (31 mg, 38%).

^{1}H NMR (400 MHz, CDCl$_3$) δ 8.73 (s, 1H), 7.99 (d, J = 8.0 Hz, 2H), 7.91 (d, J = 8.1 Hz, 2H), 7.35 – 7.24 (m, 5H), 7.16 – 7.04 (m, 2H), 7.00 (s, 1H), 4.67 – 4.53 (m, 6H), 4.51 (d, J = 6.7 Hz, 1H), 4.27 – 4.21 (m, 1H), 4.10 (d, J = 11.3 Hz, 1H), 3.82 – 3.74 (m, 2H), 3.70 – 3.64 (m, 1H), 3.40 (q, J = 6.6 Hz, 2H), 2.78 (t, J = 5.9 Hz, 2H), 2.50 (s, 3H), 2.43 – 2.33 (m, 1H), 2.19 – 2.11 (m, 3H), 1.57 – 1.50 (m, 3H), 1.32 – 1.24 (m, 4H), 1.21 (s, 6H), 0.97 (s, 9H).

^{13}C NMR (101 MHz, CDCl$_3$) δ 174.10, 172.12, 170.90, 166.56, 164.51, 154.50, 150.49, 148.14, 139.06, 138.05, 131.75, 130.74, 129.37, 128.44, 128.00, 127.57, 80.78, 69.99, 59.02, 57.73, 57.06, 53.43, 50.77, 43.10, 40.14, 36.42, 36.18, 35.27, 29.14, 28.81, 28.76, 28.73, 28.40, 26.61, 26.57, 25.82, 25.42, 15.95.

MS *m/z* (ES+) calcd for C$_{51}$H$_{68}$N$_8$O$_8$S$_2$ [M+ H$^+$] 985.3 found 985.5.

6.2 Biology

6.2.1 Proteomics

6.2.1.1 Key resources table

Reagent or resource	Source	Identifier
Biological reagents		
Dulbecco's Modified Eagle's Medium (DMEM)	Gibco	41966-029
Fetal bovine serum (FBS)	Gibco	11563397
Penicillin/streptomycin(P/S)	Gibco	11548876
Nonidet P40 Substitute (NP-40)	Sigma-Aldrich	11754599001
Triethylammonium bicarbonate buffer (TEAB)	Sigma-Aldrich	T7408
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)	Sigma-Aldrich	C4706
Iodoacetamide (IAA)	Sigma-Aldrich	I1149
Acetone	Sigma-Aldrich	650501
Trypsin	Sigma-Aldrich	T6567
TMT 11plex reagents kits	Thermo Scientific	A34808
Hydroxylamine for TMT experiments	Thermo Scientific	90115
Acetonitrile	Pierce	51101
Trifluoroacetic acid (TFA)	Honeywell Fluka	14264
Ammonium hydrogen carbonate	Supelco	533005
0.1% Formic Acid	Pierce	85174
Materials		
Sep-Pak® Plus C18 cartridges	Waters	WAT020515

6.2.1.2 Cell culture

HeLa cells were cultured in DMEM supplemented with 10% FBS and 1% P/S at 37 °C with 5% CO_2.

6.2.1.3 Cell lysis and sample preparation for MS analysis

2 mL cells (Cell density $2x10^5$ cell/ mL) were seeded in 6-well plates and let them grow overnight in the cell incubator at 37 °C and 5% CO_2. The following day, cells were incubated with the corresponding compound (or DMSO for the control sample) for 18 hours at 37 °C and 5% CO_2. At the end of the incubation, cells were washed with ice-cold PBS twice and lysed by the addition of lysis buffer consisting of 0.4% (v/v) NP-40 in PBS on ice. Lysis was finalized by four times of freeze-thaw cycles using liquid nitrogen. Cellular debris were removed by centrifugation at 25000 x g for 25 minutes at 4 °C. The protein concentration of the cell lysate was measured by DC assay (BioRad) and further diluted to a protein concentration of 2 mg/mL by the lysis buffer. For each sample, 150 µg of total proteins were further diluted in 100 mM TEAB for cysteines reduction with 200 mM TCEP for 1 h at 55 °C. 375 mM IAA was successively used for alkylation for 30 min at room temperature in the dark. After this step protein were precipitated overnight with 6 volumes of cold acetone at -20 °C. The following day the protein pellets were collected by centrifugation at 8000 × g at 4 °C for 10 min and resuspended in TEAB 100 mM before digestion. Protein digestion was performed by adding trypsin in a 1:80 enzyme: substrate ratio and left overnight at 37 °C, with gentle shaking in a ThermoMixer (Eppendorf®). Half of the sample volume corresponding to 75 µg of peptides were successively labelled with TMT 11plex reagents kits. Experimental replicates were labelled with different TMT-batches concatenated by a pooled internal reference sample for each TMT-batch, which was consisting of an additional sample obtained by pooling equal amounts of each sample

from the same TMT-batch. TMT reaction was allowed for 2 hours and quenched with 5% hydroxylamine. All the 11 samples related to each batch were pooled and dried in SpeedVac (Eppendorf EP022822993) before desalting with Sep-Pak® Plus C18 cartridges. Peptide were eluted with 40% and 60% of acetonitrile in 0.1% of trifluoroacetic acid (TFA) and dried before injection of 30 µg in the UHPLC system (Dionex U3000) for high-pH fractionation.

The separation of the peptides was carried out at a constant flowrate of 5 µl min^{-1} on a CSH C18 Acquity UPLC M-Class Peptide column, 130 Å, 1.7 µm, 300 µm x 150 mm (Waters, 186007563) using a 100 min linear gradient from 5 to 35% of mobile phase B (acetonitrile) with a subsequent 15 min gradient to 70%, before 5 min re-equilibration with 95% of mobile phase A (5mM ammonium bicarbonate, pH 10). 60 time-based fractions were pooled in 30 fractions in the collection plates. Clean-up of the fractions was performed by EvoTip according manufacturer's instructions.

6.2.1.4 LC-MS analysis

The EvoTips (EvoSep, EV2003) were loaded on the Evosep One module (EvoSep EV-1000) coupled to an Orbitrap Eclipse™ Tribid™ mass spectrometer (ThermoFisher Scientific). Peptides were loaded onto the EASY-Spray™ C18 column, 2 µm, 100 Å, 75 µm x 15 cm (ThermoFisher Scientific, ES804) using the standard "30 samples per day" Evosep method. The method eluted the peptides with a 44 min gradient ranging from 5% to 90% acetonitrile with 0.1% formic acid. The MS acquisition was performed in data dependent-MS3 with real-time-search (RTS) and a FAIMS interface switching between CVs of −50 V and −70 V with cycle times of 2 s and 1.5 s, respectively. The data dependent acquisition mode was run in a MS1 scan range between 375 and 1500 m/z with a resolution of 120000, and a normalized gain control (AGC) *target* of 100%, with a maximum injection time of 50 ms. RF Lens set at 30%. Filtering of the precursors was performed using peptide monoisotopic peak selection (MIPS), including charge states from 2 to 7, dynamic exclusion of 120 s with ±10 ppm tolerance excluding isotopes, and a precursor fit of 70% in a windows of 0.7 m/z with an intensity threshold of 5000. Selected precursors for further MS2 analysis were isolated with a window of 0.7 m/z in the quadrupole. The MS2 scan was performed over a range of 200-1400 m/z, collecting ions with a maximum injection time of 35 ms and normalized AGC target of 300% MS2 fragmentation was operated with normalized HCD collision energy at 30%. Fragmentation spectra were searched against the fasta files from the human Uniprot database (reviewed) in the RTS, set with tryptic digestion, TMT-11plex as fixed modification on Lysine (K) and N-Terminus together with cysteine (C) carbamidomethylation, and oxidation of methionine (M) as variable modification. 1 missed cleavage and 2 variable modifications were allowed with a maximum search time of 35 ms. FDR filtering was enabled with 1 as Xcorrelation and 5 ppm of precursor tolerance. Precursors identified via RTS in the MS2 scan were further isolated in the quadrupole with a 2 m/z window, maximum injection time of 86 ms and normalized AGC target of 300%. The further MS3 fragmentation was operated with a normalized HCD collision energy at 50% and fragments were scanned with a resolution of 50000 in the range of 100 to 500 m/z. The MS performances were monitored by quality control of an in-house standard of HeLa-cell lysate, both at the beginning and the end of each sample set.

6.2.1.5 Data analysis

Mass spectrometric raw files were analyzed with MaxQuant (version 1.6.2.10, https://maxquant.org/) setting all the TMT11 labels as reporter ion MS3 for lysine and N-terminus with mass tolerance 0.003 Da. The search settings included carbamidomethylation of cysteine (C) as fixed modifications, methionine (M) oxidation and acetylation of protein N-termini as variable modification and trypsin as enzyme (allowing maximum 2 missed cleavages). Peptide-spectrum match (PSM) and protein false discovery rates (FDR) were 0.01 with at least 1 minimum razor + unique peptides. Peptide sequences were searched against UniProtKB/ TrEMBL human proteome6 fasta file (downloaded on 08/08/2018,

). The protein group file generated from the MaxQuant analysis were processed for further normalization and scaling.

Contaminants and protein identified in the reverse database were removed, as well as all the proteins identified with a sum of Unique + Razor peptides below 2. Loading normalization was performed on the first 10th TMT channels, summing the intensities corrected for isotopic error for each TMT channel and calculating the respective correction factor on the average of the summed intensities. Hence, each protein intensity was normalized for the respective channel correction factor. For comparison of the experimental replicates measured in the multi-batch TMT experiment, the internal reference scaling (IRS) normalization method was applied following the procedure described in Plubell et al[214]. Briefly, the geometric mean of the channels measuring the pooled internal reference sample in the 11th channel was calculated across the replicates and an inter-batch scaling factor was calculated and applied to each protein.

Normalized intensities were further considered to calculate the average among the replicates and the fold changes (FC) for PROTACs-treated samples over the control samples. The obtained values were log2 transformed. A two-sided T-test was performed on the normalized data and obtained p-values were -log10 transformed.

Volcano plots were obtained plotting the log2 FC vs -log10 (p-value) in both Prism 5 (Graphpad Software, Inc. Version 5.3) and VolcanoseR[215].

6.2.2 Immunoblotting.

6.2.2.1 Key resources table

Antibodies

Products	Source	Identifier	Dilution
GOLIM4 Polyclonal Antibody	Invitrogen	PA5-51624	1: 1000
Anti-QSOX2 antibody	Abcam	ab121376	1: 1000
COG1 Rabbit pAb (A17594)	Abclonal	A17594	1: 1000
COG2 Rabbit pAb (A6251)	Abclonal	A6251	1: 1000
COG3 Rabbit pAb (A17785)	Abclonal	A17785	1: 1000
COG4 Rabbit pAb (A16111)	Abclonal	A16111	1: 1000
COG6 Rabbit pAb (A10319)	Abclonal	A10319	1: 1000
COG8 Rabbit pAb (A14361)	Abclonal	A14361	1: 1000
GRAMD1A Antibody	Novus Biologicals	NBP2-32148	1: 1000
Beta Actin Monoclonal Antibody	Invitrogen	AM4302	1: 5000
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	Invitrogen	31460	1: 10 000
F(ab')2-Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	Invitrogen	A24512	1: 10 000
IRDye® 800CW Goat anti-Mouse IgG (H + L)	LI-COR	925-32210	1: 15 000
IRDye® 680RD Goat anti-Rabbit IgG (H + L)	LI-COR	925-68071	1: 15 000

6.2.2.2 Cell culture

A549, HeLa and U2OS cells were cultured in DMEM supplemented with 10% FBS, 1% penicillin–streptomycin at 37 °C with 5% CO_2.

6.2.2.3 Immunoblotting analysis

For degradation tests, 2×10^5 cells/mL A549, HeLa or U2OS cells in 2 mL medium were seeded in six-well plates and incubated (37 °C and 5% CO_2) overnight. The media was then removed and replaced with the fresh media containing the required concentration of compounds dissolved in DMSO. At the specific time point, cells were washed with ice-cold PBS twice and then lysed in lysis buffer (100 mM Tris–HCl (pH 6.8), 4% (w/v) sodium dodecyl sulfate (SDS), 20% (v/v) glycerol) and sonicated by ultrasonic processor (UP100H, Hielscher). The protein concentrations of the lysate were determined using NanoDrop ND-1000 spectrophotometer (Saveen Werner AB, Limhamn, Sweden) according to the manufacturer's instructions and further diluted to 4 mg/mL. SDS–PAGE was carried out using 4-15% or 4-20% precast polyacrylamide gels (Bio-Rad, 4561086 or 4561093) and run at a constant voltage of 80 V for 5 min followed by 120 V for 1 h. Semi-dry transfer onto a PVDF membrane was performed using Bio-Rad Trans-Blot Turbo Transfer System at 1.3 A for 7 min. For Chemiluminescent detection, membranes were blocked in 5% milk in TBST (137 mM NaCl, 19 mM Tris-base, 2.7 mM KCl and 0.1% (v/v) Tween-20, blocking buffer) for 1 h at room temperature. The membrane was incubated with the primary antibody in blocking buffer overnight at 4 °C. After washing with TBST (3×5 min), the membrane was incubated with the secondary antibody in blocking buffer for 1 h at room temperature. Signals were visualized using the SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher, 34579) or the SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher, 34094) on a Li-COR Odyssey Fc. For fluorescent detection, membranes were blocked in Intercept® (TBS) Blocking Buffer (Li-COR, 927-60001) for 1 h at room temperature. The membrane was incubated with the primary antibody in blocking buffer overnight at 4 °C. After washing with TBST (3 × 5 min), the membrane was incubated with the secondary antibody in blocking buffer for 1 h at room temperature. Signals were visualized on a Li-COR Odyssey Fc.

6.2.3 In-Cell Cellular Thermal Shift Assay (CETSA).

Two T75 cell culture flasks were seeded with each 6×10^5 HeLa cells in media and incubated at 37°C and 5% CO_2 for three days. The media was then removed and replaced with 3 mL medium containing either 10 µM of the given compound or DMSO. After incubation for 1 h at 37°C and 5% CO_2 the medium was removed and the cells were washed by PBS and detached using 1.5 mL trypsin. And then the media and trypsin were removed by centrifuge at 250 x g for 3 min. After three times washing by PBS, per 1×10^6 cells were collected in 0.6 mL PBS each. Treated and non-treated cell suspensions were divided into ten aliquots, each 100 µL in PCR tubes. The aliquots were individually heated at different temperatures (Doppio 2×48 well Thermal cycler, VWR). After the heat treatment, 10 µL PBS containing 4.4% (v/v) NP-40 added to each sample and cells were lysed by freeze and thaw four times. The cell lysates were completely transferred to polycarbonate tubes and centrifuged (Micro Star 30, VWR) at 25,000 x g, 4°C for 25 min. The cleared supernatants were collected and further analyzed by Western Blot.

6.2.4 In-Cell Isothermal dose-response fingerprinting (ITDRF) experiments.

Two six well plates were seeded with each $3x10^5$ HeLa cells in 2 mL DMEM per well and incubated at 37°C and 5% CO_2 overnight. The media was subsequently removed and replaced with 0.5 mL medium containing either different concentrations (40, 20, 10, 5, 2.5, 1.25, 0.625, 0.3 µM) of the test compound or DMSO. After incubation for 1 h at 37°C and 5% CO_2, the medium was removed and cells were detached using 200 mL trypsin and collected in 0.5 mL MEM each. Treated and non-treated cell suspensions were centrifuged (350 x g, 5 min), and resuspended in 50 µL PBS in PCR tubes. The samples were heated at the given temperature (Eppendorf Mastercycler ep Gradient S). After the heat treatment, 5 µL PBS containing 4.4% (v/v) NP-40 were added to each sample and cells lysed by freeze and thaw four times and then the cell lysates were completely transferred to polycarbonate tubes and centrifuged (Micro Star 30, VMR) at 25,000 x g, 4 °C for 25 min. The cleared supernatants were collected and further analyzed by Western Blot.

6.2.5   Cell lysate Cellular Thermal Shift Assay (CETSA).

HeLa cell lysates (protein concentration 2 mg/mL) were separated in the two aliquots (1.0 mL) to be incubated for 10 min at room temperature with either 0.1% DMSO (1.0 µL) or with 10 µM of the tested compound. After incubation, each aliquot (each 100 µL) was heated at different temperatures for 3 min using PCR gradient-cyclers.

After heating, the removal of the precipitated proteins from the solution was performed by ultracentrifugation at 4 °C with a speed of 25000 x g for 25 min. The cleared supernatants were collected and further analyzed by Western Blot.

6.2.6   Isothermal dose-response fingerprinting (ITDRF) experiments

HeLa lysates (protein concentration 2 mg/mL) were separated in 9 aliquots (each 100 µL) to be incubated for 10 min at room temperature with either 0.1% DMSO or with decreasing concentration of the tested compounds (40, 20, 10, 5, 2.5, 1.25, 0.625, 0.3 µM). All the aliquots were successively heated for 3 min at the given temperature using PCR gradient-cyclers.

After heating, the removal of the precipitated proteins from the solution was performed by ultracentrifugation at 4 °C with a speed of 25000 x g for 25 min. The cleared supernatants were collected and further analyzed by Western Blot.

6.2.7   Deglycosylation assay

HeLa cell lysates were treated with PNGase F (New England Biolabs, NEB-P0704S) according to the manufacturer's instructions. Briefly, 10 µL of reaction mixture (9 µL of 1 mg/mL HeLa cell lysate and 1 µL of Glycoprotein Denaturing Buffer (10X)) was denatured by heating at 100 °C for 10 min. The denatured protein mixture was chilled on ice and centrifuged for 10 s. Then, 2 µL of GlycoBuffer 2 (10X), 2 µL of 10% NP - 40, 6 µL of H_2O were added to the previous reaction mixture to make a total reaction volume of 20 µL. Next, 1 µL PNGase F was added to the reaction mixture and the mixture was incubated at 37 °C for 2 h and the results were obtained by western blotting assay as previously described.

6.2.8   Pulldown assay

6.2.8.1   Bead preparation

The pulldown assay was performed using MyOne™ streptavidin-coated T1 magnetic beads (Dynabeads™ Streptavidin Trial Kit, Therofisher, 65801D). The beads were aliquoted (100 µL, binds to 1 nmol free biotin) and washed three times by 0.4% (v/v) NP-40 in PBS (300 µL). Then, 100 nM **C3_Bio** in 100 µL 0.4% (v/v) NP-40 in PBS was incubated with the beads for 1 h at room temperature with gentle shaking. After incubation, the solution was removed and the beads was washed by 0.4% (v/v) NP-40 in PBS (300 µL) three times.

6.2.8.2   Pulldown

The beads were incubated with 750 µL clarified HeLa cell lysate (1 mg/mL) with gentle shaking at 4 °C for 1 h. After incubation, the lysate was removed and the beads were washed three times by 0.4% (v/v) NP-40 in PBS (500 µL). After washing, the beads were dissolved in 32 µL 0.4% (v/v) NP-40 in PBS and 8 µL SDS loading buffer. Then, the beads were boiled at 100 °C for 10 min and the enriched proteins were further analyzed by Western Blot.

6.2.8.3   Competitive pulldown

5 µM **C3_Boc** was preincubated with HeLa cell lysate for 30 min and then the mixture was incubated with the beads with gentle shaking at 4 °C for 1 h. After washing, the beads were dissolved in 16 µL 0.4% (v/v)

NP-40 in PBS and 4 μL SDS loading buffer. Then, the beads were boiled at 100 °C for 10 min and the enriched proteins were further analyzed by Western Blot.

7 References

1. Boisvert F-M, Ahmad Y, Gierliński M, et al. A Quantitative Spatial Proteomics Analysis of Proteome Turnover in Human Cells*. *Mol Cell Proteomics*. 2012;11(3):M111.011429. doi:https://doi.org/10.1074/mcp.M111.011429

2. Ciechanover A, Schwartz AL. The ubiquitin-proteasome pathway: The complexity and myriad functions of proteins death. *Proc Natl Acad Sci*. 1998;95(6):2727-2730. doi:10.1073/pnas.95.6.2727

3. Komander D, Rape M. The Ubiquitin Code. *Annu Rev Biochem*. 2012;81(1):203-229. doi:10.1146/annurev-biochem-060310-170328

4. Morreale FE, Walden H. Types of Ubiquitin Ligases. *Cell*. 2016;165(1):248-248.e1. doi:https://doi.org/10.1016/j.cell.2016.03.003

5. Wang F, Gómez-Sintes R, Boya P. Lysosomal membrane permeabilization and cell death. *Traffic*. 2018;19(12):918-931. doi:https://doi.org/10.1111/tra.12613

6. Brix K. Lysosomal Proteases BT - Lysosomes. In: Saftig P, ed. Springer US; 2005:50-59. doi:10.1007/0-387-28957-7_5

7. Glick D, Barth S, Macleod KF. Autophagy: cellular and molecular mechanisms. *J Pathol*. 2010;221(1):3-12. doi:https://doi.org/10.1002/path.2697

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

9. Neklesa TK, Tae HS, Schneekloth AR, et al. Small-molecule hydrophobic tagging–induced degradation of HaloTag fusion proteins. *Nat Chem Biol*. 2011;7(8):538-543. doi:10.1038/nchembio.597

10. Banik SM, Pedram K, Wisnovsky S, Ahn G, Riley NM, Bertozzi CR. Lysosome-targeting chimaeras for degradation of extracellular proteins. *Nature*. 2020;584(7820):291-297. doi:10.1038/s41586-020-2545-9

11. Takahashi D, Moriyama J, Nakamura T, et al. AUTACs: Cargo-Specific Degraders Using Selective Autophagy. *Mol Cell*. 2019;76(5):797-810.e10. doi:https://doi.org/10.1016/j.molcel.2019.09.009

12. Békés M, Langley DR, Crews CM. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov*. 2022;21(3):181-200. doi:10.1038/s41573-021-00371-6

13. Hu B, Zhong L, Weng Y, et al. Therapeutic siRNA: state of the art. *Signal Transduct Target Ther*. 2020;5(1):101. doi:10.1038/s41392-020-0207-x

14. Vassilev LT, Vu BT, Graves B, et al. In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science (80-)*. 2004;303(5659):844-848. doi:10.1126/science.1092472

15. Schneekloth AR, Pucheault M, Tae HS, Crews CM. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorg Med Chem Lett*. 2008;18(22):5904-5908. doi:https://doi.org/10.1016/j.bmcl.2008.07.114

16. Suda H, Takita T, Aoyagi T, Umezawa H. The structure of bestatin. *J Antibiot (Tokyo)*. 1976;29(1):100-101. doi:10.7164/antibiotics.29.100

17. Itoh Y, Kitaguchi R, Ishikawa M, Naito M, Hashimoto Y. Design, synthesis and biological

evaluation of nuclear receptor-degradation inducers. *Bioorg Med Chem*. 2011;19(22):6768-6778. doi:https://doi.org/10.1016/j.bmc.2011.09.041

18. Itoh Y, Ishikawa M, Naito M, Hashimoto Y. Protein Knockdown Using Methyl Bestatin−Ligand Hybrid Molecules: Design and Synthesis of Inducers of Ubiquitination-Mediated Degradation of Cellular Retinoic Acid-Binding Proteins. *J Am Chem Soc*. 2010;132(16):5820-5826. doi:10.1021/ja100691p

19. Ohoka N, Nagai K, Hattori T, et al. Cancer cell death induced by novel small molecules degrading the TACC3 protein via the ubiquitin–proteasome pathway. *Cell Death Dis*. 2014;5(11):e1513-e1513. doi:10.1038/cddis.2014.471

20. Ito T, Ando H, Suzuki T, et al. Identification of a Primary Target of Thalidomide Teratogenicity. *Science (80-)*. 2010;327(5971):1345-1350. doi:10.1126/science.1177319

21. Lu J, Qian Y, Altieri M, et al. Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem Biol*. 2015;22(6):755-763. doi:https://doi.org/10.1016/j.chembiol.2015.05.009

22. Buckley DL, Van Molle I, Gareiss PC, et al. Targeting the von Hippel–Lindau E3 Ubiquitin Ligase Using Small Molecules To Disrupt the VHL/HIF-1α Interaction. *J Am Chem Soc*. 2012;134(10):4465-4468. doi:10.1021/ja209924v

23. Bondeson DP, Mares A, Smith IED, et al. Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol*. 2015;11(8):611-617. doi:10.1038/nchembio.1858

24. Ward CC, Kleinman JI, Brittain SM, et al. Covalent Ligand Screening Uncovers a RNF4 E3 Ligase Recruiter for Targeted Protein Degradation Applications. *ACS Chem Biol*. 2019;14(11):2430-2440. doi:10.1021/acschembio.8b01083

25. Luo M, Spradlin JN, Boike L, et al. Chemoproteomics-enabled discovery of covalent RNF114-based degraders that mimic natural product function. *Cell Chem Biol*. 2021;28(4):559-566.e15. doi:https://doi.org/10.1016/j.chembiol.2021.01.005

26. Zhang X, Luukkonen LM, Eissler CL, et al. DCAF11 Supports Targeted Protein Degradation by Electrophilic Proteolysis-Targeting Chimeras. *J Am Chem Soc*. 2021;143(13):5141-5149. doi:10.1021/jacs.1c00990

27. Li L, Mi D, Pei H, et al. In vivo target protein degradation induced by PROTACs based on E3 ligase DCAF15. *Signal Transduct Target Ther*. 2020;5(1):129. doi:10.1038/s41392-020-00245-0

28. Zhang X, Crowley VM, Wucherpfennig TG, Dix MM, Cravatt BF. Electrophilic PROTACs that degrade nuclear proteins by engaging DCAF16. *Nat Chem Biol*. 2019;15(7):737-746. doi:10.1038/s41589-019-0279-5

29. Wei J, Meng F, Park K-S, et al. Harnessing the E3 Ligase KEAP1 for Targeted Protein Degradation. *J Am Chem Soc*. 2021;143(37):15073-15083. doi:10.1021/jacs.1c04841

30. Qiu X, Li Y, Yu B, et al. Discovery of selective CDK9 degraders with enhancing antiproliferative activity through PROTAC conversion. *Eur J Med Chem*. 2021;211:113091. doi:https://doi.org/10.1016/j.ejmech.2020.113091

31. Lai AC, Toure M, Hellerschmied D, et al. Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angew Chemie Int Ed*. 2016;55(2):807-810. doi:https://doi.org/10.1002/anie.201507634

32. Yu X, Xu J, Shen Y, et al. Discovery of Potent, Selective, and In Vivo Efficacious AKT Kinase

Protein Degraders via Structure–Activity Relationship Studies. *J Med Chem*. 2022;65(4):3644-3666. doi:10.1021/acs.jmedchem.1c02165

33. Zhao L, Han X, Lu J, McEachern D, Wang S. A highly potent PROTAC androgen receptor (AR) degrader ARD-61 effectively inhibits AR-positive breast cancer cell growth in vitro and tumor growth in vivo. *Neoplasia*. 2020;22(10):522-532. doi:https://doi.org/10.1016/j.neo.2020.07.002

34. Schiedel M, Herp D, Hammelmann S, et al. Chemically Induced Degradation of Sirtuin 2 (Sirt2) by a Proteolysis Targeting Chimera (PROTAC) Based on Sirtuin Rearranging Ligands (SirReals). *J Med Chem*. 2018;61(2):482-491. doi:10.1021/acs.jmedchem.6b01872

35. Adhikari B, Bozilovic J, Diebold M, et al. PROTAC-mediated degradation reveals a non-catalytic function of AURORA-A kinase. *Nat Chem Biol*. 2020;16(11):1179-1188. doi:10.1038/s41589-020-00652-y

36. Cromm PM, Samarasinghe KTG, Hines J, Crews CM. Addressing Kinase-Independent Functions of Fak via PROTAC-Mediated Degradation. *J Am Chem Soc*. 2018;140(49):17019-17026. doi:10.1021/jacs.8b08008

37. Bai L, Zhou H, Xu R, et al. A Potent and Selective Small-Molecule Degrader of STAT3 Achieves Complete Tumor Regression In Vivo. *Cancer Cell*. 2019;36(5):498-511.e17. doi:https://doi.org/10.1016/j.ccell.2019.10.002

38. Bondeson DP, Smith BE, Burslem GM, et al. Lessons in PROTAC Design from Selective Degradation with a Promiscuous Warhead. *Cell Chem Biol*. 2018;25(1):78-87.e5. doi:https://doi.org/10.1016/j.chembiol.2017.09.010

39. Gadd MS, Testa A, Lucas X, et al. Structural basis of PROTAC cooperative recognition for selective protein degradation. *Nat Chem Biol*. 2017;13(5):514-521. doi:10.1038/nchembio.2329

40. Farber PJ, Mittermaier A. Side chain burial and hydrophobic core packing in protein folding transition states. *Protein Sci*. 2008;17(4):644-651. doi:https://doi.org/10.1110/ps.073105408

41. Neklesa TK, Tae HS, Schneekloth AR, et al. Small-molecule hydrophobic tagging-induced degradation of HaloTag fusion proteins. *Nat Chem Biol*. 2011;7(8):538-543. doi:10.1038/nchembio.597

42. Long MJC, Gollapalli DR, Hedstrom L. Inhibitor Mediated Protein Degradation. *Chem Biol*. 2012;19(5):629-637. doi:https://doi.org/10.1016/j.chembiol.2012.04.008

43. Xie T, Lim SM, Westover KD, et al. Pharmacological targeting of the pseudokinase Her3. *Nat Chem Biol*. 2014;10(12):1006-1012. doi:10.1038/nchembio.1658

44. Teutsch G, Goubet F, Battmann T, et al. Non-steroidal antiandrogens: Synthesis and biological profile of high-affinity ligands for the androgen receptor. *J Steroid Biochem Mol Biol*. 1994;48(1):111-119. doi:10.1016/0960-0760(94)90257-7

45. Gao N, Chu T-T, Li Q-Q, et al. Hydrophobic tagging-mediated degradation of Alzheimer's disease related Tau. *RSC Adv*. 2017;7(64):40362-40366. doi:10.1039/C7RA05347A

46. Banik SM, Pedram K, Wisnovsky S, Ahn G, Riley NM, Bertozzi CR. Lysosome-targeting chimaeras for degradation of extracellular proteins. *Nature*. 2020;584(7820):291-297. doi:10.1038/s41586-020-2545-9

47. Ahn G, Banik SM, Miller CL, Riley NM, Cochran JR, Bertozzi CR. LYTACs that engage the asialoglycoprotein receptor for targeted protein degradation. *Nat Chem Biol*. 2021;17(9):937-946.

doi:10.1038/s41589-021-00770-1

48. Björkhem I, Diczfalusy U. Oxysterols. *Arterioscler Thromb Vasc Biol*. 2002;22(5):734-742. doi:10.1161/01.ATV.0000013312.32196.49

49. Kloudova A, Guengerich FP, Soucek P. The Role of Oxysterols in Human Cancer. *Trends Endocrinol Metab*. 2017;28(7):485-496. doi:https://doi.org/10.1016/j.tem.2017.03.002

50. Hirayama T, Ikegami T, Honda A, et al. Differences in the Serum 4β-hydroxycholesterol Levels of Patients with Chronic Hepatitis C Virus (HCV) Infection: A Possible Impact on the Efficacy and Safety of Interferon (IFN)-free Treatment. *Intern Med*. 2018;57(9):1219-1227. doi:10.2169/internalmedicine.9479-17

51. Kulig W, Cwiklik L, Jurkiewicz P, Rog T, Vattulainen I. Cholesterol oxidation products and their biological importance. *Chem Phys Lipids*. 2016;199:144-160. doi:10.1016/j.chemphyslip.2016.03.001

52. Janowski BA, Willy PJ, Devi TR, Falck JR, Mangelsdorf DJ. An oxysterol signalling pathway mediated by the nuclear receptor LXRα. *Nature*. 1996;383(6602):728-731. doi:10.1038/383728a0

53. Lehmann JM, Kliewer SA, Moore LB, et al. Activation of the Nuclear Receptor LXR by Oxysterols Defines a New Hormone Response Pathway*. *J Biol Chem*. 1997;272(6):3137-3140. doi:https://doi.org/10.1074/jbc.272.6.3137

54. Moldavski O, Zushin P-JH, Berdan CA, et al. 4β-Hydroxycholesterol is a prolipogenic factor that promotes SREBP1c expression and activity through the liver X receptor. *J Lipid Res*. 2021;62:100051. doi:https://doi.org/10.1016/j.jlr.2021.100051

55. Hong C, Tontonoz P. Liver X receptors in lipid metabolism: opportunities for drug discovery. *Nat Rev Drug Discov*. 2014;13(6):433-444. doi:10.1038/nrd4280

56. Lee SD, Tontonoz P. Liver X receptors at the intersection of lipid metabolism and atherogenesis. *Atherosclerosis*. 2015;242(1):29-36. doi:https://doi.org/10.1016/j.atherosclerosis.2015.06.042

57. Wang Y, Kumar N, Crumbley C, Griffin PR, Burris TP. A second class of nuclear receptors for oxysterols: Regulation of RORα and RORγ activity by 24S-hydroxycholesterol (cerebrosterol). *Biochim Biophys Acta - Mol Cell Biol Lipids*. 2010;1801(8):917-923. doi:https://doi.org/10.1016/j.bbalip.2010.02.012

58. Wang Y, Kumar N, Solt LA, et al. Modulation of Retinoic Acid Receptor-related Orphan Receptor α and γ Activity by 7-Oxygenated Sterol Ligands2. *J Biol Chem*. 2010;285(7):5013-5025. doi:https://doi.org/10.1074/jbc.M109.080614

59. Lappano R, Recchia AG, De Francesco EM, et al. The Cholesterol Metabolite 25-Hydroxycholesterol Activates Estrogen Receptor α-Mediated Signaling in Cancer Cells and in Cardiomyocytes. *PLoS One*. 2011;6(1):e16631. https://doi.org/10.1371/journal.pone.0016631

60. Nelson ER, Wardell SE, McDonnell DP. The molecular mechanisms underlying the pharmacological actions of estrogens, SERMs and oxysterols: Implications for the treatment and prevention of osteoporosis. *Bone*. 2013;53(1):42-50. doi:https://doi.org/10.1016/j.bone.2012.11.011

61. Radhakrishnan A, Ikeda Y, Kwon HJ, Brown MS, Goldstein JL. Sterol-regulated transport of SREBPs from endoplasmic reticulum to Golgi: Oxysterols block transport by binding to Insig. *Proc Natl Acad Sci*. 2007;104(16):6511-6518. doi:10.1073/pnas.0700899104

62.   Adams CM, Reitz J, De Brabander JK, et al. Cholesterol and 25-Hydroxycholesterol Inhibit Activation of SREBPs by Different Mechanisms, Both Involving SCAP and Insigs*. *J Biol Chem*. 2004;279(50):52772-52780. doi:https://doi.org/10.1074/jbc.M410302200

63.   Nedelcu D, Liu J, Xu Y, Jao C, Salic A. Oxysterol binding to the extracellular domain of Smoothened in Hedgehog signaling. *Nat Chem Biol*. 2013;9(9):557-564. doi:10.1038/nchembio.1290

64.   Lum L, Beachy PA. The Hedgehog Response Network: Sensors, Switches, and Routers. *Science (80-)*. 2004;304(5678):1755-1759. doi:10.1126/science.1098020

65.   Raccosta L, Fontana R, Maggioni D, et al. The oxysterol–CXCR2 axis plays a key role in the recruitment of tumor-promoting neutrophils. *J Exp Med*. 2013;210(9):1711-1728. doi:10.1084/jem.20130440

66.   Sensi C, Daniele S, Parravicini C, et al. Oxysterols act as promiscuous ligands of class-A GPCRs: In silico molecular modeling and in vitro validation. *Cell Signal*. 2014;26(12):2614-2620. doi:https://doi.org/10.1016/j.cellsig.2014.08.003

67.   Sun S, Liu C. 7a, 25-dihydroxycholesterol-mediated activation of EBI2 in immune regulation and diseases   . *Front Pharmacol* . 2015;6. https://www.frontiersin.org/articles/10.3389/fphar.2015.00060

68.   Soffientini U, Graham A. Intracellular cholesterol transport proteins: roles in health and disease. *Clin Sci*. 2016;130(21):1843-1859. doi:10.1042/CS20160339

69.   Jaworski CJ, Moreira E, Li A, Lee R, Rodriguez IR. A Family of 12 Human Genes Containing Oxysterol-Binding Domains. *Genomics*. 2001;78(3):185-196. doi:https://doi.org/10.1006/geno.2001.6663

70.   Lehto M, Laitinen S, Chinetti G, et al. The OSBP-related protein family in humans. *J Lipid Res*. 2001;42(8):1203-1213. doi:https://doi.org/10.1016/S0022-2275(20)31570-4

71.   Levine TP, Munro S. The pleckstrin homology domain of oxysterol-binding protein recognises a determinant specific to Golgi membranes. *Curr Biol*. 1998;8(13):729-739. doi:https://doi.org/10.1016/S0960-9822(98)70296-9

72.   Loewen CJR, Roy A, Levine TP. A conserved ER targeting motif in three families of lipid binding proteins and in Opi1p binds VAP. *EMBO J*. 2003;22(9):2025-2035. doi:https://doi.org/10.1093/emboj/cdg201

73.   Du X, Kumar J, Ferguson C, et al. A role for oxysterol-binding protein–related protein 5 in endosomal cholesterol trafficking. *J Cell Biol*. 2011;192(1):121-135. doi:10.1083/jcb.201004142

74.   Yan D, Mäyränpää MI, Wong J, et al. OSBP-related Protein 8 (ORP8) Suppresses ABCA1 Expression and Cholesterol Efflux from Macrophages*. *J Biol Chem*. 2008;283(1):332-340. doi:https://doi.org/10.1074/jbc.M705313200

75.   Depta L, Whitmarsh-Everiss T, Laraia L. Structure, function and small molecule modulation of intracellular sterol transport proteins. *Bioorg Med Chem*. 2022;68:116856. doi:https://doi.org/10.1016/j.bmc.2022.116856

76.   Aw YC, Brown AJ, Wu J-W, Yang H. ORP1L, ORP1S, and ORP2: Lipid Sensors and Transporters. *Contact*. 2020;3:2515256420956818. doi:10.1177/2515256420956818

77.   Antonny B, Bigay J, Mesmin B. The Oxysterol-Binding Protein Cycle: Burning Off PI(4)P to

Transport Cholesterol. *Annu Rev Biochem*. 2018;87(1):809-837. doi:10.1146/annurev-biochem-061516-044924

78. Delfosse V, Bourguet W, Drin G. Structural and Functional Specialization of OSBP-Related Proteins. *Contact*. 2020;3:2515256420946627. doi:10.1177/2515256420946627

79. Zheng Koh DH, Saheki Y. Regulation of Plasma Membrane Sterol Homeostasis by Nonvesicular Lipid Transport. *Contact*. 2021;4:25152564211042452. doi:10.1177/25152564211042451

80. Lim C-Y, Davis OB, Shin HR, et al. ER–lysosome contacts enable cholesterol sensing by mTORC1 and drive aberrant growth signalling in Niemann–Pick type C. *Nat Cell Biol*. 2019;21(10):1206-1218. doi:10.1038/s41556-019-0391-5

81. Pietrangelo A, Ridgway ND. Golgi localization of oxysterol binding protein-related protein 4L (ORP4L) is regulated by ligand binding. *J Cell Sci*. 2018;131(14):jcs215335. doi:10.1242/jcs.215335

82. Zhao K, Foster J, Ridgway ND. Oxysterol-binding protein-related protein 1 variants have opposing cholesterol transport activities from the endolysosomes. *Mol Biol Cell*. 2020;31(8):793-802. doi:10.1091/mbc.E19-12-0697

83. Eden ER, Sanchez-Heras E, Tsapara A, Sobota A, Levine TP, Futter CE. Annexin A1 Tethers Membrane Contact Sites that Mediate ER to Endosome Cholesterol Transport. *Dev Cell*. 2016;37(5):473-483. doi:https://doi.org/10.1016/j.devcel.2016.05.005

84. Laitinen S, Lehto M, Lehtonen S, et al. ORP2, a homolog of oxysterol binding protein, regulates cellular cholesterol metabolism. *J Lipid Res*. 2002;43(2):245-255. doi:https://doi.org/10.1016/S0022-2275(20)30166-8

85. Liu X, Ridgway ND. Characterization of the Sterol and Phosphatidylinositol 4-Phosphate Binding Properties of Golgi-Associated OSBP-Related Protein 9 (ORP9). *PLoS One*. 2014;9(9):e108368. https://doi.org/10.1371/journal.pone.0108368

86. Wang P, Weng J, Anderson RGW. OSBP Is a Cholesterol-Regulated Scaffolding Protein in Control of ERK1/2 Activation. *Science (80-)*. 2005;307(5714):1472-1476. doi:10.1126/science.1107710

87. Clark BJ. The mammalian START domain protein family in lipid transport in health and disease. *J Endocrinol*. 2012;212(3):257-275. doi:10.1530/JOE-11-0313

88. Alpy F, Tomasetto C. Give lipids a START: the StAR-related lipid transfer (START) domain in mammals. *J Cell Sci*. 2005;118(13):2791-2801. doi:10.1242/jcs.02485

89. Han S, Cohen DE. Functional characterization of thioesterase superfamily member 1/Acyl-CoA thioesterase 11: implications for metabolic regulation[S]. *J Lipid Res*. 2012;53(12):2620-2631. doi:https://doi.org/10.1194/jlr.M029538

90. Torres JZ. STARD9/Kif16a is a novel mitotic kinesin and antimitotic target. *Bioarchitecture*. 2012;2(1):19-22. doi:10.4161/bioa.19766

91. Clark BJ. The START-domain proteins in intracellular lipid transport and beyond. *Mol Cell Endocrinol*. 2020;504:110704. doi:https://doi.org/10.1016/j.mce.2020.110704

92. Alpy F, Tomasetto C. START ships lipids across interorganelle space. *Biochimie*. 2014;96:85-95. doi:https://doi.org/10.1016/j.biochi.2013.09.015

93. Stocco DM. StAR Protein and the Regulation of Steroid Hormone Biosynthesis. *Annu Rev*

Physiol. 2001;63(1):193-213. doi:10.1146/annurev.physiol.63.1.193

94.   Gomes C, Oh S-D, Kim J-W, et al. Expression of the Putative Sterol Binding Protein Stard6 Gene Is Male Germ Cell Specific1. *Biol Reprod*. 2005;72(3):651-658. doi:10.1095/biolreprod.104.032672

95.   Létourneau D, Bédard M, Cabana J, Lefebvre A, LeHoux J-G, Lavigne P. STARD6 on steroids: solution structure, multiple timescale backbone dynamics and ligand binding mechanism. *Sci Rep*. 2016;6(1):28486. doi:10.1038/srep28486

96.   Sandhu J, Li S, Fairall L, et al. Aster Proteins Facilitate Nonvesicular Plasma Membrane to ER Cholesterol Transport in Mammalian Cells. *Cell*. 2018;175(2):514-529.e20. doi:https://doi.org/10.1016/j.cell.2018.08.033

97.   Laraia L, Friese A, Corkery DP, et al. The cholesterol transfer protein GRAMD1A regulates autophagosome biogenesis. *Nat Chem Biol*. 2019;15(7):710-720. doi:10.1038/s41589-019-0307-5

98.   Janowski BA, Grogan MJ, Jones SA, et al. Structural requirements of ligands for the oxysterol liver X receptors LXRα and LXRβ. *Proc Natl Acad Sci*. 1999;96(1):266-271. doi:10.1073/pnas.96.1.266

99.   Nachtergaele S, Mydock LK, Krishnan K, et al. Oxysterols are allosteric activators of the oncoprotein Smoothened. *Nat Chem Biol*. 2012;8(2):211-220. doi:10.1038/nchembio.765

100.   Adam GC, Sorensen EJ, Cravatt BF. Chemical Strategies for Functional Proteomics*. *Mol Cell Proteomics*. 2002;1(10):781-790. doi:https://doi.org/10.1074/mcp.R200006-MCP200

101.   Evans MJ, Cravatt BF. Mechanism-Based Profiling of Enzyme Families. *Chem Rev*. 2006;106(8):3279-3301. doi:10.1021/cr050288g

102.   Cravatt BF, Wright AT, Kozarich JW. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu Rev Biochem*. 2008;77(1):383-414. doi:10.1146/annurev.biochem.75.101304.124125

103.   Guo H, Li Z. Developments of bioorthogonal handle-containing photo-crosslinkers for photoaffinity labeling. *Medchemcomm*. 2017;8(8):1585-1591. doi:10.1039/C7MD00217C

104.   Sieber SA, Niessen S, Hoover HS, Cravatt BF. Proteomic profiling of metalloprotease activities with cocktails of active-site probes. *Nat Chem Biol*. 2006;2(5):274-281. doi:10.1038/nchembio781

105.   Berger AB, Witte MD, Denault J-B, et al. Identification of Early Intermediates of Caspase Activation Using Selective Inhibitors and Activity-Based Probes. *Mol Cell*. 2006;23(4):509-521. doi:https://doi.org/10.1016/j.molcel.2006.06.021

106.   Mahrus S, Craik CS. Selective Chemical Functional Probes of Granzymes A and B Reveal Granzyme B Is a Major Effector of Natural Killer Cell-Mediated Lysis of Target Cells. *Chem Biol*. 2005;12(5):567-577. doi:https://doi.org/10.1016/j.chembiol.2005.03.006

107.   Ragaini F, Penoni A, Gallo E, et al. Amination of Benzylic C   H Bonds by Arylazides Catalyzed by CoII–Porphyrin Complexes: A Synthetic and Mechanistic Study. *Chem – A Eur J*. 2003;9(1):249-259. doi:https://doi.org/10.1002/chem.200390018

108.   Chee G-L, Yalowich JC, Bodner A, Wu X, Hasinoff BB. A diazirine-based photoaffinity etoposide probe for labeling topoisomerase II. *Bioorg Med Chem*. 2010;18(2):830-838. doi:https://doi.org/10.1016/j.bmc.2009.11.048

109.   Saghatelian A, Jessani N, Joseph A, Humphrey M, Cravatt BF. Activity-based probes for the

proteomic profiling of metalloproteases. *Proc Natl Acad Sci*. 2004;101(27):10000-10005. doi:10.1073/pnas.0402784101

110.   Husaini AM, Morimoto K, Chandrasekar B, et al. Multiplex Fluorescent, Activity-Based Protein Profiling Identifies Active α-Glycosidases and Other Hydrolases in Plants. *Plant Physiol*. 2018;177(1):24-37. doi:10.1104/pp.18.00250

111.   Wang Z, Guo Z, Song T, et al. Proteome-Wide Identification of On- and Off-Targets of Bcl-2 Inhibitors in Native Biological Systems by Using Affinity-Based Probes (AfBPs). *ChemBioChem*. 2018;19(21):2312-2320. doi:https://doi.org/10.1002/cbic.201800380

112.   Sletten EM, Bertozzi CR. From Mechanism to Mouse: A Tale of Two Bioorthogonal Reactions. *Acc Chem Res*. 2011;44(9):666-676. doi:10.1021/ar200148z

113.   Lang K, Chin JW. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem Rev*. 2014;114(9):4764-4806. doi:10.1021/cr400355w

114.   Rostovtsev V V, Green LG, Fokin V V, Sharpless KB. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew Chemie Int Ed*. 2002;41(14):2596-2599. doi:https://doi.org/10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4

115.   Fang H, Peng B, Ong SY, Wu Q, Li L, Yao SQ. Recent advances in activity-based probes (ABPs) and affinity-based probes (AfBPs) for profiling of enzymes. *Chem Sci*. 2021;12(24):8288-8310. doi:10.1039/D1SC01359A

116.   Lin Z, Wang X, Bustin KA, et al. Activity-Based Hydrazine Probes for Protein Profiling of Electrophilic Functionality in Therapeutic Targets. *ACS Cent Sci*. 2021;7(9):1524-1534. doi:10.1021/acscentsci.1c00616

117.   Qian L, Pan S, Lee J-S, Ge J, Li L, Yao SQ. Live-cell imaging and profiling of c-Jun N-terminal kinases using covalent inhibitor-derived probes. *Chem Commun*. 2019;55(8):1092-1095. doi:10.1039/C8CC09558B

118.   Dai L, Prabhu N, Yu LY, Bacanu S, Ramos AD, 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(1):383-408. doi:10.1146/annurev-biochem-062917-012837

119.   Rossetti C, Laraia L. Thermal Proteome Profiling Reveals Distinct Target Selectivity for Differentially Oxidized Oxysterols. *ACS Chem Biol*. 2022;17(7):1677-1684. doi:10.1021/acschembio.2c00383

120.   Pennisi E. 'Cell painting' highlights responses to drugs and toxins. *Science (80-)*. 2016;352(6288):877-878. doi:10.1126/science.352.6288.877

121.   Bray M-A, Singh S, Han H, et al. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat Protoc*. 2016;11(9):1757-1774. doi:10.1038/nprot.2016.105

122.   Chandrasekaran SN, Ceulemans H, Boyd JD, Carpenter AE. Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nat Rev Drug Discov*. 2021;20(2):145-159. doi:10.1038/s41573-020-00117-w

123.   Hulce JJ, Cognetta AB, Niphakis MJ, Tully SE, Cravatt BF. Proteome-wide mapping of cholesterol-interacting proteins in mammalian cells. *Nat Methods*. 2013;10(3):259-264.

doi:10.1038/nmeth.2368

124. Iida T, Kakiyama G, Hibiya Y, et al. Chemical synthesis of the 3-sulfooxy-7-N-acetylglucosaminyl-24-amidated conjugates of 3β,7β-dihydroxy-5-cholen-24-oic acid, and related compounds: Unusual, major metabolites of bile acid in a patient with Niemann-Pick disease type C1. *Steroids*. 2006;71(1):18-29. doi:10.1016/j.steroids.2005.07.008

125. Sepe V, Renga B, Festa C, et al. Investigation on bile acid receptor regulators. Discovery of cholanoic acid derivatives with dual G-protein coupled bile acid receptor 1 (GPBAR1) antagonistic and farnesoid X receptor (FXR) modulatory activity. *Steroids*. 2016;105:59-67. doi:10.1016/j.steroids.2015.11.003

126. Chaurand P, Friedman DB, Caprioli RM. Chapter 22 - Mass Spectrometry in Cancer Biology. In: Mendelsohn J, Howley PM, Israel MA, Gray JW, Thompson CBBT-TMB of C (Third E, eds. W.B. Saunders; 2008:293-307. doi:https://doi.org/10.1016/B978-141603703-3.10022-6

127. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11(12):2301-2319. doi:10.1038/nprot.2016.136

128. Chahrour O, Cobice D, Malone J. Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. *J Pharm Biomed Anal*. 2015;113:2-20. doi:https://doi.org/10.1016/j.jpba.2015.04.013

129. Yu F, Qiu F, Meza J. 12 - Design and Statistical Analysis of Mass-Spectrometry-Based Quantitative Proteomics Data. In: Ciborowski P, Silberring JBT-PP and AC (Second E, eds. Elsevier; 2016:211-237. doi:https://doi.org/10.1016/B978-0-444-63688-1.00012-4

130. Zhu W, Smith JW, Huang C-M. Mass Spectrometry-Based Label-Free Quantitative Proteomics. Dorrestein PC, ed. *J Biomed Biotechnol*. 2010;2010:840518. doi:10.1155/2010/840518

131. Pappireddi N, Martin L, Wühr M. A Review on Quantitative Multiplexed Proteomics. *ChemBioChem*. 2019;20(10):1210-1224. doi:https://doi.org/10.1002/cbic.201800650

132. Pettersson M, Crews CM. PROteolysis TArgeting Chimeras (PROTACs) — Past, present and future. *Drug Discov Today Technol*. 2019;31:15-27. doi:https://doi.org/10.1016/j.ddtec.2019.01.002

133. Kawatkar A, Schefter M, Hermansson N-O, et al. CETSA beyond Soluble Targets: a Broad Application to Multipass Transmembrane Proteins. *ACS Chem Biol*. 2019;14(9):1913-1920. doi:10.1021/acschembio.9b00399

134. Gao S, Wang S, Song Y. Novel immunomodulatory drugs and neo-substrates. *Biomark Res*. 2020;8(1):2. doi:10.1186/s40364-020-0182-y

135. Yang J, Li Y, Aguilar A, Liu Z, Yang C-Y, Wang S. Simple Structural Modifications Converting a Bona fide MDM2 PROTAC Degrader into a Molecular Glue Molecule: A Cautionary Tale in the Design of PROTAC Degraders. *J Med Chem*. 2019;62(21):9471-9487. doi:10.1021/acs.jmedchem.9b00846

136. Linstedt AD, Mehta A, Suhan J, Reggio H, Hauri HP. Sequence and overexpression of GPP130/GIMPCc: Evidence for saturable pH- sensitive targeting of a type II early golgi membrane protein. *Mol Biol Cell*. 1997;8(6):1073-1087. doi:10.1091/mbc.8.6.1073

137. Bachert C, Lee TH, Linstedt AD. Lumenal endosomal and Golgi-retrieval determinants involved in pH-sensitive targeting of an early Golgi protein. *Mol Biol Cell*. 2001;12(10):3152-3160. doi:10.1091/mbc.12.10.3152

138. Natarajan R, Linstedt AD. A Cycling cis-Golgi Protein Mediates Endosome-to-Golgi Traffic. *Mol Biol Cell*. 2004;15(11):4798-4806. doi:10.1091/mbc.e04-05-0366

139. Puri S, Bachert C, Fimmel CJ, Linstedt AD. Cycling of early Golgi proteins via the cell surface and endosomes upon lumenal pH disruption. *Traffic*. 2002;3(9):641-653. doi:10.1034/j.1600-0854.2002.30906.x

140. Mukhopadhyay S, Linstedt AD. Manganese Blocks Intracellular Trafficking of Shiga Toxin and Protects Against Shiga Toxicosis. *Science (80-)*. 2012;335(6066):332-335. doi:10.1126/science.1215930

141. Mukhopadhyay S, Bachert C, Smith DR, Linstedt AD. Manganese-induced Trafficking and Turnover of the cis-Golgi Glycoprotein GPP130. *Mol Biol Cell*. 2010;21(7):1282-1292. doi:10.1091/mbc.e09-11-0985

142. Tewari R, Jarvela T, Linstedt AD. Manganese induces oligomerization to promote down-regulation of the intracellular trafficking receptor used by Shiga toxin. *Mol Biol Cell*. 2014;25(19):3049-3058. doi:10.1091/mbc.E14-05-1003

143. Tewari R, Bachert C, Linstedt AD. Induced oligomerization targets Golgi proteins for degradation in lysosomes. *Mol Biol Cell*. 2015;26(24):4427-4437. doi:10.1091/mbc.E15-04-0207

144. Venkat S, Linstedt AD. Manganese-induced trafficking and turnover of GPP130 is mediated by sortilin. *Mol Biol Cell*. 2017;28(19):2569-2578. doi:10.1091/mbc.E17-05-0326

145. P. NJ, B. KJ. Diarrheagenic Escherichia coli. *Clin Microbiol Rev*. 1998;11(1):142-201. doi:10.1128/CMR.11.1.142

146. Mukhopadhyay S, Redler B, Linstedt AD. Shiga toxin-binding site for host cell receptor GPP130 reveals unexpected divergence in toxin-trafficking mechanisms. *Mol Biol Cell*. 2013;24(15):2311-2318. doi:10.1091/mbc.E13-01-0057

147. Fraser ME, Fujinaga M, Cherney MM, et al. Structure of Shiga Toxin Type 2 (Stx2) from Escherichia coli O157:H7*. *J Biol Chem*. 2004;279(26):27511-27517. doi:https://doi.org/10.1074/jbc.M401939200

148. Masuda M, Braun-sommargren M, Crooks D, et al. Cellular Target of Manganese Exposure. 2014;67(5):205-215. doi:10.1002/syn.21632.Golgi

149. Bai Y, Cui X, Gao D, Wang Y, Wang B, Wang W. Golgi integral membrane protein 4 manipulates cellular proliferation, apoptosis, and cell cycle in human head and neck cancer. *Biosci Rep*. 2018;38(4):1-9. doi:10.1042/BSR20180454

150. Mahmood T, Yang P-C. Western blot: Technique, theory, and trouble shooting. *N Am J Med Sci*. 2012;4(9):429-434. doi:10.4103/1947-2714.100998

151. Li Z, Pinch BJ, Olson CM, et al. Development and Characterization of a Wee1 Kinase Degrader. *Cell Chem Biol*. 2020;27(1):57-65.e9. doi:10.1016/j.chembiol.2019.10.013

152. Soucy TA, Smith PG, Milhollen MA, et al. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature*. 2009;458(7239):732-736. doi:10.1038/nature07884

153. Oka T, Ungar D, Hughson FM, Krieger M. The COG and COPI Complexes Interact to Control the Abundance of GEARs, a Subset of Golgi Integral Membrane Proteins. *Mol Biol Cell*. 2004;15(5):2423-2435. doi:10.1091/mbc.e03-09-0699

154. Mauthe M, Orhon I, Rocchi C, et al. Chloroquine inhibits autophagic flux by decreasing

autophagosome-lysosome fusion. *Autophagy*. 2018;14(8):1435-1455. doi:10.1080/15548627.2018.1474314

155.   Whitmarsh-Everiss T, Laraia L. Small molecule probes for targeting autophagy. *Nat Chem Biol*. 2021;17(6):653-664. doi:10.1038/s41589-021-00768-9

156.   Tang J. Acid proteases come of age. *Nature*. 1977;266(5598):119-120. doi:10.1038/266119a0

157.   Pasquier B. SAR405, a PIK3C3/Vps34 inhibitor that prevents autophagy and synergizes with MTOR inhibition in tumor cells. *Autophagy*. 2015;11(4):725-726. doi:10.1080/15548627.2015.1033601

158.   Schrøder TJ, Christensen S, Lindberg S, et al. The identification of AF38469: An orally bioavailable inhibitor of the VPS10P family sorting receptor Sortilin. *Bioorganic Med Chem Lett*. 2014;24(1):177-180. doi:10.1016/j.bmcl.2013.11.046

159.   Quistgaard EM, Grøftehauge MK, Madsen P, et al. Revisiting the structure of the Vps10 domain of human sortilin and its interaction with neurotensin. *Protein Sci*. 2014;23(9):1291-1300. doi:10.1002/pro.2512

160.   Wang T-E, Li S-H, Minabe S, et al. Mouse quiescin sulfhydryl oxidases exhibit distinct epididymal luminal distribution with segment-specific sperm surface associations†. *Biol Reprod*. 2018;99(5):1022-1033. doi:10.1093/biolre/ioy125

161.   Adolf F, Rhiel M, Hessling B, et al. Proteomic Profiling of Mammalian COPII and COPI Vesicles. *Cell Rep*. 2019;26(1):250-265.e5. doi:https://doi.org/10.1016/j.celrep.2018.12.041

162.   Gaynor EC, Graham TR, Emr SD. COPI in ER/Golgi and intra-Golgi transport: do yeast COPI mutants point the way? *Biochim Biophys Acta - Mol Cell Res*. 1998;1404(1):33-51. doi:https://doi.org/10.1016/S0167-4889(98)00045-7

163.   Lowe M. Structural organization of the Golgi apparatus. *Curr Opin Cell Biol*. 2011;23(1):85-93. doi:https://doi.org/10.1016/j.ceb.2010.10.004

164.   Appenzeller-Herzog C, Hauri H-P. The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. *J Cell Sci*. 2006;119(11):2173-2183. doi:10.1242/jcs.03019

165.   Ma M, Burd CG. Retrograde trafficking and plasma membrane recycling pathways of the budding yeast Saccharomyces cerevisiae. *Traffic*. 2020;21(1):45-59. doi:https://doi.org/10.1111/tra.12693

166.   Zhang X, Wang Y. Glycosylation Quality Control by the Golgi Structure. *J Mol Biol*. 2016;428(16):3183-3193. doi:https://doi.org/10.1016/j.jmb.2016.02.030

167.   Frappaolo A, Karimpour-Ghahnavieh A, Sechi S, Giansanti MG. The Close Relationship between the Golgi Trafficking Machinery and Protein Glycosylation. *Cells*. 2020;9(12). doi:10.3390/cells9122652

168.   Yu I-M, Hughson FM. Tethering Factors as Organizers of Intracellular Vesicular Traffic. *Annu Rev Cell Dev Biol*. 2010;26(1):137-156. doi:10.1146/annurev.cellbio.042308.113327

169.   Ungar D, Oka T, Brittle EE, et al. Characterization of a mammalian Golgi-localized protein complex, COG, that is required for normal Golgi morphology and function. *J Cell Biol*. 2002;157(3):405-415. doi:10.1083/jcb.200202016

170.   Blackburn JB, D'Souza Z, Lupashin V V. Maintaining order: COG complex controls Golgi trafficking, processing, and sorting. *FEBS Lett*. 2019;593(17):2466-2487. doi:10.1002/1873-3468.13570

171. Oka T, Vasile E, Penman M, et al. Genetic Analysis of the Subunit Organization and Function of the Conserved Oligomeric Golgi (COG) Complex: STUDIES OF COG5- AND COG7-DEFICIENT MAMMALIAN CELLS*. *J Biol Chem*. 2005;280(38):32736-32745. doi:https://doi.org/10.1074/jbc.M505558200

172. Linders PTA, Peters E, ter Beest M, Lefeber DJ, van den Bogaart G. Sugary Logistics Gone Wrong: Membrane Trafficking and Congenital Disorders of Glycosylation. *Int J Mol Sci*. 2020;21(13). doi:10.3390/ijms21134654

173. Smith RD, Willett R, Kudlyk T, et al. The COG Complex, Rab6 and COPI Define a Novel Golgi Retrograde Trafficking Pathway that is Exploited by SubAB Toxin. *Traffic*. 2009;10(10):1502-1517. doi:https://doi.org/10.1111/j.1600-0854.2009.00965.x

174. Laufman O, Hong W, Lev S. The COG complex interacts with multiple Golgi SNAREs and enhances fusogenic assembly of SNARE complexes. *J Cell Sci*. 2013;126(6):1506-1516. doi:10.1242/jcs.122101

175. Smith RD, Lupashin V V. Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. *Carbohydr Res*. 2008;343(12):2024-2031. doi:https://doi.org/10.1016/j.carres.2008.01.034

176. Pokrovskaya ID, Willett R, Smith RD, Morelle W, Kudlyk T, Lupashin V V. Conserved oligomeric Golgi complex specifically regulates the maintenance of Golgi glycosylation machinery. *Glycobiology*. 2011;21(12):1554-1569. doi:10.1093/glycob/cwr028

177. Blackburn JB, Kudlyk T, Pokrovskaya I, Lupashin V V. More than just sugars: Conserved oligomeric Golgi complex deficiency causes glycosylation-independent cellular defects. *Traffic*. 2018;19(6):463-480. doi:https://doi.org/10.1111/tra.12564

178. Helfand BT, Mendez MG, Pugh J, Delsert C, Goldman RD. Maintaining the Shape of Nerve Cells □. *Mol Biol Cell*. 2003;14(December):5069-5081. doi:10.1091/mbc.E03

179. Zolov SN, Lupashin V V. Cog3p depletion blocks vesicle-mediated Golgi retrograde trafficking in HeLa cells. *J Cell Biol*. 2005;168(5):747-759. doi:10.1083/jcb.200412003

180. Sumya FT, Pokrovskaya ID, D'Souza Z, Lupashin V V. Acute COG inactivation unveiled its immediate impact on Golgi and illuminated the nature of intra-Golgi recycling vesicles. *bioRxiv*. Published online January 1, 2022:2022.05.24.493317. doi:10.1101/2022.05.24.493317

181. Khakurel A, Kudlyk T, Bonifacino JS, Lupashin V V. The Golgi-associated retrograde protein (GARP) complex plays an essential role in the maintenance of the Golgi glycosylation machinery. *Mol Biol Cell*. 2021;32(17):1594-1610. doi:10.1091/mbc.E21-04-0169

182. Khakurel A, Kudlyk T, Pokrovskaya I, D'Souza Z, Lupashin V V. GARP complex controls Golgi physiology by stabilizing COPI machinery and Golgi v-SNAREs. *bioRxiv*. Published online January 1, 2022:2022.07.22.501184. doi:10.1101/2022.07.22.501184

183. Elbein AD. The Use of Glycosylation Inhibitors to Study Glycoconjugate Function. In: ROBERTS DD, MECHAM RPBT-CS and EG, eds. *Biology of Extracellular Matrix*. Academic Press; 1993:119-180. doi:https://doi.org/10.1016/B978-0-12-589630-6.50009-5

184. Norris GE, Stillman TJ, Anderson BF, Baker EN. The three-dimensional structure of PNGase F, a glycosyl asparaginase from Flavobacterium meningosepticum. *Structure*. 1994;2(11):1049-1059. doi:https://doi.org/10.1016/S0969-2126(94)00108-1

185. Bagdonaite I, Malaker SA, Polasky DA, et al. Glycoproteomics. *Nat Rev Methods Prim*.

2022;2(1):48. doi:10.1038/s43586-022-00128-4

186. Louche A, Salcedo SP, Bigot S. Protein–Protein Interactions: Pull-Down Assays BT - Bacterial Protein Secretion Systems: Methods and Protocols. In: Journet L, Cascales E, eds. Springer New York; 2017:247-255. doi:10.1007/978-1-4939-7033-9_20

187. Raychaudhuri S, Prinz WA. The Diverse Functions of Oxysterol-Binding Proteins. *Annu Rev Cell Dev Biol*. 2010;26(1):157-177. doi:10.1146/annurev.cellbio.042308.113334

188. Roberts BL, Severance ZC, Bensen RC, et al. Transient Compound Treatment Induces a Multigenerational Reduction of Oxysterol-Binding Protein (OSBP) Levels and Prophylactic Antiviral Activity. *ACS Chem Biol*. 2019;14(2):276-287. doi:10.1021/acschembio.8b00984

189. Taskinen JH, Ruhanen H, Matysik S, Käkelä R, Olkkonen VM. Global effects of pharmacologic inhibition of OSBP in human umbilical vein endothelial cells. *Steroids*. 2022;185:109053. doi:https://doi.org/10.1016/j.steroids.2022.109053

190. Péresse T, Kovacs D, Subra M, et al. Molecular and cellular dissection of the oxysterol-binding protein cycle through a fluorescent inhibitor. *J Biol Chem*. 2020;295(13):4277-4288. doi:10.1074/jbc.RA119.012012

191. Ercan B, Naito T, Koh DHZ, Dharmawan D, Saheki Y. Molecular basis of accessible plasma membrane cholesterol recognition by the GRAM domain of GRAMD1b. *EMBO J*. 2021;40(6):e106524. doi:https://doi.org/10.15252/embj.2020106524

192. Andersen JP, Zhang J, Sun H, et al. Aster-B coordinates with Arf1 to regulate mitochondrial cholesterol transport. *Mol Metab*. 2020;42(July):101055. doi:10.1016/j.molmet.2020.101055

193. Wui Ng MY, Charsou C, Lapao A, et al. GRAMD1C regulates autophagy initiation and mitochondrial bioenergetics through ER-mitochondria cholesterol transport. *bioRxiv*. Published online January 1, 2021:2021.08.04.455032. doi:10.1101/2021.08.04.455032

194. Xiao X, Kim Y, Romartinez-Alonso B, et al. Selective Aster inhibitors distinguish vesicular and nonvesicular sterol transport mechanisms. *Proc Natl Acad Sci*. 2021;118(2):e2024149118. doi:10.1073/pnas.2024149118

195. Whitmarsh-Everiss T, Olsen AH, Laraia L. Identification of Inhibitors of Cholesterol Transport Proteins Through the Synthesis of a Diverse, Sterol-Inspired Compound Collection. *Angew Chemie Int Ed*. 2021;60(51):26755-26761. doi:https://doi.org/10.1002/anie.202111639

196. Yang Y. Chapter 5 - Side Reactions Upon Amino Acid/Peptide Carboxyl Activation. In: Yang YBT-SR in PS, ed. Academic Press; 2016:95-118. doi:https://doi.org/10.1016/B978-0-12-801009-9.00005-7

197. Albericio F, Bofill JM, El-Faham A, Kates SA. Use of onium salt-based coupling reagents in peptide synthesis. *J Org Chem*. 1998;63(26):9678-9683. doi:10.1021/jo980807y

198. Chan KH, Zengerle M, Testa A, Ciulli A. Impact of Target Warhead and Linkage Vector on Inducing Protein Degradation: Comparison of Bromodomain and Extra-Terminal (BET) Degraders Derived from Triazolodiazepine (JQ1) and Tetrahydroquinoline (I-BET726) BET Inhibitor Scaffolds. *J Med Chem*. 2018;61(2):504-513. doi:10.1021/acs.jmedchem.6b01912

199. Lai AC, Toure M, Hellerschmied D, et al. Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angew Chemie - Int Ed*. 2016;55(2):807-810. doi:10.1002/anie.201507634

200. Galdeano C, Gadd MS, Soares P, et al. Structure-Guided Design and Optimization of Small

Molecules Targeting the Protein–Protein Interaction between the von Hippel–Lindau (VHL) E3 Ubiquitin Ligase and the Hypoxia Inducible Factor (HIF) Alpha Subunit with in Vitro Nanomolar Affinities. *J Med Chem*. 2014;57(20):8657-8663. doi:10.1021/jm5011258

201. Steinebach C, Voell SA, Vu LP, et al. A Facile Synthesis of Ligands for the von Hippel-Lindau E3 Ligase. *Synth*. 2020;52(17):2521-2527. doi:10.1055/s-0040-1707400

202. A. Bemis T, J. La Clair J, D. Burkart M. Unraveling the Role of Linker Design in Proteolysis Targeting Chimeras. *J Med Chem*. 2021;64(12):8042-8052. doi:10.1021/acs.jmedchem.1c00482

203. Luo X, Archibeque I, Dellamaggiore K, et al. Profiling of diverse tumor types establishes the broad utility of VHL-based ProTaCs and triages candidate ubiquitin ligases. *iScience*. 2022;25(3):103985. doi:10.1016/j.isci.2022.103985

204. Petzold G, Fischer ES, Thomä NH. Structural basis of lenalidomide-induced CK1α degradation by the CRL4 CRBN ubiquitin ligase. *Nature*. 2016;532(7597):127-130. doi:10.1038/nature16979

205. Jin C, Wang Y, Sun B, Su W. A Concise Synthesis of 25-Hydroxycholesterol from Hyodesoxycholic Acid. *J Chem Res*. 2018;42(2):96-99. doi:10.3184/174751918X15184426915885

206. Iida T, Kakiyama G, Hibiya Y, et al. Chemical synthesis of the 3-sulfooxy-7-N-acetylglucosaminyl-24-amidated conjugates of 3β,7β-dihydroxy-5-cholen-24-oic acid, and related compounds: Unusual, major metabolites of bile acid in a patient with Niemann-Pick disease type C1. *Steroids*. 2006;71(1):18-29. doi:https://doi.org/10.1016/j.steroids.2005.07.008

207. Sepe V, Di Leva FS, D'Amore C, et al. Marine and Semi-Synthetic Hydroxysteroids as New Scaffolds for Pregnane X Receptor Modulation. *Mar Drugs*. 2014;12(6):3091-3115. doi:10.3390/md12063091

208. Zhang J, Che J, Luo X, et al. Structural Feature Analyzation Strategies toward Discovery of Orally Bioavailable PROTACs of Bruton's Tyrosine Kinase for the Treatment of Lymphoma. *J Med Chem*. 2022;65(13):9096-9125. doi:10.1021/acs.jmedchem.2c00324

209. Li Z, Pinch BJ, Olson CM, et al. Development and Characterization of a Wee1 Kinase Degrader. *Cell Chem Biol*. 2020;27(1):57-65.e9. doi:https://doi.org/10.1016/j.chembiol.2019.10.013

210. Steinebach C, Sosič I, Lindner S, et al. A MedChem toolbox for cereblon-directed PROTACs. *Medchemcomm*. 2019;10(6):1037-1041. doi:10.1039/C9MD00185A

211. Gruber HJ, Marek M, Schindler H, Kaiser K. Biotin−Fluorophore Conjugates with Poly(ethylene glycol) Spacers Retain Intense Fluorescence after Binding to Avidin and Streptavidin. *Bioconjug Chem*. 1997;8(4):552-559. doi:10.1021/bc970087m

212. Dölle A, Adhikari B, Krämer A, et al. Design, Synthesis, and Evaluation of WD-Repeat-Containing Protein 5 (WDR5) Degraders. *J Med Chem*. 2021;64(15):10682-10710. doi:10.1021/acs.jmedchem.1c00146

213. Zhao H-Y, Wang H-P, Mao Y-Z, et al. Discovery of Potent PROTACs Targeting EGFR Mutants through the Optimization of Covalent EGFR Ligands. *J Med Chem*. 2022;65(6):4709-4726. doi:10.1021/acs.jmedchem.1c01827

214. Plubell DL, Wilmarth PA, Zhao Y, et al. Extended Multiplexing of Tandem Mass Tags (TMT) Labeling Reveals Age and High Fat Diet Specific Proteome Changes in Mouse Epididymal Adipose Tissue*. *Mol Cell Proteomics*. 2017;16(5):873-890. doi:https://doi.org/10.1074/mcp.M116.065524

215. Goedhart J, Luijsterburg MS. VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci Rep*. 2020;10(1):20560. doi:10.1038/s41598-020-76603-3

8 Appendix

Appendix 1. NMR spectra for final compounds.

Appendix 2. Proteomics data for **C1-C5**.

Appendix 3. Proteomics data for **H1-H5**.

Appendix 4. Proteomics data for **C3_Me**, **C3_Boc** and **C3_HyT**.